

Developing a new commercial property statistical system (CPSS) for Ireland

Ottawa Group, 10-12 May 2017, Eltville, Germany

Barra Casey (CSO Ireland) and Sandra Tobin (CSO Ireland)¹

Abstract

The recent crisis in the Irish property market has highlighted the need for comprehensive, objective and independent commercial property statistics to provide greater insight and transparency into the commercial property market in Ireland. At international level, the development of a methodological framework for Commercial Property Price Indicators is one of the key priority actions identified by the G20 Data Gaps Initiative (DGI).

In Ireland, the recent development by CSO of statistics in the area of residential property has shown the extra insight that can be gained when administrative data is combined from multiple sources. For commercial property, a variety of administrative data sources already exist. The key challenge is integrating the data sources into a coherent consistent system.

The Irish Central Statistics Office (CSO) began a project in July 2016 (funded and supported by Central Bank of Ireland and the National Asset Management Agency²) to ascertain whether it was possible to develop a Commercial Property Statistical System (CPSS). The aim was that CSO would design and develop statistics on commercial property covering the overall stock, the stock under construction (referred to as pipeline in this paper) and the transactions (sales and leases) of commercial property.

A key aspect of the project was the ability to match different data sources using unique identifiers and also using a technique known as fuzzy matching. Analysing different data sources available in Ireland, it was found there is potential to build a pipeline database of commercial property and a stock/register database of commercial property. However, a multitude of problems were found in the matching of transactions (sales and leases) data; namely a lack of unique identifiers in the datasets, fundamentally different units of measurement, a lack of detail in postal addresses and uncertainty on whether we are matching like-for-like. This is in direct contrast to the CSO's success in matching these data sets for residential property. The paper briefly explores other potential data sources.

Key words: commercial property, data matching, administrative data, Ireland

¹ The authors would like to thank the members of the CPSS Steering Group - Gerard Kennedy (Central Bank), Dermot Coates (Central Bank), Jamie Bourke (NAMA) and Tony Brennan (NAMA) - for helpful comments.

² The National Asset Management Agency was set up post-financial crisis and functions as a bad bank in Ireland, acquiring property development loans from Irish banks in return for government bonds.

Introduction

CSO launched a new Residential Property Price Index (RPPI) during September 2016. The new RPPI makes innovative use of four data sources – the stamp duty returns from the tax authorities, Building Energy Regulation (BER) certification data, the Geodirectory (national register of buildings) and the Census Small Area Population Statistics. By linking these sources together through data matching, the CSO has been able to improve both the quality and detail of the RPPI and compile comprehensive additional indicators (volume, value, mean prices, median prices broken down by detailed location, type of buyer etc.) on residential property. This illustrates the potential for obtaining new statistical insights, at aggregate level, from administrative records.

The project to develop a Commercial Property Statistical System (CPSS) stems directly from the success of the RPPI and the considerable experience built up in the organisation in relation to administrative records and in particular the matching of these administrative records. The potential was there to use the same methodology to develop statistics on commercial property. In that sense, Ireland is similar to the rest of Europe in that statistics on residential property are well-developed while statistics on commercial property have lagged behind.

As well as the under-development of the statistics themselves, it would follow that the knowledge of the subject matter of commercial property within the CSO was equally less developed. Therefore, the current project attempts to address that knowledge gap by creating institutional links with market participants (public and private), liaising with international colleagues and creating internal CSO resources in the area of commercial property.

The recent literature on Commercial Property Price Indicators (CPPIs) has highlighted that it is likely that a multi-source approach will be required for the development of a headline indicator on price developments. Kanutin (2013) notes that transaction-based data are likely to be the best theoretical approach, but cannot be the sole data source as in times of stress market liquidity tends to be low³. This point is reiterated in various other studies including Passerini (2013). This is compounded by the fact that the number of transactions in a small country like Ireland would be considerably less than some of our European neighbours. It is likely that transactions data will need to be supplemented by additional estimates of price information such as valuation data from private data sources.

With this in mind, the CSO were keen to maximise the use of public data sources from administrative systems and develop our knowledge of commercial property before attempting to exploit private data sources. This project, therefore, completes an inventory of all potential data sources and assesses their potential to be matched together and used for the production of statistics. The project does not limit itself to transactions data but also explores pipeline and stock information. The project can be seen as the first step by official statistics in Ireland to fill the data gaps in the commercial property sector highlighted by the recent financial crisis.

Section 1 recalls the experience in Ireland of falling property prices post-2007 and their impact on the economy. Section 2 deals briefly with what data sources currently exist. Section 3 looks at the scope of the project and the potential statistical outputs. Section 4 deals with the methodology used for the data matching. Section 5 looks the data sources used in the project and the lessons that may

³ In Ireland, the number of commercial property sales in 2016 was 7.5 times the number of sales in 2010 (the trough in the market)

be applicable to a wider international audience. Section 6 reflects on some practical lessons from the project that may add some value to the international debate on the definitions for commercial property. Section 7 concludes.

1. Falling property values in Ireland

Ireland experienced a particularly difficult period post-2007 as property prices for both residential and commercial property fell dramatically. Figure 1 shows annual % changes in the CSO's Residential Property Price Index (RPPI) and annual % changes in two value indices related to commercial property in Ireland. The first index is from the Investment Property Databank (IPD) and uses total returns⁴ as an indicator of overall investment performance. The second index is from Jone Lang LaSalle (JLL) and only uses the capital value to estimate investment returns. The RPPI is a constant quality price index produced by the CSO using a hedonic regression. While comparing price indices and value indices is not like for like, they are instructive to examine in a general sense.

Although residential property experienced falling prices for a longer period than commercial property (sixteen quarters compared twenty-one quarters), the falls were much deeper for commercial property. By the first quarter of 2009, commercial property returns had fallen annually by 39.5% whereas residential property prices fell by a mere 20.8% annually (Q2 2009). The peak-to-trough fall for residential property was 54.4%, however commercial property fell further by 67.4%.

Likewise, as property prices recovered from 2013, it was commercial property that experienced the greater rises. This shows the extent to which commercial property is more cyclical in nature than residential property. In 2016, the annual changes in residential property prices and commercial property returns began to converge towards an annual rate of about 10%.

As one would expect, these dramatic changes in asset values had profound effects on the balance sheets of the banks. Adverse movements in property markets can generate both liquidity and solvency problems for banks. Deterioration in banks' health, due for example to a rise in non-performing loans or withdrawal of wholesale funding, in turn forces banks to take corrective actions, e.g. by tightening credit availability. This has spillover effects on the economy, which feeds back to the property market and to the financial position of the banks.

Figure 2 shows that, proportionately, the commercial property loan book in Irish domestic banks has been the most distressed asset class since the financial crisis. Despite the sizeable reduction in the proportion of impaired loans since 2014, at the end of Q3 2016, 36% of commercial property loans were still non-performing.

The National Asset Management Agency (NAMA) is a body created by the government of Ireland in late 2009, in response to the financial crisis and the deflation of the Irish property bubble. NAMA functions as a bad bank, acquiring property development loans from Irish banks in return for government bonds.

⁴ It takes account of both capital and rental elements of a property's value as the percentage capital value change plus net income accrual, relative to capital employed.

A total of €74 billion in loans had been transferred to NAMA by the five participating institutions and €31.8 billion has been paid as consideration to the institutions, an overall discount of 57%. The discount rate varied between 43% and 61% depending on the institution involved. There was a high concentration of debtors in relation to the loans transferred to NAMA with 772 debtors accounting for the full €74 billion in loans - an average loan of €96 million. The top 12 debtors accounted for over €22 billion in loans – an average loan of over €1.8 billion.

Figure 1: Residential Property Prices and Total Returns on Commercial Property (annual % changes)

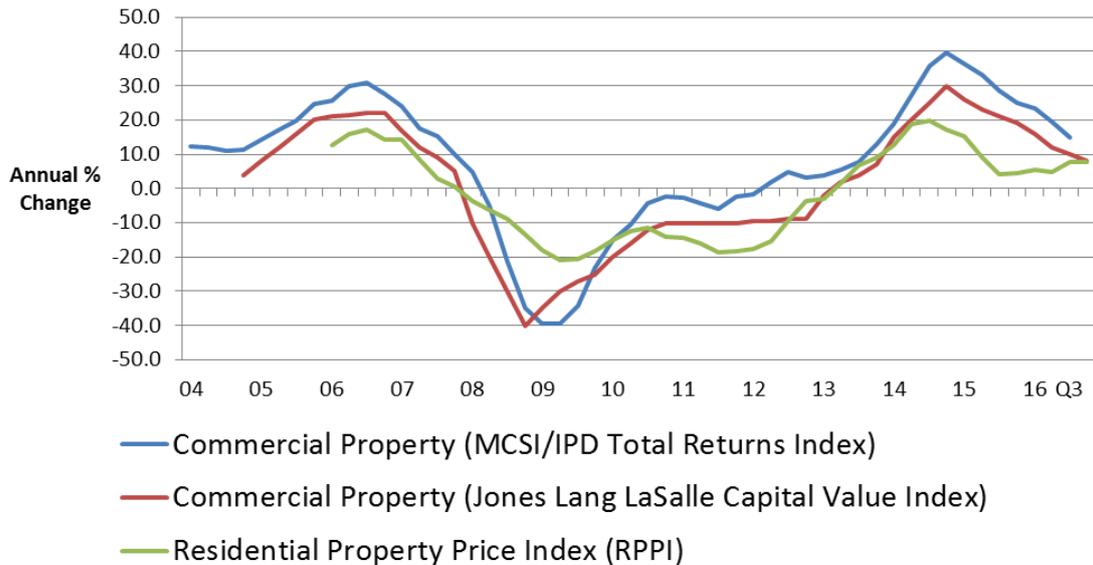
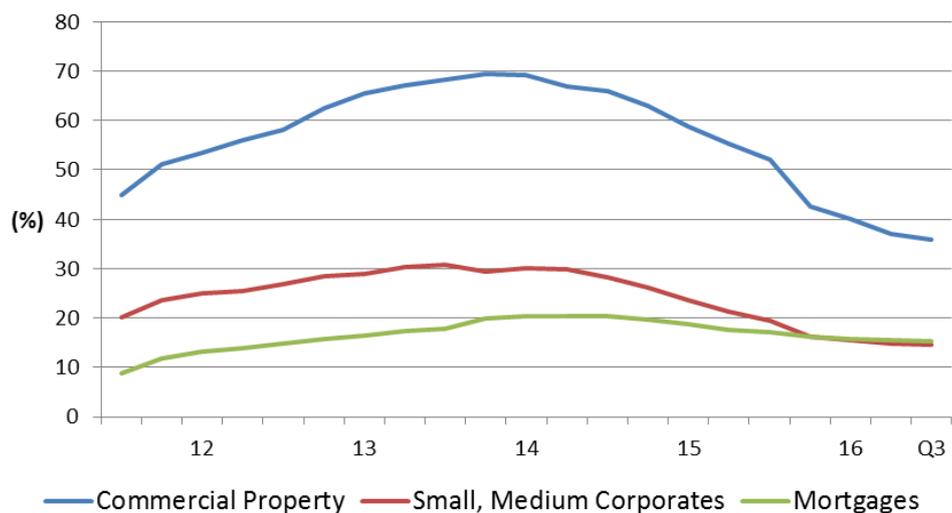


Figure 2: Domestic banks' non-performing loans as a share of loans to sector



2. Existing data sources

Like many other countries, data on physical commercial property is limited to private data sources. With the exception of limited planning permissions data, the CSO currently does not publish any official data on commercial property.

Private companies who provide value indices for Ireland include the Investment Property Databank (IPD) which is owned by Morgan Stanley Capital International (MSCI) and Jones Lang LaSalle (JLL). The JLL index has lower coverage in the market than the IPD index but both sources have followed a similar trend over the years (see Figure 1). These data sources although established for many years, are not designed to fulfil some key principals of the European Statistics Code of Practice⁵ such as impartiality, comparability and accessibility. In fact, it is unfair to judge them against these principals as the indices were created in order to provide a client-based information service which is commercially driven.

There are various other private market participants who produce analysis of the commercial property market. The information is disaggregated across different sources and it is difficult to get an overall picture of the commercial property sector. Also the definitions and coverage can vary which means that the coherency and consistency are sub-optimal.

There is a clear role here for official statistics to provide independent, harmonised statistics on commercial property.

3. Scope of the project

The data requirements for the commercial property sector are quite wide ranging. Recommendation C from the European Systemic Risk Board's (ESRB) report on closing real estate gaps outlines the need for indicators on the physical commercial property market, indicators on the financial systems' credit exposures to commercial property and indicators on lending standards. The current project aims to produce recommendations to fill the data gaps for the physical commercial property market. In the ESRB report the key indicators listed under this heading are as follows:

1. Price index
2. Rental index
3. Rental yield index
4. Vacancy rates
5. Construction starts

The current project will address the data requirements under all five key indicators. There is no scope within the current project to develop index compilation methods for indicators 1 to 3 above, this work will be completed once the data sources have been fully examined. Therefore, the objective of the project is to complete an inventory of administrative data sources and assess their

⁵ <http://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/KS-32-11-955>

capability to meet the data requirements listed above. There is also the potential that the data sources available will go beyond the indicators listed here into other relevant areas of commercial property.

Specifically there are three strands to the current project. These are as follows:

- Pipeline
- Stock and
- Transactions

The three strands of the project are laid out graphically in Appendix 1. These strands attempt to measure the different activities during the life cycle of a property.

The pipeline data captures the planning permissions issued by the national authorities and the progression from commencement of construction onto completion.

Once a new building is completed, it moves into the stock of commercial property and then is either used for commercial purposes or lies vacant. During its use (or vacancy) a building can be extended or renovated and this activity can be captured through the planning authorities.

A building can also be transacted (sold or leased) during its use or vacancy. The prices that exist in the space market (rentals) and the asset market (sales) reflect the supply and demand in both. These markets are interconnected in that the expected rental income stream from commercial properties is a key input into the valuation of that property.

Finally, a building may fall out of the stock of commercial property through demolition which frees up development land for further construction.

Due to pro-cyclical nature of commercial property, a key requirement from our users was to produce early indicators of over/under supply (and likewise strong/low demand) of commercial property in the market. This would help isolate the turning points in the market in a timely manner and provide policymakers with information to take preventative measures if required. Therefore, combining the pipeline information with the stock and transactions of commercial property would provide users with a complete picture of the physical commercial property market.

Potential outputs

Appendix 2 outlines in more detail the potential outputs from the Commercial Property Statistical System (CPSS). These represent the original aspirations of the project; however, as will be outlined later in the paper, not all these outputs will be possible in the short term.

The potential outputs are listed here:

Pipeline:

- No. of Planning Permissions Classified by Type and Location
- Area (m2) of Planning Permissions Classified by Type and Location
- No. of Commencements Classified by Type and Location
- Area (m2) of Commencements Classified by Type and Location

- No. of Completions Classified by Type and Location
- Area (m2) of Completions Classified by Type and Location

Stock:

- No. of Buildings/Units Classified by Type and Location
- No. of New Buildings/Units Classified by Type and Location
- Area (m2) Classified by Type and Location
- Vacancy Classified by Type and Location

Transactions:

- Sales of Buildings/Units Classified by Type, Location, New/Existing and Age
- Sales by Area (m2) Classified by Type, Location, New/Existing and Age
- Sale Price (per m2 or per transaction) Classified by Type, Location, New/Existing and Age
- Leases of Units/Buildings Classified by Type, Location and Age
- Leases by Area (m2) Classified by Type, Location and Age
- Rental price (per m2 or per transaction) Classified by Type, Location and Age
- Average Length of Lease Classified by Type, Location and Age

4. Methodology for data matching

In each of the strands of the project, data matching was a key element of the methodology employed. Data matching was not expected to be a major problem for the pipeline data (a planning reference number is used by planning authorities for both planning permits and building control databases) or the stock data (XY coordinates existed across datasets). However, an important question to be answered would be if data matching could be employed successfully for commercial property transactions similar to the way it was implemented for residential property transactions.

The key challenge in linking the price data on transactions (i.e. from the tax authorities) with data sources that contain detailed characteristic variables on the buildings was the lack of a common unique identifier. There was no single identifying reference number shared by the datasets. Although postcodes⁶ were introduced in Ireland in 2014 and a field was created in the relevant administrative systems to capture it, the uptake by the general public has been low and the field generally has not been populated in the relevant datasets. In the CSO, it is a key organisational priority to develop the use of postcodes across the administrative system. Postcodes or locational identifiers are one pillar of the National Data Infrastructure (NDI), along with personal identifiers and business identifiers.

⁶ Unlike some other countries, where postcodes define clusters or groups of addresses, in Ireland Eircodes (name given to postcodes) uniquely identifies an individual address – rural or urban - and shows exactly where it is located. An Eircode is a smart location code for all Irish addresses.

Therefore, there was no alternative to matching the transactions data except by using the postal addresses, a very difficult process to achieve computationally, given the considerable variation that exists in address spellings.

For the matching of addresses, an algorithm is used that measures Jaro-Winkler Distance. Jaro-Winkler Distance is a metric of similarity between two character strings, with a score of 0 implying zero commonality and a score of 1 implying an exact correspondence. A useful feature of Jaro-Winkler Distance as far as address matching is concerned is that it is possible to give extra weight to the particular characters of each string. For commercial property extra weight was given to the first four characters of the string and anywhere a number is found in the string. Urban addresses usually begin with a number (although it is less common for commercial property compared to residential property) and so it is particularly important that these should agree in the case of matching addresses.

The address string matching is a multi-step process. Firstly, addresses are 'cleaned' on the respective datasets (uncommon characters such as commas, apostrophes, full stops, etc. are stripped from the addresses etc.). Next the addresses are 'blocked' by county. Then the clean addresses are matched in the corresponding blocks across datasets. The Jaro-Winkler distance is calculated for each pairing. For residential property, the highest scoring match is automatically presumed to be a correct match, provided the score is above a threshold of 0.88. However, for commercial property it was decided to visually check all high scoring matches of greater than 0.70 and manually choose the correct one. This was possible due to the low frequency of transactions but it was also required due to the less precise nature of commercial property addresses. For example, it was often the case that a building's address would only include the street name and exclude street numbers.

5. Source data and data matching

Transactions:

Market activity in commercial property has two different aspects. There are sales of commercial property (leases over 100 years) and there are leases of commercial property (leases up to 100 years). Both the former and latter are captured on stamp duty data from the Irish tax authorities. The latter are also captured in a lease register held by the Property Services Regulatory Authority (PSRA)⁷ where additional information is captured (property type and floor area). Characteristic variables relating to properties both sold and leased are captured in the Building Energy Rating (BER) certification system.

The data sources, as outlined in Table 1, were envisaged to complement each other creating a robust dataset of commercial property market activity enabling the production of statistics on prices

⁷ The Property Services Regulatory Authority (PSRA) is the regulator of property service providers in the real estate market e.g. auctioneers, estate agents, letting agents and property management Agents. The PSRA has established a number of public registers as part of its functions including the Residential Property Price Register and the Commercial Leases Register. Both of these public registers are based on stamp duty returns from the Revenue Commissioners (tax authority in Ireland).

and rents by location, floor area and type of commercial property (office, retail, industrial). County Limerick was used as the test location.

Table 1: Data Sources: TRANSACTIONS of commercial property

	File 1	File 2	File 3	File 4
File name	Stamp Duty Form (SDR1)	Stamp Duty Form (SDR3)	Building Energy Ratings (BER) Data	Commercial Leases Register
Institution Source	Revenue Commissioners (tax authorities)	Revenue Commissioners (tax authorities)	Sustainable Energy Authority of Ireland (SEAI)	Property Services Regulatory Authority (PSRA)
Administrative Purpose	Conveyance or transfer ownership of property are chargeable to stamp duty and an SDR1 form must be filed	Leases of property are chargeable to stamp duty and an SDR3 form must be filed	Commercial buildings offered for sale or rent, require a BER certificate.	There is an obligation on the tenants of commercial leases to furnish certain information to the PSRA
Relevant legislation	Stamp Duties Consolidation Act 1999 (SDCA)	Stamp Duties Consolidation Act 1999 (SDCA)	Articles 5 and 7 of the EU Energy Performance of Buildings Directive (2002/91/EC).	Property Services (Regulation) Act 2011
Potential matching variables	Address, Eircode	Address, Eircode	Address, Eircode, BER Number	Address, BER Number
Key information variables for CPSS	Price and date of non-residential sale transaction, description of property	Price, date and term of non-residential lease transaction	Age of building, floor area and property type	Property type, floor area, date and term of lease, rent payable and rent review information
Details of test files	All stamp duty sales records (2010 – Q3 2016)	All stamp duty leases records (2010 – Q3 2016)	All non-domestic records made available	All commercial leases entered into (2010 – 2015)
Number of observations on test files	All - 27,500 (25,400 unique addresses) Limerick - 1,300 (1,200 unique addresses)	49,300 (43,500 unique addresses)	47,000 records (40,000 unique addresses)	40,700

Unfortunately, the matching of the key datasets (tax data with BER data) proved to be very difficult and resource intensive.

Data matching was not successful for the following reasons:

- There is a field on the stamp duty returns from the tax authorities for a postcode; however, this is not mandatory and invariably not completed. This is in contrast to residential property where the field is completed. The lack of a postcode (or another location identifier such as an XY coordinate) makes data matching significantly more difficult. The data matching relies entirely on the postal address in that case.

- The lack of detail in postal addresses for commercial properties makes it very difficult to match records. For example, the number of a building on a street is often omitted from the address (“Main Street” is a very common address in Ireland). One of the reasons that the stock/register data can be matched successfully (discussed in next section) is that the “trading name” is included on the file. The trading name is not relevant for transactions (only buyer and seller information available).
- Unit of measurement for the records in the stamp duty returns is fundamentally different to the records in the BER data. For example, sales records from the tax authorities typically relate to buildings or groups of buildings while BER records can relate to floors within buildings or building units (rated for rental purposes) or buildings themselves. When a match is found, it can be difficult to ensure that the correct floor area (and other characteristic variables) is being assigned to the property.
- In cases where a match is made, there can be occurrences where we are not matching like with like. For example, in the sale of a shopping complex, the anchor store or the cinema that is part of the complex may not be included in the sale and such information is not provided on the stamp duty return i.e. transparency issue. The assigned floor area in that case through data matching may be much higher than the relevant floor area of the premises sold. In other cases, there may be additional assets included in the sale of the property and again this is not identified and can result in inflated prices per square metre etc. There are also cases where a building may have separate deeds which can result or require two separate returns to be made and there is potential for counting the floor area twice in these cases.

Despite the data matching not being possible here, it is instructive to have carried out the exercise which will inform the next steps in the project. It is clear that matching stamp duty data with BER data is fraught with problems for commercial property despite how successful this method was for residential property; therefore, alternative methods to capture important characteristic variables such as floor will be necessary.

CSO are currently examining these options. Some options include 1) new surveys of market participants 2) the development of new administrative data sources with other public institutions and 3) exploring collaborative projects with institutions which hold private data sources.

Register/Stock of Commercial Property

Data from the GeoDirectory and the Valuation Office were considered for information on the existing stock as shown in Table 2.

The GeoDirectory is the most comprehensive and up-to-date source of information on commercial building stock. The Geodirectory is a service, jointly established by An Post (semi-state body for postal services) and Ordnance Survey Ireland (national mapping body) that provides a complete database of all of the residential and non-residential buildings in Ireland and their geolocation details. It holds records for approximately 1.8 million properties. In terms of location, the GeoDirectory includes the address of the property, XY coordinates and postcodes. The GeoDirectory also has a vacancy indicator so it would be possible to produce reliable vacancy rates and a NACE rev2 classification of the economic use of the properties. The GeoDirectory does not however have

direct information on the type of commercial properties. For example it does not categorise commercial property into office, retail, industrial and other uses. GeoDirectory also does not have information on the floor area (m2) of the properties.

The core business of the Valuations Office is the provision of accurate, up-to-date valuations of commercial and industrial properties to ratepayers and rating authorities⁸. The Valuation Office is currently revaluing all commercial properties in Ireland which is expected to be complete by 2020. The Valuation Office database is a particularly rich source of data for commercial property. The data includes the address of the property, the use of the property and the floor area (m2) of the property. The data has separate valuations for each floor of each building, further breakdowns of use categories (over 200 use categories captured) and separate valuations for various features, such as car spaces, etc. The data also contain XY coordinates.

The objective was to examine the matching of the GeoDirectory with the Valuations Office data to create a comprehensive data source of the stock of commercial properties in Ireland.

Table 2: Data Sources: REGISTER/STOCK of commercial property

	File 1	File 2
File name	Geodirectory	Valuations data
Institution Source	An Post Geodirectory Ltd	Valuations Office
Administrative Purpose	A register of residential, commercial and mixed use buildings	The provision of valuations of commercial and industrial properties to ratepayers and rating authorities
Relevant legislation	Not relevant – private data provider	Valuation Act 2001
Potential matching variables	Address, XY Coordinates	Address, XY Coordinates
Key information variables for CPSS	No. of buildings, number of new buildings, vacancy rates, and NACE code	Floor area and property type (category and use)
Exclusions i.e. stock not covered by data source	-	Buildings occupied by those who are exempt from paying rates
Details of test files	All available	Limerick city and county
Number of observations on test file	All – 221,000 address points (units) (151,000 buildings) Limerick – 10,000 address points (units) (6,700 buildings) Limerick – 9,500 unique addresses	Limerick - 16,200 records Records are at the level of the floor and category of use and occupier 6,300 unique addresses

The two data sources were matched using postal address and XY coordinates. The unit of measurement in both data sets is similar. The Valuation Office data is focused on the “occupier” while also recording information at a more aggregated level of the building. The data also has

⁸ The payment of rates by businesses is one of the main sources of income for local government in Ireland.

information below the “occupier” level which breaks the occupied space into different uses⁹. The Geodirectory also has the concept of buildings and units within buildings. The focus for the Geodirectory is on “address points”. The trading name of the company was also used as part of the address but the matching was also tested excluding the trading name. Limerick (county in Ireland) was used for testing purposes.

The matching process resulted in a match of greater than 85% between the Valuation Office file and the GeoDirectory file for Limerick. Approximately 30% were matched easily with either an address or an XY coordinate matching exactly. 45% were identified using the Jaro-Winkler fuzzy matching algorithm in SAS (Statistical Analysis Software) with manual acceptance/non-acceptance required for the suggested matches. The remaining 10%+ were matched using a manual search including some internet research which was a very resource intensive process.

Testing has shown that the matching of the Geodirectory data with the Valuation Office data is possible and would add value to information on the stock of commercial property. Such additional information would include characteristic variables for the properties such as floor area and property use.

Commercial properties which are not included in the Valuations Office data, due to exemptions, create a gap in the data. There is a need to assess additional sources for this data or provide for the collection of some basic information (floor area and property type) for such buildings.

A key question going forward for the project is the degree to which the assigned characteristic variables (such as floor area) from the Valuation Office are reliable and would be of publishable quality.

Pipeline/New Stock

Planning permissions, commencement and completion notices from the planning authorities and Local Government Management Agency (LGMA) was considered for information on new stock. These data sources are outlined in Table 3.

Local planning authorities and An Bord Pleanála¹⁰ send details of all planning permissions granted in each quarterly period to the CSO. Each permission granted details the planning registration number, the local authority it pertains to, a description of the permission granted, the number of units included and the floor area.

The Building Control Management System (BCMS) was implemented in March 2014 and allows property owners, builders, developers, architects and engineers to submit notifications, applications and compliance certificates online. The BCMS register includes the commencement date and the reference number of the original planning permission. The reference number of the completion certification is also given where the construction has since finished.

These two data sources were to be matched using the planning registration number available on both files. When merged, the CSO planning permission coding was expected to identify certain types

⁹ The reason for this is because rates are chargeable at different amounts depending on the use by the occupier.

¹⁰ An Bord Pleanála is national office responsible for the planning appeals process.

of relevant commercial properties, the category of building and the floor areas for works commenced or recently completed.

Although, matching of these files was not successful due to the incomplete capture of the planning reference number in the CSO, it is expected that this problem can be resolved.

A more pressing concern is the compliance rate for the completion of the building control notices i.e. commencement and completion notices etc. Compliance levels are not known however and need to be assessed.

If these problems can be rectified with the relative bodies, it is expected that valuable statistical data can be generated on pipeline commercial property. In fact, a similar process can be implemented for residential property.

Table 3: Data Sources: PIPELINE of commercial property

	File 1	File 2
File name	Planning Permissions	Building Control Management System (BCMS)
Institution Source	Local Planning Authorities and coded by the CSO	Local Government Management Agency (LGMA)
Administrative Purpose	Planning permission must be obtained from the relevant Planning Authority in relation to proposed building works. Details of the proposed building works are submitted for consideration.	Commencement and completion notices are required for Building Control Authorities
Relevant legislation	The Local Government (Planning and Development) Acts, 1963 to 2015	The Building Control Regulations, 1997 to 2015
Potential matching variables	Planning registration number	Planning registration number
Key information variables for CPSS	Floor area, number of units and property type	Commencement and completion date
Exclusions i.e. new stock not covered by data source	None	Need to assess compliance
Details of test files	Planning permissions q3 2011 - q3 2016	Commencement notices q3 2016 (part of)
Number of observations on test files	All - 86,000 permissions (q3 2011 - q3 2016)	Sample - 2,400 commencement records (q3 2016 (part of))

6. Reflecting on the definition of commercial property

Appendix 3 summarises the different categorisations of commercial property that are currently captured by administrative sources in Ireland. These are spread across the three strands of the project (stock, pipeline and transactions). The most detailed categorisation is from the Valuation Office which has over two hundred property “use” categories. In order to develop a coherent

statistical system for commercial property these categorisations will need to be harmonised in some capacity. Investment and engagement from the institutions involved will be required.

Before it is possible to embark on harmonisation of categories and definitions in a national system it would be imperative that an international definition of commercial property is firstly decided.

The Inter-secretariat Working Group for Price Statistics (IWGPS) are currently endeavouring to define commercial property in the context of preparing a document, entitled “Commercial Property Price Indicators: Sources, Methods and Issues” which will be published as a Eurostat working paper in the near future. As input into this document, the ECB, the Bundesbank and Eurostat elaborated a set of building blocks for (narrower and wider) definitions of commercial property. These are reproduced in Appendix 4. These building blocks use an “activity” concept in an attempt to stratify property – both residential and non-residential.

This approach stacks up well with the approach taken in the current project which deals with the three strands of stock, pipeline and transactions. One difficulty that would be envisaged when you analyse the categorisations from Appendix 3 is that the split between market and non-market is currently not captured. It is also not clear how to deal with public sector organisations where the building is being rented privately at market prices. In this situation the “use” of the building is non-market, however, the activity from the point of view of the investor (owner of the building) is very much market orientated.

7. Conclusion

This paper summarises the work carried in Ireland to date to develop a Commercial Property Statistical System (CPSS). The national user demand in this area, as articulated by key users, goes beyond price indices into areas such as pipeline stock coming on stream and the overall stock of commercial property stock and the related vacancy rates. The CSO has also taken note of the data needs as expressed at an international level - such as recommendations by the European Systemic Risk Board in October 2016.

To this end, the CSO has undertaken a project to assess the administrative data sources currently available to help meet these needs. The work is an ongoing project and its aim is to ultimately fill these data gaps and, crucially, to build the subject matter expertise required within the CSO to ensure that the data quality standards are met. The development of methodologies to construct price and rental indices for commercial property will be undertaken at a later stage.

A key aspect of the project was the ability to match different data sources using unique identifiers and also using a technique known as fuzzy matching or the Jaro-Winkler distance. Analysing different data sources available in Ireland, it was found there is potential to build a pipeline database of commercial property and a stock/register database of commercial property. However, a multitude of problems were found in the matching of transactions (sales and leases) data; namely a lack of unique identifiers in the datasets, fundamentally different units of measurement, a lack of detail in postal addresses and uncertainty on whether we are matching like-for-like. This is in direct contrast to the CSO’s success in matching these data sets for residential property.

Therefore, new data sources will be required to add important characteristic variables (such as floor area and type) to the transactions information captured by the tax authorities. To this end, all data sources will need to be assessed including 1) new surveys of market participants 2) the development of new administrative data sources with other public institutions and 3) exploring collaborative projects with institutions which hold private data sources.

The work is cognisant of the low frequency of transactions for commercial properties, which ultimately may hamper the development of a commercial property price index. Therefore, even if the characteristic variables can be successfully added to the transactions data, additional information may still be necessary to construct a robust commercial property price index for Ireland.

References

Kanutin, A. (2013), Ottawa Group Meeting, "ECB Progress Towards a European Commercial Property Price Index".

European System Risk Board (October 2016), "Recommendation of the European System Risk Board of 31 October 2016 on closing real estate gaps", (ESRB/2016/14)

European System Risk Board (December 2015), "Report on commercial real estate and financial stability in the EU".

Picchetti, P. (2013), Ottawa Group Meeting, "Estimating and smoothing appraisal-based commercial real-estate performance indexes".

Shimizu, C., W. E. Diewert, K. Nishimura, T. Watanabe (2013), Ottawa Group Meeting, "Commercial Property Price Indexes for Tokyo".

Balk, B. (2013), Ottawa Group Meeting, "A note on concept and measurement of property price indices".

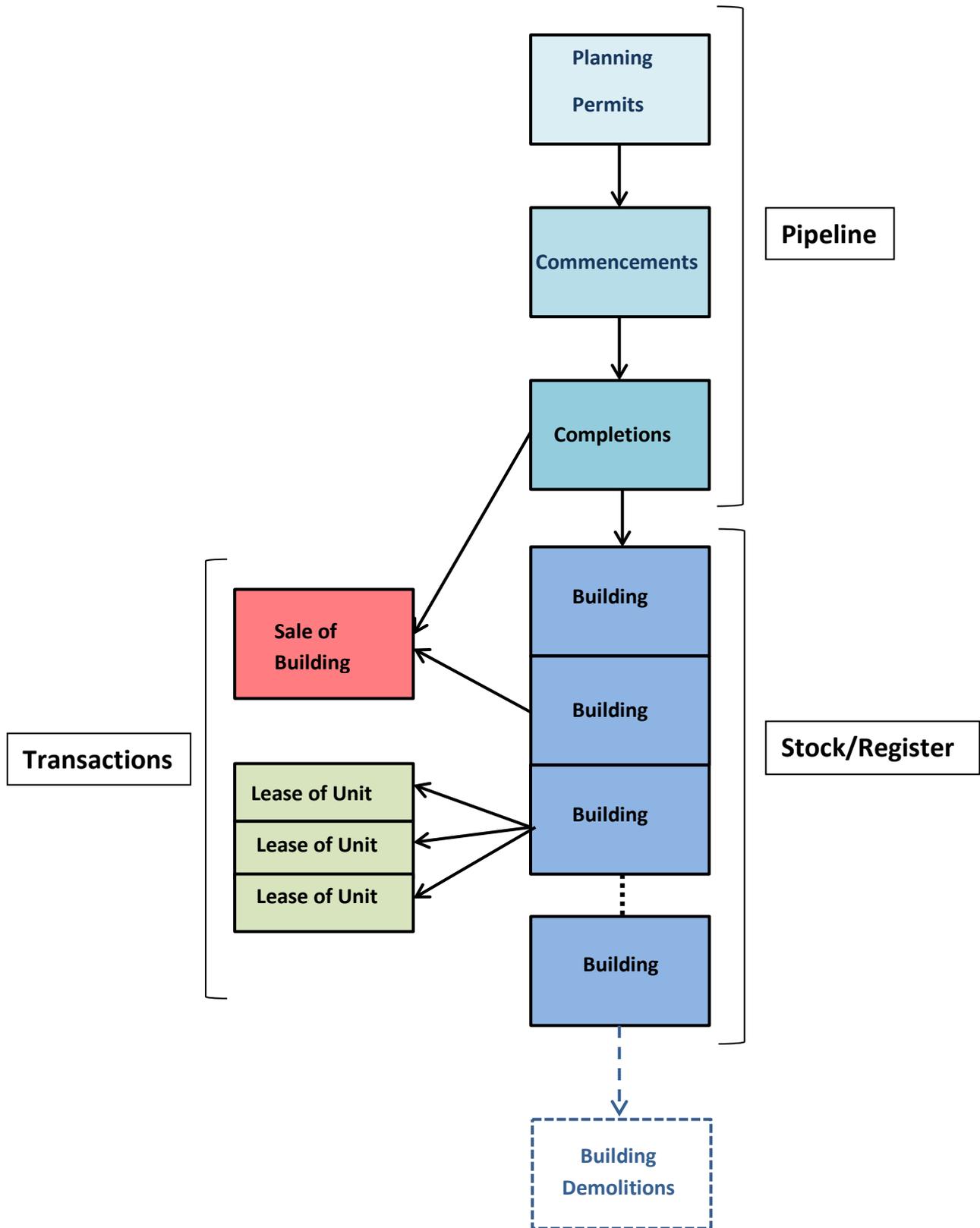
Passerini, P. (2013), Ottawa Group Meeting, "Developing Commercial Property Price Indicators".

Keiler, S. (2013), Ottawa Group Meeting, "Commercial property prices: What should be measured?"

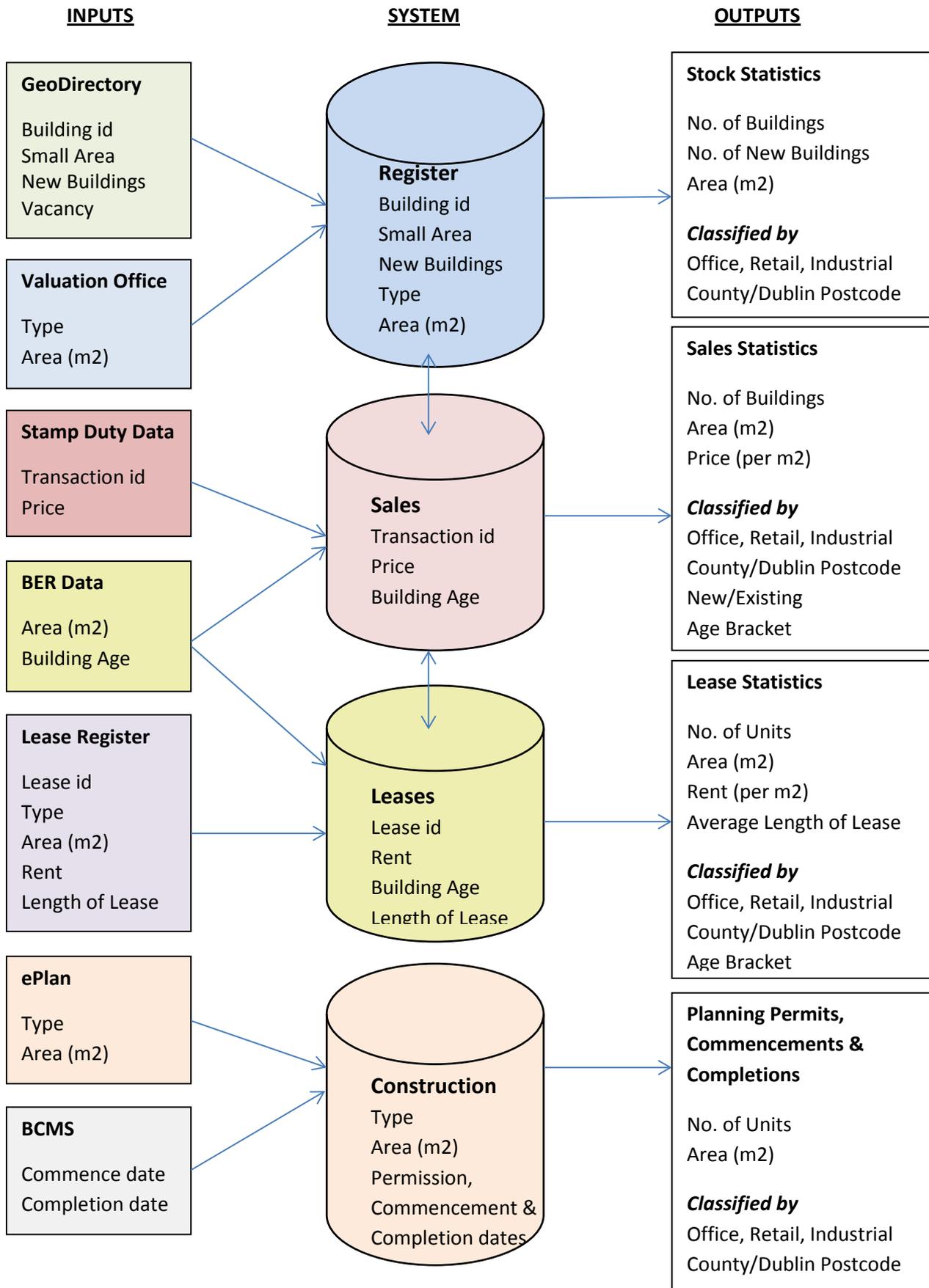
Central Bank of Ireland (2016) "Macro-Financial Review, 2016:II".

National Asset Management Agency (2010), "Annual Report and Financial Statements 2010".

Appendix 1: Commercial Property Statistical System (CPSS)



Appendix 2: Commercial Property Inputs & Outputs



Appendix 3: Categorisation of property types and uses by source

Stock		Pipeline	Transactions	
Valuations Office (Category)	Valuations Office (Use)*	Planning Permissions	PSRA	BER
Fuel/Depot	A. I. Station	Buildings for :	Agricultural	Airport terminals
Health	Abattoir	Agriculture, forestry.....	Agricultural Land	Bus station/train
Hospitality	Activity Centre	Mining, energy and water	Agricultural;Other	Community/day centre
Industrial Uses	Adult Shop	Industry and manufacturing	Industrial	Crown and county courts
Leisure	Adventure / Army /	Transport	Industrial;Agricultural	Dwelling
Minerals	Advertising Station	Trade and administration ...	Industrial;Agricultural;Other	Emergency services
Miscellaneous	Aerodrome	Health and welfare	Industrial;Other	Further education
Office	Amusement Centre	Entertainment and rec ...	New Commercial/Industrial	Hospital
Retail (Shops)	Antique Shop	Education, culture, science..	SH Commercial/Industrial	Hotel
Retail (Warehouse)	Art Gallery	Religious and funerary	Office	Industrial process building
Utility	ATM	Hotels, restaurants and cafes	Office;Agricultural	Laundrette
	Auctioneer	Public administration and ...	Office;Industrial	Libraries/museums/galleries
	Bakery	Office development, ...	Office;Industrial;Other	Miscellaneous 24hr activities
	Bank	Creche	Office;Other	Nursing residential homes
	Barber	Other non-residential	Retail	Office
	Beauty Salon/Massage		Retail;Agricultural	Primary health care buildings
	Betting Shop		Retail;Industrial	Primary school
	Bookshop		Retail;Office	Prisons
	Bridal/Formal Wear		Retail;Office;Industrial	Restaurant/public house
	Builsing Society		Retail;Office;Other	Retail
	Bulk Stores		Retail;Other	Retail warehouses
	Bus Shelter		Intangible Property	Secondary school
	Bus Station		Site Only	Social clubs
	Butcher		Non-Agricultural Land	Sports centre/leisure centre
	Butchers/Fishmongers		Other	Sports ground arena
				Telephone exchanges
				Theatres/cinemas/music halls
				Warehouse and storage
				Workshops/maintenance
* There are more than 200 use categories captured by the Valuations Office				

Appendix 4: Defining commercial property

Table A4: Building blocks for defining commercial property

	Residential property	Non-residential property
Construction of buildings (on own account, or on fee or contract basis)	Residential buildings under construction	Non-residential buildings under construction
Selling and renting of real estate (transactions)	Rentals at market prices (o/w owned by households)	Investment properties rented at market rentals
Own-use	Owner occupied housing – used as residences	Business properties – used in the production of goods & services (other than real estate) sold at market prices
Non-market	Social housing – priced below market rents	Non-business properties – used in the production of goods & services sold below market prices or provided for free