



CPB Netherlands Bureau for Economic
Policy Analysis

CPB Discussion Paper | 325

Internet searches and transactions on the Dutch housing market

Sander van Veldhuizen
Benedikt Vogt
Bart Voogt

Internet searches and transactions on the Dutch housing market

Sander van Veldhuizen, Benedikt Vogt, Bart Voogt¹

Abstract

We use Google searches of the word “mortgage” to explain monthly housing transactions in the Netherlands in the period from 2004 until 2015. Our estimates indicate that Google searches of the previous months are significantly positively associated with housing transactions in the current month. This shows evidence that internet search data can provide information about real market behavior.

JEL Codes: E2, E27, R2, R3

Keywords: Housing market; Internet search data; nowcasting; Google Trends; Google search data

¹ All authors are affiliated with CPB Netherlands Bureau for Economic Policy Analysis. We are thankful to various colleagues for valuable comments. The data and code which were used to derive the results are available upon request.

1. Introduction

In developed countries housing markets are key drivers of domestic economies. To have adequate policy responses on future housing market development, instruments are needed to understand and predict housing market development. Therefore, easily accessible and high frequency data are needed to be able to judge future developments. Nowadays search queries from the internet offer such a source.

This paper investigates if internet search data provide a source of information for housing market development in the Netherlands. In particular we use data from the service Google Trends to explain the number of transactions on the Dutch housing market. We combine administrative data on real housing transactions from Statistics Netherland with search queries retrieved from Google Trends. The idea is that future house buyers search the internet to screen the market and gather information about potential properties, locations and also financing possibilities. Google search data are accessible on a weekly basis and can be downloaded for free. Hence, search data can provide an ideal proxy of market activity also before transactions actually take place. We use a simple time series model and enhanced time series models with Google data to explain the number of transactions on the Dutch housing market.

We obtain three findings. First, Google searches of the word “mortgage” of the current and the previous month are highly significantly positively associated with housing transactions. We also find evidence that searches six and nine months ago are significantly positively associated with transactions of the current month. Third, adding Google searches, increases the explanatory power of our simple model to predict housing transactions by 4 percentage points.

We contribute to the literature in economics on the relation between online search behavior and “real world” behavior. Van Dijk and Francke (2015) investigate the relationship between market liquidity and house prices by making use of data of a big real estate online platform in the Netherlands. Their main finding is that the amount of clicks on a property can serve as proxy for housing demand. Askitas (2015) finds that searches for the word “traffic jam” predict road traffic conditions in Germany. Askitas and Zimmermann (2009) find a strong association between labor-market-related search keywords and monthly unemployment rates.² Vosen and Schmidt (2011) show that Google searches outperform survey-based indicators in predicting US consumption.

2. Data and Descriptive Statistics

We use search activity data from Google. The data provided are *relative* search data. This means that Google calculates a search index of a term relative to other searches.³ For our analysis we use the search data on the word “hypotheek”, the Dutch word for mortgage. We obtained data from January 2004 until November 2015 on a weekly basis. In a second step, we combine the Google search data with data from the Dutch housing market. This is official administrative data on a monthly basis retrieved from Statistics Netherlands which contains the number of house transactions. The last data we retrieve is from November 2015.

² Choi and Varian (2012) test the predictive value of internet search activities on various economic activities in the U.S. such as car sales, unemployment claims, consumer confidence and travel destinations.

³ Google defines the data as follows: “The numbers that appear show total searches for a term relative to the total number of searches done on Google over time. A line trending downward means that a search term's relative popularity is decreasing. But that doesn't necessarily mean the total number of searches for that term is decreasing. It just means its popularity is decreasing compared to other searches.” Source: [Google](#). Last access 14 December 2015.

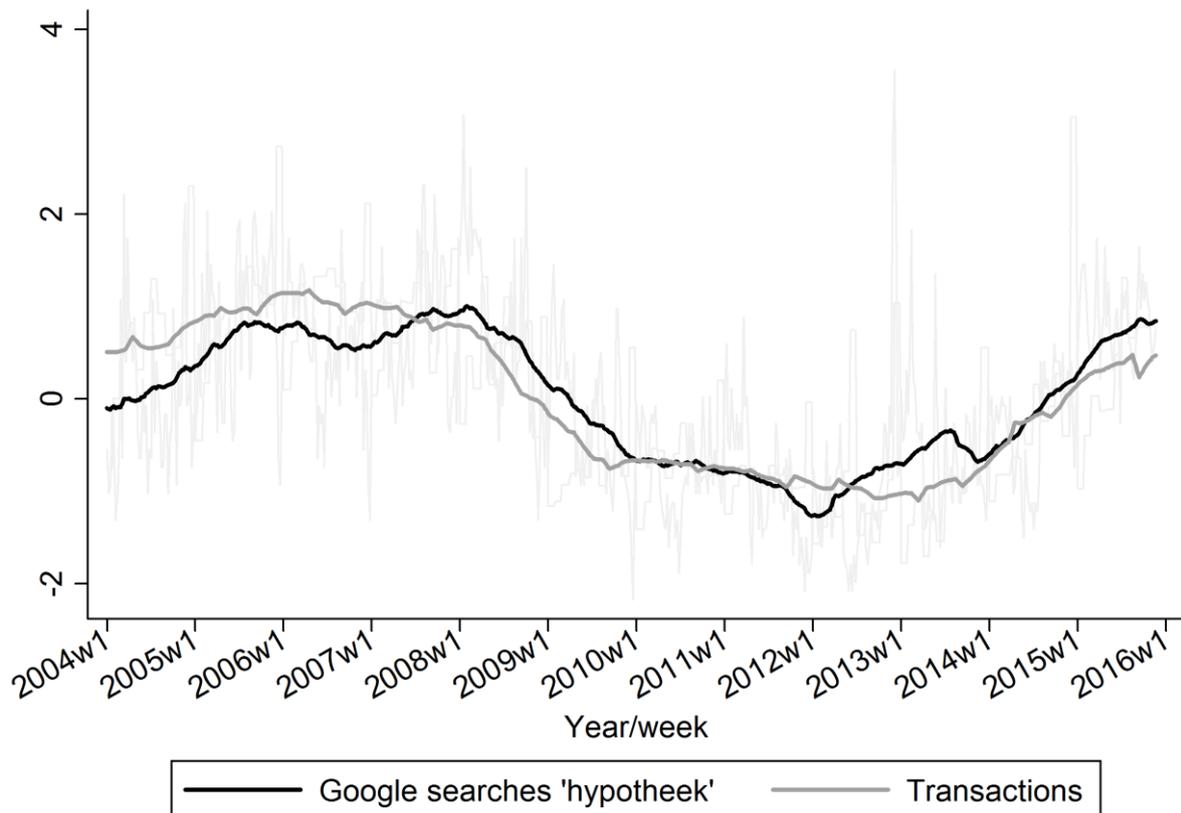


Figure 1. The evolution of housing transactions and google searches over time.

Note. The figure shows the standardized values of the search term “mortgage” and the transactions on the Dutch housing market. The data on searches can be downloaded on <https://www.google.com/trends/home/all/NL>. The data on housing transactions are retrieved from CBS statline: <http://statline.cbs.nl/Statweb/>.

Figure 1 shows the evolution of the number of transactions in the Netherlands and the search term “mortgage” between 2004 and 2015. Both variables are standardized (i.e. rescaled such that the mean equals zero and the standard deviation equals one) to make the comparison between the two measures easier. The picture that emerges from Figure 1 is that transactions and mortgage searches reveal strong seasonal trends. We make use of smoothing averages to adjust for seasonal trends.⁴ The black line shows the development of Google searches adjusted for seasonal trends. The grey line shows the development of housing transactions corrected for seasonal trends. Figure 1 reveals that Google searches and transactions co-move. Moreover, Google searches seem to be a leading indicator for future transactions.

⁴ We take the moving average from 36 weeks before, the current week and 36 weeks in advance. Different specifications yielded similar results.

3. Empirical Approach

Figure 1 suggests that there is a strong seasonal component in both time series. Seasonality in the housing market can stem from the fact that many transactions take place at a certain time in a year since for instance moving is most convenient during vacation periods (see for instance Ngai and Tenreyro, 2014). Another reason might be that house buyers make use of fiscal incentives at the end of the year before they change at the beginning of the next year. At the same time there might be a relationship between search queries and real economic behavior. This is what we are interested in.

We apply a time series estimation which allows us to predict monthly transactions on a macro level as a function of the Google search index.⁵ In a first step, we estimate a simple model of monthly housing transactions corrected for time trends.

$$y_t = \alpha + \gamma T + \epsilon_t \quad (1)$$

The depended variable y_t is the standardized number of monthly housing transactions in the Netherlands. We include year and month dummies in the matrix T to correct for seasonal trends.

In a second step, we estimate an enhanced model of monthly housing transactions including several lags of the Google search index:

$$y_t = \alpha + \beta X_t + \gamma T + \epsilon_t \quad (2)$$

The matrix X contains the standardized values of the Google search term “mortgage” of the current month and previous months. In our preferred specification we include 11 lags of the

⁵ Please note that we take the average per month of the weekly search data to deal with the mixed frequencies in our time series.

Google search index. The reason for that is that buying a house is an important decision in household consumption and might be well planned ahead of time.

4. Results

Table 1 shows the results of the two regression equations. Column (1) shows the results of equation (1). This regression only contains month and year dummies. The adjusted R^2 equals 82.9%. In column (2) we add the standardized values of the current search index. The coefficient is positive and significant at the 1 percent level. In the last column we add 11 lags of the search index.

The picture that emerges from Table 1 is that there seems to be a strong and significant relationship between the Google search term ‘mortgages’ and the number of housing transactions in the Netherlands. Especially the search index in the month of transaction and the search index one month before seem to be strong predictors for the number of transactions. Moreover, the search index of the sixth and the ninth month before the transaction month are positively and significantly associated with the number of transactions ($p < 0.1$). These findings are in line with the hypothesis that housing consumption is planned well in advance.

Most interestingly, adding the search data helps to improve the explanatory power of our model. If we compare the models in column (1) and (3) the adjusted R^2 improves by 4 percentage points from 82.9% to 86.9%. This shows that search data actually have an added value in explaining housing transactions.

Table 1. Relationship between house transactions and google searches of the word ‘mortgage’.

	(1)	(2)	(3)
Google searches		0.229*** (0.0695)	0.152** (0.0763)
Google searches t-1,			0.221*** (0.0828)
Google searches t-2			0.143 (0.0879)
Google searches t-3,			-0.00607 (0.0824)
Google searches t-4			0.0351 (0.0815)
Google searches t-5			-0.0575 (0.0803)
Google searches t-6			0.154* (0.0790)
Google searches t-7			-0.120 (0.0799)
Google searches t-8			0.0335 (0.0801)
Google searches t-9			0.141* (0.0809)
Google searches t-10			0.0543 (0.0803)
Google searches t-11			0.113 (0.0743)
Constant	0.133 (0.169)	-0.166 (0.186)	-0.440** (0.200)
Observations	130	130	130
Adjusted R-squared	0.829	0.843	0.869
Year dummies	Yes	Yes	Yes
Month dummies	Yes	Yes	Yes

Note. The table reports results of OLS regressions. The dependent variable consists of the standardized values of house transactions per month. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

5. Discussion and Conclusion

In this paper we investigate the relationship between online search activity and transactions on the housing market in the Netherlands. We find a strong and significant positive relationship between the number of transactions and search term “mortgage”. Our analysis also reveals

that searches from previous months are positively associated with the number of transactions in the current month.

The results indicate that policy makers can use internet searches to track and forecast development in the housing market. The costs to obtain the measures are low. Moreover, the data are highly up to date since the data are online available on a weekly basis. However, caution is needed when making strong conclusions about the relationship. The construction of the search data is not entirely transparent and the level of aggregation of the data is high.

References

- Askatas, N. (2015). Predicting Road Conditions with Internet Search. *IZA Discussion Paper*. Retrieved from <http://ftp.iza.org/dp9503.pdf>
- Askatas, N., & Zimmermann, K. F. (2009). Google Econometrics and Unemployment Forecasting. *Applied Economics Quarterly*, 2(55), 107–120.
- Choi, H., & Varian, H. (2012). Predicting the Present with Google Trends. *Economic Record*, 88, 2–9. <http://doi.org/10.1111/j.1475-4932.2012.00809.x>
- Ngai, L. R., & Tenreyro, S. (2014). Hot and Cold Seasons in the Housing Market[†]. *American Economic Review*, 104(12), 3991–4026.
- Van Dijk, D. W., & Francke, M. (2015). Internet search behavior, liquidity and prices in the housing market. *DNB Working Paper*. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2654170
- Vosen, S., & Schmidt, T. (2011). Forecasting private consumption: survey-based indicators vs. Google trends. *Journal of Forecasting*, 30(6), 565–578.



Publisher:

CPB Netherlands Bureau for Economic Policy Analysis

P.O. Box 80510 | 2508 GM The Hague

T (070) 3383 380

February 2016 | ISBN 978-90-5833-720-7