# Data and sample

We use data corresponding to 315 local authorities that operate at district level. Data are collected from different sources and concern the period from 2011 to 2016. We divide into three time intervals that include March 2011 – March 2013, March 2013 – March 2015 and March 2015 – September 2016.

Our main variable is the number of care homes per 1000 population that are aged 65 or over in the local authority. This variable is based on Tokunaga and Hashimoto (2011) who analyse the entry of private providers in Japanese long term care markets using a similar variable to reflect providers’ choice.

## Care homes

Information concerning the characteristics of care homes is obtained by the CQC directory of active and inactive care homes[[1]](#footnote-1). This dataset contains all the registrations of care homes that have carried out a regulated activity since 2010. The initial sample includes 24,354 records. We restrict our analysis to the entries from March 2011 onwards since a substantive proportion of the total registrations (16,054) that were carried out during 2010 and the first two months of 2011 were the result of legal requirement[[2]](#footnote-2). Our final sample is referred to 8,300 care homes.

A key advantage of this dataset is that it contains information about the entries and exits of the care homes in the market. Given that we do not have any further information available, we assume that care entry in a market since the date they are registered. Similarly, we consider they exit the market when they deregister. For determining our dependent variable, we compare the number of care homes that remain in each local authority for each time interval. In our analysis we focus on records associated with a new activity. These are different to new registrations that change the identification code due to organisational reasons (e.g. changes in the address or take overs from a different provider)[[3]](#footnote-3).

This dataset provides further information that includes the number of beds in each care home, the postcode and postal address, the city and region where the care home is located as well as the local authority that a council level is responsible for the social care services corresponding to the location of the care home. Likewise, with the exception of the number of beds, the same information is available with regards to the 3,830 providers where the care homes of our sample belong to.

In the second stage of our analysis we use information corresponding to quality ratings derived from the system implemented by the CQC since 2014. On the basis of five dimensions[[4]](#footnote-4), this new approach set a systematic method for collecting evidence that enables a more consistent assessment and comparison of the care homes’ performance. Services are rated according to four categories: *outstanding, good, requires improvement or inadequate*. For our analysis we collapse these categories into two: bad (requires improvement and inadequate) and good (outstanding and good). Because the information is only available since October 2014, this part of the analysis considers a different timeframe that involves three waves October 2014 - May 2015, May 2015 – February 2016 and February 2016 – September 2016.

## House prices

The information corresponding to prices of the properties is obtained from the price paid dataset released on a monthly basis by the Land Registry. This dataset contains all the transactions of properties carried out in England and Wales since 1995. In addition to the price paid for the transaction, the dataset includes further information such as the type of property, the address, the city, district and region where the property is located as well as whether the location was newly built and whether the property was under leasehold or freehold[[5]](#footnote-5). The information referred to transactions is collected on a daily basis. We subset the transactions that correspond to each local authority and obtain the average price.

**Instruments**

Our identification strategy uses information associated with several supply constraints. In order to capture the regulatory restrictiveness, we use the rate in change of delay of major projects which that is obtained from the Department of Communities and Local Government (DCLG). This variable compiles the number of decisions that have delayed projects for more than 13 weeks over a year. Besides, we also use the variation in the historical political composition of the local authorities. On the basis of Hilber and Vermeulen (2016), our analysis captures the historical Labour vote share at the General Election since 1983 for each local authority. The information is collected by British Election Studies Information System. In order to control for possible bias of associated this measure, we also include data on share of Labour vote corresponding to general election of June 2015. Data are obtained from the Parliament website platform.[[6]](#footnote-6)

Regarding the information associated with physical constraints that may determine the housing supply we use historical population of density in 1911. This variable is used as an instrument to correct for potential endogeneity associated with the share of developable land. Table 1 shows summary statistics for the variables of the baseline sample.

[INSERT TABLE 1 HERE]

On average, over the period of analysis there were about 1.7 care homes per 1000 population over 65. Yet, this proportion varies across the different local authorities. Figure 2 plots the spatial distribution of care homes for all the local authorites. In general, neighbouring districts present a similar proportion of care homes. There are some exceptions corresponding to some districts in the North West and Central North where the distribution of care homes is between 2 and 4 care homes per 1000 people over 65.[[7]](#footnote-7) Local authorities located in the south of the country (mainly in the South West and South East regions) have the greater amount of care homes per old population.

[INSERT FIGURE 2 HERE]

Figure 3 shows the spatial distribution of house the prices. It is possible to appreciate a clear dichotomy between local authorities situated in the north and in the south of country. Also, there some extreme cases of local authorities where the average value in the properties registers a maximum £2,170,757. Apart from this outlier, the average house price of the sample is £268,764.

[INSERT FIGURE 3 HERE]

1. This dataset is maintained by the CQC Directorate of Data and Statistics and available upon request. [↑](#footnote-ref-1)
2. Since October 2010 registration in Care Quality Commission became a legal requirement for every long term care provider wishing to carry out a regulated activity. [↑](#footnote-ref-2)
3. This situation is typically found when dealing with information contained in registries of firms. Neglecting it, apart from introducing measurement errors potentially, may lead to incorrect conclusions regarding the market dynamics and the performance of the firm. Geurts and Van Biesebroeck (2016), for instance, analyse the effect of this measurement problem on the estimations of the firm’s growth after the entry in the market [↑](#footnote-ref-3)
4. These dimensions entail the evaluation of issues related to the safety, the effectivity, the level of care and response to people’s needs as well as the management of the services. [↑](#footnote-ref-4)
5. The difference between these two types properties is based on the whether the ownership of the land or property is for a temporary (*leasehold*) or unrestricted (*freehold*) period. [↑](#footnote-ref-5)
6. Further information is provided in the following link: <http://www.data.parliament.uk/dataset/general-election-2015> [↑](#footnote-ref-6)
7. These districts are Allerdale, Lancaster, Harrogate and Scarborough. [↑](#footnote-ref-7)