ODUG Benefits Case for releasing UK Land Registry Price Paid Historical Data

Description & Request Overview

- To release the historical price paid housing database (1995 to 2012)
- The Price Paid data set contains information on residential property sales in England and Wales from 1995 to the present. The data includes property address, price paid, date of transfer, property type, whether the property is new build, and whether the property is freehold or leasehold.
- Price Paid data is the "reference" data set for information on sold house prices, and is arguably also the most authoritative source of information on basic property types at individual address level.

Data Release Rationale

- Land Registry has not significantly added value to the Price Paid data, so under PSI principles it should not be treated as a commercial product.
- Land Registry derives most its revenue from fees rather than data licensing, so open data release of this data set will have little impact on its financial model.
- Only large firms can afford to license the whole Price Paid data set on current terms.
- Open data release of recent updates alone has done little to level the playing field for
- SMEs in markets where the data is most useful.
- Current terms prevent analytic use of the data by academics and non-profits.
- According to Land Registry the data is currently used by "property price websites, media, smartphone application developers, data analysts and estate agents"; SME's and start ups; and All citizens – in terms of understanding the housing price market

Benefit Description

- Historic Price Paid data is definitive reference data set detailing property level prices paid on the vast majority of housing transactions in England and Wales since 1995. From March 2012 the preceding month's transactions have been released as Open Data under the OGL. The data includes new build flags, property types (detached, semi, terraced, flats), and a full non-PAF address including postcodes as well as the price paid.
- The Land Registry estimated a ~£600,000 loss of revenue were the data to be made open. Large commercial organisations such as Zoopla currently license the data in bulk, an estimated cost for the annual right to use the bulk data is £50,000
- The following are seen as the key benefits that would flow from releasing the entire dataset as Open Data:
    - Increased economic activity amongst house price websites and applications leading to employment, profits, taxes etc

- Increased efficiency within the housing market via the development of improved automated house price estimation algorithms
- Development of innovative household level classifications that combine census, DWP, HMRC and Land Registry data
- Access to a quasi-definitive database of household typologies at address level would enhance property modelling across a wide range of disciplines including energy usage, consumer expenditure, water conservation, insurance risk etc
- Various analytics to cross reference housing price to council tax and other economic, social and environmental factor. This could be used as an indicator for deprivation/affluence and economic growth.

Case Study – Mark Thurstain-Goodwin, Geofutures

Geofutures have previously completed a project assessing the impact of specific housing policy measures for the Audit Commission, Renew Staffordshire and Bridging Newcastle Gateshead. The foundation of the analysis incorporated a spatial regression model.

Historical price paid data was a key input into the modelling but due to the prohibitive cost of the data only the specific Local Authority was purchased and used. This led to 'edge effects' being introduced into the model.

- **Additional dimension and improved analysis** - Using the complete historical price paid dataset would have allowed macro-economic effects to be included e.g. average house price trends within 10 and 25 miles.
- **Prototype & Proof of concept** – An open price paid dataset would have allowed Geofutures to create a prototype to use in a sales pitch to win similar work from other Local Authorities.
- **Job Creation** – Geofutures believe the additional work would have required an additional FTE
- **Improved policy making leading to more effective housing renewal**

Benefits Analysis

- Quantifying the benefits of Open Data release is complicated. In particular the 'compounding' affects are dynamic, non-linear and do not yield to simplistic econometric modelling.
- However the purchase of a house is the single largest financial transaction any of our citizens are likely to complete in their lifetimes. The consequences of improved consumer decision making are enormous in both financial and social terms. Currently the pricing structure effectively excludes SME from entering the house price paid market and stifles competition. The Land Registry will also save considerable amounts of money from not having to promote, administer, and police the licensing of historical price paid data.
- The direct identifiable benefits include
  - Land Registry savings, estimated at four FTE in sales and general admin at a cost to the public purse of ~£300,000
  - Assume 10 SME's take up the challenge of producing innovative house price forecasting and reporting tools. Each SME employs one FTE on the project. The

direct, indirect and induced value of high skill modelling and IT jobs is beyond the scope of this document; however a very conservative assessment would value these 10 jobs at over £1,000,000 annually to the UK economy. These numbers are similar to the market behaviour when the census 2001 was released as open data.

- Improvements in the efficiency of the UK housing market. This market includes over £80 trillion pounds of outstanding mortgages, over £130 billion of transactions per year, and a direct tax take of over £5 billion in stamp duty. Stimulating and improving this market clearly has enormous potential for society and public finance benefits. Confidence in the market is widely recognised as a key driver of housing volumes. If improved information at no cost to the consumer was available and it only increased market activity rates by 0.1% then over £5M in additional stamp duty would directly flow to the treasury.
- Academic research into land use trends and downstream benefits of public investment in physical infrastructure (e.g. Crossrail, West Coast Main line etc) will enable public bodies to investigate fair and robust measures of the impacts of public sector investments.
- Even under the limited release and pricing models we have already seen innovative start-ups like Mouseprices enter the arena. Removing the price barrier to entry will stimulate innovations and additional business formations.
- Release of this data in conjunction with datasets such as PAF, NAG, the DVLA data, and VOA rates data will help create a national definitive spine of reference data for residential data.

Barriers and Requirements for Release

- The Land registry want to review the impact of releasing the current data before they release the historical data
- Some customers do not share the view of the Land Registry that the data is not personal data
- Land Registry believes it is too early to conclude that there have been no adverse privacy impacts.
- The Land Registry estimated a ~£600,000 loss of revenue were the data to be made open.

Recommendation

The historic price-paid data should be released as open data. The loss of revenues to the trading funds is not significant when held against the size of the markets that will function more efficiently with the data released.

The initial HM Treasury position seemed to indicate that the government's intention was for the data to be releases in bulk in its entirety. They are reviewing the release of the recent data this year before making a decision about the historical data. Following this review we would then urge the Land Registry to release all the data early in 2013, to ensure they do not act as a barrier to the much larger cross-economy benefits.