Do the Determinants of House Prices Change over Time? Evidence from 200 Years of Transactions Data

Martijn I. Dröes*[a,b,c] and Alex van de Minne†[d]

[a] *Amsterdam Business School, Plantage Muidergracht 12, 1018 TV Amsterdam, The Netherlands.*
[b] *Amsterdam School of Real Estate, Jollemanhof 5, 1019 GW Amsterdam, The Netherlands.*
[c] *Tinbergen Institute, Gustav Mahlerplein 17, 1082 MS Amsterdam, the Netherlands.*
[d] *MIT Center for Real Estate, 105 Massachusetts Ave, Cambridge, MA 02139.*

December 5, 2017

Abstract

This paper uses almost 200 years of historical data on house prices and its determinants from Amsterdam, the Netherlands. We find that before 1900 population growth, construction costs and new housing supply are the most important determinants of house price dynamics. After 1900 income starts to play a role and, with the development of the mortgage market, interest rates as well. Directly after World War II population and new housing supply are the key determinants of house prices, which is likely due to the birth of the baby boom generation and post-war reconstruction plans. The results in this paper imply that the determinants of house prices change over time, reflecting the economic state of affairs in each different era.

KEYWORDS: house prices; long-run determinants; cointegration.
JEL-codes: N9, R31, C32.

*Email: m.i.droes@uva.nl
†Corresponding author, email: avdminne@mit.edu

1 Introduction

In Europe, housing accounts for 40% – 60% of total household wealth and it is roughly 20% for the average household in the United States.[1] It should, therefore, not come as a surprise that economists and policy makers are highly interested in the fundamental determinants of house prices.

Between the mid-1980s and 2008 real house prices more or less doubled in most industrialized countries (De Wit et al., 2013). From a historical perspective, however, this is a relatively new phenomenon. Figure 1, for example, shows historical real log house price indices for the US, UK, France, and the Netherlands. Annual real house price appreciation is close to zero, or in some cases even declining (France), during most of the 20$^{\text{th}}$ century. Total real house price appreciation (averaged across the four countries) between 1900 and 1985 is 20% (on average 0.23% per year) while real house price appreciation for the period 1985 — 2010 is five times as large, about 107% (on average 7.15% per year). To understand these different growth rates, it is essential to understand how the fundamental determinants of house prices, and their impact, have changed over time.

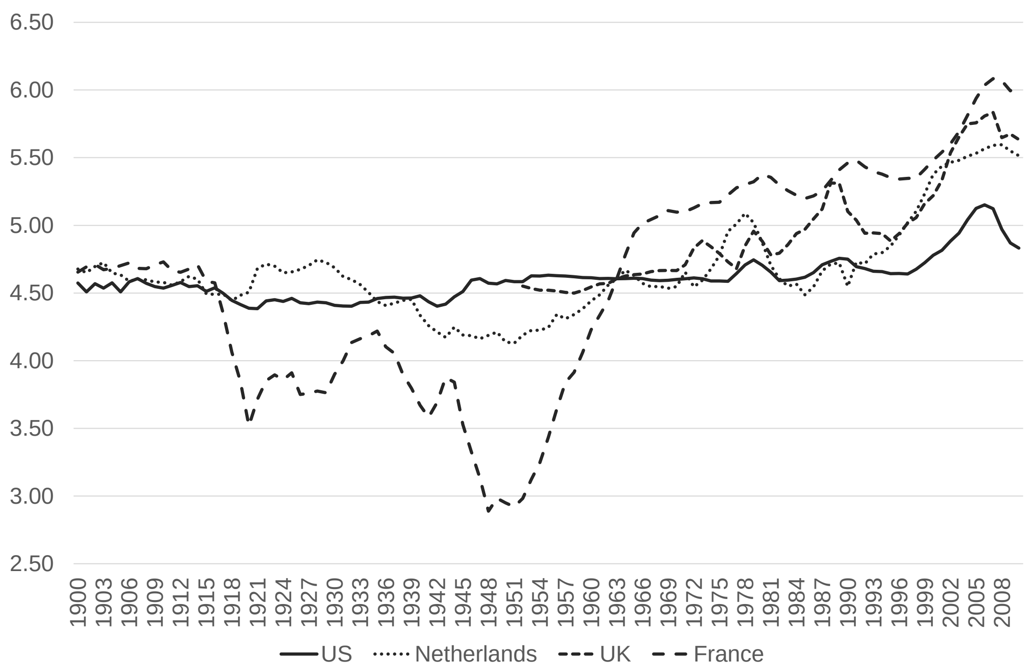
This paper examines the determinants of house prices from a long-run, historical, perspective. We use almost 200 years of house price data (1825–2012) located next to the Herengracht in Amsterdam, the Netherlands (see Eichholtz, 1997; Ambrose et al., 2013; Eichholtz et al., 2015). The Herengracht is one of the central canals in Amsterdam. Many of the original houses next to the Herengracht have survived until this day. Based on this dataset, we re-estimate the well-known Herengrachtindex (Eichholtz, 1997) using a methodology tailor-made for markets characterized by a low number of transactions (Francke, 2010). Our specific contribution is that we also add data on seven fundamental house price determinants. Housing supply, construction costs, Gross Domestic Product (GDP) per capita, the opportunity cost of capital (interest rates), labor force, unemployment, and population growth.[2] We use a rolling (window) error correction model (R-ECM) with changing covariates to examine the cointegrating relationships between house prices and its determinant and show how these relationships have changed over time. Even though it is widely acknowledged that house price determinants differ across markets, this study examines the changes in house price determinants in a single housing market over a long period of time.

We find that the long-run cointegrating relationships change over time. Population growth, unemployment, construction costs and housing supply were the main drivers of house

[1] See Statistics Netherlands and US Bureau of Economic Analysis and Statistics, respectively.

[2] Ambrose et al. (2013), but also see for example Ngene et al. (2014), do not look directly at which variables affect house prices before and after a break. More specifically, Ambrose et al. (2013) explore price-to-rent ratio's at the Herengracht for two sub-periods and Ngene et al. (2014) analyse structural breaks in long memory or fractional integration using an ARFIMA model between 1991 and 2014 in the US.

Figure 1: Log real house price indices for a selection of countries, 1900 – 2012.



Notes: The indices start at 1900, except for the UK which starts at 1952. The base year is 1963. The index for the US is taken from Freddie Mac (for the period 1975 – 2014) and is augmented by the historic data from Robert Shiller. The UK data is taken from Nationwide and for France from CGEDD. The house price data for France goes back to 1936. House price data from Paris was used to extend this time series. House prices for the Netherlands are based on our own calculations (see Section 2) until 1973 and was augmented with the price index found on the website of Eichholtz. Both the UK and UK indices are based on the 'standard' Case and Shiller (1987) repeat sales methodology. The French price index is based on (weighted) median sales prices.

prices in the 19$^{\text{th}}$ century. Our results show that the cointegrating relationships changed from more construction cost driven to more income and - especially in the end of the sample - interest rate driven variables. Mortgage market innovations and financial liberalization allowed financial intermediaries to advance higher levels of credit to consumers from the 1970s onwards (Fernandez-Corugedo and Muellbauer, 2006). Conjoined with declining interest rates this resulted in more affordable housing and subsequent increases in house prices. Moreover, the size of the effects are also time varying. For example, the effect of GDP on house prices was substantial lower during the period 1900 — 1970 than in the period 1970 — 2012. Also the coefficient of the opportunity cost of capital increased by a factor of ten between the mid 20$^{\text{th}}$ century and the end of the 20$^{\text{th}}$ century.

The fact that the size of most coefficients increased so much during the 20$^{\text{th}}$ century is also consistent with the 'Land Leverage Hypothesis' (Bostic et al., 2007). Earlier research showed that - in contrast to the structure values - land prices are mainly driven by demand side variables, like interest rates, population growth and income (Bourassa et al., 2011). Thus, if the fraction of land value tot total value is high in certain time period, the coefficients of these parameters on house prices (land plus structure value) is expected to be high as well.[3]

Further results indicate that directly after World War II population growth and investment in housing are the main drivers of house prices. This likely reflects the baby boom generation and post-war reconstruction plans. Population also had a large impact on house prices in Amsterdam during the 1970s. This is mainly due to the large scale deurbanisation taking place in that time period.[4]

Finally, we show that war has a negative impact on house prices of up to 40%, the predictability of house prices (R-squared) varies considerably over time, and the rate rate of error correction also changes. In particular, in some periods prices adjust almost instantaneously to changes in fundamentals, while in other periods it takes up to 20 years for shocks to be fully absorbed into house prices. It is well-known that housing markets are not efficient, they adjust sluggishly to new information (Case and Shiller, 1989). The results in this paper adds to this notion that housing market efficiency itself is not fixed over time.

Our results can explain some of the discrepancies in the literature about the determinants of house prices. A stylized example is the study by Englund and Ioannides (1997) versus that of Adams and Füss (2010). Both use the same OECD database and both regress house

[3]For example, given that the 'true' elasticity between land prices and income is 1 (i.e. a 1 percent increase in income will result in 1 percent higher land prices) and the fraction of land value to total value is 50 percent, we expect that the elasticity of income to *total house prices* (land plus structure value) will be 0.5, ceteris paribus. Even though we do not observe land values directly, house price appreciation was vastly superior to construction cost increases during the second half of the 20$^{\text{th}}$ century, which is indicative of relative high land values (Davis and Heathcote, 2007).

[4]It were actually the babyboomers who had families by that time that left Amsterdam for more open / green areas directly neighboring Amsterdam.

prices on a proxy for economic activity and interest rates for 15 OECD countries. However, the effect of interest rates on house prices according to Adams and Füss (2010) is a multitude of that found by Englund and Ioannides (1997). This can be explained by the fact that the study of Englund and Ioannides (1997) uses data from 1970 — 1992 and the study of Adams and Füss (2010) is based on a different time period, 1975 — 2009. Generally speaking, during the 1970s and 1980s the loan-to-value caps were a lot stricter and access to credit was relatively limited. Thus, it should come as no surprise that the effect of interest rates on house prices was lower in the pre-1990s era.

Moreover, ignoring that the cointegrating vector can change over time can easily result in house price increases to be incorrectly interpreted as a bubble (Ngene et al., 2014). In the literature it is quite standard to measure bubbles using an error correction approach. Whenever prices are above (below) equilibrium houses are overvalued (undervalued). However, the equilibrium relation (and thus the deviation from it) depends on which variables are included. In fact, a number of academic studies conducted in the early 2000s suggested that the U.S. housing market was experiencing the characteristics of a house price bubble (see Ambrose et al., 2013). However, Case and Shiller (2003) compared U.S. house price growth with income growth since 1985 and concluded that income growth could explain nearly all of the house price increases for over 40 states. In addition, McCarthy and Peach (2004) found little evidence supporting a bubble in the U.S. housing market after adjusting housing prices to account for the effects of interest rate changes. In our study, we will cope with such issues by allowing the cointegrating relationships to change over time.

The remainder of this paper is structured as follows. Section 2 provides a discussion on the historical context of the Amsterdam housing market and Section 3 describes the data used in this study. Section 4 contains the methodology. In Section 5 we report the results and Section 6 concludes.

2 House prices and the historical context of the Amsterdam housing market

To examine the long-run determinants of house prices, we start with constructing a price index for the Amsterdam housing market. We are not particularly interested in individual transaction prices, but in the price developments over time and the macro-economic determinants that can explain those changes. We used two sources to construct the long-run house price index. For the period 1825 — 1972 we exploit the same data as is used by Eichholtz (1997). The dataset covers transactions of dwellings on the Herengracht from 1628 to 1972.[5]

[5]The Herengracht is the name of the canal as well as the name of the street located directly next to the canal. In the main text, we do not distinguish between both names.

From 1973 onwards we use a separate database made available to us by the NVM.[6] The NVM is the largest brokers organization in the Netherlands. Since the 1970s more than 500 properties were sold on the Herengracht by brokers affiliated to the NVM (see Appendix A.1 for more details on both databases).

The Herengracht is one of the central canals in Amsterdam that was constructed between 1585 and 1660. By 1680, most of the lots next to the canal were developed. The population and radius of Amsterdam grew only slowly in our analyzed period, and during most of this time the Herengracht remained a mix of residential properties and offices (Geltner et al., 2014). Only in the end of the 19$^{\text{th}}$ century and during the mid of the 20$^{\text{th}}$ century did Amsterdam see a sudden expansion of its metropolitan area size.[7]

Since most of the dwellings on the Herengracht have survived until this day, we can use a repeat sales approach to construct a 'constant quality' house price index. As is typical for repeat sales models, our method does not control for capital expenditures (including large scale renovations) and depreciation (Harding et al., 2007). Especially in our case, with almost 200 years of data, many structures will have been altered completely. However, the most important characteristics of the properties remain the same: Location, land size and property type.[8] Moreover, pairs with an (absolute) average return of 50% per year were omitted from the data. These 'abnormal' returns are probably caused by large scale changes to the property between sales (Goetzmann and Spiegel, 1995; Clapp and Giaccotto, 1999). Also, in our specific application, we will only look at 30-year windows (see Section 4). The effect of (occasional) renovations on a house price index is less over 30-year windows than over the entire 200 years. Therefore, we expect that the repeat sales index will still be a good approximation of house price appreciation even in the long-run and we will use the index as basis for our analysis.

One particular complication with this type of historical data is that at some periods in time there are not many or even no sales (see Appendix A.1). In addition, it may also take a long period of time between transactions of the same house. This is something we explicitly have to take into account when constructing the house price index. For this research, we have therefore used a structural time series approach to estimate a local linear trend (i.e. house price index) model using the repeat sales methodology described by Francke (2010),

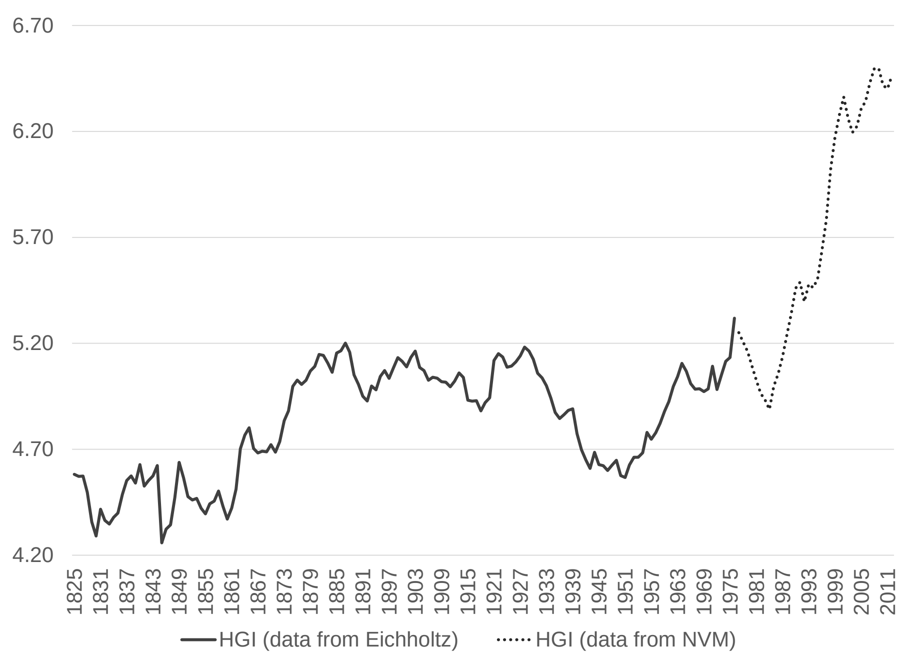
[6] Previous studies on the Herengracht, like Ambrose et al. (2013) and Eichholtz et al. (2015) actually use a Dutch national index after 1973.

[7] During the mid 19$^{\text{th}}$ century, many households moved out of the old city centre to live in one the new neighbourhoods, either next to the Vondelpark (for wealthier households) or in the Pijp (small and cheap housing). In the 1930s an extensive construction plan ('plan Zuid') was executed and supervised by the famous Dutch architect Berlage. After the Second World War, and with help of the Marshall-plan, Amsterdam expended to both the West and East.

[8] Especially in large cities (like Amsterdam), it is known that the location is an important price determinant as the land value takes up a large proportion of the total house price value (Glaeser et al., 2008).

instead of the more standard dummy variable approach (Bailey et al., 1963). This approach has the following benefits. Firstly, the model is tailor made for thin markets and is able to cope with the often large time between sales (which can be decades in our case). Secondly, because we estimate a stochastic (local linear) trend to construct a house price index, the resulting price index should be less sensitive to (short-run) outliers and more sensitive to macro-economic (long-run) shocks. A detailed description of the construction of the house price index is available in Appendix A.2 and A.3.

Figure 2: Log real house price index (1825 – 2012) and its sources.



Notes: Price index of the Herengracht (HGI) is based on the Herengracht micro data before 1973, using the Bayesian Repeat Sales model described in Appendix A.2 and A.3. After 1973 we use the NVM transaction database, using the same methodology.

Figure 2 contains the estimated real (log) price index from 1825 until 1973. The price index used in the analysis is deflated using the Consumer Price Index (CPI) which is directly available from the website of Statistics Netherlands. The CPI is the only deflator available to us for the entire time period.

Some of the explanatory variables in this research will be Amsterdam specific. Other explanatory variables, like GDP per capita, and unemployment rates, are on a (Dutch) aggregate level since this data is not available to us for Amsterdam for such a long time period. We are comfortable making the assumption that these macro variables will still affect Amsterdam house prices, as the Dutch economy is heavily intertwined since the Renaissance (Geltner et al., 2014). In addition, in case of the construction costs and interest rates, there is no reason to believe that there is a difference between the nationwide and Amsterdam specific time series. In that regard, it is also important to note that the Netherlands is comparable in terms of population and land size to a large Metropolitan Statistical Area (Dröes and Hassink, 2013). The Netherlands has a clear urban core (of which Amsterdam is part of) and a surrounding periphery, which accords with the definition of a MSA.[9]

3 The determinants of house prices

In long-run equilibrium, new building developments are determined by production costs and the costs of land. When prices go up, because of an increase in demand and a temporary shortage of houses, there is an incentive to construct new houses (Francke et al., 2009). The supply of these houses will bring house prices down to a new equilibrium (DiPasquale and Wheaton, 1994). Since we are interested in this long-run equilibrium, house prices should be examined by a macroeconomic housing model where supply and demand factors are both considered. In this study we focus on the following seven fundamental determinants of house prices: Housing supply, construction costs, Gross Domestic Product (GDP) per capita, the opportunity cost of capital (i.e. interest rates), population growth, unemployment rate, and the working age population as percentage of total population. These variables are typical in studies which focus on explaining house prices (from a long-run perspective), see Table 1.[10] In this Section, we discuss why these determinants are important, the different data sources that have been used, and the descriptive statistics.

Table 2 contains the data sources used in this paper. Most of the macro-economic factors are available from Statistics Netherlands (CBS). The interest rates, used to compute the opportunity cost of capital, are taken from the NVM and Homer and Sylla (2005). House prices, construction costs, and GDP are index values. Housing supply is measured as the number of houses in Amsterdam.[11] The labor force share, the opportunity cost of capital, and employment are in percentages. Population is the total number of inhabitants in Amsterdam. Table 2 also contains a broad classification of the macro-economic factors into (housing) demand

[9]The Netherlands is comparable in terms of population with large Metropolitan Statistical Areas (MSAs) such as the New York MSA and has the same GDP as the Los Angeles MSA.

[10]Unfortunately, there are no time series on the non-housing wealth of households for the total studied time period in the Netherlands.

[11]Homes built outside of Amsterdam can also affect house prices inside Amsterdam. We therefore also tested the effect of Dutch total housing supply on the Herengracht price index. The results are comparable.

Table 1: House price determinants according to a selected number of studies.

	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)	(VIII)	(IX)
Construction costs		✓	✓			✓		✓	
Housing or land supply		✓	✓	✓	✓				
GDP / income	✓	✓	✓	✓	✓	✓	✓	✓	
Wealth		✓			✓				
Interest rates / User costs	✓	✓		✓	✓	✓	✓	✓	✓
Population / Labor force	✓	✓	✓		✓	✓			
Country	US	UK/US	US	US	US	NL	NL	OECD	NL
Frequency	Y	Q	Y	Y	Y	Q	Y	Q	Y
Sample start	1979	1969	1979	1982	1981	1975	1965	1975	1650
Sample end	1996	1996	1995	2004	2003	2002	2009	2007	2005

I = Malpezzi (1999), **II** = Meen (2002), **III** = Capozza et al. (2002), **IV** = McCarthy and Peach (2004), **V** = Verbruggen et al. (2005), **VI** = Gallin (2006), **VII** = Francke et al. (2009), **VIII** = Adams and Füss (2010). **IX** = Ambrose et al. (2013).
US = United States, **NL** = the Netherlands, **UK** = the United Kingdom, **OECD** = the countries participating in the Organisation for Economic Co-operation and Development.
Y = Yearly, **Q** = Quarterly.
All studies mentioned in Table 1 used an error correction framework.

and supply factors (including their expected sign). Table 3 and 4 reports the descriptive statistics of the macro-economic factors respectively in levels and in log first-differences. All time series (except population, housing supply, labor force and unemployment) are deflated using the CPI. The time series (figures) in log-levels are given in Appendix A.4.

Table 2: Data sources: Macro-economic variables and house prices.

Variable	**Aggregation**	**Source**	**Unit**	**Type of determinant**	**Expected sign**
House price index	Amsterdam	Eichholtz (1997), NVM	index		
Housing supply	Amsterdam	OIS	units	Supply	-
Construction costs	Ams./Neth.	CBS, Neha	index	Supply	+
GDP per capita	Netherlands	CBS	index	Demand	+
Labor-Force	Netherlands	CBS	%	Demand	+
Opp. cost of capital	Netherlands	NVM, Homer and Sylla (2005)	%	Dem./Sup.	-/+
Unemployment	Netherlands	CBS	%	Demand	-
Population ($\times 1,000$)	Amsterdam	OIS	total	Demand	+
Consumer Price index	Netherlands	CBS	index		

CBS = Statistics Netherlands, Neha = Dutch Historical Archives, OIS = Research, Information and Statistics, City of Amsterdam and NVM = the Dutch Association of Realtors.

Real GDP per capita is seen as a proxy for economic activity and/or income (Englund and Ioannides, 1997). An increase in income is expected to have a positive effect on housing demand and, consequently, house prices. GDP has been increasing over time, on average, by 2.8 percent each year (see Table 4). Note that GDP is missing for the First and Second

Table 3: Descriptive statistics (real, levels): Macro-economic variables and house prices.

Variable	Mean	Std. Dev.	Min.	Max.	Skewn.	Kurt.	P-value
House price index	175.38	126.47	72.38	678.59	2.73	6.86	0.352
Housing supply ($\times 1,000$)	188.12	103.01	80.00	397.46	0.56	1.93	0.999
Construction costs	171.08	101.27	64.83	405.14	0.77	2.18	0.038
GDP per capita	6,257	11,888	97.00	45,569	2.07	6.20	0.579
Labor-Force	40.27%	2.17%	36.75%	47.59%	1.28	5.52	0.116
Opp. cost of capital	6.12%	2.59%	1.00%	12.79%	5.11	2.84	0.000
Unemployment	4.80%	2.90%	0.80%	17.40%	2.10	8.88	0.233
Population ($\times 1,000$)	556.54	237.81	192.33	872.43	-0.30	1.49	0.707
Consumer Price index	693.89	953.05	96.00	4,572.00	1.68	4.44	0.939
Number of observations	187						
Sample period	1825-2012						

Note. The reported P-values are the significance levels at which you can reject the null hypothesis of a unit root (Augmented Dickey Fuller test). All ADF tests were done with a constant and a trend. Critical values are taken from MacKinnon (2010). The test is conducted on the log of the variable.
The lag lengths differ per variable and are based on the Akaike Information Criterion.

Table 4: Descriptive statistics (real, ln first differences): Macro-economic variables and house prices.

Variable	Mean	Std. Dev.	Min.	Max.	Skewn.	Kurt.	P-value
House prices	0.010	0.071	-0.365	0.239	-0.370	3.729	0.000
Housing supply	0.009	0.011	-0.020	0.080	1.510	11.717	0.002
Construction costs	0.006	0.067	-0.338	0.278	-0.254	7.732	0.000
GDP per capita	0.028	0.057	-0.190	0.187	-0.318	3.899	0.000
Labor-Force	0.001	0.006	-0.017	0.022	0.528	4.780	0.324
Opp. cost of capital	-0.005	0.284	-0.847	0.999	0.142	5.066	0.000
Unemployment	0.001	0.207	-1.211	0.606	-0.813	9.969	0.000
Population ($\times 1,000$)	0.008	0.015	-0.068	0.080	-0.034	9.890	0.000
Consumer Price index	0.019	0.058	-0.145	0.360	0.739	9.094	0.000
Number of observations	186						
Sample period	1826-2012						

Note. The reported P-values are the significance levels at which you can reject the null hypothesis of a unit root (Augmented Dickey Fuller test). All ADF tests were done with a constant and a trend. Critical values are taken from MacKinnon (2010).
The lag lengths differ per variable and are based on the Akaike Information Criterion.

World War periods (see Appendix A.4).

Population growth is another typical demand-side variable. Between 1825 and 1970 population of Amsterdam steadily grew from less than $200,000$ to almost $870,000$ inhabitants (see Table 3). If supply, at least in the short-run, is fixed due to the time it takes to construct buildings (Harter-Dreiman, 2004) or legislation and lack of available space (Hilber and Vermeulen, 2012), an increase in population is expected to have a positive effect on house prices. Between 1970 and 1985 the population of Amsterdam shrunk with almost $200,000$ inhabitants due to large scale deurbanisation in that period. Glaeser and Gyourko (2005) found that population decline has a disproportionate effect on house prices, because the durability of housing means that it can take decades for negative urban shocks to be fully reflected in housing supply levels. During the 1990s and 2000s the population of Amsterdam grew to almost $800,000$.

Alternatively, working age population as percentage of total population might also have a positive effect on house prices (Case and Shiller, 2003). In essence, having a job is a precondition for owning a house. It typically, conjointly with income, determines house price dynamics (see also Chan, 2001). The working age is defined as the percentage of population aged between 20 and 65. The percentage working age population to total population during the period 1825–2012 has been between 36 to 47 percent (see Table 3). We only have data on the percentage working age population on a Dutch aggregate level.

Several studies also show that unemployment negatively affects house prices (see for example De Wit et al., 2013; Adams and Füss, 2010; Abraham and Hendershott, 1996). On average unemployment levels have been relatively low in the Netherlands (4.8 percent, see Table 3). However, in the 1930s - during the Great Depression - unemployment peaked at 17 percent.

The 5-year-annuity (nominal) mortgage interest rate, from 1973 onwards, is taken from the NVM. We use an index of the long-term Dutch government bond yields to proxy for the (mortgage) rates before this period (taken from Homer and Sylla, 2005). Subsequently, the real opportunity cost of capital is calculated by (see Williams, 2009)

$$OCC_t = (N_t - \mathrm{E}\big[\Delta cpi_t\big]) + 2\%, \tag{1}$$

where N_t is the nominal rate and cpi_t the log of the CPI in year t. As is usual when computing the opportunity cost of capital we take the expected inflation instead of inflation itself, by using a simple (7-year) Moving Average filter. We add 2 percent to measure (imputed) rental returns minus maintenance expenditures and other costs. During the 19^{th} century inflation was at times extreme. The time series of the opportunity cost of capital is very volatile (see Table 4). Only after the Second World War does the opportunity costs of capital seem to stabilize. We introduce a (lower bound) opportunity cost of capital cap of 1%, since we want to circumvent taking the log of a negative value and generally to filter out extreme values. This happened in 9 (consecutive) periods, with 5 of them being during the Second World War. The extremely low opportunity cost of capital in the 1970s is not surprising given the high inflation during this period (oil crises). The opportunity cost of capital can be interpreted as a demand and a supply-side factor. In particular, higher out-of-pocket costs (in case of increasing oppurtunity cost of capital) will decrease the demand for housing resulting in decreasing house prices. Alternatively, a higher interest rate may also have a negative effect on the ability of construction companies to obtain a loan, which decreases the supply of new housing and, consequently, increases house prices (DiPasquale and Wheaton, 1994; Capozza et al., 2002). The effect of the real interest rate is, therefore, mainly an empirical question.

The value of a property can be interpreted as the value of land plus the value of the structure (Bourassa et al., 2011). The construction cost of a property measures the replacement value of a structure and, therefore, is typically capitalized into house prices (see Case and Shiller, 1990; Davis and Palumbo, 2008). Furthermore, any given positive economic shock will be easier for an area to absorb if the housing stock can be increased at low cost. Therefore, we hypothesize that variables proxying for the cost of increasing the supply of housing should affect house prices (Capozza et al., 2002). A construction cost index for the Netherlands from 1913 onwards is directly available from Statistics Netherlands. We used data from the Dutch Economic Historical Archives (in Dutch abbreviated as Neha) as a basis to construct a measure for the construction costs index before 1913. The Neha reports the costs of all building materials in Amsterdam on a yearly basis from 1800 to 1913. Using the expert opinion of a Dutch architect specialized in 19$^{\text{th}}$ century buildings in Amsterdam we constructed a 'standard' home for this era. Next, the materials needed for this 'standard' home have been multiplied by the costs given by Neha. Based on the historic data from Statistics Netherlands, we assume that the material costs is constant at 70 percent of the total cost for new housing. For the remaining 30 percent, the costs are indexed by the national wage index (also obtained from Neha). Finally, the construction cost index is deflated by the CPI. The construction cost index quadrupled during our sample period (see Table 3). Also note the large increase in construction costs during World War I. This was mainly driven by the scarcity of (building) materials during this period.

Finally, we use Amsterdam specific housing supply to measure new housing construction. More specifically, we use the total number of housing units (rental and owner-occupied) in Amsterdam. This data is from Research, Information and Statistics, City of Amsterdam (OIS).[12] We use an index of the number of buildings to proxy for the number of housing

[12]There are many other potential measures of housing construction, such as the number of building permits issued (Hilber and Vermeulen, 2012; Paciorek, 2013) or new housing starts (Mayer and Somerville, 2000),

units before 1900. There is no data available on housing supply between 1900 and 1908 and for the first 10 years (1825 – 1835) of the data. In both instances we assume the housing stock is fixed.

During the 1930s the number of housing units sky-rocketed in Amsterdam. To this day, this is reflected in the substantial share of social housing in the Amsterdam (Van Ommeren and Koopman, 2011). We expect that new housing supply has a negative effect on house prices. Differences in supply elasticity have been argued to explain differences in house price levels and volatility across US metropolitan statistical areas (MSAs) (Green et al., 2005; Glaeser et al., 2008; Paciorek, 2013; Wheaton et al., 2014). In part, this may also reflect differences in regulation and space constraints (Hilber and Vermeulen, 2012). To the extent that those changes also occur over time, the size of the housing supply effect can change over time.

4 Methodology

The effect of the macro-economic determinants on house prices reflects both short-run fluctuations and long-run trends. These price dynamics can be captured by an error correction model (ECM), in which the dynamics are captured by a combination of current and past shocks and a gradual adjustment towards equilibrium. This model is based on the idea that the included time-series are, although non-stationary, cointegrated: Linear combinations of the variables are stationary. These linear combinations can be interpreted as equilibrium relationships. Therefore, it should be no surprise that error correction models are a popular tool in analysing long-run house prices. The standard error correction model is given by:

$$p_t = \beta + x_t'\delta + \varepsilon_t, \tag{2}$$

$$\Delta p_t = \sum_{k=0}^{n} \lambda_k \Delta p_{t-k} + \sum_{k=0}^{n} \Delta x_{t-k}'\theta_k + \alpha\big(p_{t-1} - p_{t-1}^*\big) + \eta_t, \tag{3}$$

where Eq. (2) represents the long-run equilibrium relation and Eq. (3) represents the short-run relation. Variable p_t is the (log) house price index at time t, p^* are the fitted values of Eq. (2) and x is a vector of macro-economic covariates (i.e. population growth, housing supply, labor force, construction cost, unemployment, opportunity cost of capital, and GDP per capita). Parameter α, in the short-run relation (Eq. (3)), measures the degree of mean reversion and is estimated from the data. The series $\big(p_t - p_t^*\big)$ is the error correction term. If the series $\big(p_t - p_t^*\big)$ is stationary, then $\big(p_t - p_t^*\big)$ is the co-integrating relation. In this study,

however the actual housing supply is generally regarded as the most appropriate measure (Paciorek, 2013).

we are especially interested in whether the parameters β and δ (the long-run cointegrating relationship) change over time. Moreover, we examine which set of covariates belongs to the cointegrating relationship at each point in time.

There are three key reasons why the parameters of Eqs. (2) – (3) might change over time. First, the parameters are likely time-varying due to real estate cycles. The main reason for real estate cycles to occur is because developers tend to overbuilt if developers' future projection of demand (usually measured by macro-economic variables) is positive (Pyhrr et al., 1999). Especially delivery lags and illiquidity worsens the ability of developers to respond quickly to changes in demand. Too much supply suppresses prices (Mayer and Somerville, 2000) for a considerable time, until aggregate demand and supply are in equilibrium again. Unfortunately, literature in the field of real estate cycles is not unambiguous on the average length of a typical cycle. The 'average' real estate cycle is somewhere between 18 years (Rabinowitz, 1980) and 60 years (Kaiser, 1997).

A second reason why the parameters are likely to be time-varying relates to regime shifts. The change of, or shift in, political and economic regimes usually occurs when a smooth change in an internal process (feedback) or a single disturbance (external shocks) triggers a completely different system behaviour. Common examples in the real estate literature are changes in legislation and innovations in the construction or mortgage market (Fernandez-Corugedo and Muellbauer, 2006).

Finally, the parameters can change because the fraction of land value to total value likely changes over time. Bourassa et al. (2011) showed that the determinants of the price of land and structure value are not the same, with changes in interest rates, income and population driving land values, while building values are primarily determined by construction costs. Thus, it is expected that the effect of the aformentioned fundamentals on the sum of both prices (i.e. total house prices) will depend on the fraction of land value. Even though the fraction of land value to total value is stationary in the very long run (as it must be between 0 and 1), it is expected that certain time periods are characterized by relatively high or low land values. Although we cannot exactly identify this effect - we do not have data on land to structure values - we will discuss the implications of this effect in more detail in the Section 5.

The challenge is to recognize when a 'cycle' starts and ends, which variables are part of the cointegrating relationship, and when there has been a shift in regime. In practice this is very difficult to identify. A yearly rolling regression with changing combinations of covariates is attractive in this regard, because it allows us to estimate a series of parameters without imposing any particular structure on the way in which conditional covariates change over time (Rossi, 1996). To simplify the procedure we estimate the error correction model in a (rolling) 2-step Engle-Granger framework (Engle and Granger, 1987). Since we estimate the long-run equation (first step, Eq. (2)) separate from the short-run equation (second step, Eq. (3)) and because no lags are included in the long-run equation, the total number of possible combinations of covariates reduces considerably.[13] To simplify the procedure even further, we use the same variables in the short-run equation as we use in the long-run equation.

Consequently, the specification of the rolling error correction model becomes:

$$p_t = \beta_r + x_t'(r)\delta_r + \varepsilon_t, \tag{4}$$

$$\Delta p_t = \lambda_r \Delta p_{t-1} + \Delta x_t'(r)\theta_r + \alpha_r \left(p_{t-1} - p_{t-1}^*\right) + \eta_t, \tag{5}$$

for $t = r, \dots, r+n-1$ and $r = 1, \dots, T-n+1$. The dependent and independent variables are of fixed length for any regression and represents the n periods (denoted *the window length*) immediately preceding period t. The function r reflects different combinations of covariates (one such combinations could be: Population and construction costs) per window. Thus, we get estimates for δ in every window for every combination of covariates r. This estimation procedure provides consistent estimates of the β and δ values - provided that p and x are cointegrated (Lütkepohl and Krätzig, 2004). In total the estimation and selection procedure consist of four steps.

Firstly, one important requirement is that the variables are integrated of order 1 (I(1)). If a variables is I(0) in a particular time window, the variable is excluded from the regression in that particular window (see Appendix A.5). Secondly, we regress all remaining combinations of covariates (r) in every window on house prices, using Eq. (4). We choose to fix the window length n at 25 years, as this is roughly the average length of a real estate cycle found in literate. For robustness, we also tried window lengths of 20, 30, 35, 40 and 45 years. However, the key results did not change. In addition, the number of cointegrating relationships was largest using windows of 25 years. In certain consecutive periods were we could not find any cointegrating relationships we used the results from the model with fixed window length $n = 30$ years in stead, which happened twice: During the turn of the 19$^{\text{th}}$ century and directly after the Second World War.

In the third step we establish which combination of covariates (x_r) are cointegrated. The most important requirement is that the error correction term $\left(p_{t-1} - p_{t-1}^*\right)$ is stationary. If multiple combinations of x are cointegrated in one window, the combination of variables which are cointegrated at the 1% level are reported in favor of the combination of variables which are cointegrated at the 5% level. If the degree of cointegration is similar for more than one combination of covariates in the same window, we chose the one with the highest

[13] Instead, estimating the R-ECM in a dynamic way, in an ADL framework or by the Johansen trace/eigenvalue test would acquire lags of the dependent and independent variables. The number of lags can be different for every variable. Different combinations of lag structures can result in different cointegrating relationships.

adjusted R^2. Another requirement is that the parameter estimates are at least significant at the 10% level. To account for the missing data for GDP per capita during the two war periods and, more in general, to estimate the effect of the wars on house prices a dummy for these periods is added to every regression in which one or both wars are within the window length.

In the fourth and final step, the model is re-estimated in first-differences, with the inclusion of the error correction term (lagged one period) estimated from the first step, see Eq. (5). Both (long- and short-run) models are estimated with OLS. In our application, we end up with 163 windows. Our 7 variables give a total of 127 possible combinations of covariates. Thus, we estimate more than $40,000$ different models (total long-run *and* short-run regressions). The next Section summerizes our findings.

5 Results

The results are summarized in Table 5 and Figure 3. Table 5 reports the most important variables during certain time periods, based on how many times the variable was part of a cointegrating relationship. The final two columns of Table 5 also give the average coefficient estimate and the average total effect over the entire period respectively. The total effect is calculated by multiplying the corresponding parameter coefficient by the min-max difference of the variable (per window).

The main drivers of house prices in the 19^{th} century were construction costs, housing supply, unemployment and population. After 1900 the Gross Domestic Product per capita starts to play a role and after the 1970s interest rates as well. After World War II population and housing supply are key additional determinants of house prices.

In over 60% of all windows at least one cointegrating relationship was found. In most cases the number of variables in the cointegrating relationship is between 2 and 4 (including a constant), see Figure 3. Moreover, each year is at least part of one period in which there is a cointegrating relationship.

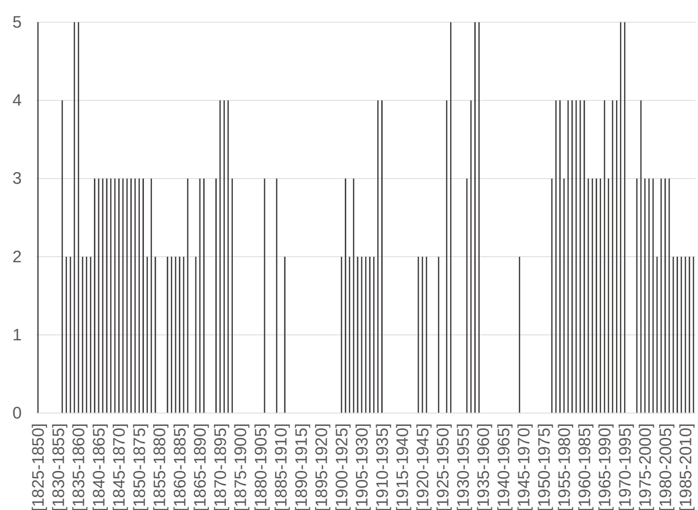
We did not find a cointegrating relationship in in all periods. One reason is that deviations from equilibrium do not adjust to the long-run equilibrium level when a housing market is in crisis. Indeed, both Hall et al. (1997) and later Nneji et al. (2013) found evidence for this effect using error correction Markov switching models for the UK and the US, respectively. The durable nature of housing and 'anchoring' of home-buyers are the main reasons explaining the absence of cointegration. Moreover, housing supply usually does not adjust negatively in considerable quantities (especially not in the short- or even long-run) in case demand for housing goes down (Glaeser and Gyourko, 2005). In addition, in time of crisis households tend to have too high reservation prices due to negative (home) equity or because of loss aversion

(Genesove and Mayer, 1997, 2001). In general, periods with a low number of cointegrating relationships are typically period of crises. For example, the least number of cointegrating relationships are in the period 1900 – 1945. During this period, the First World War (1910s), the Great Depression (1930s) and the Second World War (1940s) all affected the Dutch economy severely.[14]

Table 5: Most important variables per subperiod.

Variable	**1825 – 1900**	**1900 – 1945**	**1945 – 1970**	**1970 – 2012**	**Times part of cointegrating relationship**	**Mean coefficient effect**	**Mean total effect**
Housing supply	✓		✓		32	-1.51	-0.35
Construction costs	✓				33	0.78	0.33
GDP per capita		✓	✓	✓	35	0.63	1.29
Labor force		✓			17	5.63	0.13
Opp. cost of capital			✓	✓	18	-0.51	-1.24
Unemployment	✓				22	-0.63	-0.50
Population	✓		✓		40	2.09	0.73
Total					197		

Figure 3: Number of variables within a cointegrating relationship.



In all cases a constant is included.

From Table 5 it is also evident that unemployment and labor force are not part of a

[14] Other examples of crises in the 20th century are the Oil crisis in the 1970s and the dotcom and the Financial crisis in the beginning of the 21th century. In the 19th century the Belgian revolution (1830*s*) and two large crop failures (first one around 1850 and the second one at the end of the 19th century) resulted in an economic crisis.

cointegrating relationship in many windows. The opportunity cost of capital also is not part of a cointegrating relationship in many cases, however it is still of interest to us as the cointegrating relationships with opportunity cost of capital in it are all concentrated in the final period (whereas the others are more spread-out).

In the remainder of this Section we are going to discuss the results in more detail. Section 5.1 contains the time-varying, long-run, estimates of the demand side factors, and Section 5.2 those of the supply side factors. In Section 5.3, the model diagnostics, the error correction term estimates (i.e. the adjustment parameter), and the impact of the war time period are discussed. For expositional reasons, the other short-run, second step, estimates are not reported.

5.1 Demand side determinants of house prices

Figure 4 shows the rolling window point-estimates of population on house prices and its effect on house prices. If an estimate is only statistically significant at the 10 percent significance level this is denoted by * (otherwise it is significant at the 5 percent level). The horizontal axis gives the time window for which the error correction model was estimated. The left vertical axis gives the value of the parameter estimate (marginal effect) and the right vertical axis gives the total (maximum) effect of population on house prices.

The population variable is part of the cointegrating relationship in many of the rolling windows. Interestingly, population growth was mainly part of the cointegrating relationship in the 19$^{\text{th}}$ century and during the 1950s and 1960s. The effect of a one percent population increase has an average positive effect of about one to two percent on house prices. The total effect is between 5 percent and 160 percent. Population probably started to affect house prices in the beginning of the 1950s, due to the high birth rates (baby boomers) after the Second World War. The population parameter increases steadily from the 1950s onwards. This might be indicative of increasing land values as this period is characterized by high increases in real house prices as well, see Figure 2.

Figure 5 contains the estimates and total effect of the labor force on house prices. The effect of changes in the labor force is only part of the cointegrating relationship in 17 windows and it is only highly statistically significant in four cases. A one percent increase in the working age population as percentage of the total population has a positive 2 to 14 percent effect on house prices. However, the min-max range of labor force has been relatively low per window. Therefore, the total impact on house prices (averaged across periods) has only be around 13 percent (see Figure 5).

In Figure 6 the unemployment effects are depicted. Similarly to the effect of labor force, there are not many windows in which unemployment has a large effect on house prices.

Figure 4: Effect of population growth.

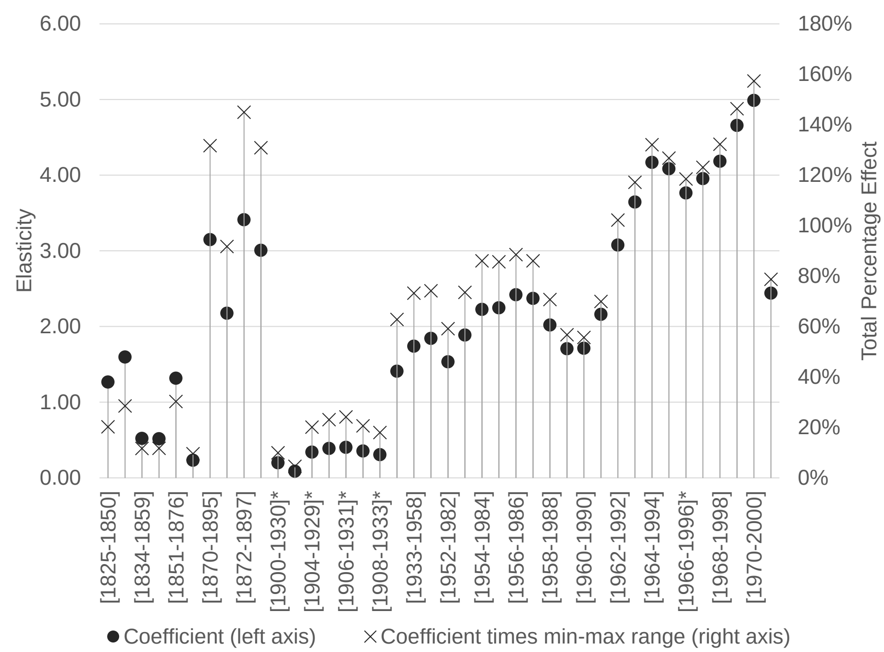
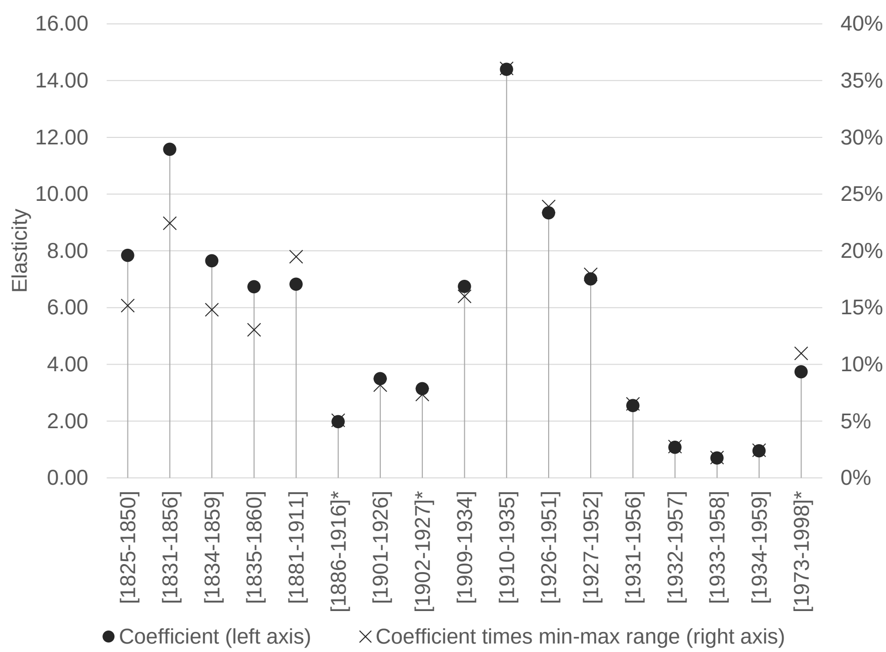
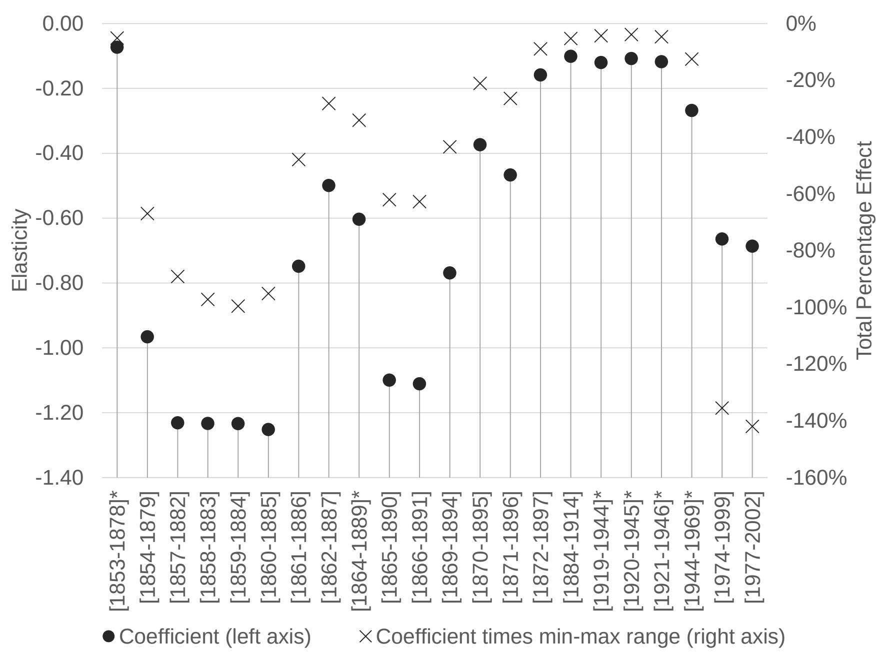


Figure 5: Effect of labor force.



Especially before the turn of the 19th century, there seems to be some effect of unemployment. The total effect (averaged across periods) of a change in unemployment on house prices is 50 percent.

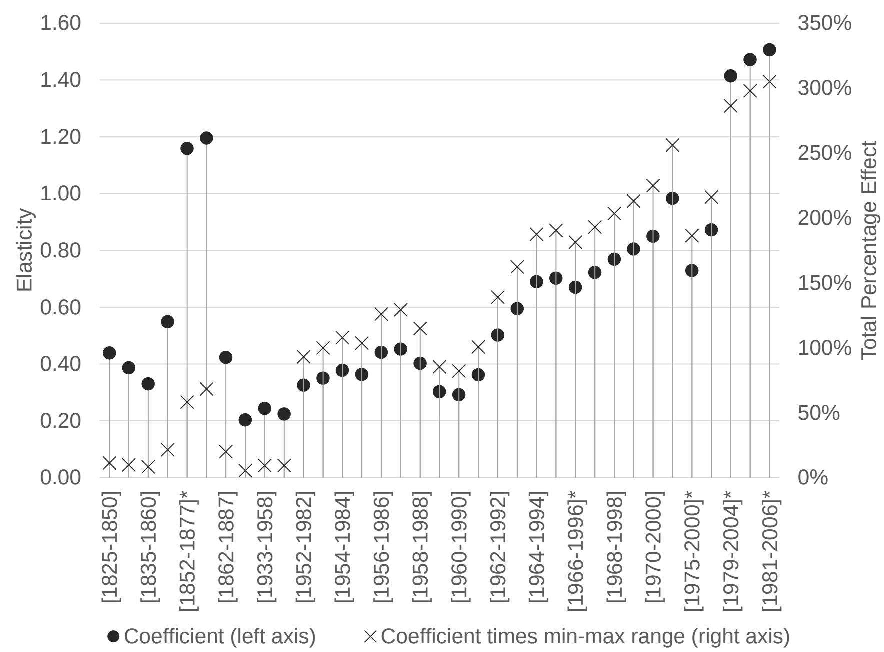
Figure 6: Effect of unemployment.



The effect of GDP per capita and the opportunity cost of capital are presented in Figure 7 and Figure 8, respectively. There are two striking similarities between these figures. First, both variables are part of the same cointegrating relationship in many of the time windows. Second, the coefficient estimates increase in size especially from the 1970s onwards. Although house prices were affected by GDP per capita during most of the 20th century, the coefficient is relatively small, less than 0.4 percent after a one percent increase in GDP per capita in most cases, before the 1960s. During this period most houses were financed by own savings. Instead, during the 1970s financial innovation and liberalization, combined with tax benefits on mortgage debt, made the use of mortgage debt more popular (Fernandez-Corugedo and Muellbauer, 2006) and, consequently, the impact of GDP on house prices increased. Interestingly, during the 1980s the Loan-to-Value cap increased to over 100 percent, a feature of the Dutch housing market which persists until this day (Andrews et al., 2011).

The amount of mortgage debt a household can borrow is not only determined by income but also by interest rates. It is, therefore, not surprising that both variables have jointly determined house prices after 1970. A one percent increase in GDP per capita had an effect on house prices between 0.2 and 1.4 percent. A percent decrease in the opportunity cost of capital (again the variable is in logs) has had a positive effect of 0.2 to 1.1 percent. Although we argued that the opportunity cost of capital can also be viewed as a supply-side factor affecting housing construction, our empirical estimates suggest that the demand side impacts are dominating. The total maximum effect of the opportunity costs of capital after 1970 is over a 100 percent on average. Apart from the mortgage market innovations during this period - and similar to population growth - increased land prices might also explain the ever increasing effect of both GDP and the opportunity cost of capital on house prices.

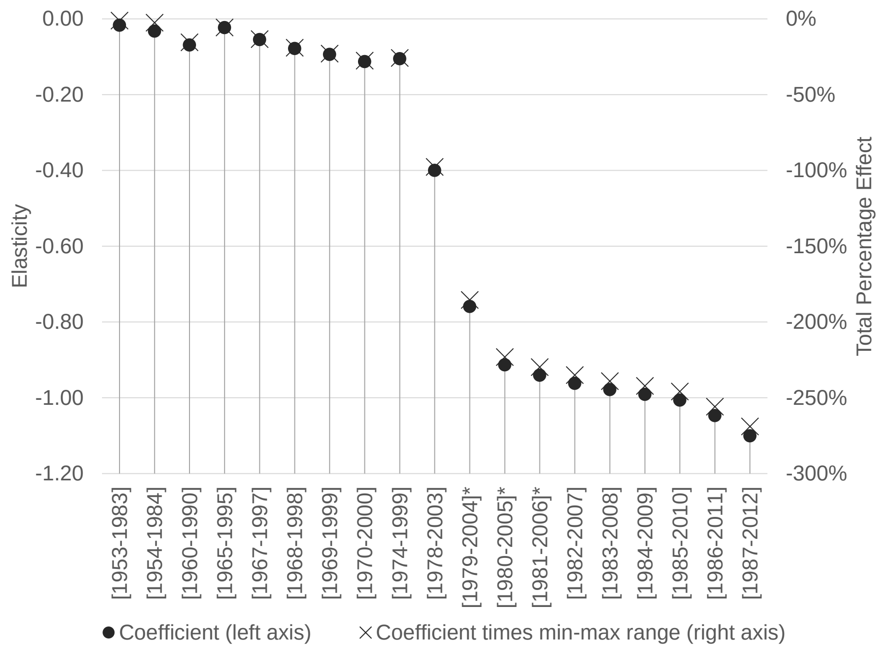
Figure 7: Effect of GDP per capita.



5.2 Supply side determinants of house prices

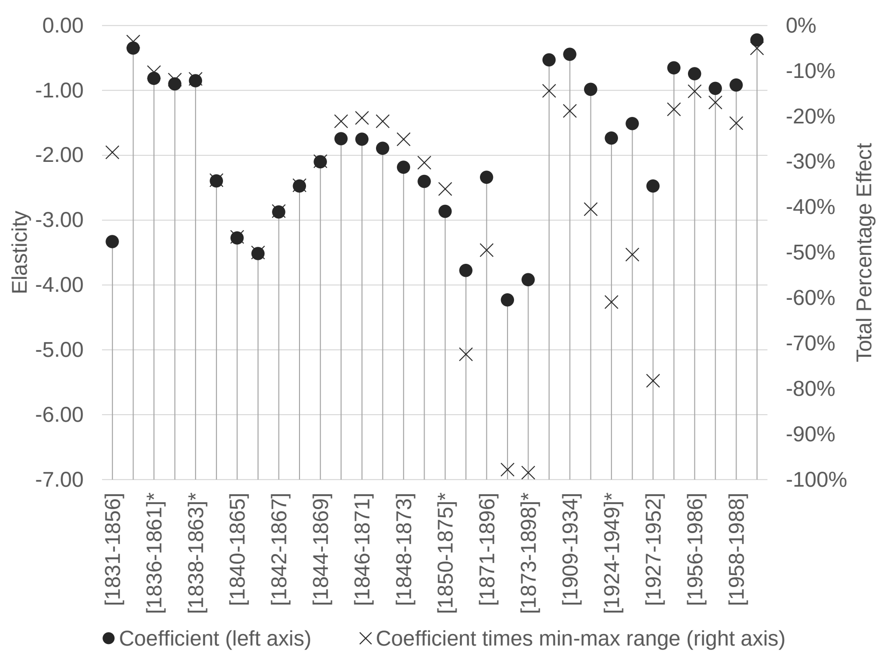
Housing supply and construction costs are considered supply side determinants of house prices. Figure 9 shows the long-run coefficient estimates for the housing supply variable. The effect of a one percent increase in housing supply results in a 1.5 percent house price

Figure 8: Effect of opportunity cost of capital.



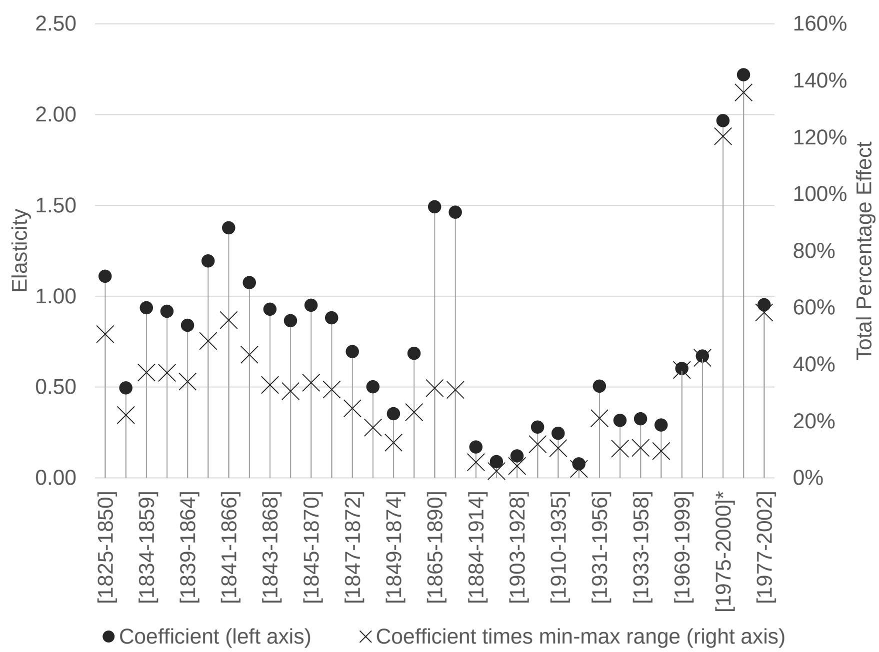
drop on average. Between 1840 and 1870 the coefficient of housing supply on house prices is relatively large, compared to the other periods. However, the min-max range of housing supply during this period is low, which implies that the total effect on house prices is low. Interestingly, during the end of the 19th century and during the 1930s there is a relatively large effect of housing supply on house prices. We also find an effect directly after the second world war. This likely reflects the large scale new construction during these periods.

Figure 9: Effect of housing supply.



The construction cost index is used to proxy for changes in structure values and for the rate at which constructors can add new housing supply to the market. Figure 10 shows the effect of construction costs on house prices. In most cases the elasticity is close to one during the 19th century. This suggest that house prices have mainly increased because the construction cost of houses increased. In a well-functioning market this is what one would expect. However, as mentioned earlier in Section 5.1, during most of this period population and unemployment was also part of the cointegrating relationship. If housing markets would be efficient, however, supply should adjust immediately if prices increase and population should not have an effect on house prices. The maximum total effect of construction costs on house prices ranges mostly between zero and 60 percent.

Figure 10: Effect of construction costs.



5.3 Model diagnostics, error correction, and the War Time period

This section discusses some remaining issues regarding model diagnostics, the adjustment parameter (α), and the effect of war on house prices. The upper left panel of Figure 11 presents the effect of the war time periods (World War I, World War II) on house prices. The Figure shows that the war time period had a negative effect on house prices of about 5 to 40 percent. This is likely an underestimate since the price index is only based on those houses which were actually sold.

The adjusted R-squared of the estimated regression models are depicted in the upper right panel in Figure 11. The average adjusted R-squared is quite high, about 0.6. There is quite some variation around the average. During the mid (end) of the 19$^{\text{th}}$ century, and during the 1930s and 1970s, the adjusted R-squared is above 0.80. There are also periods in which a standard house price model does not teach us much about house price dynamics. During these periods price changes seem to be determined by house price shocks. This results thus suggests that the predictability of house prices changes over time. A standard house price model has a varying degree of success in predicting house prices.

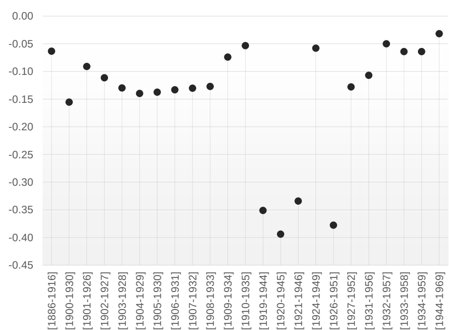
Finally, the lower panel of Figure 11 presents the rolling point-estimates of the coefficient of the error correction term in the short-term model, α_r in Eq. (5). The average effect of the error correction term is 0.33. This suggests that shocks out of equilibrium are absorbed within 3 years. However, there is large variation in this adjustment parameter. For example, during the beginning of the 19$^{\text{th}}$ century and 20$^{\text{th}}$ century there are several periods were shocks out of equilibrium are corrected almost instantaneously. Instead, in some periods the parameter estimate is as low as 0.05. At that rate shocks out of equilibrium are only absorbed after 20 years. This suggests that the housing market is more efficient in some periods than in others.

6 Concluding remarks

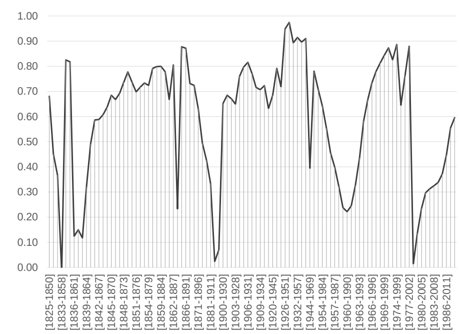
This paper has examined the determinants of house prices using almost 200 years of data from the Amsterdam housing market, the Netherlands. The results show that at different points in time there are different key determinants of house prices (cointegrating relationships).

During the 19$^{\text{th}}$ century, population, unemployment, housing supply, and construction costs are the main drivers of house price dynamics. At the start of the 20$^{\text{th}}$ century income starts to play a role. After World War II there are a few decades in which housing supply and especially population growth determine house prices. This reflects the post-war reconstruction efforts in the Netherlands and increases in housing demand as a result of the birth of the baby-boom generation and subsequent decrease in demand due to the adult baby-boomers

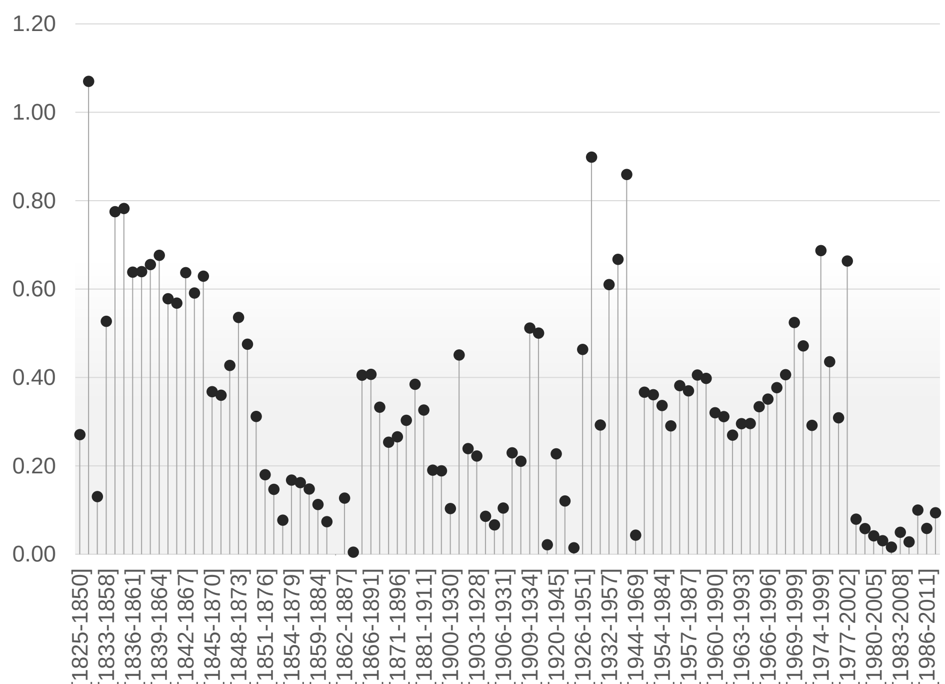
Figure 11: Model diagnostics and the War dummy.



(a) War dummy



(b) Rolling adjusted R^2



(c) Error Correction Term

leaving the city for greener and more spacious areas neighboring Amsterdam. Finally, from the 1970s onwards, income and interest rates start to have a large impact on house prices, most likely due to financial innovation and liberalization. Starting from the 1970s financing a house through mortgage debt became more popular. It also signals the beginning of a remarkable period in time in which house prices start to increase rapidly in many countries (i.e. the 1990s). We also showed that the predictability of house prices and the long term error correction mechanism (market efficiency) changes over time.

The results in this paper can explain why in some instances the key determinant of house prices differ across studies, even if those studies focus on the same housing market, and it provides a more long-term perspective on the fundamentals of house prices. The rapid increase in house prices can, for example, easily be mistaken as a bubble if it is unclear what the impact of different determinants are and how such determinants have changed over time. This paper has provided an analysis from the perspective of the Amsterdam/Dutch housing market. It would be interesting to see a similar analysis for other countries to examine to what extent the broad trends discussed in this study are generalizable across housing markets.

Acknowledgements

We thank Marc Francke, Rob van Alderen, and participants of the international AREUEA 2015 conference in Washington DC. for useful comments. Piet Eichholtz, the NVM and Marcel Theebe are gratefully acknowledged for providing the transaction data.

References

Abraham, J. M. and P. H. Hendershott (1996). Bubbles in metropolitan markets. *Journal of Housing Research 2*, 191–207.

Adams, Z. and R. Füss (2010). Macroeconomic determinants of international housing markets. *Journal of Housing Economics 19*(1), 38–50.

Ambrose, B. W., P. M. A. Eichholtz, and T. Lindenthal (2013). House prices and fundamentals: 355 years of evidence. *Journal of Money, Credit and Banking 45*(2-3), 477–491.

Andrews, D., A. C. Sánchez, and A. Johansson (2011). Housing markets and structural policies in OECD countries. Technical report, OECD Publishing.

Bailey, M. J., R. F. Muth, and H. O. Nourse (1963). A regression method for real estate price index construction. *Journal of the American Statistical Association 58*(304), 933–942.

Bostic, R. W., S. D. Longhofer, and C. L. Redfearn (2007). Land leverage: Decomposing home price dynamics. *Real Estate Economics 35*(2), 183–208.

Bourassa, S., M. Hoesli, D. Scognamiglio, and S. Zhang (2011). Land leverage and house prices. *Regional Science and Urban Economics 41*(2), 134–144.

Capozza, D. R., P. H. Hendershott, C. Mack, and C. J. Mayer (2002). Determinants of real house price dynamics. Technical report, National Bureau of Economic Research.

Case, K. E. and R. J. Shiller (1987). Prices of single family homes since 1970: New indexes for four cities. *New England Economic Review*, 45–56.

Case, K. E. and R. J. Shiller (1989). The efficiency of the market of single-family homes. *The American Economic Review 79*, 125–137.

Case, K. E. and R. J. Shiller (1990). Forecasting prices and excess returns in the housing market. *Real Estate Economics 18*(3), 253–273.

Case, K. E. and R. J. Shiller (2003). Is there a bubble in the housing market? *Brookings Papers on Economic Activity 2*, 299–342.

Chan, S. (2001). Spatial lock-in: do falling house prices constrain residential mobility? *Journal of Urban Economics 49*(3), 567–586.

Clapp, J. M. and C. Giaccotto (1999). Revisions in repeat-sales price indexes: Here today, gone tomorrow. *Real Estate Economics 27*, 79–104.

Davis, M. A. and J. Heathcote (2007). The price and quantity of residential land in the United States. *Journal of Monetary Economics 54*(8), 2595–2620.

Davis, M. A. and M. G. Palumbo (2008). The price of residential land in large US cities. *Journal of Urban Economics 63*(1), 352–384.

De Wit, E. R., P. Englund, and M. K. Francke (2013). Price and transaction volume in the Dutch housing market. *Regional Science and Urban Economics 43*(2), 220–241.

DiPasquale, D. and W. C. Wheaton (1994). Housing market dynamics and the future of housing prices. *Journal of Urban Economics 35*, 1–27.

Dröes, M. I. and W. H. J. Hassink (2013). House price risk and the hedging benefits of home ownership. *Journal of Housing Economics 22*(2), 92–99.

Eichholtz, P. M. A. (1997). A long run house price index: The Herengracht index, 1628–1973. *Real Estate Economics 25*(2), 175–192.

Eichholtz, P. M. A., R. Huisman, and R. C. J. Zwinkels (2015). Fundamentals or trends? a long-term perspective on house prices. *Applied Economics 47*(10), 1050–1059.

Engle, R. F. and C. W. J. Granger (1987). Co-integration and error correction: Representation, estimation, and testing. *Econometrica 55*, 252–276.

Englund, P. and Y. M. Ioannides (1997). House price dynamics: an international empirical perspective. *Journal of Housing Economics 6*(2), 119–136.

Fernandez-Corugedo, E. and J. Muellbauer (2006). Consumer credit conditions in the United Kingdom. *Bank of England Working Paper No. 314*.

Francke, M. K. (2010). Repeat sales index for thin markets: a structural time series approach. *Journal of Real Estate Finance and Economics 41*, 24–52.

Francke, M. K., S. Vujić, and G. A. Vos (2009). Evaluation of house price models using an ECM approach: the case of the Netherlands. Technical report, Ortec Finance.

Gallin, J. (2006). The long-run relationship between house prices and income: Evidence from local housing markets. *Real Estate Economics 34*, 417–438.

Geltner, D. M., N. G. Miller, J. Clayton, and P. M. A. Eichholtz (2014). *Commercial Real Estate, Analysis & Investments* (4th Edition ed.). Cengage Learning.

Genesove, D. and C. Mayer (1997). Equity and time to sale in the real estate market. *American Economic Review 87*, 255–269.

Genesove, D. and C. Mayer (2001). Loss aversion and seller behavior: Evidence from the housing market. *The Quarterly Journal of Economics 116*(4), 1233–1260.

Glaeser, E. L. and J. Gyourko (2005). Urban decline and durable housing. *Journal of Political Economy 113*(2), 345–375.

Glaeser, E. L., J. Gyourko, and A. Saiz (2008). Housing supply and housing bubbles. *Journal of Urban Economics 64*(2), 198–217.

Goetzmann, W. N. and M. Spiegel (1995). Non-temporal components of residential real estate appreciation. *The Review of Economics and Statistics 77*, 199–206.

Green, R. K., S. Malpezzi, and S. K. Mayo (2005). Metropolitan-specific estimates of the price elasticity of supply of housing, and their sources. *The American Economic Review 95*(2), 334–339.

Hall, S., Z. Psaradakis, and M. Sola (1997). Switching error-correction models of house prices in the United Kingdom. *Economic Modelling 14*(4), 517–527.

Harding, J. P., S. S. Rosenthal, and C. F. Sirmans (2007). Depreciation of housing capital, maintenance, and house price inflation: Estimates from a repeat sales model. *Journal of Urban Economics 61*, 193–217.

Harter-Dreiman, M. (2004). Drawing inferences about housing supply elasticity from house price responses to income shocks. *Journal of Urban Economics 55*(2), 316–337.

Hilber, C. A. L. and W. Vermeulen (2012). The impact of supply constraints on house prices in England. *London School of Economics Spatial Economics Research Centre Discussion Paper* (119).

Homer, S. and R. Sylla (2005). *A History of Interest Rates* (4th Edition ed.). John Wiley & Sons, Ltd.

Johansen, S. (1995). Likelihood-based inference in cointegrated vector autoregressive models. *OUP Catalogue*.

Kaiser, R. W. (1997). The long cycle in real estate. *Journal of Real Estate Research 14*(3), 233–257.

Lütkepohl, H. and M. Krätzig (2004). *Applied Time Series Econometrics.* Cambridge University Press.

MacKinnon, J. G. (2010). Critical values for cointegration tests. *Queen's Economic Department working paper 1227.*

Malpezzi, S. (1999). A simple error correction model of house prices. *Journal of Housing Economics 8*, 27–62.

Mayer, C. J. and T. C. Somerville (2000). Land use regulation and new construction. *Regional Science and Urban Economics 30*(6), 639–662.

McCarthy, J. and R. W. Peach (2004). Are home prices the next bubble? *FRBNY Economic Policy Review 10*(3), 1–17.

Meen, G. (2002). The time-series behavior of house prices: A transatlantic divide? *Journal of Housing Economics 11*, 1–23.

Ngene, G. M., C. A. Lambert, and A. F. Darrat (2014). Testing long memory in the presence of structural breaks: An application to regional and national housing markets. *The Journal of Real Estate Finance and Economics*, 1–19.

Nneji, O., C. Brooks, and C. W. R. Ward (2013). House price dynamics and their reaction to macroeconomic changes. *Economic Modelling 32*, 172–178.

Paciorek, A. (2013). Supply constraints and housing market dynamics. *Journal of Urban Economics*.

Pyhrr, S. A., S. E. Roulac, and W. L. Born (1999). Real estate cycles and their strategic implications for investors and portfolio managers in the global economy. *Journal of Real Estate Research 18*(1), 7–68.

Rabinowitz, A. (1980). *The Real Estate Gamble: Lessons from 50 Years of Boom and Bust*. Amacom.

Rossi, P. H. (1996). *Modelling stock market volatility: bridging the gap to continuous time*. Academic Press.

Van Ommeren, J. and M. Koopman (2011). Public housing and the value of apartment quality to households. *Regional Science and Urban Economics 41*(3), 207–213.

Verbruggen, J. P., H. Kranendonk, M. van Leuvensteijn, and M. Toet (2005). Which factors determine the house-price development in the Netherlands? Technical Report 81, Centraal Planbureau (Netherlands Bureau for Economic Policy Analysis).

Wheaton, W. C., S. Chervachidze, and G. Nechayev (2014). Error correction models of MSA housing "supply" elasticities: Implications for price recovery. *MIT working paper series No. 14-05*.

Williams, D. (2009). House prices and financial liberalisation in Australia. *University of Oxford Working Paper*.

A Appendix

A.1 Transaction Data Herengracht

The data description is provided in Tables A1 – A3 and Figure A.1 for both the Herengracht data (henceforward HR data) and the NVM data. The descriptives are based on the data after filters are applied. If a home was converted to an office or if the home was combined with a neighboring home the sale was removed from the data.

It is apparent that there are some differences between the datasets. Obviously the NVM data has less observations than the HG data, as the NVM data only spans 40 years compared to more than 300 years for the HG data.[15] As a results there are more than 10 times as many observations in the HG data compared to the NVM data. We also observe more unique properties in the HG data (580) compared to the NVM data (200), see Table A1. This also has an effect on how many times a unique property is sold on average. Most homes in the HG data were sold 6 times in our data. One specific home was even sold 17 times during the 350 year period (Table A2). On average homes are sold 'only' 3 times in the NVM data.

The average number of transactions per year is approximately 10 for the HG database and 12 for the NVM database. However, in some years there are no transactions (see Figure A.1). The biggest differences between the datasets is the average time between sales and the return statistics. The average years between sales in the HG data is over 30 years, whereas it is only 8 years in the NVM data. This can partly be explained by realizing that long-run occupants are observed more frequently in datasets spanning over a longer period. Also, as the latter part of the 20$^{\text{th}}$ century was characterized by high house price appreciation (also see Table A3), a positive price-volume correlation might also be an explanation (De Wit et al., 2013). On average the house price appreciation between 1973 – 2014 is four times as high compared to the period 1628 – 1972.

Table A1: Descriptive statistics Herengracht data.

Description	**HG data**	**NVM data**
Number of transactions	3,416	501
Number of transactions (with at least two sales)	2,953	304
Number of different homes	580	197
Minimum year of sale	1628	1973
Maximum year of sale	1972	2014
Average years between repeat sales	34	8

[15] We have estimated the price index using data starting from the 17$^{\text{th}}$ century onwards to increase the accuracy of our estimates. The analysis in the remainder of this study, however, is based on the price index from 1825 onwards since the explanatory macro factors are only available as of 1825.

Figure A.1: Number of transactions per year, Herengracht data.

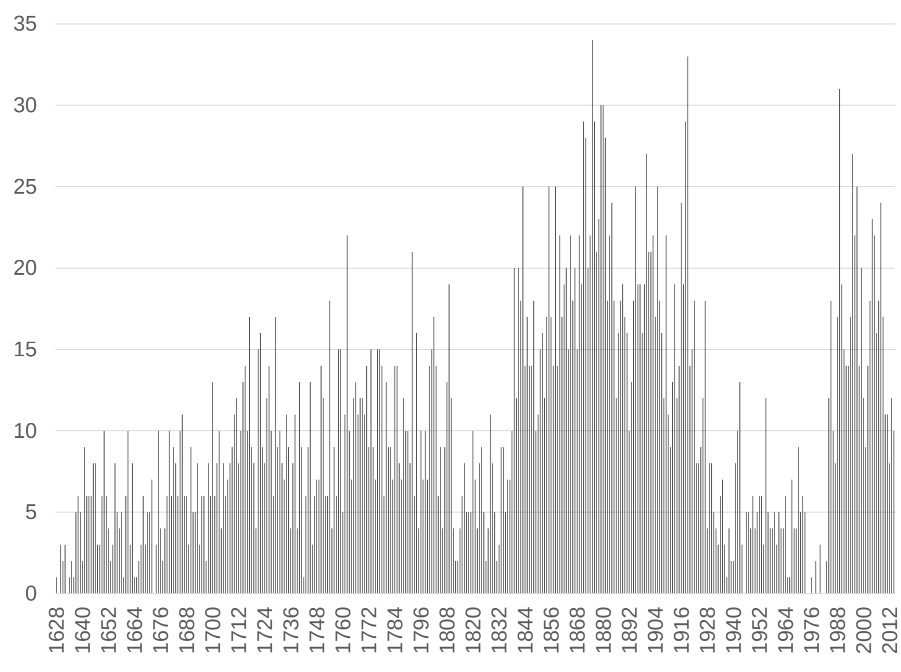


Table A2: Number of sales of the same property

Number of sales	**HG data**	**NVM data**
2	66	135
3	54	37
4	62	18
5	68	4
6	95	3
7	61	1
8	55	0
9	41	0
10	24	0
11	16	0
12	12	0
13	5	1
14	4	0
15	2	0
17	1	0

Table A3: Log nominal return statistics

	Annual log nominal returns	
Quantile	**HG data**	**NVM data**
average	0.018	0.074
0.025	-0.052	-0.203
0.050	-0.033	-0.092
0.100	-0.019	-0.023
0.500	0.005	0.075
0.900	0.059	0.194
0.950	0.104	0.270
0.975	0.185	0.396

Unfortunately we do not observe addresses in the Herengracht data, so we cannot merge the two datasets. As a results we estimated two separate indices (using the same methodology) and subsequently conjoined them.[16]

[16] Alternatively, we also augmented the Herengracht index after the 1970s with house price data from the Statistics Netherlands (CBS) and the NVM, covering the period from 1965 through 2005. The numbers denote median house prices for the year. However, after re-estimating the models the results and conclusions remain the same, albeit that the coefficients in the end of the sample were lower. This should come as a surprise as price appreciation on the Herengracht was threefold that of the Netherlands in general, in nominal terms.

A.2 Constructing a House Price Index

The methodology to estimate the price index using the Herengracht data is described extensively in Francke (2010). Here we give a brief description of the model and some descriptive statistics of the data. The 'standard' Case and Shiller repeat sales is given by:

$$\ln\left(\frac{P_{i,t}}{P_{i,s}}\right) = \beta_t - \beta_s + \alpha_{i,t} + \epsilon_{i,t}, \qquad \epsilon_{i,t} \sim N(0, \sigma_\epsilon^2), \tag{6}$$

$$\alpha_{i,t+1} = \alpha_{i,t} + \eta_{i,t}, \qquad \eta_{i,t} \sim N(0, \sigma_\eta^2), \tag{7}$$

for $t = 1, \ldots, T$ and $i = 1, \ldots, M$, where T is the number of years and M is the number of houses. P are house prices sold at time t (sale) and s (buy), with $t > s$. Subscript i is for the individual properties. The coefficient β_t is the logarithm of the cumulative price index at time t. The random walk component (α) is the cumulative idiosyncratic drift of each property (Case and Shiller, 1987), since the variance of the error term is related to the interval between time of sales.

In the repeat sales model it is typically assumed that the β_t's are fixed unknown parameters. In the methodology described by Francke (2010), it is assumed that β_t is a scalar stochastic trend process in the form of a local linear trend model, in which both the level and slope can vary over time. The local linear trend model is given by:

$$\beta_{t+1} = \beta_t + \kappa_t + \zeta_t, \qquad \zeta_t \sim N(0, \sigma_\zeta^2), \tag{8}$$

$$\kappa_{t+1} = \kappa_t + \xi_t, \qquad \xi_t \sim N(0, \sigma_\xi^2). \tag{9}$$

The local linear trend model 'in differences' is estimated with the Bayesian procedure described in Francke (2010), avoiding the somewhat more usual ad hoc two-step procedure. The model can be expressed as a linear regression model with a prior for β, induced by the local linear trend model. Estimates of parameters are obtained by maximizing the likelihood of the 'differenced' data.

A.3 Estimation Results

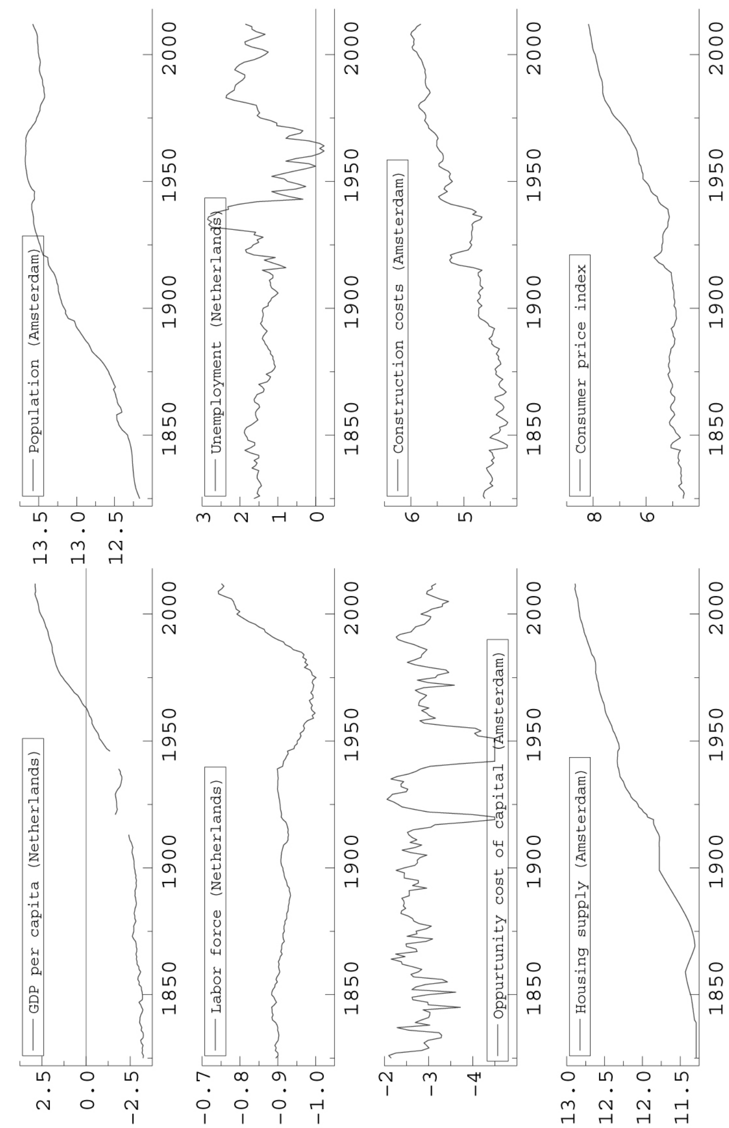
The results of the model on both datasets is given in Table A4. The standard error of regression (σ) is relatively high, approximately 30 percent for the HG data. This could be because of the large average time between sales (34 years, Table A1). Taking into account the large average time between sales and the large variance in time between sales (Table A2), it is surprising to note that the standard deviation for the individual random walks (σ_η) is near 0 in the HG data. The individual random walk hyperparameter (σ_η) is larger and significant for the NVM data. This also helps explain why the noise is a lot lower in the NVM data compared to the HG data. Both models behave more or less as a random walk, as the stochastic drift hyperparameter (σ_ξ) is near 0 and/or not-significant in both cases. The signal (σ_ζ) is slightly higher for the model estimated on the NVM data.

Table A4: Estimation results

	Herengracht data			
	Estimate	**Log estimate**	**Std. error.**	**t-value**
σ	0.310	-1.171	0.015	84.28
σ_η	0.000	-10.471	1.688	6.20
σ_ζ	0.083	-2.492	0.162	15.39
σ_ξ	0.003	-5.936	2.325	2.55
	NVM data			
	Estimate	**Log estimate**	**Std. error.**	**t-value**
σ	0.164	-1.809	0.064	28.190
σ_η	0.052	-2.962	0.208	14.245
σ_ζ	0.107	-2.233	0.187	11.916
σ_ξ	0.000	-9.431	49.520	0.190

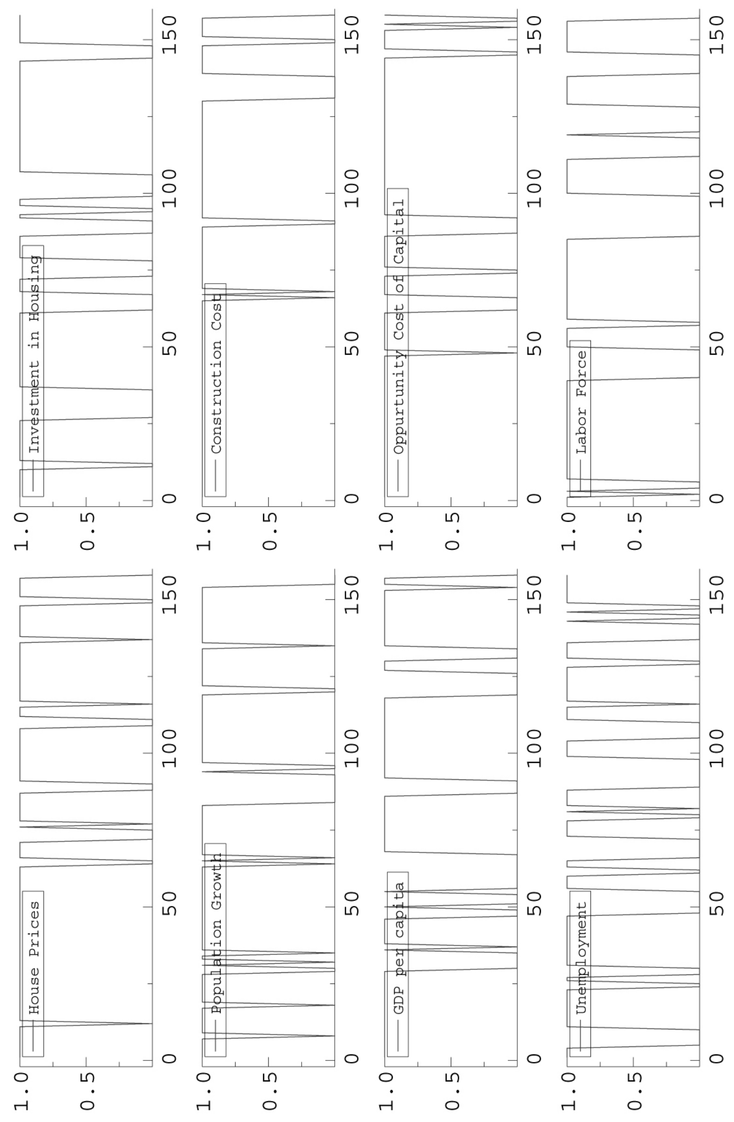
A.4 Long-run time series: Macro Determinants

Figure A.2: Time series in Log levels.



A.5 Order of Integration

Figure A.3: Order of integration of the variables.



Note. Order of integration: 1 means I(1), which is necessary for the R-ECM framework. We only use I(1) variables in the regressions. The x-axis is the number of years starting from 1825.