

Paper prepared for  
the International Conference on Real Estate Statistics  
Luxembourg, 20<sup>th</sup> – 22<sup>nd</sup> of February 2019

**Combining administrative and market data in the development of new commercial real estate indicators**

By Cajsa Mølskov and Sigrid Krogstrup Jensen

For additional information please contact:

Sigrid Krogstrup Jensen  
Statistics Denmark, Prices and Consumption  
Sejrøgade 11, DK-2100 Copenhagen, Denmark  
E-mail: [sij@dst.dk](mailto:sij@dst.dk)  
Phone: +45 3917 3456

**Abstract**

This paper will discuss the advantages and disadvantages of different types of data sources in relation to producing commercial real estate indicators. The discussion will be followed up with examples that will illustrate the perspectives in using the different data to produce new commercial real estate indicators and to expand existing indicators.

The discussions and conclusions reached in this paper will be based on a newly established register for Danish commercial real estate. In 2016 Statistics Denmark was given the assignment to collect detailed data on Danish commercial real estate for the Danish tax authorities for them to use for improvement of their valuations of commercial properties for taxation purposes. Since then, Statistics Denmark has constructed a register based on both administrative and private sources. This register is largely available for statistics production.

## 1. Introduction

In Denmark authorities define commercial property or commercial real estate as property that generates a profit for its owner, and thus, must be valuated and taxed accordingly. This definition covers many types of property; housing for rent, office buildings, stores, warehouses, production facilities, private institutions and specialized property such as harbors, fish farms, etc. Furthermore, commercial property can be owned by private persons, companies, public authorities, housing organizations, pension funds, etc. The table below shows the number of commercial properties in Denmark, as of 2017, according to geographic region and type:

**Table 1 Number of commercial properties in Denmark according to geographic region and type, 2017**

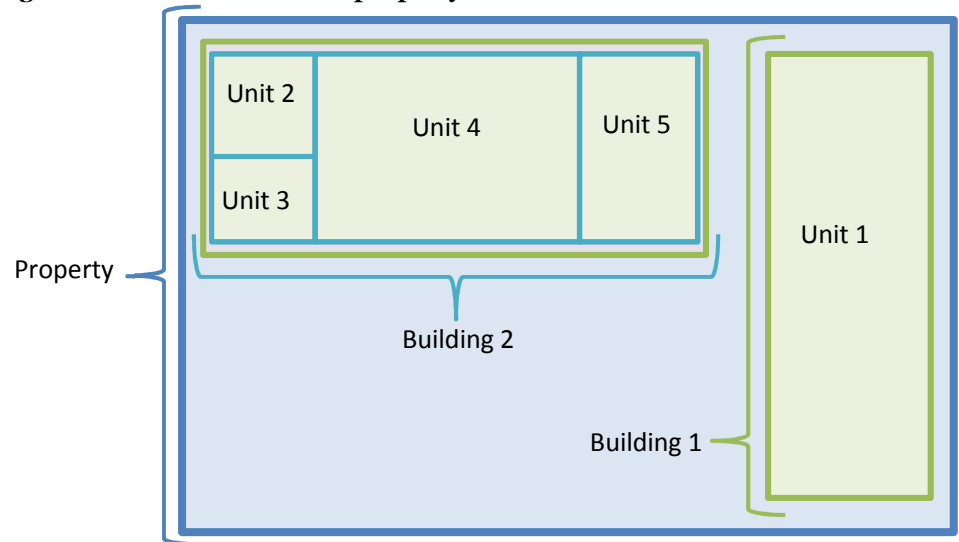
	<b>Business</b>	<b>Mixed housing and business</b>	<b>Production and warehouse</b>	<b>Private institutions</b>	<b>Specialised property</b>	<b>Other</b>	Total
<b>Capital region</b>	844	2.744	1.005	72	82	325	5.072
<b>Sealand region</b>	1.225	514	827	129	268	449	3.412
<b>Southern Denmark region</b>	1.822	735	1.171	189	595	1.037	5.549
<b>Central Jutland region</b>	1.813	1.054	1.709	130	732	743	6.181
<b>Northern Jutland region</b>	1.044	646	735	119	625	710	3.879
<b>All of Denmark</b>	6.748	5.693	5.447	639	2.302	3.264	24.093

Source: Statistics Denmark - [www.statistikbanken.dk/EJDFOE1](http://www.statistikbanken.dk/EJDFOE1)

There are 24.093 pieces of commercial property in Denmark, rather evenly distributed across the five administrative regions in the country. As can be expected, there are slightly more properties in the three regions which hold the three largest cities in Denmark, i.e. the Capital region, the Central Jutland region and the Southern Denmark region. The capital region, by far, holds the most property in the category mixed housing and business, whereas the majority of commercial properties in the Central Jutland region and the Southern Denmark region are pure business properties. However, these two regions combined also hold more than half of the total number of properties for production and warehouse. Private institutions and specialized property is mostly situated outside the capital region. This is probably due to the fact, that the capital region contains no rural areas, as is the case for the remaining regions.

The 24.093 properties can be split into several buildings and the buildings can be split into even more units, illustrated by figure 1 below:

**Figure 1 – The three levels of property**



Thus, number of units and building pertaining to commercial property surpass the number of commercial properties itself to a great extent.

Over the course of 2017 and 2018 Statistics Denmark has been developing a register containing market data on these commercial properties. The primary purpose of the register is to form the basis for the Danish tax authorities' new system for valuating commercial properties for taxation purposes. Secondly, data can be used for statistical purposes, i.e. producing commercial real estate indicators. The data in the register is collected from six different sources; both from public authorities in the form of administrative data and from commercial sources, in the form of market data. Statistics Denmark has built a system that validates the data and matches it to the Danish Buildings and Dwellings Register (BBR).

The purpose of this paper is to inspire others with regards to the choice of data sources, data collection and data validation methods concerning data for the production of commercial real estate indicators. Furthermore, it is to gather feedback from others regarding the same aspects. The paper will describe the data sources in the new register, the data, the validation and the match to the BBR. Furthermore, it will discuss the advantages and disadvantages of different types of data sources in relation to producing commercial real estate indicators. Finally, it describes the possibilities for producing commercial real estate indicators on the basis of data from the new register.

## **2. Data sources and data**

As mentioned above the data sources in the register for market data on commercial property fall into two categories; public authorities that hold administrative data and commercial sources which hold market data. Table 2 below gives an overview of the data sources and the key variables contained in data:

**Table 2 Overview of data sources and variables in the register for market data on commercial property**

		Administrative data			Market data		
		The Danish FSA	The Danish Buildings and Property Agency	Landsbyggefonden	The Danish Property Federation	Ejendomstorvet	Lokalebasen
<b>Rent (net/gross)</b>	<b>Market rent</b>	X			X	X	X
	<b>Actual rent</b>	X	X	X	X		X
<b>Operating costs (net/gross)</b>	<b>Market operating costs</b>					X	X
	<b>Actual operating costs</b>	X	X	X	X		X
<b>Yield</b>		X			X	X	
<b>Location</b>	<b>Property ID (BBR)</b>	X	X	X	X		
	<b>Address</b>	X	X		X	X	X
<b>Type of property</b>		X	X	X	X	X	X
<b>Type of transaction</b>	<b>Sale</b>					X	X
	<b>Rent</b>			X		X	X
<b>Date</b>		X	X	X	X	X	X
<b>M2</b>		X	X		X	X	X
<b>Value</b>	<b>Price</b>					X	X
	<b>Valuation</b>	X			X		
<b>Vacancy</b>					X		

The sources are described in further detail in the following.

## 2.2 Administrative data

The administrative data for the register is collected from three different public authorities. As Statistics Denmark is also a public authority the data has been supplied to Statistics Denmark voluntarily. However, it has been necessary to legally obtain access to data from the Danish FSA by stating it in the law on property valuations. For all sources contracts have been written to ensure the confidentiality of data.

### The Danish Financial Supervisory Authority (FSA)

The main task of the FSA is the supervision of the Danish financial institutions, including mortgage institutions. The four largest Danish mortgage institutions voluntarily report all lending offers they produce to the FSA. These data contain, among other variables, mortgage lending values, physical characteristics for the objects valuated, yield and rent. However, it is not possible to identify whether an offer has been realized or not. Primarily properties valuated at €67.000 or more are included in data. Data from the FSA is a central source in the Danish tax authorities' new valuation system as it contains a large number of observations and is geographically diverse. Data covers the following types of property;

housing, office, store and production/warehouse. There are very few missing values and the variables are reported consistently. The data consist of 1.804.875 units spread over 676.533 buildings. The geographical coverage of data is described in the table below:

**Table 3 Geographical coverage of data from the FSA**

<b>Region</b>	<b>Province</b>	<b>Number of unique buildings in the new register*</b>	<b>Share of total commercial buildings in province, pct.</b>
<b>Capital region</b>	<b>City of Copenhagen</b>	24.469	19,3
	<b>Copenhagen Suburbs</b>	49.162	20,6
	<b>Northern Sealand</b>	72.761	19,5
	<b>Bornholm</b>	7.279	11,9
<b>Sealand region</b>	<b>Eastern Sealand</b>	24.277	20,9
	<b>Roskilde</b>	10.646	20,6
	<b>Western and Southern Sealand</b>	84.244	12,5
<b>Southern Denmark region</b>	<b>Funen</b>	43.429	12,6
	<b>Odense</b>	21.161	18
	<b>Southern Jutland</b>	52.042	12,3
	<b>Esbjerg</b>	12.983	15,5
	<b>Kolding</b>	12.332	16,1
	<b>Vejle</b>	13.631	14,9
<b>Central Jutland Region</b>	<b>Eastern Jutland</b>	65.456	15,2
	<b>Randers</b>	11.108	15,6
	<b>Aarhus</b>	31.460	21,9
	<b>Western Jutland</b>	55.596	12,1
<b>Northern Jutland Region</b>	<b>Northern Jutland</b>	60.419	12,7
	<b>Aalborg</b>	24.078	17,1

\*As FSA data is based on lending offers there may be more than one lending offer per building. Thus, the data in this table is based on unique lending offers and not buildings as opposed to the remaining tables.

Statistics Denmark receive data on a quarterly basis, 40 or less working days after the end of the quarter. Unfortunately, data can only be used for the primary purpose; valuating commercial properties for taxation purposes, and not for statistical purposes, which is stated explicitly in the law on property valuations.

### **The Danish Buildings and Property Agency**

The Danish Buildings and Property Agency is the property enterprise and developer for the Danish state. The agency manages all property used for governmental purposes, thus, data e.g. contains information on physical characteristics for these properties, rent and operating costs. Statistics Denmark receive these data from the Danish Buildings and Property Agency on a monthly basis. There are very few missing values and the variables are reported consistently. However, the data contains very few observations. It covers 339 unique units spread over 40 unique buildings, where only office buildings are represented. The geographical coverage of these buildings is shown in the table below:

**Table 4 Geographical coverage of data from the Danish Buildings and Property Agency**

<b>Region</b>	<b>Province</b>	<b>Number of unique buildings in the new register</b>	<b>Share of total commercial buildings in province, pct.</b>
<b>Capital region</b>	<b>City of Copenhagen</b>	16	0,0
	<b>Northern Sealand</b>	1	0,0
	<b>Bornholm</b>	1	0,0
<b>Sealand region</b>	<b>Roskilde</b>	1	0,0
<b>Southern Denmark Region</b>	<b>Funen</b>	3	0,0
	<b>Odense</b>	5	0,0
	<b>Southern Jutland</b>	1	0,0
<b>Central Jutland region</b>	<b>Randers</b>	1	0,0
	<b>Aarhus</b>	1	0,0
	<b>Western Jutland</b>	7	0,0
<b>Northern Jutland region</b>	<b>Northern Jutland</b>	3	0,0

#### **Landsbyggefonden**

Landsbyggefonden is a foundation established by the Danish social housing organizations. Social housing is non-profit housing where the rent 100 pct. reflects the operation costs. Social housing organizations consist of housing departments and each department owns and operates their own property. Landsbyggefonden supports and develops the construction of new buildings for social housing in Denmark and it manages the register for rents in social housing. Statistics Denmark has gained full access to this register and some of the information it contains is included in the new register for commercial property; physical characteristics for the rented objects, rent, deposits, and operating costs. The information primarily concerns housing. The data is of a good quality as the variables are reported consistently, however, they are biased towards housing of a more moderate quality than what is the case for the private rental market. Furthermore, the data contains unique administrative keys which can be used to match the observations to the BBR. Data consists of 575.160 unique units dispersed onto 86.517 unique buildings. Table 5 below describes the geographical coverage of data:

**Table 5 Geographical coverage of data from Landsbyggefonden**

<b>Region</b>	<b>Province</b>	<b>Number of unique buildings in register</b>	<b>Share of total commercial property in province, pct.</b>
<b>Capital region</b>	<b>City of Copenhagen</b>	2.463	1,9
	<b>Copenhagen Suburbs</b>	9.671	4,1
	<b>Northern Sealand</b>	6.396	1,7
	<b>Bornholm</b>	1.119	1,8
<b>Sealand Region</b>	<b>Eastern Sealand</b>	1.850	1,6
	<b>Roskilde</b>	1.010	2
	<b>Western and Southern Sealand</b>	8.159	1,2
<b>Southern Denmark region</b>	<b>Funen</b>	5.612	1,6
	<b>Odense</b>	3.821	3,2
	<b>Southern Jutland</b>	10.688	2,5
	<b>Esbjerg</b>	2.567	3,1
	<b>Kolding</b>	1.765	2,3
	<b>Vejle</b>	1.622	1,8
<b>Central Jutland region</b>	<b>Eastern Jutland</b>	6.085	1,4
	<b>Randers</b>	1.487	2,1
	<b>Aarhus</b>	5.043	3,5
	<b>Western Jutland</b>	8.435	1,8
<b>Northern Jutland region</b>	<b>Northern Jutland</b>	5.345	1,1
	<b>Aalborg</b>	3.379	2,4

Data is received monthly – at the latest on the 15<sup>th</sup> in the month following the month that the data concerns.

## 2.3 Market data

The market data for the register is supplied by two private companies who operate in the market for commercial properties and by the Danish Property Federation. All three sources have been paid to supply data and contracts have been written to ensure the confidentiality of data.

### The Danish Property Federation

The Danish Property Federation is a professional organization for owners, managers and administrators of commercial property in Denmark. The members of the organization cover about 85 pct. of the Danish professional, commercial property market measured in property value. It is from these members, that data is collected. The Danish Property Federation is the only data source in the register that is paid to collect new data. This means that it is possible to order collection of data e.g. from a specific region and thus increase representativity in areas or types of property that are scarcely covered by the remaining data sources. This in relation to geography, size, type, etc. Data contains, among other things, information on rent, operating costs, m<sup>2</sup> and vacancy for housing, hotels/restaurants, offices, warehouses and properties for production. The periodicity of data depends on what the current order asks for at a given point in time.

Currently, data for 56.345 unique units has been collected. These units are spread over 5.237 unique buildings. The geographical coverage of data is stated in the table below:



**Table 6 Geographical coverage of data from the Danish Property Federation**

<b>Region</b>	<b>Province</b>	<b>Number of unique buildings in the new register</b>	<b>Share of total commercial property in province, pct.</b>
<b>Capital region</b>	<b>City of Copenhagen</b>	625	0,5
	<b>Copenhagen Suburbs</b>	503	0,2
	<b>Northern Sealand</b>	464	0,1
	<b>Bornholm</b>	34	0,1
<b>Sealand region</b>	<b>Eastern Sealand</b>	96	0,1
	<b>Roskilde</b>	126	0,2
	<b>Western and Southern Sealand</b>	494	0,1
<b>Southern Denmark region</b>	<b>Funen</b>	517	0,1
	<b>Odense</b>	203	0,2
	<b>Southern Jutland</b>	346	0,1
	<b>Esbjerg</b>	50	0,1
	<b>Kolding</b>	56	0,1
	<b>Vejle</b>	130	0,1
<b>Central Jutland region</b>	<b>Eastern Jutland</b>	159	0,0
	<b>Randers</b>	48	0,1
	<b>Aarhus</b>	263	0,2
	<b>Western Jutland</b>	297	0,1
<b>Northern Jutland region</b>	<b>Northern Jutland</b>	150	0,0
	<b>Aalborg</b>	676	0,5

### **Ejendomstorvet**

Ejendomstorvet is an online portal for listing properties for rent and/or sale. It is the largest of its kind in Denmark gathering listings from more than 200 brokers from all over the country. Data contains physical characteristics for the objects offered for rent or sale, yield percentage, rental or sales price and operating costs. Housing, hotels/restaurants, offices, warehouses and properties for production are all represented in data.

Ejendomstorvet is paid to supply their data. Statistics Denmark receive data on a monthly basis – at the latest on the 15<sup>th</sup> in the month following the month that the data concerns. Data so far contains 45.860 unique units dispersed onto 8.837 unique buildings. The table below states the geographical coverage of data:

**Table 7 Geographical coverage of data from Ejendomstorvet**

<b>Region</b>	<b>Province</b>	<b>Number of unique buildings in the new register</b>	<b>Share of total commercial property in province, pct.</b>
<b>Capital region</b>	<b>City of Copenhagen</b>	1.511	1,2
	<b>Copenhagen Suburbs</b>	1.202	0,5
	<b>Northern Sealand</b>	1.168	0,3
	<b>Bornholm</b>	52	0,1
<b>Sealand region</b>	<b>Eastern Sealand</b>	396	0,3
	<b>Roskilde</b>	324	0,6
	<b>Western and Southern Sealand</b>	1.437	0,2
<b>Southern Denmark region</b>	<b>Funen</b>	509	0,1
	<b>Odense</b>	812	0,7
	<b>Southern Jutland</b>	1.065	0,3
	<b>Esbjerg</b>	323	0,4
	<b>Kolding</b>	373	0,5
	<b>Vejle</b>	375	0,4
<b>Central Jutland region</b>	<b>Eastern Jutland</b>	1.060	0,2
	<b>Randers</b>	370	0,5
	<b>Aarhus</b>	1.076	0,7
	<b>Western Jutland</b>	883	0,2
<b>Northern Jutland region</b>	<b>Northern Jutland</b>	768	0,2
	<b>Aalborg</b>	656	0,5

#### **Lokalebasen**

Lokalebasen is also an online portal for listing objects for rent and/or sale. Lokalebasen gathers listings for commercial properties in Denmark using web scraping and direct advertising on their website. Housing, hotels/restaurants, offices, warehouses and properties for production are all represented in data.

Lokalebasen have been paid to supply their data. They have supplied one data set for listings and one for realized sales/leases. As most of the property on the market is listed on Ejendomstorvet.dk, very little new information is provided by Lokalebasen. Thus, from July 2018 and onwards no new data will be collected from Lokalebasen. Listings data contains 106.794 unique units spread over 8.035 unique buildings. Data on realized sales and leases consists of 4.483 unique units dispersed onto 1.014 unique buildings. Table 8 and 9 below show the geographical coverage of data for listings and realized sales and leases, respectively:

**Table 8 Geographical coverage of listings data from Lokalebasen**

Region	Province	Number of unique buildings in the new register	Share of total commercial property in province, pct.
<b>Capital region</b>	<b>City of Copenhagen</b>	960	0,8
	<b>Copenhagen Suburbs</b>	634	0,3
	<b>Northern Sealand</b>	673	0,2
	<b>Bornholm</b>	41	0,1
<b>Sealand region</b>	<b>Eastern Sealand</b>	226	0,2
	<b>Roskilde</b>	170	0,3
	<b>Western and Southern Sealand</b>	751	0,1
<b>Southern Denmark region</b>	<b>Funen</b>	331	0,1
	<b>Odense</b>	414	0,4
	<b>Southern Jutland</b>	607	0,1
	<b>Esbjerg</b>	187	0,2
	<b>Kolding</b>	184	0,2
	<b>Vejle</b>	177	0,2
<b>Central Jutland region</b>	<b>Eastern Jutland</b>	581	0,1
	<b>Randers</b>	195	0,3
	<b>Aarhus</b>	561	0,4
	<b>Western Jutland</b>	585	0,1
<b>Northern Jutland region</b>	<b>Northern Jutland</b>	423	0,1
	<b>Aalborg</b>	335	0,2

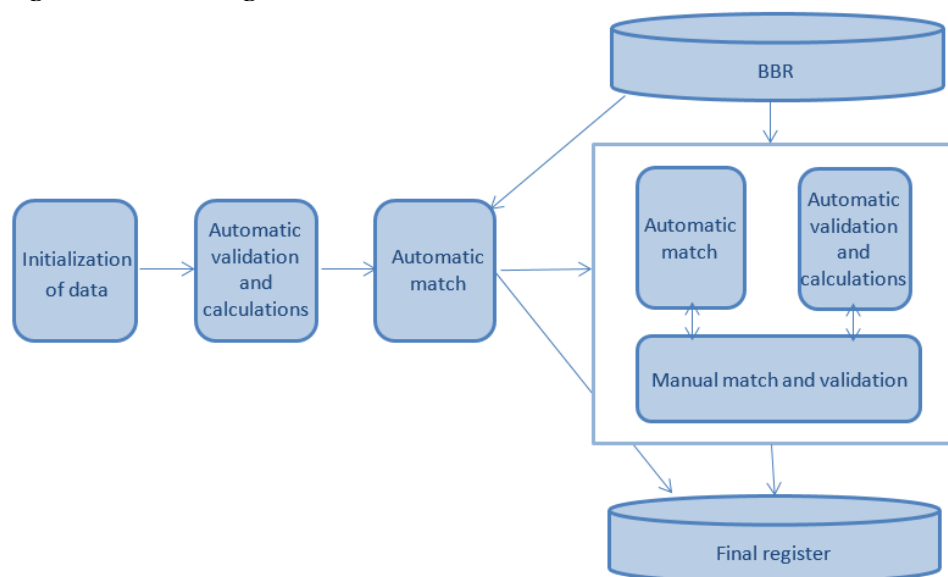
**Table 9 Geographical coverage of realized sales/rentals data from Lokalebasen**

Region	Province	Number of buildings in register*	Share of total commercial property in province, pct.
<b>Capital region</b>	<b>City of Copenhagen</b>	271	0,2
	<b>Copenhagen Suburbs</b>	196	0,1
	<b>Northern Sealand</b>	122	0,0
<b>Sealand region</b>	<b>Eastern Sealand</b>	45	0,0
	<b>Roskilde</b>	36	0,1
	<b>Western and Southern Sealand</b>	85	0,0
<b>Southern Denmark region</b>	<b>Funen</b>	23	0,0
	<b>Odense</b>	94	0,1
	<b>Southern Jutland</b>	49	0,0
	<b>Esbjerg</b>	14	0,0
	<b>Kolding</b>	37	0,0
	<b>Vejle</b>	29	0,0
<b>Central Jutland region</b>	<b>Eastern Jutland</b>	44	0,0
	<b>Randers</b>	15	0,0
	<b>Aarhus</b>	131	0,1
	<b>Western Jutland</b>	51	0,0
<b>Northern Jutland region</b>	<b>Northern Jutland</b>	36	0,0
	<b>Aalborg</b>	35	0,0

### 3. Data validation and match to the Danish Buildings and Dwellings Register

As the data structure is unique for each source it has been necessary to develop separate individual systems for validation and enrichment for each of them. However, they all follow the same template. The figure below outlines the general data flow through the system.

**Figure 2 Outline of general data flow**



#### **Initialization**

Data is initialized into the system, all raw data is stored in a separate table, and duplicate records are removed. Furthermore, observations are flagged if they have empty or invalid fields and it is noted whether a record has been received at an earlier point in time. All unique observations are then forwarded to validation and further calculations.

#### **Automatic validation and calculations**

The validation process is adapted to each data source, but generally data is reviewed for extreme values, and logical checks are made to ensure that the values are reasonable. It is e.g. checked if the rental price is reasonable compared to the size of the rental object and its operating costs. Negative signs are changed to positive.

Calculations are made to create some uniformity across the data sources. E.g. rent per m<sup>2</sup> is calculated if it does not already exist in data. All corrections and calculations are flagged.

#### **Automatic match**

In the next step it is attempted to match all observations to the Danish Buildings and Dwellings Register (BBR). This is done to further validate the data and to enrich it with data from the BBR. Data can be matched on three levels; unit, building and/or property as illustrated by figure 1 in chapter 1.

If an administrative key is available at unit level, this will be the primary match variable. However, in most cases the address is used as the primary match variable. Before any attempt to match, the address is validated using Danish Addresses Web API (DAWA). DAWA is a web service provided by The Danish

Agency for Data Supply and Efficiency and it uses a database that contains all valid addresses in Denmark. When an address is sent to DAWA, it will return an official address ID if possible. These address IDs will correspond to an ID in the BBR. It will also return an estimate of how precise the ID matches the address queried.

The type of property and the m<sup>2</sup>, which are both included in the BBR, are used as secondary match variables to validate the match. I.e., when a match is found at the unit or building level using the address, the reported m<sup>2</sup> of the observation is compared to the value stated in the BBR. If the m<sup>2</sup> are consistent within 10 pct., the match is said to be correct. If the m<sup>2</sup> reported for the observation is between 10 and 20 pct. larger than the m<sup>2</sup> stated in the BBR, the match is preliminarily accepted and can only be deemed as a correct match manually. If a match is deemed correct, either automatically or manually, the reported type of property is then compared between the original data and the BBR. If these are consistent the quality of the match is characterized as 'good'. If there are minor deviations the quality is characterized as 'medium', whereas major deviations result in a characterization as an 'inferior' match.

If the reported m<sup>2</sup> is 20 pct., or more, larger than the m<sup>2</sup> in the BBR, the match cannot be accepted and it is attempted to match at a higher level, i.e. building or property level. If an automatic match is not possible on any level, the observation will be treated manually.

#### **Manual match and validation**

In the manual treatment there are three possible outcomes:

1. The cause of the failure to match correctly is corrected and the matching process is repeated
2. The data cannot be corrected but the match, however, is assessed to be correct and the match is forced through
3. The data cannot be corrected, the match is assessed to be incorrect and the match is given up

If there are no match variables available for the observation the match is given up automatically and the observation will not be treated manually.

#### **Preliminary match results**

Throughout the development of the register, Statistics Denmark has received historical data from the sources for different time periods. These data have been initialized, validated, however only automatically, and matched. Table 10 below shows the match results pertaining to unique observations in data.

**Table 10 Results of match process for historical data without manual validation of data – unique observations**

Data source	Time period	No. of unique observations	Match percentages			
			Unit level	Building level	Property level	No match
<i>Administrative data</i>						
The Danish FSA	2012-2018	1.804.875	69,1	1,2	29,4	0,3
The Danish Building and Property Agency	2017-2018	339	10,3	19,5	68,1	2,1
Landsbyggefonden	2013-2018	575.160	99,6	0,0	0,0	0,4
<i>Market data</i>						
The Danish Property Federation	2017-2018	56.345	52,3	19,6	28,1	0,0
Ejendomstorvet	2014-2018	45.860	15,0	16,4	36,6	32,4
Lokalebasen - supply	2014-2018	106.794	14,7	17,5	44,3	23,5
Lokalebasen - realised	2014-2018	4.483	16,8	12,5	25,5	45,2

## 4. Discussion

The following discussion will fall in three parts; one part concerning the advantages and disadvantages of using administrative data sources for producing commercial real estate indicators and a second part concerning the same issues for market data sources and a final part describing the possibilities for producing new commercial real estate indicators on the basis of data from the new register.

### 4.1 Advantages and disadvantages of administrative data sources

There are some obvious advantages connected to using administrative data. As the data generally is collected, stored and maintained by public authorities, they will often be available to the NSI free of charge and it is relatively easy to gain access. However, in the case for data from the FSA regarding the mortgage institutions' lending offers, Statistics Denmark has not been granted permission to use the data for statistical purposes. This precaution has, most likely, been taken to ensure the continuous supply of data from the mortgage institutions as data can be of high sensitivity to them.

Administrative data is collected voluntarily and regularly, or even automatically through administrative IT-systems, and the variables are consistent over time as they follow strict definitions. Administrative data often covers the total population of what it is registering, as is the case with the data from Landsbyggefonden and The Danish Building and Property Agency, which cover all rental property within the public housing sector and all property used for governmental purposes, respectively. Furthermore, there will be very few missing observations in administrative data.

Moreover, this type of data will hold administrative keys which can be used to link data to other administrative data and thus expand the possible uses of the data beyond the register itself. This is the case for all three administrative data sources described above, bringing the match percentages with BBR close to 100 pct. as is depicted in table 10. This match to BBR is an important factor in validating and enriching data correctly.

For statistical purposes, in the case of commercial real estate indicators, there are also some disadvantages to using administrative data. It will sometimes be the case that the population described by the administrative data does not completely mirror the target population for the statistics. E.g. of the three data sources described above only one contains information on actual commercial properties, i.e. the FSA. The remaining sources, Landsbyggefonden and the Danish Building and Property Agency, hold information that can be used for approximating the dynamics of the commercial real estate market. And of these two, the Danish Building and Property Agency have too little data to be useful on its own. However, data from Landsbyggefonden is expected to produce a decent proxy for rent per m<sup>2</sup> as the developments for rent in social housing and rent in commercial property can be expected to follow the same trend although the data from Landsbyggefonden will be biased towards housing of a more moderate quality than what is the case for the private rental market.

## **4.2 Advantages and disadvantages of market data sources**

As described above Statistics Denmark collect market data from two different types of sources; online portals for listing property for rent or sale and the Danish Property Federation, a professional organization for owners, managers and administrators of commercial property. There are some differences in the advantages and disadvantages for these two types of data sources.

Generally, the main advantage of using market data for calculating commercial real estate indicators is, that these data directly reflects the market for commercial properties. Two of the three market data sources hold listings data that contain market rents and market operation costs as estimated by the property suppliers. Thus, these data should give a good indication of the dynamics of the commercial property markets.

There are, however, some disadvantages. Firstly, as the data is owned by private companies it will often cost money to obtain access to data, unless the NSI can claim this access legally. This could, however, be an advantage in disguise as it in some cases will give the NSI, as a paying customer, an opportunity to make some demands with regards to data, as is the case for the Danish Property Federation. In this case Statistics Denmark has been involved in specifying which data should be collected (including administrative keys), how the variables are defined, level of validation before it is transferred to Statistics Denmark, geographical coverage and the types of property represented in data. This gives Statistics Denmark a way to ensure representativity of data.

In the case for Lokalebasen and Ejendomstorvet, the data is collected via property suppliers listing their property for rent or sale, and data might be inconsistent across observations, as it is up to the supplier to interpret the definitions set out by the online portals. E.g., it is often impossible to identify whether a rent has been reported as net or gross. Furthermore, market data does not hold any administrative keys which makes it difficult to link it to administrative registers. This is reflected in the low match percentages to the BBR stated in table 10 for Lokalebasen and Ejendomstorvet. For these data sources addresses are used for matching the data to the BBR. However, as the addresses stem from listings they will have spelling errors and the like making the match to the BBR uncertain. This in turn makes it difficult to validate and enrich data, thus, large amounts of data may be lost due to the lack of quality with regards to production of statistics.

## 5. Indicators to be developed

### 5.1 The demand for new indicators

Demands for commercial real estate indicators have been increasing in the wake of the financial crisis of 2018. In their *2015 Report on Commercial Real Estate and Financial Stability in the EU*, the European Systemic Risk Board (ESRB) expressed a need for more and better data on commercial real estate markets. This is supported by the ECB and the European Commission, amongst others. One of the priority areas are assessing the dynamics of supply and demand<sup>1</sup>. This demand can in part be met using the data from the new register on market data for commercial property. The ESRB describes the cycles on the commercial real estate market in four phases; recovery, boom, decline and bust.

In the **recovery phase** the demand for commercial property is rising. It can only partially be covered by existing property, thus the remaining demand must be met by constructing new property. This creates a lag in the supply of properties which causes both the rental and sales prices to rise and the lending criteria to tighten. Vacancy rates are low.

During a **boom phase** the demand for commercial property is high. The supply is increasing and so are the prices. Sales prices are increasing faster than rental prices. The rising sales prices are substantiated by a large lending supply and the banks and mortgage institutions are competing by loosening their lending criteria.

In the **decline phase** the demand for commercial property is decreasing whilst the supply continues to increase due to the construction of new property. This leads to decreasing prices on commercial property. As a result the banks and mortgage institutions reduce their lending supply and lending criteria are tightened which leads to further price reductions.

During the **bust phase** the tendencies from the decline phase are continued. The demand is still too low to absorb the supply which results in decreasing prices and high vacancy rates. Mortgages and loans are scarce which makes new investments in commercial properties risky.

Thus, indicators describing the dynamics of supply and demand on the commercial property market could be:

- i. Commercial property prices
- ii. Rental prices on commercial property
- iii. Vacancy rates
- iv. Commercial property for rent or sale
- v. Building permits for commercial property
- vi. Lending supply and criteria

### 5.2 Development work in Statistics Denmark

For the coming two years the plan is to start developing indicators that match the needs listed above. This means that Statistics Denmark will develop new indicators and substantiate existing statistics further, among other things, by implementing the use of the new register for market data on commercial property.

---

<sup>1</sup> European Systemic Risk Board (2015): Report on commercial real estate and financial stability in the EU, December



For this work Statistics Denmark has received grant funds from the EU and take part in the Eurostat task force, TF CREI.

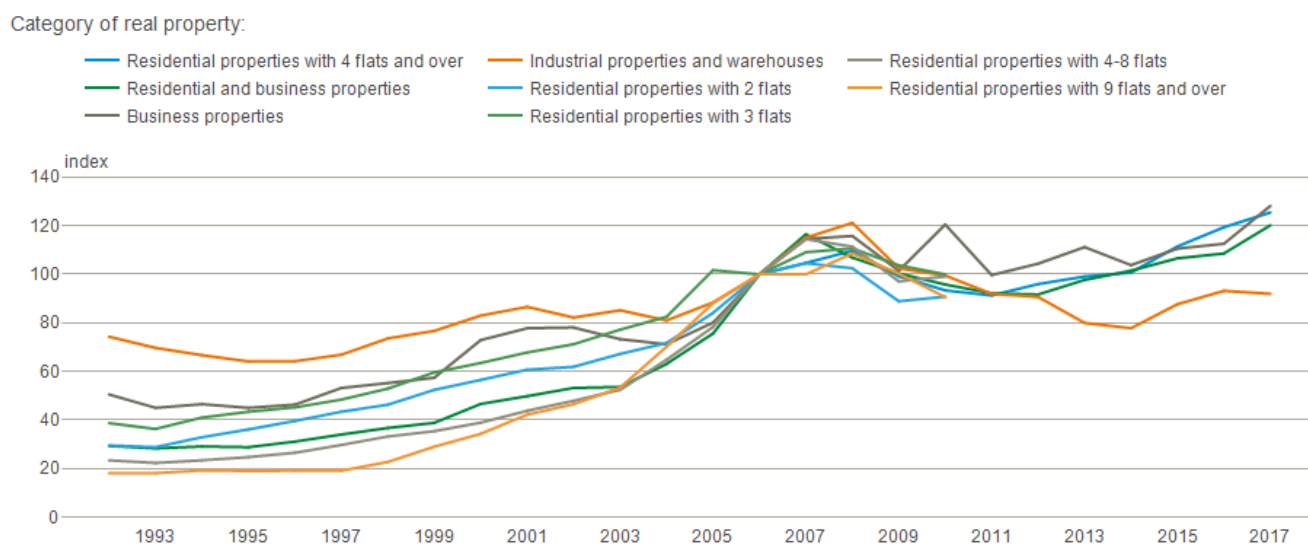
The development of these new indicators listed above highly depends on the level of representativity that can be obtained in data and the estimated consistency of variables, both across and within the data sources. These analysis have yet to be carried out. In some cases the coverage of data may have to be expanded further. Thus, some work will be put into exploring new data sources.

It is the intention that all new indicators developed will be included in the statistical program for Statistics Denmark and published.

### i. Commercial property prices

Currently, Statistics Denmark is producing a CPPI using data from the Danish Land Registry. The CPPI is published quarterly and yearly, and the time series goes back to 1992 with 2006 as base year.

**Figure 3 Price index for sales property, 1992-2017 (2006=100)**



Source: Statistics Denmark - [www.statistikbanken.dk/EJEN6](http://www.statistikbanken.dk/EJEN6)

The land registry data covers about 85 pct. of the transactions of commercial properties. The remaining 15 pct. consist of transactions of commercial property within transactions of companies. It has not yet been investigated, but the new register on market data for commercial property may contain data that might help approximate the price developments for these remaining 15 pct. of the transactions. If the data is available it will also be necessary to obtain weights in order to construct a total index. It should be established what measure would be the most fit for weighting the indices, e.g. number of transactions, m2 transacted, the value of the properties transacted or a fourth measure.

### ii. Rental prices on commercial property

Currently, Statistics Denmark produce a yearly price index for the rent per m<sup>2</sup> on housing as part of the consumer price index. This index will be expanded by using data from the register for market data on commercial property and it will become a quarterly statistic from 2020. Furthermore, it is expected that the new register will make it possible to construct an index for the rental prices commercial properties as a whole and potentially stratified according to geography and sector. The register contains relevant monthly data from 2014 and onwards.

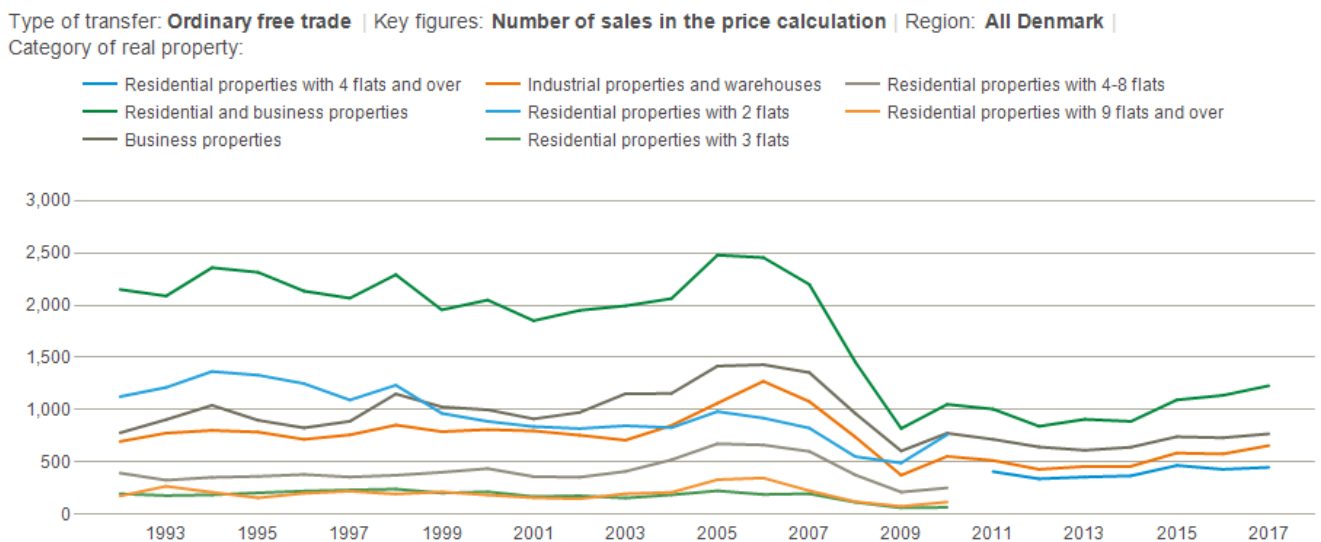
### iii. Vacancy rates

The Danish Property Federation currently produce a vacancy rate statistics. Preliminary plans are made for Statistics Denmark to overtake this production. The rates are calculated on the basis of data reported by both major, medium and minor property investors covering about 50.000 leases. The rates are produced quarterly for both financial and areal vacancy. The vacancy rates are calculated as the share of vacant leases of total available leases at the time of calculation. The Danish Property Federation stratify the statistics according to both geography and sector. It has not yet been decided how the production will be implemented in Statistics Denmark but this is expected to be established during 2019.

### iv. Commercial property for rent or sale

For the time being, Statistics Denmark produce quarterly statistics on the number of sales of commercial property stratified according to geography and sector.

**Figure 4 Sales of real property, 1992-2017**



Source: Statistics Denmark - [www.statistikbanken.dk/EJEN88](http://www.statistikbanken.dk/EJEN88)

The statistics is based on the same data as described for the CPPI. Thus, if it is possible to identify the remaining 15 pct. of transactions of commercial property, this statistics will too be substantiated further.

Moreover, it is the aim to produce statistics for the supply of properties for both sale and rent by using data from the register on market data for commercial property. The register contains relevant monthly data from 2014 and onwards.

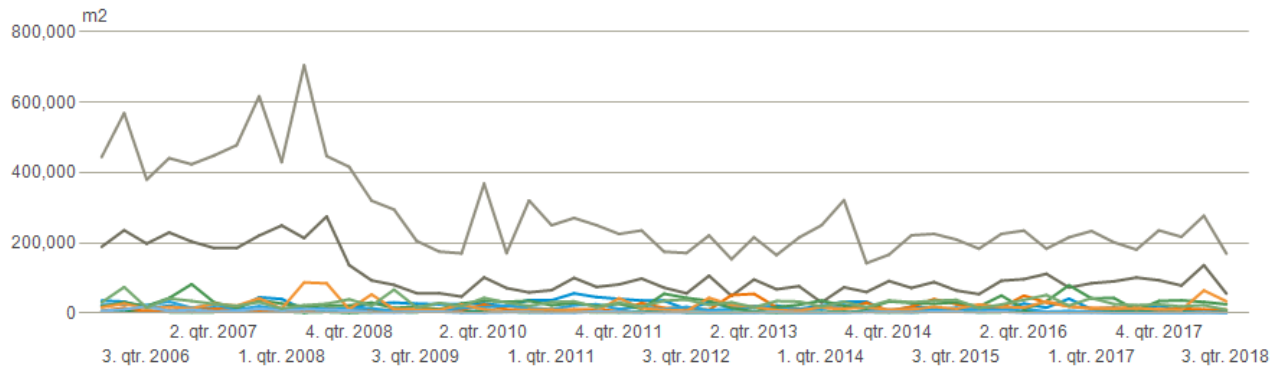
### v. Building permits for commercial property

Statistics Denmark calculate construction statistics on a quarterly basis including number of building permits stratified according to use. The statistics is based on data from the BBR and goes back to 2006Q1.

**Figure 5 Permitted construction, 2006Q1-2018Q3**

Phase of construction: **Permits** | Region: **All Denmark** | Use:

- Residential buildings for communities
- Other residential building
- Factories, workshops, etc.
- Power stations, gaswork, etc.
- Other building used for production
- Transportation or garage
- Office, trade, inventory, incl. public administration
- Hotel, restaurant, hair dresser and other services
- Unspecified transport and trade
- Sports centres, club houses



Source: Statistics Denmark - [www.statistikbanken.dk/BYGV11](http://www.statistikbanken.dk/BYGV11)

**vi. Lending offers and lending criteria**

For the time being, calculation of statistics on the lending supply and criteria is out of scope for the development work planned in Statistics Denmark as the data is not available through the new register for market data on commercial property or in Statistics Denmark in general. In Denmark the National Bank is the producer of statistics on Money and Credit Markets.

**6. Closing remarks**

The above is a possible take on how to accommodate the demands for new commercial real estate indicators. However, inputs on other indicators to be produced are very welcome. It has been four years since the ESRB came with their recommendations and the demands may have shifted or expanded since then. In the development process of these statistics national authorities such as the National Bank and the Ministry of Finance and others such as the Confederation of Danish Industry, will also be invited to give their input with regards to new data sources and their wishes for new indicators in the field of commercial property.