Data-driven model for analyzing the need for more affordable rental housing

Final report

by

Domenico Altieri & Giovanni Branca
domenico.altieri@supsi.ch - giovanni.branca@supsi.ch

ISAAC, University of Applied Sciences and Arts of Southern Switzerland, Mendrisio, Switzerland, CH

Submitted to the Federal Housing Office, Hallwylstrasse 4
CH-3003 Berna
info@bwo.admin.ch

Contents

1. Introduction

Nowadays, there is a growing need for approaches capable of analyzing the local rental market, identifying possible needs for more affordable rents. The housing market is not always ready to respond adequately to changes in demand [1, 2, 3, 4] and a scarce availability of affordable housing for the most disadvantaged sections of the population not only affect individual households but also environmental factors such as employment levels, built environment use, etc [5].

In general, the social impact of a tool for calibrating sustainable rents should be identified primarily in the improvement of the quality of life of a specific segment of the population and in a mitigating effect on the relative poverty rate. Favoring a redistribution of rent by maximizing the absorptive capacity of the supply has as a secondary impact also a reduction of social inequalities.

The need of a dedicated analysis appears even more relevant in a situation like the one in Ticino, which in recent years has been characterized by a continuous increase in the share of vacant housing units in parallel with a marked phase of economic contraction [6, 7]. The important increase of non-rented dwellings and the need to promote the offer of lower rents require ad-hoc monitoring tools, able to model an extremely heterogeneous and volatile field of analysis. In this perspective, the main objective of the proposed computational approach is to offer a new analysis model focused on the regional level (or an aggregation of municipalities) and calibrated on the actual territorial distribution of incomes by type of household. The same result cannot be reached by simply adapting the currently available building-centric *hedonic price methods* (hereafter, HPM)[8] to the new scale of analysis and a major effort is therefore required especially in the admissible rent definition and in the data acquisition process. In fact, there is a lack of dedicated approaches aimed at implementing significant variables that take into account the actual households financial capacity, which to varying degrees is able to directly affect the estimate of optimal rents at a macro scale, larger than the single property.

The report is organized as follows: the first section is focused on a synthetic overview of the situation in Ticino, with reference to demographic and built environment aspects. Next, the reader is introduced to the proposed methodology for estimating the need for sustainable rents. More specifically, additional details are provided for the characterization of both the *Available* and *Admissible* rents required

to identify any areas of market imbalance. In the final part two additional computational approaches are introduced, which address other aspects of the phenomenon under investigation. The first one uses a Monte Carlo simulation on the basis of simulated samples in order to study the practised effort rate per location and household. The latter aims to quantify reliable sustainable rent levels by employing information from the available-admissible rents comparison. Finally, the possible development of approaches based on machine learning methods capable of explaining the contribution of different environmental variables (e.g.mobility) and performing predictive analyses is discussed.

2. Demography and built environment in Ticino

2.1. Demography

Analysis of demographic trends is an important part of the process of identifying potential demand for rental housing.

From 2010 onwards, the population of the Canton Ticino has followed a less regular evolution compared to the Swiss average and to other cantons taken as a reference. In particular, the population expanded strongly until 2015 (350'363), after which the growth slowed down rather markedly, reaching a total of 353'709 inhabitants in 2018. Moreover, if we further characterize this trend, it shows how the decrease in demographic growth has mainly affected the lower age ranges, thus favoring a progressive ageing of the population structure.

Looking at the evolution of the total balance it is possible to highlight that its strong decrease is mainly due to a significant reduction in the balance of *international migratory movements*. Leaving aside the balance of *intercantonal movements*, which has always been negative, the Canton of Ticino was able to attract population from abroad at least until 2013, after which this phenomenon gradually faded until it pushed the total balance into negative levels. International movements are an important issue for Ticino's demographic development on several fronts. It is clear how the phenomenon of cross-border commuting (increased by about 20% in 5 years ,2012-2017) influences in various ways the labour, integration and development policies of the Canton in the short and long term.

Finally, when analyzing the households size, compared to the Swiss average in Ticino there is a slightly higher incidence of households with one or three people. Considering, instead, the evolution in 6 years (2012-2018), the only significant increase (about 3%) has once again involved the number of one-person households,

with the remaining compositions basically maintaining a percentage either constant or with a slight decrease.

2.2. Buildings and housing

An introductory framework is provided to characterize the existing buildings stock in terms of supply, space, type and level of occupancy.
In 2018, the Ticino reality recorded a total of 111'867 residential buildings (1'748'477 in Switzerland) and 165'161 dwellings (3'784'230 in Switzerland), source BDS[1].
The building stock picture shows a higher availability of flats with 4 rooms, a trend confirmed to a lesser extent for 2 and 3 rooms, together with a relevant percentage of dwellings occupied by a single person. More specifically, a lower housing density (occupants/sq.m.) is showed if we focus on small to medium-sized flats, which represents an important aspect to account for towards a more sustainable occupancy to be encouraged in the promotion of low rents.
Another strategic aspect for the objective of the project is the analysis of the rate of empty proprieties. Since 1997, the total number of empty dwellings in Ticino has been in continuous decline until reaching a plateau in 2004 ($\approx 0.7\%$), which was confirmed until 2014, the year after which there was a marked increase in the percentage of empty dwellings, so much so that it far exceeded the national average in 2019 ($\approx$ 2.2%). It is therefore clear that in Ticino this phenomenon plays an important role in the definition of future housing strategies.
In terms of the age of the building stock, in 2018 buildings constructed before 1980 in Ticino accounted for 74% of the total, well above the 63% recorded in Switzerland. The age of the existing stock is an important factor in defining future projections of more affordable rents requirements and effective strategies that can be implemented immediately. A housing stock characterized mainly by old buildings will necessarily be affected by renovation/reconstruction works, which may be reflected in rent increases in different ways and to different degrees. A proportion of today's moderately priced housing stock (public or private) is therefore likely to fall out of this range within a few years.

2.3. Rental market

The analysis of the already rented proprieties (hereinafter referred as *practised rental market*), in relation

[1] `https://www.bfs.admin.ch/bfs/en/home/statistics/construction-housing/surveys/gws2009.html`

to the offered ones (hereinafter referred as *offered rental market*), plays an important role in identifying territorial realities with specific needs for lower rents. The characterization of the practised market is based on a three-year data pool (2015-2016-2017) obtained from the Structural Survey (SS)[2], while the data on the offered market were obtained by extrapolating the information of interest from real estate advertisement web platforms.
Some relevant aspects of the comparison between *offered* and *practised* housing market in Ticino are summarized below:

- Three-room apartments account for the largest share of available properties on the offered market, at around 40%, followed by two and four-room apartments, both at around 20%. On the other hand, three and four rooms apartments account in total for about 70% of the practised market. It can be seen in the offered market an increase in the percentage of smaller properties, mainly at the expense of four-room apartments, if compared to the practised market. This aspect, even if not yet validated on the basis of historical series, indicates a trend in the offered market towards a more sustainable occupation, aimed at responding to a demand in which small households predominate.

- A comparative analysis of the annual rents per square meter shows substantial differences between the offered and practised market. Smaller properties are the ones that have seen a significant increase in the offer price compared to the rental contracts of previous years, a delta that can even reach 100% in some cases. This is an important finding, especially when compared with the general trend in rents in Ticino, which shows substantially constant values in recent years (SS 2015-2017). In addition, the analysis of the available data shows that as the number of rooms increases, the rent gap (delta *offered market-practised market*) decreases sharply, settling at around $20 - 30\%$.

- In terms of living space, smaller dwellings in the offered market are further reducing in size (with respect the practised

[2] `https://www.bfs.admin.ch/bfs/en/home/statistics/population/surveys/se.html`

market), whereas, as the number of rooms increases, the net surfaces offered also show an upward trend. It is therefore the studios and large dwellings on offer that show the greatest differences in living space sizes compared to the practised rental market. The average interval of living space on offer varies from about 36 m^2 for a studio to 176 m^2 for a large apartment (+6 rooms), while the range is much narrower for the already occupied properties [48 - 124 m^2].

3. Rental market analysis

The proposed methodology aims to identify the realities of the Ticino rental market that fail to meet the demand for lower rents by the most disadvantaged segments of the population. This was obtained after a data collection process to characterize the rental market together with the distribution of household income for a given location and size (number of persons), thus assuming income data as a descriptive parameter of the economic capacity of an inhabited area. The ability to model in a direct way the relationship 'rents - households economic power', would ensure that the distribution of rents offered is calibrated so that it can be more likely absorbed by the local demand.

In Figure 1 inputs/outputs of the analysis model are identified, while the entire process of defining a distribution curve for available and admissible rents, as well as the comparison methodology for the analysis of the rental market, will be detailed later on.

As shown in Figure 1, in addition to the main data on rents and income, results be declined according to location and household size. This turns out to be a relevant aspect from the point of view of a sustainable occupation of the living space, capable of guaranteeing minimum surfaces but still adequate to avoid the attainment of low rents by means of too high housing densities. The possibility of rapidly varying surfaces, as well as the maximum admissible effort rate, but also the type of rent (net or gross), offers flexibility of analysis to study multiple scenarios of interest.

In summary, the proposed methodology aims to achieve the following three main objectives:

- Directly compare *admissible rents* with *available rents* by evaluating multiple scenarios of housing density, number of people and effort rate. The results obtained manage to take into account the rental market offered in its entirety, incorporating both occupied and currently vacant housing on the market,

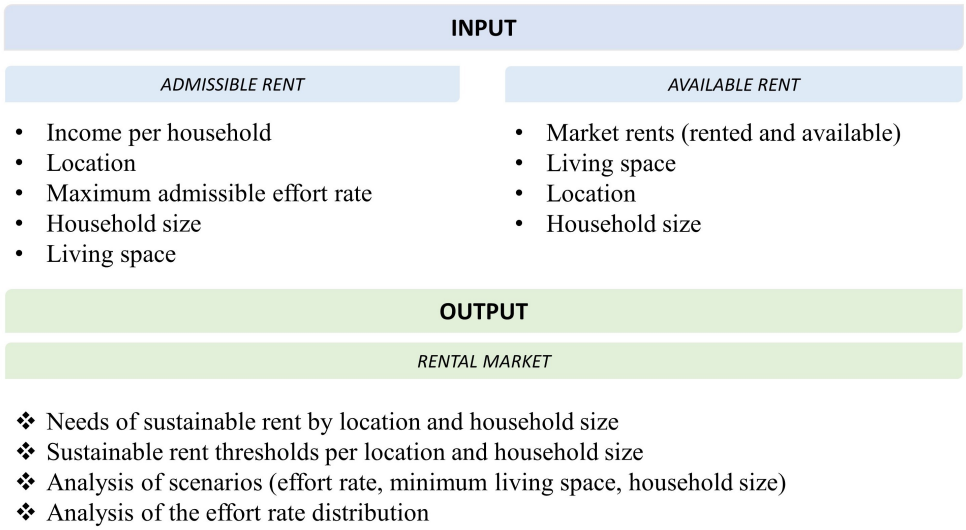

Figure 1: Input and output.

- A second analysis aims to identify sustainable rent levels based on market imbalance areas.

- Finally, in the last part, the practised effort rate is analyzed, without assumption on the housing density of occupancy. This objective is achieved using a statistical model to generate artificial samples that simulate likely associations between *income* and *rent paid* in function of location and household size.

3.1. Available rents

3.1.1. Data source

The available rents are described by means of a probabilistic distribution calibrated on the occurrences of rents in the rental market practised (apartment already occupied) and offered (empty apartments on the market). Therefore, the need arises to identify two different data sources, one to characterize the occurrences of rental contracts already in existence and one to analyze the rents offered for lodgings on the market that are currently vacant. As already mentioned, this was possible by using both the Structural Survey and online dedicate platforms (e.g. Homegate.ch, Comparis.ch). The analysis of rents is by its nature more immediate and direct than working with income data, which instead requires intermediate steps to be transformed into an *admissible rent*. In detail, the characterization of the Ticino's

practised rental market is based on a total of 22'632 observations, while the dataset on rental properties available on the market reaches a total of 21'104 observations recorded from 1/3/2019 to 1/3/2020. Both datasets underwent careful data processing analysis to ensure the use of reliable and consistent information.

3.1.2. Data analysis

The workflow required to define a probabilistic distribution to fit the offered and practised rents is summarized in Figure 2.

In detail, once the location and size of the household of interest is identified, a minimum guaranteed living space (MGA) must be associated with that household in order to consider only housing units with an eligible living space. The Housing Assessment System (HAS) prepared by the Federal Housing Office, represents a relevant tool for the housing assessment with regard to location, dimension, use and function. The policy of supporting sustainable renting must also be a promoter of an efficient and optimal supply of living space, considering also a wise use of building land.

The HAS 2000 edition, in this regard, defines the minimum net living area, which corresponds to the useful area defined by the SIA 416 standard, as the size of the household varies. The HAS 2015 edition, on the other hand, defines minimum net living area requirements per room number, losing the direct relationship with housing density. Referring then to the HAS 2000 we derive the following relationship for calculating the minimum living area per household dimension(n_p):

$$MGA = 40m^2 + (n_p - 1) * 10m^2 \quad (1)$$

Once the MGA of interest has been defined, we move on to the extraction of data on rents from the SS (rental market practised) and from the web platforms (rental market offered) for properties with a living space comparable with the selected MGA. The occurrences of the rental market practised, are in general characterized by rental costs on average more competitive than those offered today, due to rental contracts already in existence for many years that have benefited from market conditions no longer reproducible. Figure 3 shows an example of the rental market offered and practised for an apartment with a surface area of 70 m^2. In this specific case it can be seen that the peak of occurrences in the offered rental market occurs at a cost per m^2 per year about 20% higher than that practised (250 CHF/m^2 VS 210 CHF/m^2).

3.2. Admissible rents

Identifying the admissible rents distribution for a given location and household size represents the real

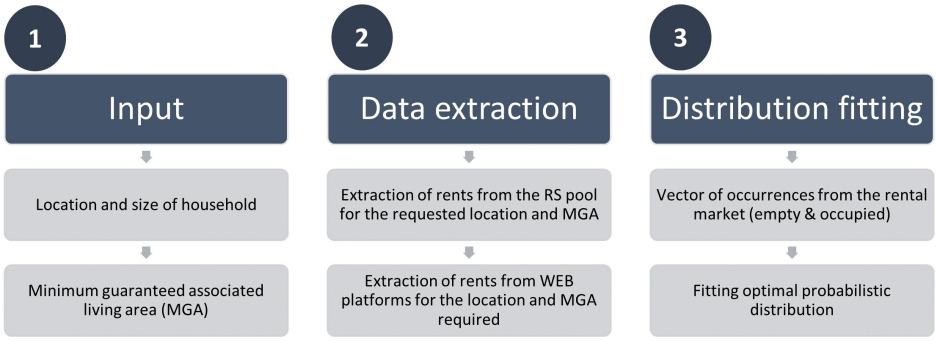

Figure 2: Workflow for calculating the distribution of available rents.

challenge for a reliable estimate of sustainable rent needs. As already argued in the previous paragraphs, the adoption of the equivalent disposable income (EDI) allows to localize imbalances between the distribution of rents and the local households spending power, moving on to a more quantitative analysis of the rental market as function of the income territorial distribution. The use of the equivalised disposable income has the great advantage of being able to model the lower spending power of large households in the analysis by means of corrective factors, which would otherwise be difficult to reproduce with other types of income. The EDI is calculated from the disposable income, taking into account the number of persons using the *household equivalence scale*.

More specifically, the disposable income is computed by considering expenses and earnings reported in Figure 4, while to account for the household size (a household of five people does not spend five times as much as a single person to have the same standard of living), a value of 1.0 is assigned to the oldest person in the household, a value of 0.5 to each of the other persons aged 14 and over, and 0.3 to each child under 14; the "*equivalent size*" of the household corresponds to the sum of the assigned values.

The use of equivalised disposable income brings with it the great advantage of being able to model the lower spending power of large households obtaining a reliable comparison between households of different sizes, something otherwise difficult to reproduce with other types of income.

3.2.1. Data source

The process of acquiring representative income data is by its nature complex, given the degree of sensitiv-

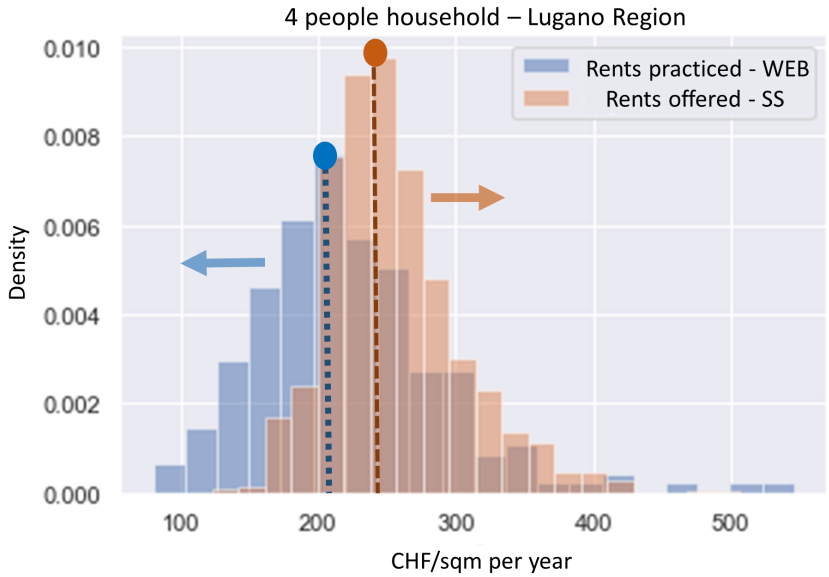

Figure 3: Offered and paid rents in the rental market for accommodations with an adequate MGA for a 4-people household in the Lugano Region.

ity of the subject matter. The current study had the approval of the Contribution Division and the Cantonal Statistics Office for the use of the database adopted for the *Poverty Study*, a work aimed at providing a framework of the social and economic situation on the basis of tax data from 2015, the most recent year with acceptable statistical coverage. The final database has a reference population of about 280'000 individuals and 130'000 households, excluding taxpayers with a taxable income of more than CHF 200'000 and a taxable wealth of more than CHF 1 million. *Source Taxation* [3] is also excluded. In order to guarantee data protection, the data was always provided in aggregated and anonymous form, only for municipalities with a population of more than 1'500 inhabitants and only in terms of occurrences by income interval and household size.

In order to use more recent data, it was decided to integrate the 2015 tax data with the 2017 direct federal tax data. The latter database is available and accessible to all and is about two years behind the current year. This data source also only reports the

[3] Tax is deducted directly from the employee's salary and passed on to the state. This form of taxation mainly concerns foreign residents without a C permit.

Income (+)	Expenditure (-)
• Labour income	• Cantonal tax
• Assets income (e.g. movable and immovable properties)	• Federal tax
• Income from transfers (e.g. pensions, social insurances, etc)	• Municipal tax
• Income from minors	• Ecclesiastical tax
• Social aids	• Alimony payments
	• Health insurance
	• Mandatory social contributions
	• Professional Expenses

Figure 4: Income and expenditure considered in the calculation of disposable income.

number of total occurrences for certain income intervals, with no link to household type and treating net income rather than equivalent disposable income. The calibration phase of the 2017 data based on the 2015 supply was conducted by parameterizing the two income distributions (Federal Tax VS Cantonal database) and optimizing the match between them for each analyzed municipality. By doing so, a vector of conversion coefficients is obtained, thanks to which it is possible to estimate the most probable distribution of equivalent disposable income in 2017 for each municipality of interest from the federal tax database. This analysis procedure will make it easier to update future income data, ensuring a more direct data acquisition flow. In the next section, the steps required to obtain an admissible rent distribution to compare with the offered and practised rents will be analyzed.

3.2.2. Admissible rents distribution

The proposed methodology aims to detail the estimation of sustainable rent needs in relation to the household dimension, thus linking it to the offered living space. It is therefore necessary the definition of different distributions of equivalent disposable income as the number of people in the household under consideration changes.

When working with federal tax data, it is not possible to derive a direct estimate of the various admissible rent distributions for increasing household sizes, given the lack of detailed data. To overcome this problem, the pro-

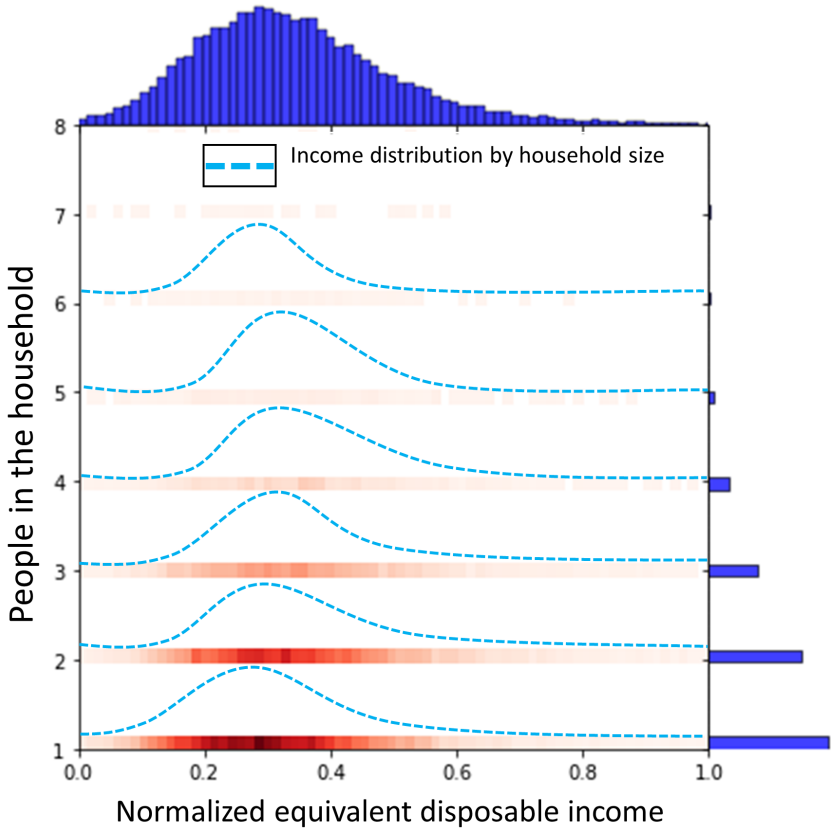

Figure 5: Copula-based PDF of normalized equivalent disposable income (2017) by household size.

posed method of analysis implements a statistical calculation tool, called Copula [9](Cherubini et al. 2004), which makes it possible to identify the probability distribution of income conditional on the number of persons in the household, starting from the two associated marginal distributions (Figure 5). The procedure for defining the admissible rents distribution, whose steps are shown in Figure 6, starts with the calibration stage of the income data from the 2017 direct federal tax. The next steps involve declining the distribution according to the location and size of the household of interest. The copula must then be calibrated in order to obtain an equivalent disposable income distribution for each household size considered. The final step aims at obtaining an admissible rent by using the concept of effort rate and minimum guaranteed living space.

In order to convert an income into an admissible rent, a maximum admissible effort rate (MAER) must be used to calculate the maximum sustainable expenditure by households on housing, which must then be related to the minimum guaranteed living space to arrive at a cost per square meter per year. Below is the equation for calculating the admissible rent (AR), for income level i and household k:

$$AR_{i,k} = \frac{EDI_i * MAER}{MGA_k} \quad (2)$$

The selection of the optimal MAER is still a matter of debate that does not yet show clear and unanimously validated directions of choice. The adoption of a MAER equal to 30% seems to be the most consolidated choice in the literature ([4],[10]), if, however, we adopt gross income as the basic variable for estimating admissible rents. The use of equivalised disposable income in addition to the declination of the result by location and household size, raises the question of which specific MAER to adopt given that a univocal choice would be difficult to sustain. In order to consider the uncertainty implicit in this issue, we will proceed by scenarios, calculating the needs for different values of MAER (40% - 45% - 50%), finally providing an average result.

3.3. Admissible rents VS Available rents

After obtaining the available and admissible rent distribution curves, the final step involves a direct comparison between them in order to highlight specific areas of sustainable rent requirements. Figure 7 takes the case of a 4-person household in a Ticino region as example, with an MGA of 70 m^2 , according to equation 2. In Figure 7A the probabilistic distribution of the admissible rent $f(x)$ (PDF) is compared with the PDF of the available rent. The functions of the distributions adopted to approximate the real occurrences (HISTs in Figure 7) are selected by minimising the fitting error. In particular, when varying the boundary variables (household details, location, MGA, MAER) the calculation model always chooses the optimal distribution for both available and admissible rents, selecting it from a predefined set of potential distributions [4]. Figure 7B shows $\Delta f(x)$ computed as difference between the two distributions, required to highlight imbalance areas that are representative of the needs of low-income households that do not

[4]The set of potential distributions consists of: *Gaussian distribution - Lognormal distribution - Gumbel distribution - exponentially modified Gaussian distribution.*

Figure 6: Estimation process of the admissible rents distribution.

see enough affordable housing on the actual available rents despite reaching the MAER. From the analysis of the area of need (red area in Figure 7B) it is possible to derive detailed information on the selected case. In particular, in the reported example of a 4-person household with a MAER of 40%, the area of need begins at point Ω, which has abscissa equal to 213 CHF/m^2 per year, which for a 70 m^2 flat is equivalent to about 1'240 CHF per month gross, thus including accessory costs. Thus, the proposed approach allows to quantify the sustainable rent by differentiating it by location and target, overcoming the limitations of the classic methodologies that set constant values depending on the offered rents.
Finally, the number of 4-people households in need of rent below CHF 1'240 is provided directly by the area of need highlighted in Figure 7B.

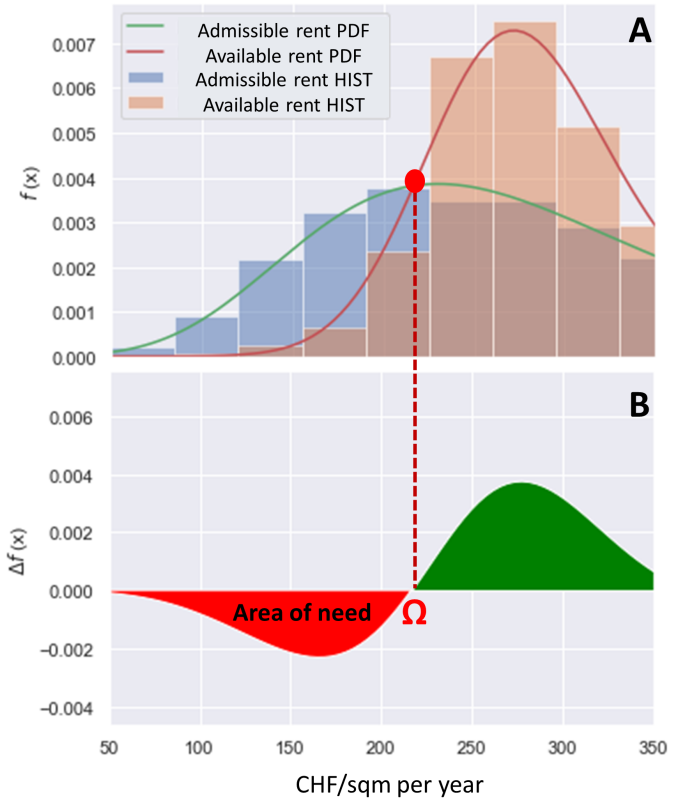

Figure 7: Comparison between the distribution of available and admissible rents for a MAER of 40% and a 4-person household with a MGA of 70 m^2.

3.4. Sustainable rent levels analysis

The quantification of sustainable rent levels represents another important aspect that has to be analyzed, an objective that cannot be separated from the local incomes distribution. In common practice, a rent that is equal to or less than a given percentile of the distribution (e.g. 1^{st} quartile) is considered sustainable. Obviously, this approach is simpler and more immediate, but it does not guarantee any real connection with the distribution of income and, therefore, defines values that are unreliable and not necessarily truly sustainable. A connection with a more in depth analysis of the rental market is therefore required.

Returning to the comparison in Figure 7, once the area in red of need has been defined, the rent value below which this can be considered sustainable is that identified by the point Ω. This definition assumes that a market imbalance is needed to identify a

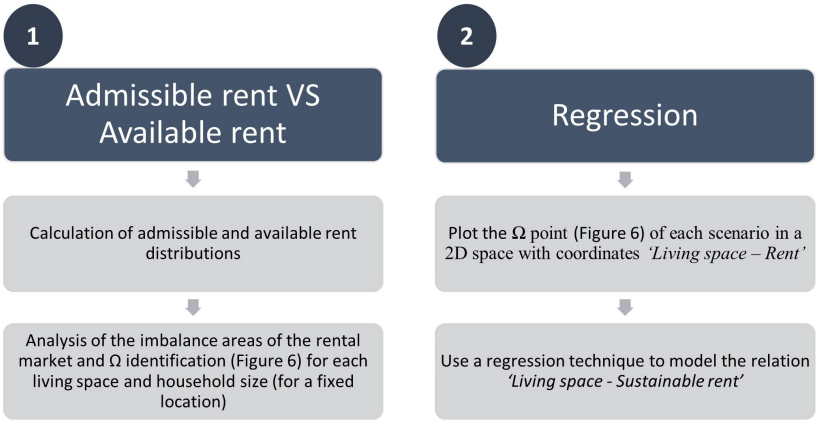

Figure 8: Sustainable rent definition per location and living space.

sustainable rent value. This assumption is completely reasonable, given that in the presence of a perfectly balanced market, all existing rents would already be sustainable and therefore there would be no need to define lower rents to reduce the effort rate. Starting from this assumption, Figure 8 shows the steps required for identifying sustainable rental levels, depending on the location and the living space. More specifically, a cloud of Ω points have to be identified for each fixed location by varying the living space and thus the target household size. Finally, a regression model allows to model the variability of the sustainable rent levels for different property dimensions, in a given location.

3.5. Practised effort rate analysis

This additional analysis is aimed at providing a picture of the current situation considering the actual occupancy of housing units and thus the *practised effort rate* (PER). No assumptions are therefore made about the living space to be associated with a given household size, in order to match the current practised rental market.

The copula tool makes it possible to link the rent paid by a specific household with the most likely income received by that same household, a connection otherwise impossible considering the available databases.

By comparing the municipalities for which we have obtained income data, we can estimate a correlation between *income* and *paid rent* (both expressed through local average values) and use it to calibrate the selected copula. After that, the practised effort rate can be analyzed in probabilistic terms, by generating random artificial samples

from the copula model. Each analyzed sample represents one household at which it is associated an income and the most probable paid rent, following the related distribution functions of the analyzed location. The main steps are summarized in Figure 9, while in Figure 10 is reported one preliminary output. More specifically, in Figure 10A are reported artificial samples with coordinates *EDI - PER*, generated from the copula model, that can be employed to compute the conditional probability of exceeding a fixed PER for increasing EDI levels (Figure 10B).

4. Results

4.1. Admissible rents VS Available rents

The estimate of sustainable rent requirements is calculated on three MAER values ($35\% - 40\% - 45\%$) and provided in terms of average values by region and housing occupancy scenario.

As mentioned, the living space to be ensured for each household is an initial input variable, which must be selected based on assessments related to housing quality criteria. In the HAS 2000 scenario, the areas are distributed as per equation 2 (see Scenario A in Table 1); three more scenarios will be added to this first one (Table 1) in order to evaluate the imbalance of the rental market according to different housing density targets. Figure 11 shows the share of households that, despite practising an effort rate equal to the MAER, cannot find a sufficiently large living space and therefore, in order to meet housing costs, must sustain a high effort rate. The results are reported in percentage terms with respect to the total number of households in rented properties in the whole Canton Ticino, always considering a percentage value calculated as

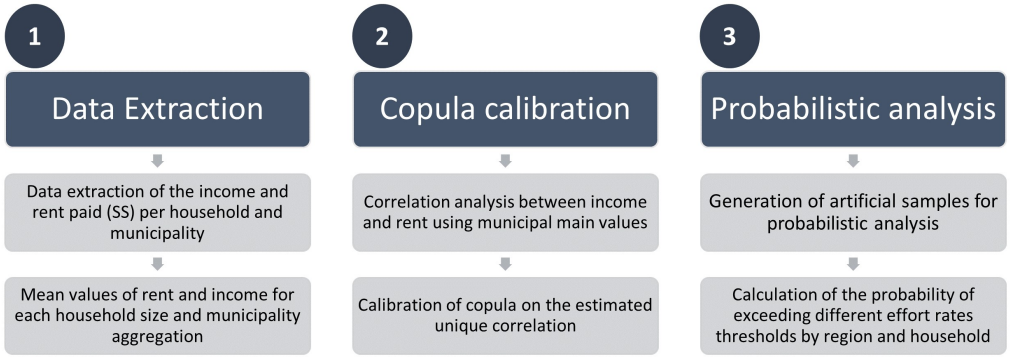

Figure 9: Practised effort rate analysis process.

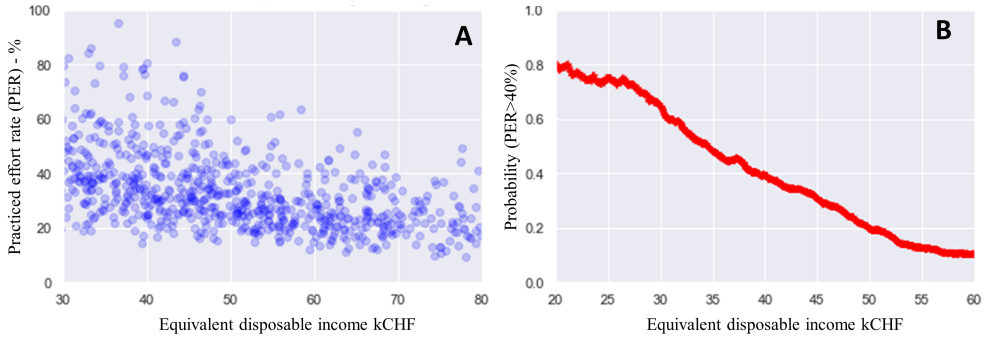

Figure 10: Scatter plot of the practised effort rate as a function of income (A) and probability of exceeding a fixed effort rate (40%) as income changes (B), for the case of a 2-person household in a Ticino region (A and B)

Scenario	Household size - Number of people				
	1	2	3	4	5
A	40 m^2	50 m^2.	60 m^2	70 m^2	80 m^2
B	50 m^2	60 m^2	70 m^2	80 m^2	90 m^2
C	60 m^2	70 m^2	80 m^2	90 m^2	100 m^2
D	70 m^2	80 m^2	90 m^2	100 m^2	110 m^2

Table 1: Housing scenarios as a function of household size and living space.

the average between the results corresponding to three MAER scenarios ($40\% - 45\% - 50\%$).
Relative totals for the Canton start from about 1.4% ($\approx$ 900 households in need of sustainable rent) in scenario A, to about 9 % ($\approx$ 6′000 households in need of sustainable rent) in scenario D, in which relatively large living spaces are ensured if compared to the minimum net values indicated in the HAS 2000 ($+30m^2$ per person). In absolute regional terms (Figure 11), it is the Lugano area that shows the highest values, rising from about 400 households in scenario A to about 3'250 in scenario D. The Bellinzonese and Locarnese show comparable absolute values, rising from a minimum of about 220 to a maximum close to 950 households (Scenario A vs Scenario D), while the Tre Valli region shows almost negligible absolute values, with finally the Mendrisiotto characterised by low absolute values compared to the more populous regions ($\approx$ 35 Scenario A, $\approx$ 390 Scenario D).
In Figure 12 the result is declined per household size (1 to 5 people).

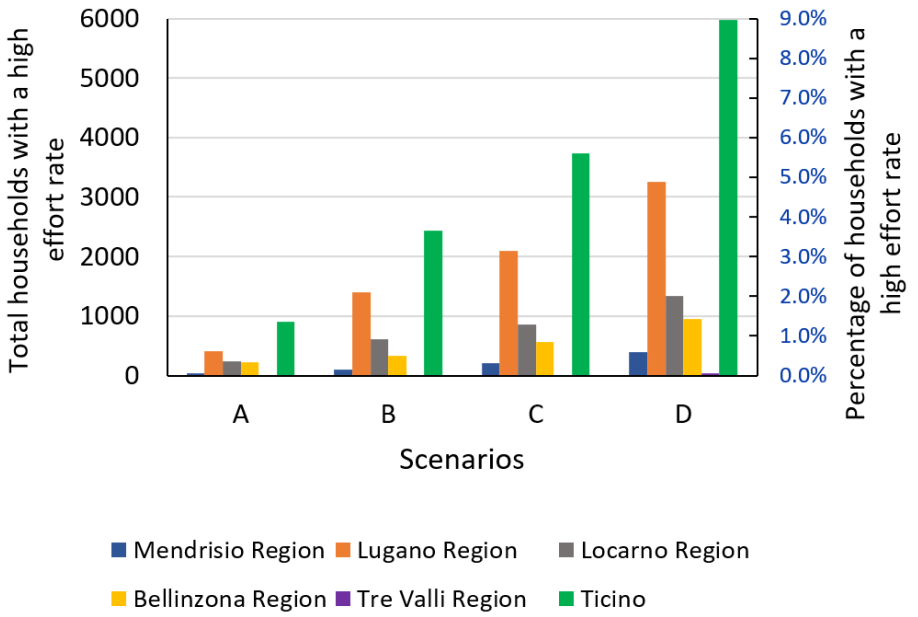

Figure 11: Percentage of households requiring more sustainable rents for different locations and living space scenario. Percentage calculated with respect the total number of households in rental properties.

In general, as the area occupied increases, all the household sizes experience an increase in the relative proportion of high effort rates. This was an expected trend, since larger living space require increasing expenditures on housing; note, however, that the rate of growth is much more pronounced for single persons, for whom in some Ticino regions it goes from 0 (assuming a 40 m^2 property) to around 2'000 (assuming a 70 m^2 property). This is mainly due to the higher increase in price when moving from a 40 m^2 apartment to a 70 m^2 one, compared to an increase of 30 m^2 on an apartment that is already larger than a studio apartment. This greater variation is therefore reflected in the offer (rents) and consequently in the estimate of the need. The total sustainable rent requirement can be declined not only to household size and location but also to housing size. In this regard, Figure 13 shows the distribution of sustainable rent needs (for the entire Canton) as a function of available living space. It can be seen that the share of "small" lodgings (40 m^2. - Scenario A; 50 m^2 - Scenario B; 60 m^2 - Scenario C; 70 m^2. - Scenario D) increases considerably (+18%) going from Scenario A (10%) to D (28%). This is obviously

a direct consequence of what has already been seen for the households of one person in Figure 12, and highlights how very small housing currently offered in the rental market, as well as those in the practised market, present a lower probability of being associated with high practised effort rate and thus, in fact, more sustainable and well calibrated with respect the admissible rents distribution, always within the limit of admissible criteria of adequacy of housing. Finally, in scenario D, which in practice comes closest to what is the current housing situation, it can be seen that the shares of sustainable rent requirements are distributed almost evenly across all the living spaces considered (with the exception of 110 m^2 housing).

4.2. Sustainable rent levels

The analysis of various scenarios allows to characterize a series of combinations that, for each analyzed living space and location, identify slightly different values of sustainable rent (Figure 14). After that, through a regression model, that can be more complex (e.g. taking into account the uncertainty levels) or simpler (linear, as shown in Figure 14), it is possible to provide an estimate of how much the sustainable rent changes as the location and living space vary. Therefore, for each fixed location we can obtain a cloud of points and calibrate the regression model. Figure 15 shows a comparison of sustainable rent values for different living spaces and locations. These values are usually well below the median of the local proposed rents, demonstrating the need of additional measures to support a rental supply accessible to low incomes without reaching high effort rate. It must be specified that these sustainable rent values consider the entire regional territory, without providing detailed variations as one moves from major urban centers to more peripheral areas. This aspect risks underestimating sustainable rent levels in urban centers and, on the contrary, overestimating them in peripheral areas. For this reason, future analytical efforts should focus on providing evolutionary models of sustainable rent as a function of reference urban character as well.

4.3. Practised effort rate

Starting from the probabilistic analysis of the PER in the various Ticino's regions, Figure 16 shows the probability that the latter exceeds a limit value (here fixed at 40%) as the equivalent disposable income and the size of the household analyzed vary. The figure reported is therefore conditional on income level, and the estimated probabilities always refer only to households living in rented properties in the region under examination

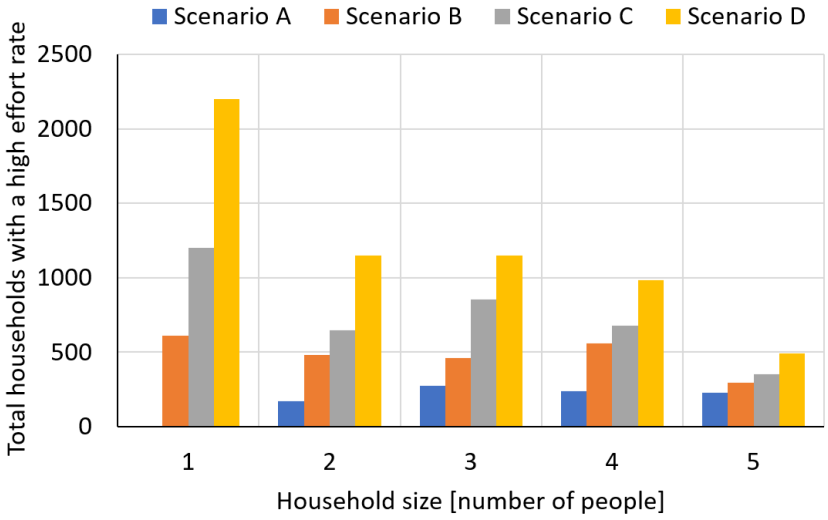

Figure 12: Percentage of households requiring more sustainable rents as living space and number of persons vary (1 to 5 people). Percentage calculated with respect the total number of households in rental properties having the same size.

and having that specific size.

These results, equal to those obtained in the comparison between available and admissible rents, show a general trend whereby relative need grows almost linearly with household size. It should be specified that these probabilities refer to the total households of a given size, while Figure 12 shows the absolute quantitative values. For example, a 10% exceedance probability calculated on households of 3 persons means that out of 100 households of 3 persons, 10 are likely to suffer from high effort rates.

Results clearly show how larger households have higher probabilities of exceeding. This is actually an expected result, which depends primarily on the type of income used and thus, in this case, on the corrective coefficients that define equivalent income, homogenizing households of different sizes. In the transition from disposable income to equivalent disposable income, in fact, it is the large households that are "penalized" the most, which must also sustain higher average housing costs due to larger living spaces.

Although the results in Figure 16 would already allow deriving information of interest, extrapolated in this case for a single MAER (40%), an analysis of the incidence of high PER should be conducted not as a function of the level of income received but rather on the unconditional prob-

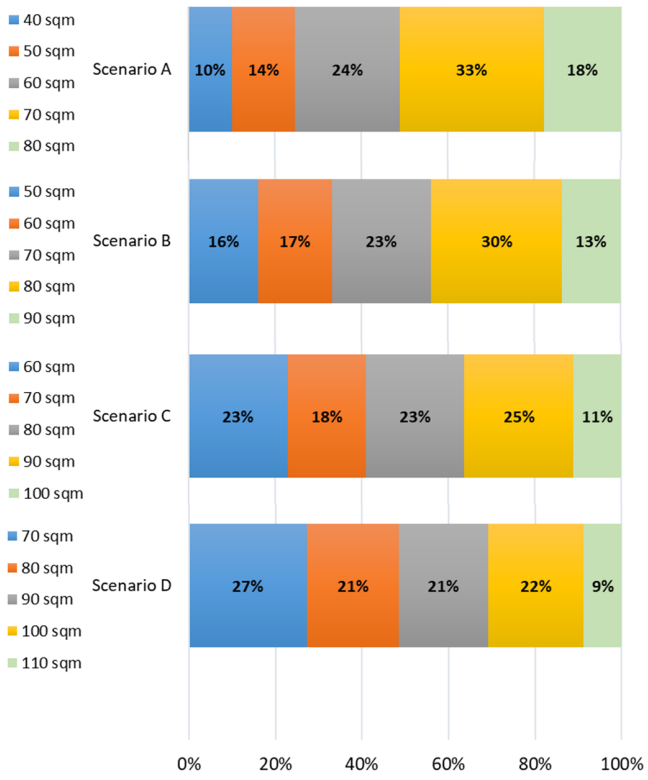

Figure 13: Sustainable rent need percentage by size of rental property and analyzed scenario.

abilities of exceedance (thus encompassing each income bracket). These can be calculated using, for example, the *low of total probability*[5]:

$$P(PER > LER) = \sum_{k=1}^{n} P(PER > LER|EDI_k) * P(EDI_k) \quad (3)$$

In the reported equation, the unconditional probability that the practised effort rate (PER) exceeds a threshold value (LER) is equal to the summation of the same probability of exceedance, conditional on a specific level k of equivalent disposable income (EDI), multiplied by the probability of occurrence of the same income value.

In doing so, we obtain the curves reported in Figure 17. These show the probability that the PER exceeds different values of threshold effort rates (values in the abscissa), considering

[5]The law of total probability is used to calculate the probability of an event, $P(A > B)$, through the total conditional events $P(A > B|C)$ in the same sample space.

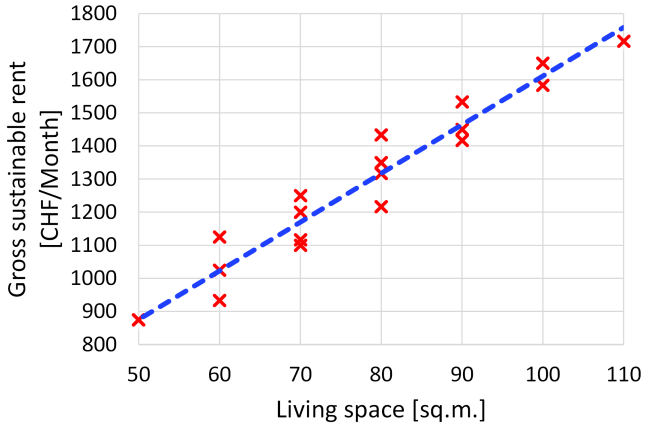

Figure 14: Scatter plot of multiple Ω points (*Rent-Living space*)for a given location and linear regression.

all income brackets. As was to be expected, as the threshold increases, the relative probability of exceedance decreases. This is to emphasize again how the maximum considered admissible effort rate (MAER) has an important impact on the final estimates of sustainable rent requirements.

Figure 18 shows the reduction in percentage terms of the total number of households in difficulty (practced effort rate>MAER) in going from a MAER of 35% to 50%. Less pronounced reductions highlight a problem of sustainable rent, in some respects more serious, with the presence of a greater share of households who are unable to access housing except through extremely high PER. It should be emphasized that the proposed analysis of the practised effort rate as a variable not conditioned to the living space, is not able to incorporate a comprehensive description of the actual rental market. More specifically, only the already occupied properties are considered, since a connection between an empty offered property and an household income is not feasible. This aspect inevitably leads to an inaccurate estimate of the amount of households in need, for two specific reasons: 1 - the spending needs of a portion of these households could already be satisfied through the rental properties offered by the market (not occupied) 2 - The already occupied properties are on average characterized by more advantageous rents.

In addition, this picture fails to decline the needs as a function of housing density, an important issue from the perspective of sustainable rent.

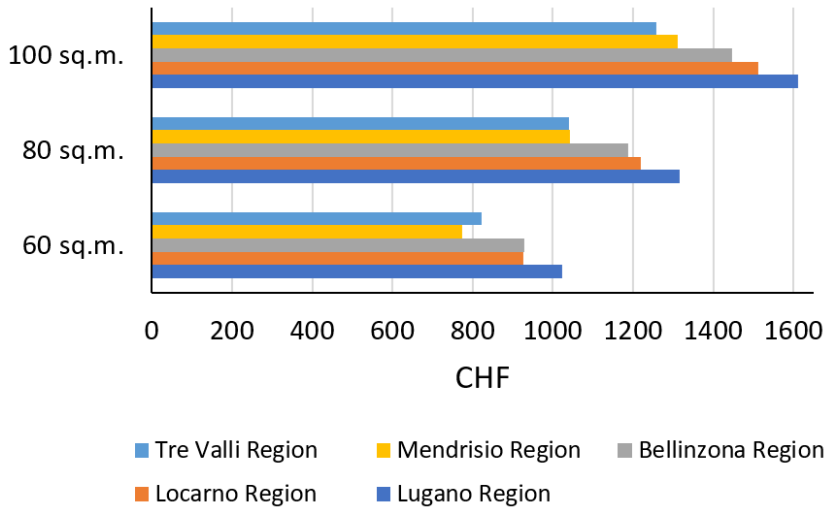

Figure 15: Sustainable rent levels for different living spaces and locations.

Hence why this additional analysis is perfectly suitable for analyzing the real practised effort rate, while it is required to move towards a more comprehensive rental market analysis to quantify the local need of sustainable rents.

5. Machine Learning Regression Algorithm and future developments

The implementation of robust predictive analyses aimed at forecasting the likely evolution of sustainable rent requirements in the coming years, as well as the levels of sustainable rents, is an extremely complex objective, the achievement of which requires additional developments.
Obtaining robust predictive trends poses several challenges, both in terms of the type and volume of data and the algorithms used. The achievement of this objective involves the construction of a dedicated database for the calibration of a model capable of:

- Providing an explicit descriptive analysis of the evolution of the rental market as a function of boundary variables (e.g. mobility and services);
- Finding direct evolutionary relationships between these variables and the sustainable rent requirement;
- Providing an estimate in probabilistic terms of future scenarios;

The construction of the training

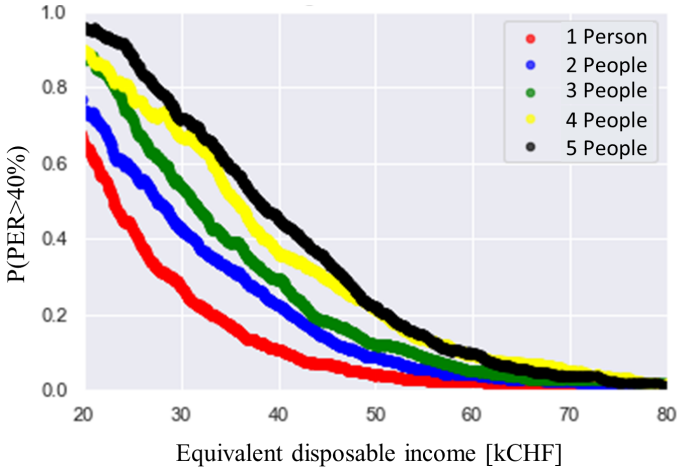

Figure 16: Probability of exceeding a fixed effort rate threshold conditioned to the income level and the chosen location.

database will be based on observations obtained by combining municipalities with adequate data coverage in different ways and measures. This allows to considerably increase the size of the database starting from a relatively small number of municipalities. In fact, on the basis of a simple combinatorial calculation, from a total of n municipalities available we will be able to generate $2^n - n - 1$ observations, if we decide to exclude aggregations with only one municipality. Thus, for example, with just 20 municipalities we could obtain maximum $1'048'555$ observations.

A development in this direction also foresees a partial and progressive independence from data on households income, which by their nature are difficult to access and not always up-to-date. This is theoretically possible by incorporating indirect indicators, such as GDP, unemployment rate, etc., into the calibration database. In addition, it is necessary to characterize the available rents distribution dynamically, e.g. by analyzing the evolution of the share of housing with long-term leases, always within the limits dictated by the available databases.

There are several algorithms that can be used for forecasting analyses of this type, which offers the opportunity not only to define the evolution of a variable of interest, but also to identify the associated level of uncertainty, so as to highlight the most probable evolutionary scenario. The discussion of strategies for implementing robust predictive analy-

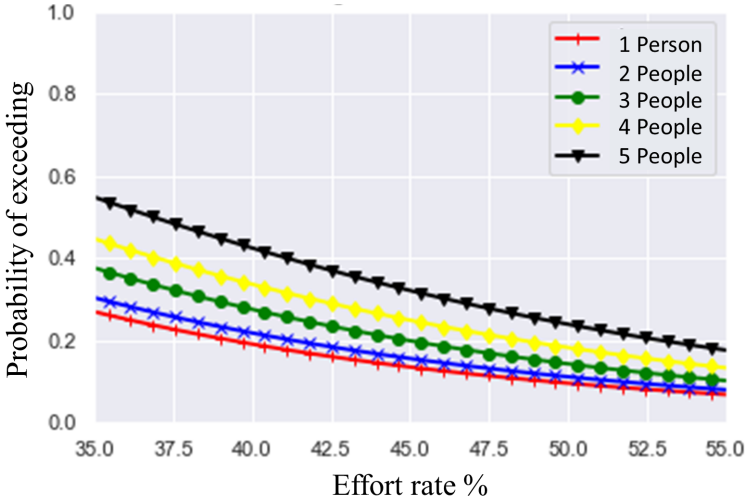

Figure 17: Unconditional probability of exceeding increasing thresholds of practised effort rate for a chosen location.

ses will not be the subject of this report, but it is worth pointing out that improvements to the model that will go in this direction have already been planned, and subjected to a dedicated feasibility analysis.

In spite of this, the main steps required to structure and conduct a predictive analysis of the local rental market are reported in Figure 19, not only from the point of view of sustainable rent, but also in terms of the multiple aspects that characterize the demand and supply of housing today.

6. Discussion

Three types of analysis have been presented, which to different degrees and in different ways investigate three issues related to the rental market, i.e., areas of imbalance related to sustainable rent, the practised effort rates distribution and finally, the identification of robust values of sustainable rent. Figure 20 aims to summarize the key aspects of each approach by providing an overview of some innovative as well as weak points, which will however be discussed in the following section. In this regards it is worth highlighting and briefly summarizing the advantages and limitations of the proposed approach.

6.1. Advantages

- Simplified access to the analysis of multiple scenarios concerning the rental market. The analysis model is extremely flexible with respect to different input

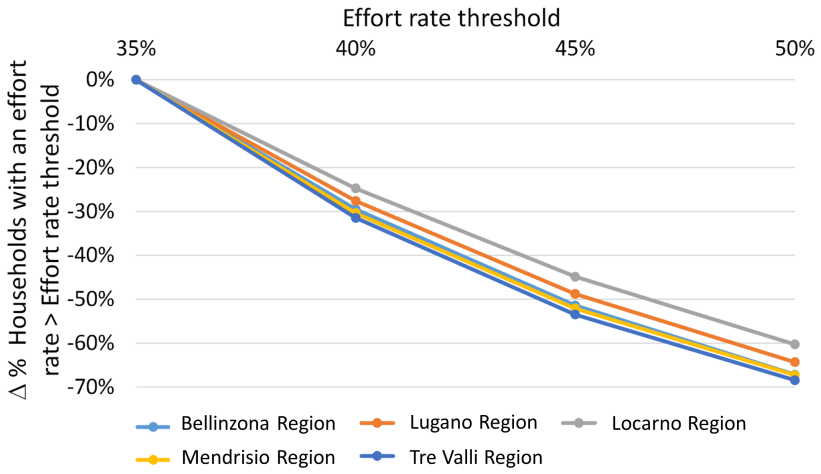

Figure 18: Reduction of the percentage of households requiring more sustainable rents by increasing the MAER.

variables. It is, for example, easy to analyze the economic sustainability of the rental market in case of multiple effort rates and/or guaranteed living spaces. This factor makes the approach suitable also in case the interest shifts towards the estimation of future patterns.

- Identification of sustainable rent values defined on the basis of real areas of imbalance in the rental market. Until now, the quantitative estimate of sustainable rent was always calibrated from a percentile/quartile threshold of the available rents distribution (that should not be exceeded). This leads to an approximate quantification of sustainable rent levels, which were often not representative of the actual rental market.

- The possibility of identifying an admissible rents distribution allows to quantitatively analyze the imbalance areas within the rental market and thus to calculate the real need per location and household size.

- The model, appropriately calibrated from the point of view of the evolution of the effort rate for high incomes, also lends itself to uses outside those of sustainable rent. For example, it is also possible to quantify the current supply of luxury housing in relation to real de-

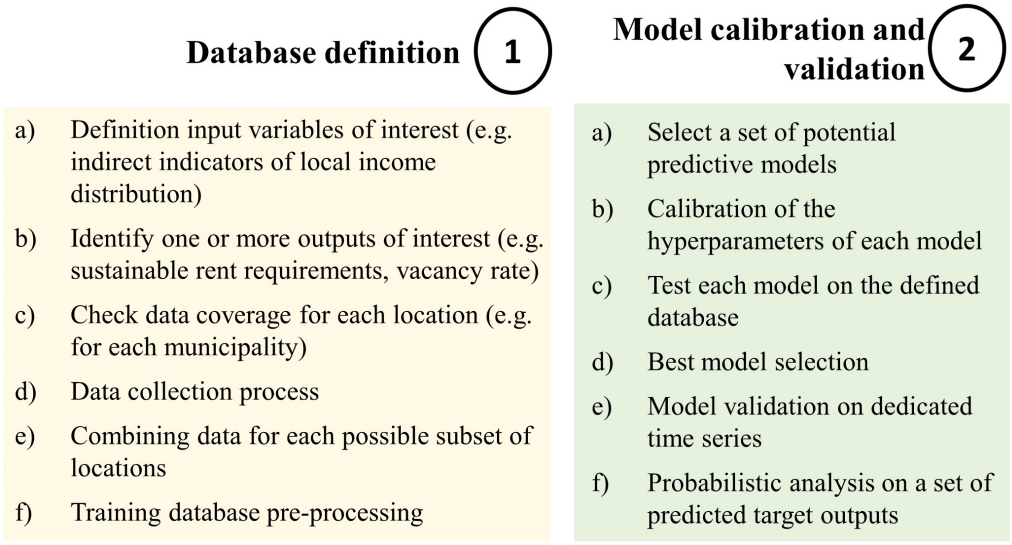

Figure 19: Main steps for the definition of a predictive algorithm working on the local rental market.

mand, or alternatively, to identify the guaranteed living space for which the market is self-balancing.

- In general, a more detailed needs definition allows for more targeted and thus effective interventions and measures. In particular, if we consider the two macro-areas of support for sustainable rent, such as *stone aid* rather than *personal aid*, these have a direct impact on the distribution curves of the available (stone aid) and admissible (personal aid) rents. The analysis model therefore also offers the possibility of quantifying more complex action strategies and monitoring their effects on the market.

6.2. Limitations

On the other hand, analyzing the current limitations of the proposed model, it should be noted that:

- The characterization of needs by households requires sensitive data that are not easily accessible. Access to data appears to be a fundamental aspect of the proposed approach, which is therefore constrained by a flow of information acquisition that is not always continuous and easy to trigger.
- Although there is the possibility of varying the rate of ef-

	Analysis		
	Admissible rent vs Available rent	**Practiced effort rate**	**Sustainable rent level**
Main Objective	Analyze areas of imbalance in the rental market in terms of sustainable rent needs	Analyze the distribution of the local effort rate	Identify robust levels of sustainable rent
Approach	• Direct comparison between admissible and available rent by means of probability distributions • Copula tool used to obtain income distributions for different household sizes	• Copula tool used to connect '*household size-paid rent-income*' for each analyzed location • Probabilistic analysis of the practiced effort rate by means of artificial samples	• Based on the comparison between admissible and available rent • Identifying the starting point of the imbalance area in the rental market
Innovation points	• Full characterization of the available rents (both occupied and available properties on the market) • Detail income data to characterize demand • Results per location, household size and living space	• Practiced effort rate based on the copula connection between income and paid rent • Analysis of the probability of exceeding of multiple threshold of effort rate per location and household size	• Sustainable rents levels computed from a comprehensive analysis of rents and incomes • Sustainable rents levels computed as function of the location and the living space
Limitations	• Income data not easily accessible • the impact of additional variables (e.g. mobility) is not explicitly considered in the analysis • Predictive analysis is not supported	• The analysis does not capture the sustainable rent needs due to incomplete supply definition • Income data not easily accessible	• A time consuming analysis as preliminary step is required • Income data not easily accessible

Figure 20: Summary outline of the proposed analyses.

fort at will, the proposed approach requires setting a specific MAER against which to estimate the number of households still struggling to access adequate housing. It is therefore still necessary to define clear economic and social parameters that can identify a unique acceptable effort rate compatible with the proposed analysis. To this end, one could use disposable income and amounts defined by social policies to estimate minimum monthly needs according to the type of household, in order to estimate a maximum amount that can be allocated to housing costs.

- Identifying the admissible rent curve presents some limitations due to the extreme heterogeneity of the problem. Some economically well-off households could, for example, occupy inexpensive housing units, generating an overestimation of the rental market imbalance. The approach taken is however regionally calibrated, thus limiting the effect of localized cases. Greater efforts must however be

made to obtain more detailed data in order to be able to take extreme cases into account.

- The proposed approach still requires an additional development step for the implementation of time series and thus future estimates or, possibly, for the inclusion of additional variables describing other environmental factors (e.g. mobility, building conditions).

7. Conclusions

The proposed research project seeks to address the need for new data-driven approaches aimed at studying and analyzing the adequacy of sustainable rental supply.
The analysis of the rental property market poses several challenges, such as the correct identification of the *admissible rents* distribution to be compared with the actual distribution of the *available rents*. The *admissible rents* curve aim to identify the total number of households still suffering in finding a sustainable rental property despite reaching the maximum admissible effort rate for an adequate living space. The adopted concept of *admissible rent*, defined and used in the proposed market analysis, is therefore aimed at identifying the maximum spending power of households in a given location that can be used for housing costs.
In addition, two development steps are proposed in order to analyze the distribution of the practised effort rate, without therefore making assumptions on the living space, and to provide a calculation model more suitable for predictive analyses and additional scenarios implementation. The Monte Carlo approach allows capturing all the cases in which households can, for example, decide to live in a larger apartment reaching a higher effort rate just for having access to a higher quality of life. However, this second method is not able to account for the available rents and consequentially it should not be intended as a valid alternative to estimate the needs of sustainable rents. Finally, information obtained from the direct rental market analysis have been used to identify sustainable rent levels as location and living space varied, finding out how these levels are often well low the median of the local housing supply.

Acknowledgements
We would like to thank the Canton Ticino for providing the necessary data for the implementation of the proposed methodology and for supporting the project with a dedicated mandate.

List of Abbreviations

MAER Maximum admissible effort rate

MGA Minimum guaranteed living space

AR Admissible rent

HAS Housing Assessment System

EDI Equivalent disposable income

SS Structural Survey

BDS Buildings and Dwellings statistic

PDF Probability density function

PER Practised effort rate

Bibliography

[1] M. A. Adabre, A. P. Chan, Critical success factors (csfs) for sustainable affordable housing, Building and Environment 156 (2019) 203–214.

[2] X. Gan, J. Zuo, P. Wu, J. Wang, R. Chang, T. Wen, How affordable housing becomes more sustainable? a stakeholder study, Journal of Cleaner Production 162 (2017) 427–437.

[3] E. Mulliner, V. Maliene, An analysis of professional perceptions of criteria contributing to sustainable housing affordability, Sustainability 7 (2015) 248–270.

[4] A. S. Del Pero, W. Adema, V. Ferraro, V. Frey, Policies to promote access to good-quality affordable housing in oecd countries (2016).

[5] E. Mulliner, V. Maliene, Criteria for sustainable housing affordability (2011).

[6] M. Bigotta, V. Giancone, Struttura ed evoluzione dei salari in ticino (2020).

[7] I. Simic, Evoluzione del mercato immobiliare in Ticino dal 2000: Impatto del basso livello dei tassi di interesse, Ph.D. thesis, Scuola Universitaria Professionale della Svizzera Italiana, 2019.

[8] Y. Xiao, Hedonic housing price theory review, in: Urban morphology and housing market, Springer, 2017, pp. 11–40.

[9] U. Cherubini, E. Luciano, W. Vecchiato, Copula methods in finance, John Wiley & Sons, 2004.

[10] Y. Bochsler, F. Ehrler, T. Fritschi, N. Gasser, C. Kehrli, C. Knopfel,

R. Salzgeber, Wohnversorgung in der schweiz: Bestandsaufnahme über haushalte von menschen in armut und in prekären lebenslagen (2015).