

OPEN DATA'S IMPACT

HM LAND REGISTRY: THE UK'S TRADING FUNDS, AND TWO FUTURES FOR OPEN DATA



OPEN DATA

by Becky Hogge

January 2016

OPEN DATA'S IMPACT

HM LAND REGISTRY: THE UK'S TRADING FUNDS, AND TWO FUTURES FOR OPEN DATA

by Becky Hogge
January 2016

www.odimpact.org



This work is licensed under a Creative Commons
Attribution-NonCommercial-ShareAlike 4.0 International License



HM LAND REGISTRY: THE UK'S TRADING FUNDS, AND TWO FUTURES FOR OPEN DATA

Dimension of Impact

- ✓ Solving Public Problems
- ✓ Data-Driven Engagement

Summary

HM Land Registry began a phased release of its data on property transactions—the Price Paid Dataset—in March 2012, and by November 2013 the entire historic record dating back to 1995 was released. The data provides much-needed transparency in a historically “murky” business, and is already being used extensively by some traditional players in the property market. Additionally,

new players are consolidating around the field of proptech, developing digital tools to bring buying and selling property “out of the Stone Age”. Proptech startups attracted an estimated \$1.4 billion in investment globally in 2014. PI Labs, an incubator for proptech startups, opened in London in late 2014.

Key Takeaways

- HM Land Registry is a trading fund, a type of government agency that meets its outgoing costs from the money it charges for services. Along with the UK's other trading funds, its future has been a sticking point in the UK's journey towards open data.
- In contrast to some of the other datasets studied in this report (e.g., TfL), HM Land Registry's data is delivering value most where it is cross-referenced with other, often proprietary, datasets.
- Releasing Price Paid data as open data has had a positive effect on data quality. HM Land Registry report an increase in the number of notifications about inaccurate data by members of the public since the Price Paid data was released.
- HM Land Registry is a potential

candidate for privatisation. The government set a poor precedent when it privatised the Postcode Address File alongside Royal Mail (see page 12). Selling off HM Land Registry

data would likely lead to a slowdown in the positive impacts described in this case: market transparency and proptech innovation.

The case in numbers

Number of transactions captured by HMLR's Price Paid Dataset: **20 million**

Number of downloads of HMLR Price Paid Dataset between January 2012 and March 2013: **78,000**

Global investment in proptech startups in 2014: **\$1.4 billion**

"We're the Land Registry's number one fan. This is what the government should be doing to make efficient businesses. They should open up data."

Adrian Black, Founder, YOUpower

Background

HM Land Registry (HMLR) was created in 1862 to register the ownership of land and property in England and Wales. It is responsible for maintaining the Land Register, where more than 24 million titles (evidence of land and property ownership) are documented.

In 1993 it was established as a trading fund—a type of government agency that has the authority (under the Government Trading funds Act 1973) to meet its outgoing costs from the money it charges for services. Other trading funds include Companies House, The Met Office and Ordnance Survey (OS). HMLR's receipts come mainly from fees it charges individuals and organisations to register land, and to register changes in title to land and properties.

The Land Register is the world's largest property database¹. At the height of the UK property boom in 2007, it processed around £1 million worth of property every minute in England and Wales². However, the credit crunch saw a slowdown in the UK market. HMLR lost £220m over three years to January 2011 and was forced to increase fees to users by 30%³. Those fees have since decreased, as part of HM Land Registry's strategy to encourage more people to register

¹ Land Registry. (2006, April). Internet Archive copy of webpage: Welcome to Land Registry's 10-year strategic plan. Retrieved from Internet Archive: <http://web.archive.org/web/20070602044143/http://www.landregistry.gov.uk/strategy/10yearplan/>

² Land Registry. (2006, April). Internet Archive copy of webpage: Welcome to Land Registry's 10-year strategic plan. Retrieved from Internet Archive: <http://web.archive.org/web/20070602044143/http://www.landregistry.gov.uk/strategy/10yearplan/>

³ KPMG. (2011). HMLR Feasibility Study.

transactions online.

The trading fund model has played a controversial role in the UK's journey to open data, since at least in the cases where agencies are funded by receipts from reselling data (for example OS) the two approaches are in conflict. Indeed, a previous study by the author of the UK's journey to open data, records "strong resistance" from trading funds and in particular OS, to a future of open government data:

The Power of Information Review had recommended a detailed cost-benefit analysis of the Trading fund model, which was commissioned jointly by HM Treasury and the then Department for Business, Enterprise and Regulatory Reform, and published in February 2008. Despite that study finding strongly in favour of abandoning the Trading fund model, and despite the Power of Information Taskforce recommending reform on the back of its findings, positions against re-examining the Trading fund model within government remained entrenched.⁴

In November 2010, the government announced plans to create a Public Data Corporation (PDC) to coordinate the release of data from the trading funds that deal in data and information (OS, The Met Office, Companies House, and HMLR). The plans represented "a wholesale reorganisation of the largest trading funds", but did not go into detail on how much of the data released by the PDC would be open, and also included the possibility of the PDC being part-privatised⁵.

In August 2011, the Department for Business, Innovation and Skills (BIS) consulted with the public on the PDC's data policy⁶. In November 2011, Chancellor George Osborne made commitments on open data in his Autumn Statement that included data releases from Companies House, HMLR, and The Met Office, a limited data release from OS, and the establishment of a Public Data Group to share best practice that would supersede the PDC⁷. In March 2012 BIS published a response that included terms of reference for a Data Strategy Board that would "seek to maximise the value of data from the Public Data Group ... for long-term economic and social benefit, including through the release of data free of charge"⁸.

In July 2014 the UK's coalition government abandoned plans to privatise or part-privatise

-
- 4 Hogge, B. (2011, May). *Open Data Study: New Technologies*. Retrieved from Transparency and Accountability Initiative: http://transparencyinitiative.theideabureau.netdna-cdn.com/wp-content/uploads/2011/05/open_data_study_final1.pdf
 - 5 Arthur, C. (2010, November 19). Met Office and Ordnance Survey to be part of 'public data corporation'. Retrieved from The Guardian: <http://www.theguardian.com/technology/2010/nov/19/government-public-data-corporation>
 - 6 Department for Business, Innovation & Skills and Cabinet Office. (2011, August 04). A consultation on data policy for a Public Data Corporation: government response. Retrieved from gov.uk: <https://www.gov.uk/government/consultations/data-policy-for-a-public-data-corporation>
 - 7 Cabinet Office. (2011, November 29). Policy paper: Open data measures in the Autumn Statement 2011. Retrieved from gov.uk: <https://www.gov.uk/government/publications/open-data-measures-in-the-autumn-statement-2011>
 - 8 Department for Business, Innovation & Skills and Cabinet Office. (2011, August 04). A consultation on data policy for a Public Data Corporation: government response. Retrieved from gov.uk: <https://www.gov.uk/government/consultations/data-policy-for-a-public-data-corporation>

HMLR, after Liberal Democrat business secretary Vince Cable blocked the move⁹. However, after the May 2015 elections returned a majority Conservative government privatisation may be back on the table: Chancellor George Osborne has outlined plans to privatise £23bn worth of government assets as an attempt to bolster public finances¹⁰, and campaigners against privatisation identify HMLR as one of the targets¹¹.

In April 2015, the functions of the Public Data Group were integrated into a new BIS-run board, the Digital Culture, Services Platforms, and Data Board, and the PDG met for the last time¹².

The data

HMLR's Price Paid Dataset contains over 20 million transactions going back to 1995 for residential properties sold for full market value. It was this dataset that Chancellor George Osborne committed the Land Registry to releasing as open data in his 2011 Autumn Statement. In March 2012, HMLR began releasing monthly files of transactions. 2013 saw the release of historic data in two phases: Records of transactions between 2009 and 2012 were released in June 2013; and in November 2013 the full historical record from January 1995 to the current month was released. That year also saw the release of the price paid database as linked data. In 2014, HMLR worked with a third-party supplier to create an “easy to use” front end, the Price Paid Data Report Builder tool¹³.

The data is not a full picture of land ownership in England and Wales—it only includes details of properties bought and sold between January 1995 and the present (see box on page 32). It is released under an Open Government Licence and as such it conforms with the open definition. It is updated every month with the new transactions that have been recorded, and corrections to older entries.

The quality—and timeliness—of the data relies in part on those people (homeowners and their solicitors) who register it with HMLR. A sample monthly update file showed approximately 2,500 entries against transactions older than a year, which might be assumed to be corrections to existing transactions. Of the (roughly) remaining 79,000 transactions assumed to be new transactions, 3,345 (4.2%) were more than six months old.

Before its release as open data, Price Paid data had been made available in bulk to paying customers as part of HMLR's commercial activities. As such, HMLR had sought the view of the Information Commissioner's Office (ICO) in 2004 as to whether Price Paid data (which includes

⁹ Syal, R. (2014, July 14). Land Registry privatisation plans abandoned by ministers. *The Guardian*.

¹⁰ Quinn, J. (2015, May 20). George Osborne to privatise £23bn of taxpayer assets.

Retrieved from The Telegraph: <http://www.telegraph.co.uk/finance/newsbysector/banksandfinance/11618980/George-Osborne-to-privatise-23bn-of-taxpayer-assets.html>

¹¹ We Own It. (n.d.). Don't sell off our top trumps. Retrieved September 30, 2015, from: <http://weownit.org.uk/take-action/dont-sell-off-our-top-trumps>

¹² Public Data Group. (n.d.). Public Data Group. Retrieved from gov.uk: <https://www.gov.uk/government/groups/public-data-group>

¹³ Interview, Lynne Nicholson, Head of Data Products and Services, HM Land Registry

addresses) constituted biographical information, and the ICO had indicated that it did not. Before its release in 2012, HMLR carried out a Privacy Impact Assessment¹⁴, and reviewed the assessment again in 2013¹⁵. Both reviews concluded that Price Paid data was deemed not to be biographical or personal in nature. The pre-release review flagged an increase in companies using the data for direct marketing purposes as a potential issue; the 2013 review concluded that there had been no increase in direct marketing as a result of the release.

What is in HM Land Registry's Price Paid dataset, and what isn't

HM Land Registry's Price Paid Dataset, the subject of this study, is a subset of the information HM Land Registry holds on land and property ownership in England & Wales. The over 20 million transactions in the current version of the dataset appear to relate to 11,738,465 unique addresses¹⁶. In total, the Land Register consists of 24 million titles.

Even the Land Register itself is not a complete picture of land and property ownership in England & Wales, since compulsory registration of land and property is only triggered by a property transaction of some kind. And although the Land Registry was created in 1862, compulsory registration across England & Wales only arrived in the 1990s. Since then, notifying the Land Registry has been compulsory when land or property is bought or sold, and (more recently) when any mortgage is taken out on it, or when land or property is inherited. Some land and properties have remained in the same family or organisation's ownership for generations, and have therefore never been registered. In July 2012, HM Land Registry announced that 80% of the land mass of England & Wales was now registered¹⁷.

The Price Paid Dataset is a subset of the information on the 80% of the land mass that HM Land Registry does hold (just under 50% of the titles it holds that in turn relate to 80% of the land mass of England & Wales). The Price Paid Dataset also does not contain information on the individuals who own the title, or the title ID. The address, postcode, and transaction value is included, as well as some basic information about the property, such as whether it is a new build. Further information may be obtained for a fee of £3 per title.

While it's easy to understand why the Land Registry don't include the registered owners of all the properties in the Price Paid Dataset, some of the dataset's other exclusions are harder to understand. Price Paid only includes property transactions between individuals where property

¹⁴ Land Registry. (2012). Privacy Impact Assessment Report: Making price paid data available through publication in a machine readable and reusable format. Retrieved from Land Registry:
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/343616/ppd_pia.pdf

¹⁵ Land Registry. (2013). Privacy Impact Assessment Review: Price paid data, transaction data and historical price paid data. Retrieved from Land Registry:
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/343604/PIA_Report_6_13.pdf

¹⁶ Price Paid Dataset does not include title IDs, so unique addresses were used as a proxy. The author downloaded the dataset on 21 September 2015, and queried it for the number of entries with a unique Postcode/PAON/SAON combination.

¹⁷ <http://archive.is/pFEzX#selection-1665.0-1665.251>

has been sold for full market value. It does not include any information (for example) that the Land Registry holds on properties that have been inherited, or on properties owned by corporations, or on land transactions. Land Registry does make available in bulk its data on land and property owned by corporations (excluding private individuals, overseas companies, charities, and trustees), but only to paying customers. This database consists of 3.2 million title records.

In July 2015 David Cameron announced that HM Land Registry would release, as open data, data on property and land owned by corporate entities registered overseas, including the name and correspondence address of the legal owner of the company. This followed a series of reports from Private Eye magazine using similar data, obtained from the Land Registry under Freedom of Information laws. The data released to Private Eye consisted of details of 96,441 titles¹⁸.

Using this data, Private Eye were able to uncover a number of public interest stories, including the frequency with which property and investment companies use overseas corporate vehicles to own significant London landmarks, thereby achieving capital gains tax and stamp duty advantages.

An interactive map that displays the data is available on the Private Eye website¹⁹. In order to build the map, Private Eye used a dataset of title plans released by HM Land Registry under the EU's INSPIRE scheme—the INSPIRE index polygons. Although the INSPIRE index polygons are released by HMLR under the OGL, the dataset contains third-party intellectual property that belongs to Ordnance Survey (OS). HMLR states in its conditions of use²⁰ that this means anyone making the polygons available to third parties, or using them for anything other than "personal, non-commercial or commercial or non-commercial use within your organisation" needs to "contact Ordnance Survey for the relevant licence conditions". Christian Eriksson, the freelance journalist who led the project on Land Registry data, says that neither he nor Private Eye contacted Ordnance Survey before publishing the interactive map, believing his use of the INSPIRE index polygons constitutes fair dealing under Section 30(2) of the Copyright Designs and Patent Act 1988²¹.

The path to open

Who was calling for HMLR Price Paid data to be released as open data? The fact that HMLR is a trading fund means it was a general target for open data enthusiasts. In responses to the government's PDC consultation, HMLR data was rated as the third most requested dataset to

¹⁸ An FOI request for the same dataset is available at (Bowden 2015)

¹⁹ Private Eye. (2015). In the back: Selling England by the offshore pound. Retrieved from Private Eye: <http://www.private-eye.co.uk/registry>

²⁰ Land Registry. (2014, August 14). INSPIRE Index Polygons spatial data > conditions of use. Retrieved from gov.uk: <https://www.gov.uk/guidance/inspire-index-polygons-spatial-data#conditions-of-use>

²¹ legislation.gov.uk. (n.d.). Copyright, Designs and Patents Act 1988. Retrieved September 30, 2015, from legislation.gov.uk: <http://www.legislation.gov.uk/ukpga/1988/48/section/30>

be “free for use and reuse” after Ordnance Survey data and Address/Postcode information²². Lynne Nicholson, Head of Data Products and Services at HMLR, says the data release responded to “a general call for government to release more data²³”, and did not mention any particular organisations who were calling for the data to be released, beyond the Open Data Institute, which in any case was only created at the same time as the announcement that the Price Paid Dataset was to be made available as open data.

Data based on HM Land Registry data, including Property Sales²⁴ and Median House Prices²⁵ had been published intermittently on data.gov.uk by the Department for Communities and Local Government since 2010, and had fed into a small number of apps, notably the parking app JustPark. Before the open data release, HMLR Price Paid data had been available as downloadable .txt files (and also as bespoke reports) to a small group of organisations on a commercial basis. A business case for releasing the data, drafted before the Chancellor’s autumn Statement and released under FOI in February 2012, reveals that HMLR were attracting a modest £600,000 of income from this activity per year. It notes that:

Existing customers for this data are, in the main, companies that publish the data on their own websites. Land Registry has 30 existing subscribers who are website based businesses and 7 others which they class as consultant/research users. In a typical month they also receive ad hoc requests for specific areas and time periods of data from a small group of customers, most of which are Local Authorities—on average 5 requests a month. They also receive ad hoc requests from Estate agents and solicitors, usually 1-2 a month from each sector²⁶.

During interview, HMLR did not wish to reveal the identities of its paying customers prior to the data release²⁷. A source in the industry speculated such customers might include property portals (such as Rightmove, Prime Location, and MousePrice), business-to-business publications (such as Hometrack) and nationwide property consultancies (such as Savilles). Attempts by the author to interview one such potential customer (Rightmove) were unsuccessful.

The same business case document asserts that, across all the trading funds considered, it would be existing customers of the data that would capture a significant proportion of the initial

22 As well as these categories of data, which all relate to location, meteorological data, transport data, health, education, and crime statistics were all cited by respondents as having particular societal benefits. Department for Business, Innovation & Skills and Cabinet Office. (2011, August 04). A consultation on data policy for a Public Data Corporation: government response. Retrieved from gov.uk: <https://www.gov.uk/government/consultations/data-policy-for-a-public-data-corporation>

23 Interview, Lynne Nicholson, Head of Data Products and Services, HM Land Registry

24 Department for Communities and Local Government. (2010, February 09). Property sales based on Land Registry data. Retrieved from data.gov.uk: https://data.gov.uk/dataset/property_sales_based_on_land_registry_data

25 Department for Communities and Local Government. (2010, September 09). Median house price. Retrieved from data.gov.uk: https://data.gov.uk/dataset/median_house_price

26 HM Treasury. (2012, 02 13). Public Data Group Business Case. Retrieved from gov.uk: <https://www.gov.uk/government/publications/open-data-public-data-group-business-case>

27 Interview, Lynne Nicholson, Head of Data Products and Services, HM Land Registry

benefit of releasing the data for free. This notwithstanding, when asked to identify stakeholders who were against the release of the Price Paid Dataset as open data, Lynne Nicholson, Head of Data Products and Services at HMLR, pointed to concerns from existing commercial customers “that opening up the data would allow more competition into the market”²⁸.

Indeed, the response to a consultation on government open data strategy put forward by the Cabinet Office in 2011, includes one from Landmark Information Group (who own Mouseprice.com). It argues strongly in favour of a presumption for open data, but notes that:

Due consideration should be made of the private sector organisations who may already offer services around any data sets which currently attract a charge from the public sector data providers which may be candidates for being re-designated as ‘open’. Thus avoiding any potential negative affects to the existing markets.²⁹

Users and Outcomes

Lynne Nicholson is keen to underscore that, now the Price Paid Dataset is released openly, HMLR cannot track who is using it:

It's open data. We put it out there. We don't put a registration system up there. So we have absolutely no idea who's downloading it. We can put analytics around it to say how many downloads, but we simply don't know who's downloading it. That's the whole point of open data. For the Price Paid data we had roughly 30 customers [before the open data move]. Obviously they're going to be downloading it. But our downloads are, you know, thousands a month. We have no idea who the others are or what they're doing with it.³⁰

According to the second privacy impact assessment, between January 2012 and March 2013, the Price Paid data was downloaded a total of nearly 78,000 times³¹.

One data re-user I spoke to described how portals listing Price Paid data have proliferated on the web:

28 Interview, Lynne Nicholson, Head of Data Products and Services, HM Land Registry

29 Cabinet Office. (2011, August 04). Making Open Data Real.

Retrieved from gov.uk: <https://www.gov.uk/government/consultations/making-open-data-real>

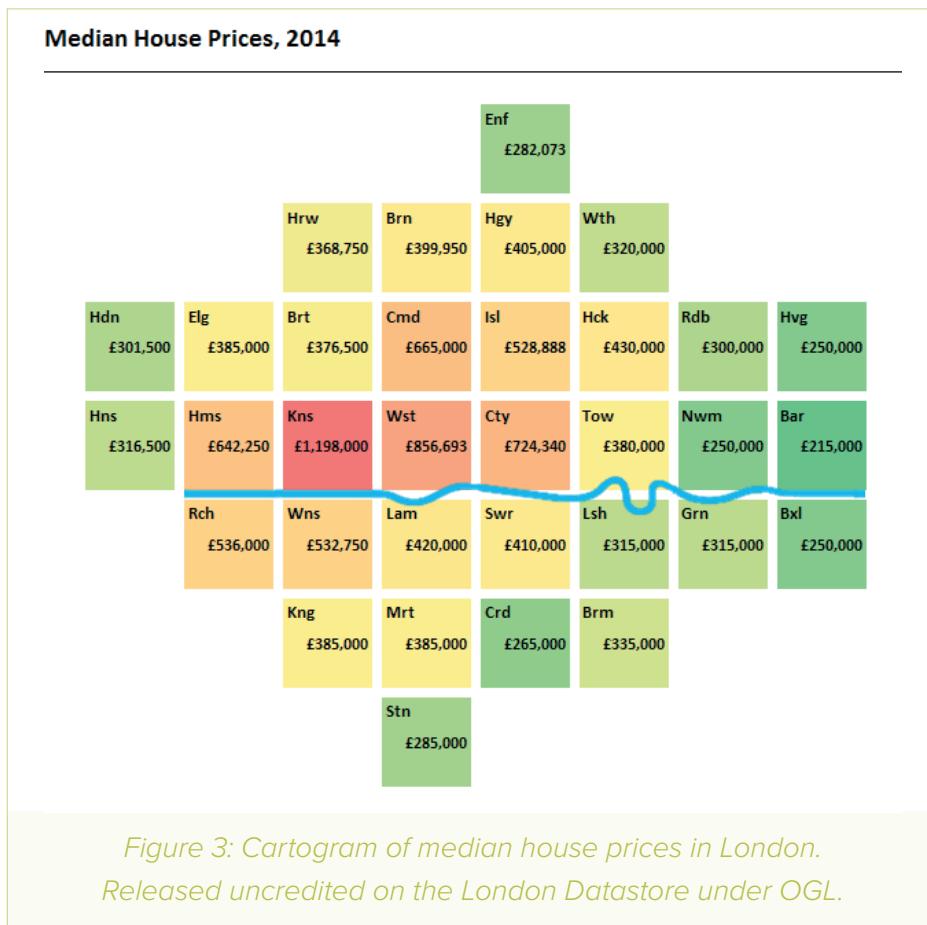
30 Interview, Lynne Nicholson, Head of Data Products and Services, HM Land Registry

31 Land Registry. (2013). Privacy Impact Assessment Review: Price paid data, transaction data and historical price paid data.

Retrieved from Land Registry: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/343604/PIA_Report_6_13.pdf

Now you can download a .csv and have your own [HM Land Registry] property price portal up in a couple of hours If you Google your own house address you will see there are dozens of sites showing sold prices for your house.

Nicholson also notes that government departments were attracted by the open data offer, then entered into confidential data-sharing agreements involving richer HMLR datasets as a result. Nicholson mentions “significant use by the media and researchers,” making an example of the data and what could be achieved with it. The cartogram in Figure 3 (which appears uncredited on the London Datastore and was picked up by bloggers at UCL’s Centre for Advanced Spatial Analysis³²), shows the median house prices in different London boroughs.



HMLR hosted an event in July 2014 for users of its data³³. At the event, Manuel Timita from Illustrists—a platform that takes open data from a variety of sources (including the Office for National Statistics, OS, Police.uk, Environmental Agency, Department for Transport, as well

32 Mapping London. (2015, March 19). House Prices: A Borough Cartogram. Retrieved from Mapping London: <http://mappinglondon.co.uk/2015/house-prices-a-borough-cartogram/>

33 Ramsay, F. (2014, July 03). Telling stories with open data and maps. Retrieved from Land Registry: <http://blog.landregistry.gov.uk/using-open-data-map-future/>

as HMLR) and displays it on a map to help people decide where they want to live—showed a timeline of London house prices, matched against key events surrounding the credit crunch (see Figure 4)³⁴.

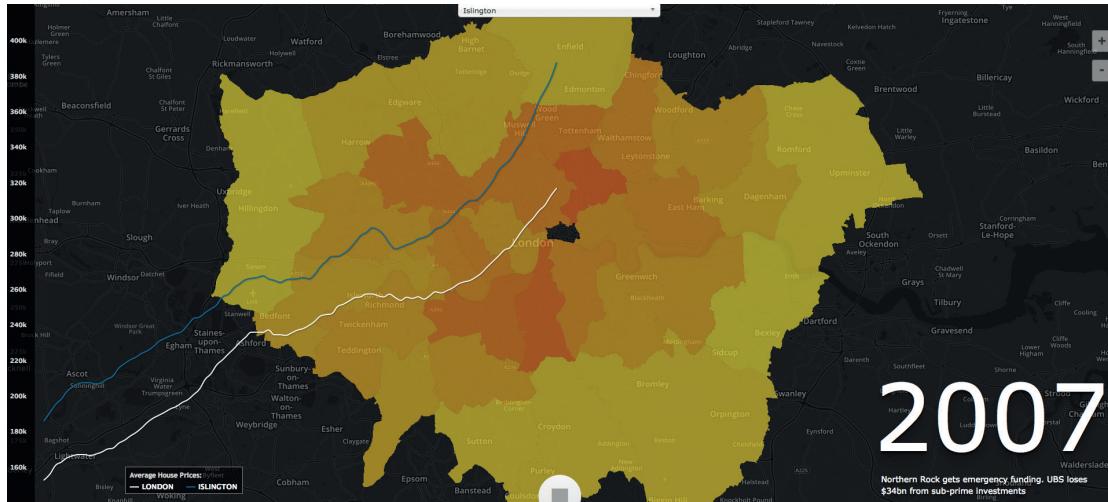


Figure 4: Screenshot of *Illustreets*' video timeline of rise in London house prices.

Copyright Manuel Timita. Reproduced with permission.

Also at the event was Johnny Morris from estate agents Hamptons International, who, according to the Land Registry's report of the event, "divulged how [Hamptons International] are enriching their market view of the 'downsizer', through mixing their own proprietary data with Land Registry Price Paid Data and Ordnance Survey open data"³⁵. However, it is not clear whether Hamptons were an original customer of the HMLR Price Paid data, or only began using the data once it was released openly.

One estate agent who is using the data now it's been released by HMLR is Adrian Black, founder and director of YOUpoint. He moved to estate agency after 10 years working in technology for Goldman Sachs, drawn to a business that looked ripe for technological change:

In every agency business you have diminishing returns and you have to become more productive and efficient and also pick up revenue from other sources. That's not happening in estate agency. Fee rates are coming down, but they're coming down slowly and the increase in property prices means agents are actually earning more. [With online portals] reaching buyers is getting easier, and yet revenues per sale are increasing.³⁶

³⁴ Timita, M. (2014, February 27). London house prices: evolution over 13 years.

Retrieved from *Illustreets*: <http://illustreets.co.uk/blog/maps-and-apps/london-house-prices-evolution-over-13-years/>

³⁵ Ramsay, F. (2014, July 03). Telling stories with open data and maps.

Retrieved from Land Registry: <http://blog.landregistry.gov.uk/using-open-data-map-future/>

³⁶ Interview, Adrian Black, Founder and Managing Director, YOUpoint

YOUhome is using Land Registry data and other proprietary sources to make their estate agency business more efficient, and more attractive to potential clients.

Sellers typically look to engage agents who offer the highest valuation and the lowest fees. But that leads to low conversion rates.³⁷

Black says seeing actual data can persuade sellers to go for a more realistic asking price, can get buyers comfortable more quickly, and can achieve a quicker sale at a fair price. And it's working for their business. They report that typical central London agents charge 2-3% rates, whereas their central London rate is 0.8%. And in Bournemouth their brokers transact on average 50 properties per year per agent, against an industry average of around 25 (15 in London). "We're the Land Registry's number one fan," says Black:

This is what the government should be doing to make efficient businesses. They should open up data.³⁸

Property expert Henry Pryor anticipates that more and more groups will find ways to make use of HMLR's data:

Retailers are already making use of the value of property as part of the matrix that they look at when trying to segment their market. Clearly, they're trying to target different individuals with different products. To know that somebody lives in a million pound house as opposed to a hundred thousand pound house, you can make assumptions based on that information.³⁹

The release of HMLR's Price Paid Data has coincided with the rise of the new field of proptech: businesses using data and technology to innovate in the property sector. One report quotes figures from CrunchBase showing investors worldwide put a record \$1.4bn into proptech startups in 2014, although most of the money went to US-based firms⁴⁰. In late 2014 Pi Labs, the UK's first proptech incubator, was set up by estate agents Cushman & Wakefield and venture capitalists Spire Ventures.

"It's kind of buzzy at the moment and so a lot of different things fall under [proptech]," explains one HMLR data re-user who does not wish to be identified. A video filmed to promote a

³⁷ Interview, Adrian Black, Founder and Managing Director, YOUhome

³⁸ Interview, Adrian Black, Founder and Managing Director, YOUhome

³⁹ Interview, Henry Pryor, Independent Residential Property Expert

⁴⁰ Assetti. (2015, June 14). The Rise of the New Property Technology (Prop Tech). Retrieved from assetti.pro: <http://assetti.pro/en/2015/06/14/the-rise-of-the-new-property-technology-prop-tech/>

roundtable on the future of the property market in May 2015⁴¹ showcases companies including Splittable (which helps tenants of shared houses split their bills), Hubble (a web platform for making renting office space easier), Homeshift (a platform through which agents, tenants, and housebuyers can communicate during a house move), Fixflo (a tool for tenants and property managers), Land Technologies (a data aggregator that identifies land that's suitable for building on), MoveBubble (an online agent that acts on behalf of tenants), and We Are Pop Up (an agency for short-term retail lets).

Peter Thum-Bonanno sold his first proptech startup, the property portal Find Properly, to search engine Nestoria in November 2014. Although he is unwilling to ascribe the rise of proptech entirely to the release of the HM Land Registry dataset, he says there is a link:

I don't know if it's grown or if I've just become more aware of it, but there's definitely a rise in property technology companies who call themselves proptech companies. I think a lot of it is because property seems to still be stuck in the Stone Age. The way things are done still seems to be very antiquated. People think technology could help make that more efficient. And I think, compared to a lot of other industries, there's a huge amount of data out there and the Land Registry is mainly to thank for that.⁴²

Thum-Bonanno's new venture is a tool to assess the efficiency of agents by using data including asking prices, sold prices, and conveyancing times. GetAgent uses a mix of public and private data sources, but HMLR data "plays a big part in that".⁴³ As well as using HMLR Price Paid data, it uses a non-open data API supplied by property portal Zoopla⁴⁴. This is not uncommon, and the developers and entrepreneurs I spoke to all agreed with Vasanth Subramanian of Splittable, that HMLR's Price Paid Dataset was just "one tool in the toolbox":

It's when it's combined with other data, expertise or products you or your organisation might possess that the real value of the dataset is realised.⁴⁵

A developer who did not wish to be identified feared that because Price Paid data was creating value only in conjunction with other datasets, arguments for it to remain open during any privatisation process would be weakened. This would be a mistake, according to Vasanth Subramanian:

41 Fowler, R. (2015, June 23). *Looking at the future of #proptech*. Retrieved from Tech City News: <http://techcitynews.com/2015/06/23/looking-at-the-future-of-proptech/>

42 Interview, Peter Thum-Bonanno, Co-founder and CTO, GetAgent

43 Interview, Peter Thum-Bonanno, Co-founder and CTO, GetAgent

44 <http://developer.zoopla.com/>

45 Interview, Vasanth Subrahmanian, Co-founder and CTO, Splittable

It would be a big problem if the Price Paid got sold off with HMLR. Even if the dataset [is not providing amazing value in and of itself], it's a vital building block and you'd see a slow down in proptech innovation.⁴⁶

Impact

How can we measure the impact of HMLR releasing their Price Paid data? Henry Pryor believes one of the main benefits will be in bringing transparency to the housing market, and likens the change to the introduction of screen-based trading on the London Stock Exchange in the 1980s:

This data shines a light into a murky world and provides transparency which in turn provides trust and confidence in transaction values associated with the purchase and sale of residential property. We don't have it in the rental market and the market suffers as a result. But just as we saw with Big Bang in the city and this idea of effectively real-time pricing, buyers and sellers can get far more confidence and as a result you get a far firmer, more mature market where confidence is high.⁴⁷

This echoes Adrian Black's experience on the ground: that homeowners use data to feel comfortable about the price of the property they are buying or selling, helping the market move more quickly.

But the state of the UK housing market is subject to many more forces than the state of the estate agency business, however “murky” or “Stone Age” those involved in it believe it to be. Prior to the UK’s 2007 credit crunch, precipitated by the run on mortgage lender Northern Rock, housing transaction volumes had peaked at around 150,000 a year. 2015 figures (see Figure 5) show they have now stabilised at around 100,000 a year⁴⁸. Although the steady upward climb towards this figure from a low of around 50,000 a year in December 2008 appears to accelerate after March 2012, given the complexity of the market and the availability of the data to market participants through HMLR’s paying customers prior to this date, it seems wise to regard HMLR’s open data release as a correlation, and not a cause.

⁴⁶ Interview, Vasanth Subrahmanian, Co-founder and CTO, Splittable

⁴⁷ Interview, Henry Pryor, Independent Residential Property Expert

⁴⁸ HMRC. (2015, September 22). UK Property Transaction Statistics. Retrieved from gov.uk:

https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/461354/UK_Tables_Sep_2015_cir_.pdf

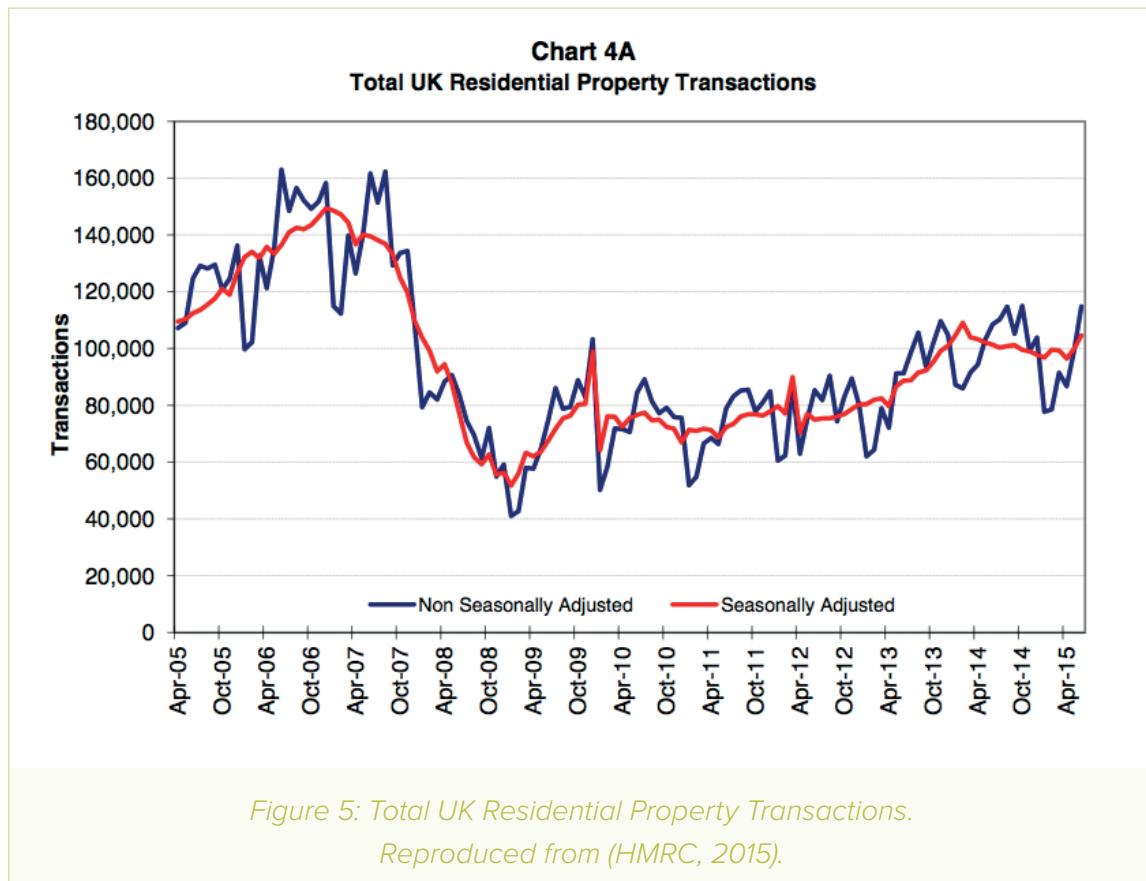


Figure 5: Total UK Residential Property Transactions.

Reproduced from (HMRC, 2015).

The OECD ranks the UK in the top 10 countries with over-valued housing markets, one of six in this group where house prices are still rising⁴⁹. The affordability of homes in the UK is a significant political issue. Transaction volumes are also distorted by the attempts of successive governments to both maintain an economy that is back-stopped for many voters by the (inflated) value of their property, and also help first-time buyers afford a home⁵⁰.

Compared to ascertaining the release of open data's effect on the housing market, its effect on innovation in the property sector looks slightly more straightforward. Here, the rise of the proptech sector, and the opening of Pi Labs, look like good indicators for innovation, and many of the developers and entrepreneurs involved are quick to recognise the contribution of HMLR. But all of them qualify this with statements that indicate that HMLR data works best in conjunction with other datasets, be they other open datasets (as with Illustreets) or, more often, closed data sets (as with Peter Thum-Bonanno's GetAgent). This stands in contrast to TfL, where the end products created by innovators are far more likely to rely on TfL data alone. Any contribution that proptech makes or goes on to make to the UK economy will be hard to pick

49 OECD. (n.d.). Focus on house prices. Retrieved September 30, 2015, from OECD:
<http://www.oecd.org/eco/outlook/focusonhouseprices.htm>

50 HMRC. (2015, September 22). UK Property Transaction Statistics. Retrieved from gov.uk:
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/461354/UK_Tables_Sep_2015_cir_.pdf

apart: It will be hard to say how much of this is thanks to HMLR's open data policy.

One area where it is easiest to say that the open data policy has had impact is data quality. Lynne Nicholson states that increased exposure has also helped HMLR to maintain data quality and that they are reacting to an increased number of queries since release:

Now we're publishing information, it's more accessible, more people are viewing it, so you're bound to get more queries and questions, and requests too. I'll give you one example. Say we've captured a property as an end terrace. The owner might ring up and say 'I regard my property as semi-detached, not an end terrace'. So those are some of the types of questions we get.⁵¹

HMLR report an increase in the number of notifications about inaccurate data by members of the public since the Price Paid data was released, saying they received 2,256 issues related to incorrect or missing Price Paid information from the release of the data up to the publication of their second privacy impact assessment in May 2013. However, no figure for complaints about inaccurate data prior to the Price Paid release is given for comparison⁵².

Discussion

Lynne Nicholson explains how open data has shifted the focus of HMLR:

Our focus is changing from commercial activity to the release of data so that organisations can develop products and services from our data. Our remit is not to compete with the private sector and not to compete with organisations who are downloading our data.⁵³

HMLR is committed to making the data it releases easier to digest, and has invested in a reporter tool for Price Paid data that it believes is driving more non-expert users to its data.

As for the proptech sector, for the moment innovation rather than consolidation looks set to rule. "You've got lots of startups like ours that are nibbling away at small bits of the value chain," says Peter Thum-Bonanno:

⁵¹ Interview, Lynne Nicholson, Head of Data Products and Services, HM Land Registry

⁵² Land Registry. (2013). Privacy Impact Assessment Review: Price paid data, transaction data and historical price paid data. Retrieved from Land Registry:

https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/343604/PIA_Report_6_13.pdf

⁵³ Interview, Lynne Nicholson, Head of Data Products and Services, HM Land Registry

So you have companies that are trying to make getting a mortgage easier and more straightforward. We're trying to help you find an estate agent There's no one yet who is doing the whole process for you from finding an estate agent, finding you a buyer, helping you move, helping you manage your home. It seems like that's the aspiration of all these small proptech startups and it will be interesting to see if there will be a consolidation in the industry, allowing you to manage your entire home online, or if it's going to continue to be piecemeal in the future. I'm not sure.⁵⁴

HMLR's data is delivering value to existing stakeholders in property price information (estate agents and homeowners) and stimulating innovation that looks set to contribute to UK economic growth. Putting an exact figure on the impact of the data, however, may never be possible.

In the absence of quantitative methods, it is likely that future data release policies will be governed by other factors. Attitudes to HMLR's current data release policy, and those of other trading funds such as Ordnance Survey, varied wildly between the data users interviewed. Innovators ranged from being "very grateful" to HMLR for the data release, to feeling frustrated that more data was not being either collected (e.g., number of bedrooms, square footage) or released (e.g., full land ownership records). Meanwhile, at the more traditional end of the spectrum, Henry Pryor did not see why businesses who were able to monetise HMLR data were not required to pay for it.

This contrast in attitudes persists to the highest levels of government. Claudia Arney, chair of the Public Data Group between June 2012 and its final meeting in April 2015, describes trading funds navigating their open data strategies in an environment of competing pressures: strong pressure to open up data; and equal and opposite pressure to provide evidence that any change to the status quo represented true value for money and maintained the integrity of the data in question⁵⁵. She believes the integration of the Public Data Group into the Digital Culture, Services Platforms, and Data Board, which will include representatives from the Cabinet Office and HM Treasury, will allow trading funds to plot their future path with greater certainty. But the election of a Conservative government and the broader focus on the health of public finances in UK politics may yet shift the balance.

Any sell-off of HMLR may not be as swift as anti-privatisation campaigners fear. The replacement as Chief Land Registrar of former Student Loans Company Chief Executive Ed Lester with the former Chief Executive of Thurrock and Brentwood councils, Graham Farrant, could be read as a subtle shift away from immediate privatisation towards a strategy of delivering efficiency savings before any sell-off. Confirming Farrant's appointment in a written statement to the House of Commons, Matt Hancock (then a minister at BIS and now Minister for

⁵⁴ Interview, Peter Thum-Bonanno, Co-founder and CTO, GetAgent

⁵⁵ Interview, Claudia Arney, Chair, Public Data Group 2012-2015

the Cabinet Office) wrote “Graham has the skills and knowledge to manage the organisation through its transformation into a modern, digital organisation.⁵⁶”

What will be of consequence is less if or when the sell-off comes, than whether HMLR data is included in the package. Regrettably, the last government set a precedent here when it included the Postcode Address File in the sell-off of Royal Mail, a move that attracted strong criticism from the open data community⁵⁷ and from Parliament’s Public Administration Select Committee⁵⁸ (see page 12).

Calls to action

For policymakers

- Policymakers need to recognise the contribution HM Land Registry’s Price Paid data is making both to the traditional property market and to innovation in the sector. Regardless of the future of HM Land Registry as a public body, this data should remain in public hands.
- The impact of HM Land Registry’s Price Paid Dataset suggests that more of HM Land Registry’s data, such as records of land and properties owned by corporations registered in the UK, should be considered for release as open data.

For open data advocates

- HM Land Registry may be a target for privatisation. The open data community needs to pressure the UK government to prevent the bundling of HM Land Registry data with any sale. In particular, the open data community should ensure that government is aware of all the stakeholders in HM Land Registry data as open data, including estate agents and proptech innovators such as those interviewed for this study.

For funders

- Funders could encourage open data advocates to network more with innovators in the proptech sector, in order to cement alliances against data privatisation at HM Land Registry.
- Any campaign against data privatisation will be greatly supported by economic analysis that speaks the language of those making the decisions at HM Treasury. However, it is also important to recognise that open data reforms are in some sense a leap of faith (see TfL study).

⁵⁶ Aston, S. (2015, March 13). Graham Farrant to replace Ed Lester as Land Registry chief. Retrieved from Civil Service World: <https://www.civilserviceworld.com/articles/news/graham-farrant-replace-ed-lester-land-registry-chief>

⁵⁷ Hope, C. (2013, April 19). Everyone’s postcodes to be privatised in Royal Mail flotation, despite objections from Sir Tim Berners-Lee. Retrieved from The Daily Telegraph: <http://www.telegraph.co.uk/news/uknews/royal-mail/9994741/Everyones-postcodes-to-be-privatised-in-Royal-Mail-flotation-despite-objections-from-Sir-Tim-Berners-Lee.html>

⁵⁸ Arthur, C. (2014, March 17). MPs and open-data advocates slam postcode sell-off. Retrieved from The Guardian: <http://www.theguardian.com/technology/2014/mar/17/mps-and-open-data-advocates-slam-postcode-selloff>