

The Total Return and Risk to Residential Real Estate

Piet Eichholtz

Maastricht University

Matthijs Korevaar

Erasmus University Rotterdam

Thies Lindenthal

University of Cambridge

Ronan Tallec

Université Paris 2 Panthéon-Assas

We estimate total returns to rental housing by studying over 170,000 hand-collected archival observations of prices and rents for individual houses in Paris (1809–1943) and Amsterdam (1900–1979). The annualized real total return, net of costs and taxes, is 4.0% for Paris and 4.8% for Amsterdam and entirely comes from rental yields. Our returns weakly correlate with the implied returns in Jordà et al. (2019) and are substantially lower. We decompose total return risk at the individual asset level and find that yield risk becomes an increasingly important component of property-level risk for longer investment horizons. (*JEL* G11, G12, N20, R30)

Received February 29, 2020; editorial decision January 21, 2021 by Editor Stijn Van Nieuwerburgh. Authors have furnished an Internet Appendix, which is available on the Oxford University Press Web site next to the link to the final published paper online.

Housing is the world's largest asset class, but with some exceptions, it did not have much institutional investor interest in the decades before the Great Recession. Since then, however, housing markets all over the world have been

We thank the staffs of the Paris and Amsterdam city archives for smoothly dealing with our incessant requests for archival materials. Furthermore, we thank Bruno Deffains and Will Goetzmann for their encouragement and two anonymous referees, Anne Alonzo, Jacques Friggit, Dmitry Kuvshinov, Gilles Postel-Vinay, Sébastien Pouget, Jean-Laurent Rosenthal, Carolin Schmidt, Christophe Spaenjers, Stijn van Nieuwerburgh (editor), and Michael Visser for their helpful comments. The Netherlands Organization for Scientific Research funded Korevaar under the Research Talent scheme [grant no. 406.16.552]. A Cambridge Humanities Research Grant and a Lincoln Institute Dissertation Fellowship funded part of the data collection for this paper. Supplementary data can be found on *The Review of Financial Studies* web site. Send correspondence to Matthijs Korevaar, korevaar@ese.eur.nl.

The Review of Financial Studies 34 (2021) 3608–3646

© The Author(s) 2021. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

doi:10.1093/rfs/hhab042

Advance Access publication April 12, 2021

booming, and so has investor interest. Both private and institutional investors are putting capital into rental housing (Bracke 2021; Mills, Molloy, and Zarutskie 2019). No doubt, their interest has been spurred by the recent performance of housing markets, with high levels of house price growth observed across the globe in the past few decades (Knoll, Schularick, and Steger 2017).

In a recent paper, Jordà et al. (2019) aim to determine the total rate of return to housing and to compare it to the performance of stocks and bonds all over the world. Their results—based on secondary data sources—suggest that housing returns are surprisingly high given their risk. Indeed, in a follow-up paper, Jordà et al. (2019) point out an unsolved risk premium puzzle for housing investments.

However, the housing returns data on which these papers are based suffer from several measurement problems, which led to a debate on whether the reported high returns to housing are real, or the result of mismeasurement (Chambers, Spaenjers, and Steiner 2021; Eislefeldt and Demers 2018; Dimson, Marsh, and Staunton 2018). The main aim of our paper is to shed light on this issue.

The crucial piece of information that is typically lacking in studies that aim to compute long-term housing returns and risk, both at the aggregate and at the individual levels, is an accurate assessment of the rental yield. Early papers aiming to assess the total return to housing used an imputed rent (Flavin and Yamashita 2002) or based it on national accounts data (Piazzesi, Schneider, and Tuzel 2007). More recently, the common approach to deal with this issue is to estimate yields based on actual house price and rental series, but these tend to pertain to different housing market segments (e.g., Brounen et al. 2014; Eislefeldt and Demers 2018; Giglio et al. 2018; Jordà et al. 2019).

This issue goes beyond measuring housing returns themselves. Without data on actual yields, the use of implied rent-price ratios or rental returns has been standard in much of the literature on housing markets (e.g., Himmelberg, Mayer, and Sinai 2005; Sinai and Souleles 2005; Campbell et al. 2009; Ambrose, Eichholtz, and Lindenthal 2013; Sommer, Sullivan, and Verbrugge 2013; Favilukis, Ludvigson, and Van Nieuwerburgh 2017). Our paper overcomes this hiatus, and our first main contribution is to calculate total returns and risks to residential real estate at aggregate and individual property level by studying primary historic data on house prices and rents for the same homes. We do so for two important housing markets: Paris (1809–1943) and Amsterdam (1900–1979). We then compare these novel return estimates to existing measures based on the Jordà et al. (2019) implied returns, which cover the two cities for overlapping time periods.

The second main contribution of our paper lies in a better understanding of the idiosyncratic risk to housing investments. Our data for Amsterdam allow us to decompose total risk into idiosyncratic risk and market risk at the property level, and to do that for different holding periods. This provides a unique picture of the role of idiosyncratic risk – and of the role of the yield component of the return therein. The market return and risk to housing investments are of

limited relevance for investment performance, as many residential property investors hold highly concentrated portfolios, due to the indivisibility of assets, their capital intensity, and high transaction costs.¹ For markets in which full diversification is unattainable, both theory (see, for instance, Levy 1978; Merton 1987) and empirical works (e.g., Fu 2009; Eiling et al. 2019) suggest that idiosyncratic risk and expected returns are linked in the cross-section. Existing work has looked at idiosyncratic capital gains risk in residential properties (e.g., Merton 1987; Peng and Thibodeau 2017; Giacomelli 2021; Eiling et al. 2019). However, this ignores yield risk resulting from changes in the rental values of properties. To the best of our knowledge, our paper is the first to study property-level changes in yields and the contribution of property yield risk to total idiosyncratic risk, and to do that over increasing investment horizons.

Given housing's important role in the economy and investment portfolios, it is important to establish the annual total return and risk to residential real estate in a way that avoids measurement problems as far as possible, using a data set that is large and representative enough for reliable inference. Our paper aims to measure total returns to residential real estate as accurately as possible and to assess the risk of that investment at the aggregate level and property level.

A longer-run perspective is essential, since holding periods for rental housing are typically long: in our sample, median holding periods are about 10 years, and we aim to study the relative role of idiosyncratic and systematic risk for holding periods of up to 20 years. Despite the enormous size of the housing market, data limitations regarding capital appreciation, gross rental yields, and taxes and costs have so far rendered it difficult to make accurate estimates of the total return and risk on long-term residential real estate investments.

Three recent studies have specifically attempted to construct total return and risk estimates to real estate investments. Eisfeldt and Demers (2018) study total returns to rental housing investments between 1986 and 2014 for a panel of cities in the United States. Although they do not observe actual yields in this period, they construct implied yields by extrapolating a city-specific hedonic pricing model for rental properties to owner-occupied properties. Both within and across cities, they find that rental yields decline in price tiers. Cities with lower net yields experience higher price appreciation, but have lower Sharpe ratios because capital gains are more volatile than yields. However, within cities, this logic reverses, as low price tier areas experience higher total returns due to both higher yields and higher capital gains.

Longer observation periods are important to establish the time-series properties of aggregate housing risk and returns over different market conditions and economic cycles. A recent influential paper by Jordà et al. (2019) constructs total return indexes from a great number of existing house price and rent indexes,

¹ In the Netherlands, for example, data from CBS (Statistics Netherlands) show 47% of the private rental stock is owned by individuals. Of these buy-to-let investors, 80% own a single property and only 4% own more than five properties.

based on construction methods that vary over time and across countries. Given the ambition of their paper, that is, to assess housing investment returns and risks for a large cross-section of countries between 1870 and today, this is understandable. Their data collection is momentous as it is. However, their series may suffer from measurement error in all dimensions of the total return, that is, the capital appreciation, the gross rental yield, and in taxes and costs. Moreover, although their indexes aim to pertain to nations as a whole, they mostly use data on the level of a nation's main city or cities for the early parts of their indexes and switch to national data at different moments for each index. All this might make inferences based on their findings unreliable, and a key motivation for our paper is to investigate the extent to which such long-term implied return series are affected by measurement errors.

Chambers, Spaenjers, and Steiner (2021) have made an impressive effort to construct estimates of total real estate returns for England, using the archives of four prominent 'Oxbridge' colleges between 1901 and 1983. The archival ledgers allow to precisely track property-level annual rental income and costs over time and can be matched to transaction prices when properties are purchased or sold. This enables them to measure returns much more precisely, and they find that long-run real estate investment is less profitable than suggested by Jordà et al. (2019), with a real return to housing of 2.3%, and 5.4% for agricultural property. Given its aim and sample period, their paper is complementary to our study. Relative to our paper, their study also covers agricultural land and commercial real estate and specifically investigates the role of costs in driving asset-level returns and risk. However, because colleges infrequently transacted property, they cannot directly measure aggregate capital gains and total returns. To derive a total return statistic for the entire period, they use changes in the U.K. house price index of Knoll, Schularick, and Steger (2017) and adjust it to match the yields they do observe.

In this paper, we construct long-term annual total return indexes and provide a picture of the return and risk to rental housing investment at the aggregate and the property level, and for different holding periods. To do so, we study two previously unexplored primary data sets of house prices and rents on individual homes for Paris (1809–1943) and Amsterdam (1900–1979). These data sets are large: in total, we hand-collected 171,740 observations of rents, sales prices, and property-level taxes and costs, covering a representative sample of about 40,000 different properties. We have enough repeated price observations to estimate capital gains using repeat-sales regressions to control for changes in housing quality.

The key innovation of our database is that it includes rents and prices for a large representative set of properties. For Paris, we can link property-level sales prices to rent prices retrieved from rental contracts or inheritance records, registered in the years before or after the sale. For Amsterdam, we observe property-level rents and prices concurrently, and in part of the data even repeatedly. In total, we have 63,575 observations of property-level gross yields.

Most existing literature uses implied yields from other series, and studies that do measure actual yields rely on small samples from a limited set of investors (Bracke 2015; Chambers, Spaenjers, and Steiner 2021).²

We also have property-level information on taxes and costs, but these data only cover a subset of properties, so that we cannot study property-level costs with the same level of detail as Chambers, Spaenjers, and Steiner (2021). To construct annual cost series, and convert our gross yields to net yields, we combine our data with city-level data on taxes, costs, and vacancies, similar to Eisfeldt and Demers (2018).

We find total net geometric returns to rental housing of 6.3% for Paris and 8.0% for Amsterdam, with index-level standard deviations of 8.6% and 10.3%, respectively. In real terms, geometric average returns amount to 4.0% per annum in Paris and 4.8% in Amsterdam. In Jordà et al. (2019), who use data from Paris and Amsterdam for 90% of the period in which our samples overlap, French and Dutch returns are, respectively, 1.4% and 2.3% higher than our estimates and are also less volatile. This results in their Sharpe ratio estimates being approximately 40% higher than ours. The long-term real return to housing can entirely be attributed to the rental yield, with real capital gains around zero. We find that our series of net rental yields are uncorrelated to the implied yield series in Jordà et al. (2019), showing that it is very difficult to accurately estimate long-term yields from secondary data. Correlations between our total return series and those of Jordà et al. (2019) are only 0.3–0.4.

Our findings imply that the long-run risk-return assessment of rental housing investments are less rosy than previously assumed. Evidence for the housing risk premium puzzle is much weaker than suggested by Jordà et al. (2019), especially when property transfer taxes are taken into account, and the much lower Sharpe ratios we find for Paris and Amsterdam close most of the gap with equities that Jordà et al. (2019) document.

We find that our real returns for Paris rank lowest among all countries for which Jordà et al. (2019) provide total returns in the comparable period, while only three of their countries have lower reported returns than our findings for Amsterdam, with real returns on average 1.1% per year higher in other countries. Given the fact that Paris and Amsterdam have both experienced substantial economic and population growth over our sample periods, one would not expect such comparatively low housing returns for these two cities. This suggests that the returns for other countries reported in Jordà et al. (2019) might be prone to the same errors. This conclusion is supported by the findings of Chambers, Spaenjers, and Steiner (2021) for the United Kingdom that also point at much lower housing returns than reported by Jordà et al. (2019).

These results are all at the portfolio level, but, when we account for idiosyncratic risk, the Sharpe ratios are reduced even further. Our analysis of

² Bracke (2015) studies 1,922 yields from London for 2006–2012, and Chambers, Spaenjers, and Steiner (2021) use 1,359 distinct transactions to measure yields, of which 549 are for residential real estate.

idiosyncratic risk is based on a subset of Amsterdam data for which we observe yields and sales prices repeatedly. These data allow us to study idiosyncratic risk in residential real estate investment, highlighting the role of yields as a risk component to housing investments. We document significant persistence in property-level yields over time, even up to 20-year holding periods, and even after accounting for differences in yields across neighborhoods. This implies that properties purchased at above-market yields will continue to earn above-market yields many years after the purchase.

In aggregate, we find idiosyncratic risk to be substantial, contributing to about half of gross total return risk for holding periods of 15 years and above. However, the composition and importance of idiosyncratic risks change over time. In the short term, nearly all total return risk is idiosyncratic and comes from capital gains volatility. Because yields are persistent, yield covariance becomes an increasingly important component of total risk for longer holding periods. Idiosyncratic capital gains risk becomes less important relative to systematic risk, given the flat structure of idiosyncratic capital gains risk across holding periods (Giacoletti 2021). Although we cannot account for property-level costs in our analysis of idiosyncratic risk, our findings suggest that ignoring yield risk substantially underestimates the role of idiosyncratic risk for housing investments, especially over long holding periods. For a 10-year holding period, the total volatility of property-level gross housing returns is approximately 50%, whereas the volatility of 10-year gross portfolio returns is around 32%. For that holding period, including property-level risk thus implies a reduction in Sharpe ratios of about 35%.

1. Data and Historical Context

We employ two main archival data sources to construct indexes of actual rental yields and house prices for Paris and Amsterdam, which we complement with data from other sources. Table 1 presents a brief overview of the data sources and the number of observations employed in this paper. Importantly, all our data contain observations on the level of rents and sales prices of an entire property. In both Paris and Amsterdam, properties typically contain several housing units. We provide an overview of our data collection approach for these two cities in the remainder of this section. Internet Appendix A offers a more detailed description.

1.1 Paris

We extract the Paris housing data from a land register: the *Sommier foncier*. It covers the period from 1809 until 1943 and is part of the wider French administration responsible for collecting taxes on legal acts, the *Enregistrement*. The *Sommier foncier* provides information on all property transfers and corresponding prices in Paris. Most transaction prices in Paris are based on regular sales, but about one-third of transaction prices originate from properties

Table 1
Sources and sample sizes property-level data

Data type	Period	# obs.	Source
<i>Paris</i>			
Sale prices	1806–1943	38,168	Sommier Foncier
Rent prices	1806–1943	44,379	Sommier Foncier
Matched yields	1809–1943	27,722	Sommier Foncier
Taxes	1809–1926	4,474	Sommier Foncier, Tax Registers Sainte-Avoie
Asking yields	1872–1940	10,052	Le Figaro
Realized yields	1883–1939	1,060	Cote des Terrains et Immeubles, Le Temps
<i>Amsterdam</i>			
Sale prices	1840–1979	35,519	Brouwer & Zn.; Eichholtz 1997; Verwey 1943
Rent prices	1900–1979	25,834	Brouwer & Zn.
Matched yields	1900–1979	24,741	Brouwer & Zn.
Taxes	1917–1979	9,798	Brouwer & Zn.
Costs	1889–1967	2,454	Burgerweeshuis, Doopsgezinde Gemeente

Paris transaction prices are based on auction prices (37%) and regular sales (63%), while Amsterdam data are auction prices (48%), regular sales (8%), and appraisals (44%). Amsterdam rental prices are based on lease contracts (88%) and appraised rental values (12%).

sold in public auctions. Most properties in Paris were investment properties, which frequently sold in auctions. A small subset of these auction sales corresponds to foreclosure sales.

We also obtain data on rents from property donations during lifetime or inheritances after death (*successions*). By law, the taxable value on these transfers was equal to twenty times the sum of all leases on the property.³ The first two series of the *Sommier foncier*, covering the period until 1880, also contain information about rental contracts and corresponding prices on these properties.

The combination of rental prices and house prices for the same properties allows one to compute property-level gross yields, for a period covering more than a century, and for the entire city. To the best of our knowledge, this has been hitherto impossible, even with modern data.

We collect data from a random sample of 327 registers, containing properties throughout the entire city. In total, we digitize data for approximately 20,000 different residential properties. For each property, we list the street address, the type of legal act, the registered price or value, and the date of registration and transfer. Incomplete observations are removed, and the reduced sample covers 82,547 registrations: regular property sales, auction sales, leases for entire properties, inheritances, and donations.

After 1870, our rent data are almost entirely based on observations of declarations of successions and donations, and one might worry that these are not accurately measuring rental prices. This concern might be particularly severe after 1918 when the formal legal link between the declared values and rental prices disappears. To construct an alternative series of yields, we, therefore, collect 10,052 announcements of property sales from *Le Figaro*, a

³ Internet Appendix B explains how we recovered the actual rental prices from these observations.

French newspaper, spanning the period from 1872 to 1940. One limitation of this database is that it is based on asking prices and self-reported rents, which might deviate from actual rental prices and sales prices. To estimate the difference between minimum bids and realized prices, we collect 1,060 observations from data in two real estate news bulletins, the *Cote des Terrains et Immeubles* and *Le Temps Immobilier*.

To obtain estimates of taxes, we collect data on paid property taxes for 2,094 observations in the first register of the *Sommier*. For each of these observations, we also know the sale or rental price such that we can compute a property tax rate. To obtain tax rates after 1860, we collect 1,704 property-level tax observations for a sample of streets in *Sainte-Avoye*, a neighborhood in Paris.⁴ We match these to our data from the *Sommier*. It is not necessary to diversify this tax sample since the law prescribed the equivalence of property tax rates across the city.

1.2 Amsterdam

The city of Amsterdam had and still has a unique history of selling property for investment purposes in public auctions. Such auctions have been organized since the 1600s and still take place today. The format of these auctions has changed very little over time.

The market for auctioned property in Amsterdam was large. For the period between 1900 and 1942, we gather statistics on the number of properties put up for sale from the yearbooks of the Amsterdam statistical office (*Gemeente Amsterdam* 2018), which indicate that on average about 1,000 properties per year were put up for auction. This was about 2% of the housing stock. Most of these properties were auctioned voluntarily (*vrijwillige verkoop*), but some properties were sold after foreclosure or bankruptcy (*executoriale verkoop*). Selling investment property in auctions was the norm, and foreclosed properties were sold in the same auctions as regular properties.

Before the advent of modern house price indexes, auctions gave market participants important information about market prices and yields. Because most properties were purchased for investment purposes, information on rents and taxes was presented for nearly every property for sale. Properties were typically sold with tenants at current rental prices. If a property was not rented out, the auctions typically listed the assessed rental value of the property. The auctions were public, so individuals could record and register this information. Our data regarding these auctions stem from the archives of the *Firma Jan Brouwer & Zn.*, who developed a unique card system to store information on sales prices, rents, and taxes of properties sold in these auctions. This system covers the period from 1900 to 1979 and contains information on 19,786 properties.

⁴ Archives de Paris, D13P2/17, 67 to 69.

Beyond registering prices and rents for auctioned properties, the realtor also registered information on the appraisal value of these properties. These appraisals were requested by banks and other mortgage providers. The appraisers assessed both the market value of the property and the rental value because property developers and investors used rental cash flows to pay mortgage interest and amortization (Smid 2019). If the property was vacant or newly constructed, the appraisers estimated the rental value of the property rather than using actual rent contracts. In a small number of cases, the realtor also listed data on regular sales prices.

Excluding observations from outside of Amsterdam (4% of the data), this data collection results in 25,834 observations of rents (88%) or appraised rental values (12%), 30,528 transaction prices (48%) or appraisals (52%), and 9,798 observations of taxes. In 24,741 cases, we have both a rental price and a price observation for the same home in the same year. In 8,579 cases, we can also adjust this yield for taxes. To complete our database of prices, we augment it with 2,480 repeated transaction prices from Verwey (1943) for property auctions between 1840 and 1940, and 2,826 transaction prices from the Herengracht index database of Eichholtz (1997), covering the 1840–1972 period. Removing duplicate observations across databases, the total number of prices is 35,519, and 93% of these price observations concern the 1900–1979 period.

To provide estimates of nontax costs, we compile data on actual costs from the archives of two institutional investors: the Amsterdam Orphanage (*Burgerweeshuis*) and the Doopsgezinde Gemeente, an Amsterdam church. For social institutions like the *Burgerweeshuis* and the *Doopsgezinde Gemeente*, commercial property investments provided the largest part of their funding. Eichholtz, Korevaar, and Lindenthal (2019) use information on rental contracts from both institutions to construct multiple series of market rent prices, and we refer to their paper for information on the investment activity of these investors. Chambers, Spaenjers, and Steiner (2021) use similar data from Oxbridge colleges to obtain estimates of costs.

From its archives, we collect property-level information on rental income and expenses for 100 rental units, covering the period from 1937 to 1969.⁵ For the *Doopsgezinde Gemeente*, we obtain data on 30 different properties spanning the period from 1889 to 1924. Ledgers are incomplete before and after these periods, but these properties likely stayed in their hands for decades or even centuries. The mentioned costs include expenses on maintenance and renovation, taxes, insurance, and management costs (only at property level), lost rents due to rent arrears and vacancies, and water use (if not paid by the

⁵ Since the 17th century, the *Burgerweeshuis* has been among the largest institutional investors in the Amsterdam residential real estate market. Although the *Burgerweeshuis* reduced its property portfolio over time, it still owned 60 properties until the mid-20th century, containing over 100 rental units. Most of these properties already had been acquired in the 16th and 17th centuries.

tenant). In short, this database provides all asset-level costs. In total, this resulted in 2,454 property-level observations of rental prices and corresponding costs.

2. Components of Total Housing Returns

In this section, we estimate the components of the total returns of residential real estate in Paris and Amsterdam. The return to rental housing investments for a property (or a portfolio of properties) consists of both capital gains and net rental yields (Equation (1)).

$$Return_{i,t} = \frac{P_{i,t} - P_{i,t-1}}{P_{i,t-1}} + \frac{R_{i,t}(1 - c_{i,t} - \tau_{i,t})}{P_{i,t-1}}. \quad (1)$$

$P_{i,t}$ and $R_{i,t}$ equal the market price and rent of the same property i at time t . To estimate our total returns as precisely as possible, and to establish how measurement error could contribute to wrongly specified returns, we split this equation into three parts. First, we estimate capital gains. The main challenge here is to adequately control for housing quality, as well as to have sufficiently large and representative samples of housing transactions since we only observe market prices for properties when they transact. Second, we look at gross rental yields: the current or estimated rent divided by the sales price.⁶ Third, we study the implications of costs and taxes on yields, with a particular focus on property-level taxes (τ). We proxy for nontax costs (c) using institutional cost data for Amsterdam and findings from other studies and employ time series of vacancy rates to assess vacancy costs.

2.1 Capital appreciation

The literature has employed a wide set of methods to estimate house price indexes, some aiming to control for changes in quality of the underlying housing stock, and some not. Of the former, the two most commonly used are the repeat-sales method (Bailey, Muth, and Nourse 1963) and the hedonic method (Rosen 1974). In a standard framework, the log price of a transaction (p_i) at time t can be written as the sum of a “quality” component (α) and a time-varying market value component (β) plus a transaction error (ε).

$$p_{i,t} = \alpha_i + \beta_t + \varepsilon_{i,t} \quad (2)$$

Crucial to both methods is that they attempt to separate improvements in the quality of homes from increases in market prices. Because the quality of the housing stock has increased throughout the 20th century, inadequate quality control will result in indexes with an upward bias (Eichholtz, Korevaar, and Lindenthal 2019). In the repeat-sales method, which we employ for both cities

⁶ Note that, formally, rental yields slightly differ from the rental returns defined in Equation (1). The latter expresses the rent price relative to the market price in the previous period. Our rent observations specify the annual rental price at the time of the transaction: we thus assume this is equal to the rental price for the upcoming year.

in this paper, this is accomplished by focusing on repeated transactions of the same properties. Because we do not have observations on actual housing quality for the properties in our sample, using the hedonic alternative is not feasible.

For both Paris and Amsterdam, we estimate a standard repeat-sales index, controlling for the type of sale observed in the data. For the Parisian index, we use data from 1806 to 1943 to estimate the index but only report on its development from 1809 to 1943 when the number of repeat-sales is large. For Amsterdam, we include price observations for the entire period from 1840 to 1979 but only report on the index development from 1900 to 1979, the period for which we have yield data besides transaction prices.

To reduce the sensitivity of our index to extreme outliers, which may signal unobserved changes in quality or cases where only a part of a property was sold (but not indicated), we exclude price pairs that have a log price difference exceeding 1.95 or below -1.95 after correcting for log changes in the rent price index from Eichholtz, Korevaar, and Lindenthal (2019) ($> 600\%$ or $< -86\%$). For Paris, this removes 1,456 price pairs, leaving us with 17,770 price pairs to estimate the house price index. For Amsterdam, we use 15,125 transaction pairs to estimate the index, excluding 161 outliers.

Table 2, panel A, provides summary statistics of the new house price indexes. We find comparable figures for both cities. For Paris, we find a geometric average capital gain of 2.4% (arithmetic: 2.8%), with a standard deviation of 8.7%. Adjusted for inflation, the capital gain is 0.3% per year.⁷ For Amsterdam, we find a geometric average annual log capital gain of 2.6% (arithmetic: 3.1%) and a standard deviation of 10.3%. Adjusting for inflation, the real log capital gain averages -0.6% per year.

2.2 Yields

2.2.1 Gross yields. To estimate the gross annual rental yield for the two cities, we divide the summed rental prices of properties in the sample by their summed sales prices for each year. For Amsterdam, all yields are based on the rental price in the year of the sale. For Paris, we typically do not observe the value of the rent price (R) in the same year as the sales price (P), and we adjust for this using Equation (3). On average, we observe two rent price observations and two house price observations for each property. To be able to compute yields, we link each property sale to the nearest rent observation on the property before the sale (at time $t-x$) and after the sale (at time $t+z$), with x and z limited to 30 years. We adjust these rent observations for changes in market rent prices, which we estimated using a repeat-rent index (RPI) based on observations of rental contracts, successions, and donations, estimated using Equation (2). To compute the final yield, we apply linear interpolation, so that rent observations

⁷ We use the CPI indexes assembled in Eichholtz, Korevaar, and Lindenthal (2019) for both Paris and Amsterdam. These are city-specific CPI indexes pooled from various sources.

Table 2
Capital gains, rental yields, and total returns

Index	Period	Geometric		Arithmetic		Real geom.		Sharpe
		Mean	SD	Mean	SD	Mean	SD	
<i>A. Capital gains</i>								
Paris	1809–1943	2.4%	8.7%	2.8%	8.9%	0.3%	10.3%	
Amsterdam	1900–1979	2.6%	10.3%	3.1%	10.6%	−0.6%	10.3%	
<i>B. Gross yields</i>								
Paris	1809–1943	6.9%	1.1%	7.2%	1.2%			
Amsterdam	1900–1979	9.9%	2.0%	10.5%	2.3%			
<i>C. Net yields</i>								
Paris	1809–1943	3.9%	0.7%	4.0%	0.7%			
Amsterdam	1900–1979	5.4%	1.2%	5.5%	1.2%			
<i>D. Total returns</i>								
Paris	1809–1943	6.3%	8.6%	6.8%	8.9%	4.0%	10.2%	0.25
Amsterdam	1900–1979	8.0%	10.3%	8.7%	10.6%	4.8%	10.3%	0.35

This table reports the headline estimates for capital gains, gross yields, net yields, and total returns for both Paris and Amsterdam in geometric, arithmetic, and real geometric terms. Sharpe ratios for total returns are computed using long-term bond rates for France and the Netherlands.

closest to the sales price get the most weight, and divide these by the sales price at time t :

$$Yield_{i,t} = \frac{z}{x+z} \times \frac{R_{i,t-x}}{P_{i,t}} \times \frac{RPI_t}{RPI_{t-x}} + \frac{x}{x+z} \times \frac{R_{i,t+z}}{P_{i,t}} \times \frac{RPI_t}{RPI_{t+z}}. \quad (3)$$

At the annual level, the estimates on portfolio yields can be sensitive to extreme observations in the data. Most importantly, very large properties or properties with extreme yields can distort the yields at the annual level. To combat this, we remove observations that have log yields deviating more than 1.39 from the median log yield in the sample (more than 300% larger or -75% smaller). Second, we remove observations with extremely high rent levels (> 800% of the median house rent). For Paris, there remain 24,827 gross yields in the sample, and for Amsterdam 25,058 yields. Again, this procedure removes a larger fraction of observations for Paris (10% of data, 2,895 yields) relative to Amsterdam (3% of data, 683 yields), due to the larger amount of noise in the Parisian data.

Table 2, panel B, shows a Parisian gross log portfolio yield for rental housing of 6.9%, with a standard deviation of only 1.1%. For Amsterdam, the average gross log portfolio yield equals 9.9% with a larger standard deviation of 2.0%. We plot the series in Figure 1. For Paris, the gross yield moves in a rather limited range, roughly between 5% and 10%, with relatively high yields in the Napoleonic Era, after the Siege of Paris in 1870, and during the Great Depression. The picture is more volatile for Amsterdam than for Paris. Amsterdam yields are quite stable until 1965, and then start increasing substantially in the late 1960s, with the gross yield peaking at more than 20%

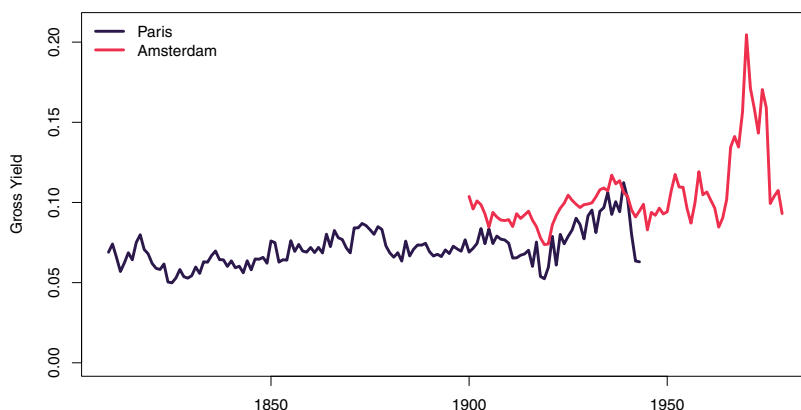


Figure 1
Gross housing yields, Paris and Amsterdam

This graph provides gross yields to rental housing for Paris (1809–1943) and Amsterdam (1900–1979). For Paris, the gross yield moves in a rather limited range, roughly between 5% and 10%. For Amsterdam, the average gross yield is 10%, with a peak of 20% in the 1970s. In the periods in which the Paris and Amsterdam series overlap, they correlate closely (correlation = .75).

in the 1970s. The evolution of yields in the overlapping period appears to be similar, with yields declining at the end of World War I, but increasing in the 1920s and 1930s. The correlation between the gross yields for Paris and Amsterdam in that period is 0.75.

2.2.2 Costs. For Paris, we compute the average annual tax rate directly for a subset of 2,094 transaction prices or rents between 1809 and 1854 for which we have information on the level of the annual property tax, expressing it as a fraction of total rent. Between 1855 and 1917, we match tax payments on the properties in Sainte-Avoys to the rental prices coming from successions, donations, and rental contracts in the *Sommier*. To match rental prices to tax payments, we use the same procedure that we employ to match rental prices to sales prices, by finding the nearest rent price on the property and adjusting these for changes in the market price. We compute the average tax rate based on these matched observations and interpolate it for years in which data are missing. Because tax rates were very stable in this period, this likely does not introduce major errors. After 1917, we use data from Duon (1946) who computed the fraction of property taxes borne by the property-owner annually, expressing it as a fraction of gross rent.

For Amsterdam, one-third of the rental yield observations include the required property-level taxes (9,798 observations). The most important of these were direct property taxes, street taxes, and a fee for the use of water. For properties with leaseholds, we also register land lease costs. From 1924 onward, sufficient observations are available to estimate the level of taxes as a fraction of

total rents.⁸ To do so, we compute the average tax rate in each year, controlling for differences in tax rates due to the presence of land leases. To estimate the level of tax yields before 1924, we estimate a repeated-tax index based on 635 annual observations of taxes for properties of the Doopsgezinde Gemeente between 1900 and 1924. We use the 1924 tax rate to splice these to the tax rate series from the yield database. For periods of missing data (1915–1916), we interpolate the tax rate.

In Amsterdam, annual taxes account on average for 14.8% of property rental value, with a volatility of 5.6%. In Paris, tax rates amount on average to 10.7% of the rental value with a volatility of 4.3%. Internet Appendix C offers a more detailed discussion of the different taxes and their evolution over time.

To measure vacancy rates, we make use of city-level statistics. For Amsterdam, we compute the annual vacancy rate by dividing the number of vacant housing units in the market by the total number of housing units in Amsterdam. For most years, this is reported in the statistical yearbooks of the Municipality of Amsterdam (Gemeente Amsterdam 2018) or in the national housing censuses. In case data are missing, we linearly interpolate vacancy rates (1900–1908, 1945–1946, 1948–1955, and 1957–1965). Vacancy rates in Amsterdam were low, with an average of 2%, a low of 0.3% right after World War II, and a high of 5.5% in 1935. For Paris, we use estimated vacancy rates based on the statistical yearbooks of Paris and data on the number of vacant accommodations given by Faure and Lévy-Vroelant (2007) and Duon (1946). Prior to 1869, no vacancy data are available, and we use the 1869 number for this period. We linearly interpolate data in periods with missing data, mostly in the 1870s and 1880s. Vacancy rates average 3.1% of rental value in Paris.

One limitation of our sample is that we do not have asset-level information on costs other than taxes, implying that we have to make estimates of these costs based on other sources of data. The approach we take in this paper is to make a well-informed estimate of the total amount of costs other than taxes and vacancies, using estimates from our Amsterdam cost data and findings from other literature.

First, we estimate the average fraction of nontax and nonvacancy costs in our institutional cost sample from Amsterdam. Excluding vacancy costs and taxes, but including maintenance costs, management costs, and insurance costs, the cost fraction on residential properties averages 31.8% in our sample. This value reduces to about 26.5% when we control for location, taking properties along the main canals as baseline, and the relative value of properties. Larger properties and properties in expensive locations tend to have lower cost fractions, while the properties in the sample are a bit less expensive than average. Internet Appendix D provides a more detailed discussion of our cost data and analysis.

⁸ The rental tax rate can be estimated more precisely than the tax rate as a fraction of property value.

In comparable studies, the fraction of nontax costs typically amounts to approximately 30%–35% of rental value. For the United States, Eisfeldt and Demers (2018) apply a fixed fraction of about 35% of the total rental value as costs, and add time-varying local property tax costs and vacancy costs to this estimate, as we do in this paper.⁹ Chambers, Spaenjers, and Steiner (2021) have full information about property-level costs and study a comparable time period as we do. They find actual costs for residential real estate of 32% of gross rental income, including taxes, but excluding vacancy costs. The cost fractions they identify fluctuate substantially over time but only apply to a very specific group of investors: Oxford and Cambridge colleges. For Paris, the tax authorities assumed a 25% maintenance cost fraction since the French Revolution, including management and insurance costs (Duon 1946). Jordà et al. (2019) also discuss the evolution of maintenance costs historically, and find that maintenance, management, and insurance costs constitute about 30% of gross rent without strong time trends.

Based on our findings in our cost analysis and other estimates in the literature, we apply a fixed cost fraction to our gross yields of 30% of the rental value for both Paris and Amsterdam, excluding costs for vacancies and taxes.

2.2.3 Net yields. The next step is to convert the gross rental yields reported in Table 2, panel B, to net rental yields, using our estimates of costs, taxes, and vacancy rates. These are reported in panel C of Table 2, and we observe a net yield for Paris of 3.9% for the full 1809–1943 sample period, and a net yield of 5.4% for Amsterdam for the 1900–1979 period.

Given some uncertainty surrounding the true level of maintenance and management costs, and their evolution over time, these numbers might be different if cost fractions differed from our estimates. For example, if the true cost fraction excluding taxes and vacancies was 25% or 35% of the rental value, instead of 30%, net yields would increase or, respectively, decrease by 0.35% in Paris and 0.5% in Amsterdam.

3. Returns to Rental Housing: Results and Discussion

This section first provides aggregate statistics on total returns for Paris and Amsterdam, both in nominal and in real terms, and will then go into a discussion of our results in the light of existing research on this topic.

3.1 Total return to residential real estate

We first deflate nominal returns into real returns, for which we use the CPI indexes assembled in Eichholtz, Korevaar, and Lindenthal (2019) for both

⁹ They assume nontax and nonvacancy costs are the sum of 2.13% of property value and 6.63% of rental value, which equates to about 35% of the property value based on the gross yields in their sample.

Paris and Amsterdam. These are city-specific CPI indexes pooled from various sources. We combine the housing capital value index reported in Section 2.1 with the net yield development in Section 2.2.3 by directly applying Equation (1) to get the total net return to rental housing in Paris and Amsterdam. Table 2, panel D, provides statistics for these series.

First, the geometric average net return to rental housing is 6.3% for Paris (arithmetic: 6.8%) and 8.0% for Amsterdam (arithmetic: 8.7%). In real terms, geometric returns in both cities are more similar: 4.0% in Paris and 4.8% in Amsterdam. The lack of any real capital gains on housing for both Paris and Amsterdam also implies that the real total long-term returns on housing over time accumulate from rental cash flows rather than capital gains. The standard deviations of the nominal total return are 8.6% for Paris and 10.3% for Amsterdam, almost equalling the standard deviation of the capital return for these cities.

To assess risk premiums and Sharpe measures, we use series of long-term bonds returns that we can apply consistently over time since bill rates are not available for all time periods. For France, we use the bond yield on French 5% annuities before 1833, and the 3% annuity between 1833 and 1943. For the Netherlands, we take the long-term Dutch government bond yield from 1900 to 1979, also reported in Jordà et al. (2019). Relative to long-term government bonds, housing earned a risk premium of about 2.1% in Paris and 3.5% in Amsterdam. Combining this with the standard deviations of the geometric total returns gives a Sharpe ratio of 0.25 for Paris and 0.35 for Amsterdam.

Figure 2 reports the evolution of real cumulative returns over time for both Paris and Amsterdam. Real total returns dwarf capital gains but the disproportionate influence of capital gains on total return volatility is evident, especially for Paris after 1914. The consequences of World War I scarred the performance of real estate investments in Paris, as France experienced significant inflation and introduced nominal rent controls. This was much less the case in Amsterdam, as the Netherlands was neutral during World War I. In Amsterdam, real total returns accumulated steadily up to the 1930s, then went into a 20-year hiatus following the Great Depression and World War II, before picking up pace again in the 1950s.

3.2 Comparing our indexes to previous work

Having established our main estimates for housing returns in both Paris and Amsterdam, this remainder of this section aims to discuss these estimates in more detail. First, we want to address how our estimates compare to the results reported in Jordà et al. (2019) and in other related literature and discuss the likely role of measurement error in the differences between our estimates and previous series. This section also discusses how our improved estimates for Paris and Amsterdam change our view regarding the performance of housing as an asset class, particularly relative to equities. We will end the subsection

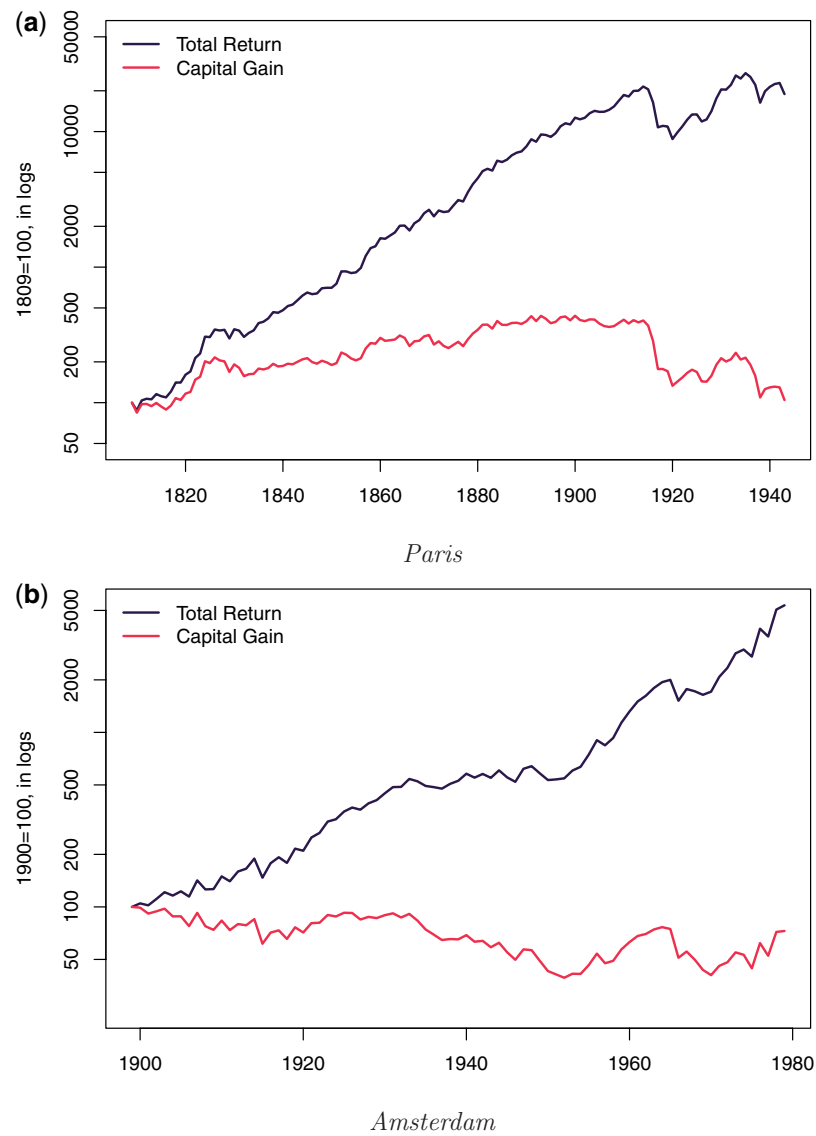


Figure 2
Real total returns and capital gains

These graphs depict inflation-adjusted cumulative total returns and capital gains for Paris and Amsterdam. Real capital gains are dwarfed by total returns, but the price volatility is very visible, especially for Paris after 1914. World War I scarred the real estate investment performance in Paris but did not affect Amsterdam much. In Amsterdam, real total returns steadily accumulated up to the 1930s, then went into a 20-year hiatus before picking up pace again in the 1950s.

with a short discussion of potential limitations and issues with our estimates, and provide the results of the robustness checks.

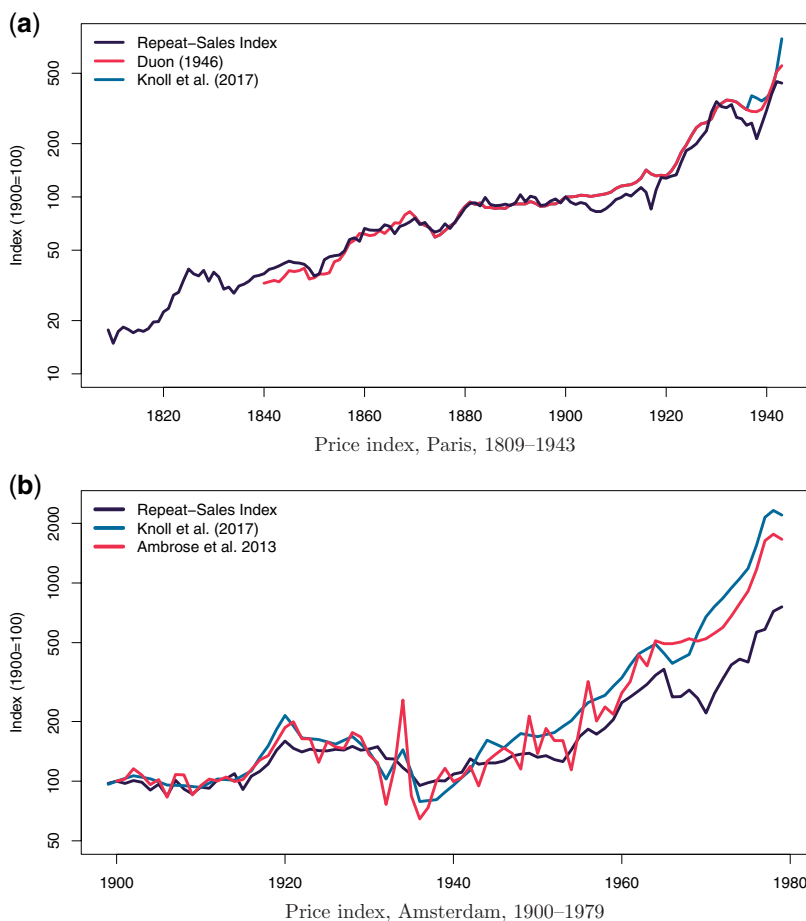


Figure 3
House price indexes

The plots compare our new repeat-sales indexes for Paris and Amsterdam with earlier repeat-sales indexes that rely on smaller samples, and sometimes apply smoothing techniques. For Paris between 1870 and 1935, Knoll, Schularick, and Steger (2017) use the index of Duon (1946) but switch to national data from 1935 onward, so that the index of Knoll, Schularick, and Steger (2017) diverges from Duon (1946). For Amsterdam, Knoll et al. (2017) switch to national data from 1970, while the index of Ambrose, Eichholtz, and Lindenthal (2013) uses national data from 1965.

3.2.1 Capital gains. We start our comparison of total housing returns by looking at changes in house prices for Paris and Amsterdam. Figures 3a and 3b depict our newly estimated repeat-sales indexes in nominal terms and compare the nominal indexes with other housing series, most specifically those used by Knoll, Schularick, and Steger (2017), on which the total return estimates in Jordà et al. (2019) are based. Table 3 provides numerical comparisons.

For Paris, Knoll, Schularick, and Steger (2017) used the index of Duon (1946) for the period between 1870 and 1935. This index is likely the world's oldest

Table 3
Comparing return estimates

Index	Period	Geometric		Arithmetic		Real geom.		Corr.
		Mean	SD	Mean	SD	Mean	SD	
<i>A. Capital gains</i>								
Paris	1840–1943	2.4%	8.6%	2.8%	8.9%	−0.6%	10.3%	
Paris (Duon 1946)	1840–1943	2.7%	5.8%	3.0%	6.0%	−0.2%	8.5%	0.39
Paris	1871–1943	2.4%	9.3%	2.9%	9.7%	−1.5%	11.3%	
Paris (Knoll, Schularick, and Steger 2017)	1871–1943	3.2%	7.6%	3.5%	8.6%	−0.7%	9.2%	0.31
Amsterdam	1900–1979	2.6%	10.3%	3.1%	10.6%	−0.6%	10.3%	
Amsterdam	1900–1979	3.5%	26.0%	7.0%	27.3%	0.4%	25.5%	0.31
(Ambrose, Eichholtz, and Lindenthal 2013)								
(Knoll, Schularick, and Steger 2017)	1900–1979	3.9%	10.9%	4.6%	11.2%	0.8%	9.5%	0.40
Amsterdam								
<i>B. Net yields</i>								
Paris	1871–1943	4.3%	0.6%	4.4%	0.6%			
Paris (Jordà et al. 2019)	1871–1943	4.9%	0.7%	5.0%	0.8%			0.11
Amsterdam	1900–1979	5.4%	1.2%	5.5%	1.2%			
Amsterdam (Jordà et al. 2019)	1900–1979	6.3%	2.1%	6.5%	2.3%			−0.08
<i>C. Total returns</i>								
Paris	1871–1943	6.7%	9.3%	7.3%	9.6%	2.8%	11.4%	
Paris (Jordà et al. 2019)	1871–1943	8.1%	7.8%	8.6%	8.8%	4.2%	9.6%	0.30
Amsterdam	1900–1979	8.0%	10.3%	8.7%	10.7%	4.8%	10.3%	
Amsterdam (Jordà et al. 2019)	1900–1979	10.2%	10.6%	11.1%	10.9%	7.1%	9.2%	0.39

This table compares our results on capital gains, net yields, and total returns to previous studies. The correlation coefficient is computed as the correlation in capital gains, yields, or log returns with our baseline series. The house price indexes from Knoll, Schularick, and Steger (2017) are used to compute capital gains and corresponding total returns in Jordà et al. (2019). The price index of Duon (1946) for Paris was smoothened using a moving average of unknown length. This explains the low volatility and correlation of this series and that of Jordà et al. (2019), who use Duon's index until 1935. Until 1965, the Ambrose, Eichholtz, and Lindenthal (2013) index for Amsterdam is based on a small number of transactions along the Herengracht, Amsterdam's most affluent canal, resulting in very high volatility and low correlation. Knoll, Schularick, and Steger (2017) use a smoothened version of this index.

repeat-sales index, although it is unclear how this index was constructed exactly. Duon traced all previous sales for the 4,389 homes sold in Paris between 1941 and 1944 and used this to estimate a house price index. Duon then smoothed his index using a moving average of unknown length (Duon 1943), to cope with the low number of transaction observations.

Given the overlap in methodology, this index unsurprisingly displays a very similar long-term development to ours, but because of smoothing, it is less volatile than our index. This is visible in the graph and in the numbers: the standard deviation of annual capital gains is 5.8%, relative to 8.6% for our index in the same period, as can be observed in panel A of Table 3. This table also shows a correlation of 0.39 between house price changes according to the Duon index and our Paris house price index.

Knoll, Schularick, and Steger (2017) use this index until 1935, and then splice it to a national repeat-sales house price index. However, house prices bottomed out in 1935, and because the index of Duon (1946) is smoothed, the splicing takes place at an overly high index level, resulting in an underestimation of the fall in house prices that took place in the 1930s. As a result, Knoll, Schularick, and Steger (2017) find much higher average house price growth in this period. So because the index of Duon (1946) is a smoothed index,

Knoll, Schularick, and Steger (2017) substantially underestimate volatility in the 1870–1935 period, and then overestimate price growth afterward due to incorrect splicing. The correlation between the Knoll, Schularick, and Steger (2017) index and our Paris index is only 0.31.

For Amsterdam, we compare the capital return part of our index to the indexes in Ambrose, Eichholtz, and Lindenthal (2013) and Knoll, Schularick, and Steger (2017). Both indexes are primarily based on the biannual Herengracht index of Eichholtz (1997). Their index employs repeated sales of 17th–18th century properties along Amsterdam's best-known canal, ensuring constant quality. Ambrose, Eichholtz, and Lindenthal (2013) estimated an annual version of the Herengracht index, while Knoll, Schularick, and Steger (2017) annualize the original index by applying the biannual observation only for the first of the 2 years. To interpolate the second year, which is now missing, they take simple averages of the previous and next observation.

These differences in index construction and sample result in substantial differences between our indexes and these alternative estimates. Without adjustment, the low number of observations in the annual Herengracht index of Ambrose, Eichholtz, and Lindenthal (2013) results in unrealistically high levels of volatility. The volatility in Knoll, Schularick, and Steger (2017) is more realistic and closer to ours, but this seems coincidental given the rather forceful smoothing technique they use. While smoothing can bring volatility down to levels that are seemingly more realistic, it completely changes the process that generates annual capital gains and will result in volatility estimates that are biased and inconsistent. By estimating annual indexes—both for Paris and Amsterdam—that are based on much larger sets of repeat sales, our new indexes strongly mitigate these estimation problems.

We find that capital gains in Amsterdam are substantially lower in our index relative to Ambrose, Eichholtz, and Lindenthal (2013) and Knoll, Schularick, and Steger (2017). The Herengracht database, on which their indexes are based, has a very low number of observations from the mid-1960s onward, resulting in very high but even more uncertain growth rates of house prices. Knoll, Schularick, and Steger (2017) switch to a national house price index from 1970 built on median prices, whereas Ambrose, Eichholtz, and Lindenthal (2013) already do so from 1965. The difference in capital gains is completely caused by this switch to an annual index that does not control for quality. Low index quality, as defined by the degree to which the index adjusts for the changing quality of the underlying housing stock, results in apparently higher house price increases, as Gatzlaff and Ling (1994) and Eichholtz, Korevaar, and Lindenthal (2019) show.

The issues discussed above result in relatively low correlations between the capital gains in our new house price index for Amsterdam and those in Ambrose, Eichholtz, and Lindenthal (2013) and Knoll, Schularick, and Steger (2017): we find levels of 0.31 and 0.40, respectively. Because the volatility of capital gains is much higher than that of yields (Table 2), these differences in index quality

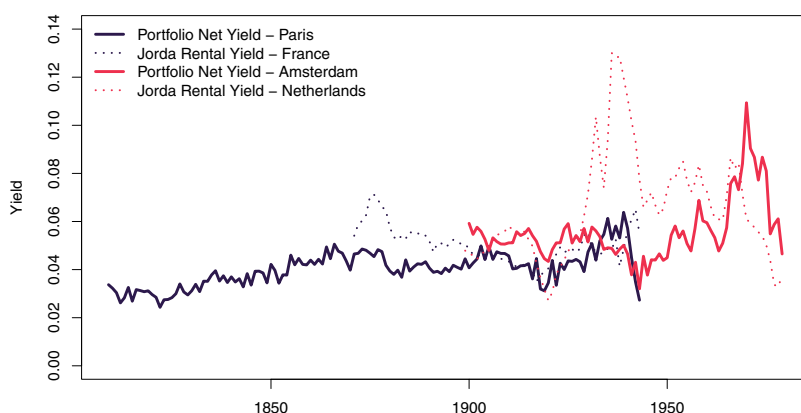


Figure 4
Net yields, Paris and Amsterdam

Net yield estimates for both cities are compared to the implied yield underlying the Jordà et al. (2019) total return calculations, which show a different evolution, in particular for Amsterdam.

also will be the main cause of the low correlation in total returns between our series and Jordà et al. (2019).

We do not think these issues are unique to these series for Paris and Amsterdam but apply to historical series of house prices and rent prices more generally. Given limitations in data availability, long-term series of house prices and rents still frequently build on relatively thin databases and often splice together indexes constructed with different methods, from different localities, and based on different housing quality segments. For example, for rent prices in the United Kingdom, Chambers, Spaenjers, and Steiner (2021) suggest the estimates of Jordà et al. (2019) diverge from those in their paper and in Eichholtz, Korevaar, and Lindenthal (2019) because of inappropriate index splicing and insufficient control for quality changes in the underlying housing stock. For house prices in the United States, Fishback and Kollmann (2014) provide an extensive comparison with the well-known index of Shiller (2005) for the years between 1920 and 1940, when the Shiller index relies on self-reported repeat-house values rather than actual transaction prices. They show that the resultant index substantially underestimates the magnitude of the boom and the bust in this period relative to several alternative measures they develop.

3.2.2 Yields. Figure 4 plots the development in net rental yields at the portfolio level for both Paris and Amsterdam relative to the yields in Jordà et al. (2019). Comparing our net yields to those of Jordà et al. (2019) for the periods when both samples overlap, we find that our annual net yields are about 1% lower. Correlations between our net yields and those of Jordà et al. (2019) are only 0.11 for Paris and -0.08 for Amsterdam.

The low correlations with Jordà et al. (2019) are partly caused by the fact that we adjust for time-varying taxes and vacancies, while they do not. However, when we look at correlations between the gross yields (not reported in the table) we observe that they are slightly higher but still low: 0.22 for Paris and 0.21 for Amsterdam.¹⁰

These low correlations show it is difficult if not impossible to accurately derive the evolution of rental yields over time from series of house prices and rents that relate to different sets of dwellings or use different methodologies. The French yield series in Jordà et al. (2019) are primarily based on quality-controlled series of Parisian prices and rents but originate from different sources and housing market segments, and use different methodologies. For their Dutch yield estimates, Jordà et al. (2019) combine quality-controlled series of Amsterdam house prices with national rent price series that do not control for quality. Brounen et al. (2014) use similar data sources and are therefore prone to the same error. These considerations also hold for papers that combine more recent house price and rent data, such as Eisfeldt and Demers (2018) and Giglio et al. (2018).

3.2.3 Total returns. Relative to Jordà et al. (2019), we estimate lower total returns: the difference in annual geometric returns is 1.4% for Paris, and 2.2% for Amsterdam. For both cities, about 60% of the total real return difference between our returns and those of Jordà et al. (2019) is caused by lower capital gains, and the remaining 40% can be attributed to lower yields. While the volatility estimates for Amsterdam are comparable to Jordà et al. (2019), we find higher estimates of volatility for Paris. For the periods where our samples overlap, we find substantially lower Sharpe ratios. For Paris, we find a Sharpe ratio of 0.31 instead of their 0.54 and 0.35 instead of 0.55 for Amsterdam (Sharpe ratios not reported in the table).¹¹

We find low correlation coefficients in annual log returns between our total return series and those of Jordà et al. (2019), even as both series attempt to track the same asset base. For Paris and Amsterdam, the correlations are 0.30 and 0.39, respectively.

In the previous subsections, we have already provided a number of reasons why the series used in Jordà et al. (2019) result in distorted estimates of capital gains and yields over the short term, and we will point at some limitations in our own series in the robustness section. Some of these distortions, such as the use of smoothing in the underlying capital gains series in Jordà et al. (2019), will reduce correlations in short-term return estimates, but vanish in importance

¹⁰ For Paris, we also find very low correlations with the Jordà et al. (2019) yields using alternative measures of yields, with a value of 0.02 when we use the gross yield series based on Le Figaro data.

¹¹ Using bill rates instead of bond yields whenever they are available, our Paris Sharpe ratio is 0.42 against 0.69 in Jordà et al. (2019), and, for Amsterdam, these numbers are 0.45 and 0.65, respectively.

Table 4
Longer horizons

	Length	New index		Jordà et al.		
Index	(Years)	Mean	SD	Mean	SD	Corr.
Paris, 1871–1943						
log yields	1	4.3%	0.6%	4.9%	0.7%	0.11
log yields	3	12.9%	1.6%	14.7%	2.1%	0.17
log yields	5	21.5%	2.4%	24.4%	3.3%	0.20
log yields	10	42.4%	4.0%	48.1%	5.7%	0.16
Nominal log returns	1	6.7%	9.3%	7.9%	7.5%	0.30
Nominal log returns	3	20.3%	17.4%	22.8%	14.5%	0.59
Nominal log returns	5	33.4%	22.5%	37.5%	17.5%	0.81
Nominal log returns	10	65.4%	28.2%	74.1%	22.5%	0.88
Real log returns	1	2.8%	11.4%	4.0%	9.4%	0.55
Real log returns	3	9.2%	21.8%	11.7%	22.2%	0.78
Real log returns	5	15.6%	29.6%	19.8%	31.9%	0.91
Real log returns	10	32.2%	41.1%	41.0%	44.7%	0.95
Amsterdam, 1900–1979						
log yields	1	5.4%	1.2%	6.3%	2.1%	−0.08
log yields	3	16.2%	3.4%	19.1%	6.1%	−0.12
log yields	5	27.0%	5.5%	32.2%	9.7%	−0.16
log yields	10	53.6%	9.7%	65.6%	17.4%	−0.22
Nominal log returns	1	8.0%	10.4%	10.2%	10.7%	0.39
Nominal log returns	3	23.4%	19.6%	31.1%	24.3%	0.65
Nominal log returns	5	38.8%	27.1%	51.1%	33.1%	0.75
Nominal log returns	10	75.3%	37.8%	101.8%	51.3%	0.78
Real log returns	1	4.8%	10.4%	7.1%	9.2%	0.31
Real log returns	3	14.3%	17.9%	21.8%	18.4%	0.46
Real log returns	5	23.3%	23.9%	35.8%	21.7%	0.52
Real log returns	10	44.9%	33.9%	71.4%	30.3%	0.49

This table reports average total geometric returns across investment horizons of 1, 3, 5, and 10 years, comparing estimates in Jordà et al. (2019) to our new estimates. We compute these by summing log returns over the respective horizons. Note that the sample periods used to compute annual and longer-term horizons do not fully overlap, because we need more prior return observations to compute returns over longer horizons.

at longer horizons. We thus might expect higher correlations when computing returns at longer horizons.

To better understand how total returns developed and correlated over medium-term horizons, Table 4 shows the level of average total log yields and returns for different horizons of up to 10 years, both in nominal and in real terms.

For both Paris and Amsterdam, we find that the low correlations between the yields in our series and those in Jordà et al. (2019) persist when we compute them based on summed log yields across increasing medium-term horizons. This is concerning for housing return estimates based on implied yields since in the long run most of the real return to housing originates from yields rather than capital gains.

Over medium-term horizons, the volatility of capital gains remains the dominant component that is driving total return volatility. Comparing the volatility of total yields in Table 4 to those of total returns shows that yield volatility remains minor relative to total return volatility at all horizons. This is particularly so for Paris.

We find stronger correlations in nominal and real returns at longer horizons, particularly for Paris. This is consistent with the underlying data. For 93% of the overlapping sample period, the French capital gains data in Jordà et al. (2019) is based on a representative but smaller set of Parisian repeat-sales. Given the limited importance of yields in driving total return volatility, the French total return estimates in their paper should correlate highly with our series over longer horizons, when the applied smoothing techniques only have a limited impact on total returns.

For Amsterdam, the increase in correlation for lengthening time horizons is much less pronounced, particularly in real terms. First, yield volatility in Amsterdam is much higher compared to Paris, and therefore also significantly contributes to total return estimates over longer horizons. This should decrease correlations with Jordà et al. (2019), since implied yield series do not correlate with actual yield series, also over longer horizons. Second, the capital gains series in Jordà et al. (2019) switch to a national house price index that does not adjust for quality from 1970 onward, which diverges substantially from the capital gains based on our quality-adjusted series for Amsterdam. Finally, most of the increase in correlation in nominal terms appears to be driven by high inflation rates that are reflected in both series, since the increase in correlation between 3- and 10-year horizons is negligible in real terms.

Because small errors in rent or price series can lead to persistent under- or overestimation of yields, we find that implied return and yield series have low correlations with more accurately constructed return series, both in the short and in the long terms. So only when the underlying series very precisely track the evolution of rents and market prices and cover the same location and housing market segments, can return estimates from implied returns be informative about the evolution of longer-term returns. Unfortunately, long-term series having such properties are rare.

Until now we have compared our return estimates to other estimates for Paris and Amsterdam. A second, more complicated question is how these returns compare to estimates for other regions. When we compare our estimates to the numbers Jordà et al. (2019) report for other countries, we find that our real return index for Paris ranks lowest among all countries for which Jordà et al. (2019) provide return numbers in the 1870–1943 period.¹² For all countries that Jordà et al. (2019) consistently report on in the 1900–1979 period, eight countries have higher returns and three countries have lower returns than Amsterdam, with real returns on average 1.1% per year higher in other countries.

Without a proper counterfactual, it is difficult to identify whether these differences are driven by measurement error or by actual differences in returns across regions. However, the comparatively low return estimates for Paris and Amsterdam, both capital cities that grew substantially over time—economically

¹² Their sample includes only four countries in 1870 but increases to 11 countries by 1900. Average returns in all these 11 countries are higher in the 1870–1943 period than in our estimates for Paris.

and population-wise—suggest that measurement errors might affect their return series for other countries, leading to an overestimation of the total return to rental housing investments for these countries as well. This conclusion is supported by the findings of Chambers, Spaenjers, and Steiner (2021) for the United Kingdom. They also find substantially lower rental housing returns than Jordà et al. (2019).

However, Chambers, Spaenjers, and Steiner (2021) look at a portfolio of urban and rural properties, whereas our own analysis and that of Jordà et al. (2019) are based on urban rental housing. In that regard, a more representative comparison might be to look at a consistent urban sample. Although precise return numbers are not available for other cities, implied discount rates of urban housing cash flows could be informative of long-term housing returns. Bracke, Pinchbeck, and Wyatt (2018) use leasehold data to measure housing discount rates in London for different horizons. They find discount rates of housing cash flows in the short term in the range of 5%–6%. For horizons of 10–75 years, discount rates hover around 4%–5%, similar to the real return and net yields we document in this study. For very long-horizons of over 100 years, Bracke, Pinchbeck, and Wyatt (2018) find discount rates of 2%–3%. This number is in line with Giglio, Maggiori, and Stroebe (2015), who estimate very long-run discount rates based on modern leasehold and freehold data from London and Singapore. Again, these findings are much lower than those reported in Jordà et al. (2019).

3.2.4 Comparison with equities. One important finding in Jordà et al. (2019) is that rental housing earns substantially higher risk-adjusted returns as compared to equities. Across all countries they study, they find geometric returns to housing to be on average 2% higher than returns on equities, but with only half the volatility. In a follow-up paper, Jordà et al. (2019) specifically point to a housing risk premium puzzle.

To make the comparison with equities for Paris rental housing investments, we can employ two high-quality historic series on French equity returns. For the period between 1809 and 1854, we use the index constructed by Arbulu (1998), and for the years from 1854 to 1943, we use the blue-chip stock market index created by Le Bris and Hautcoeur (2010). Both compute total returns weighted by market capitalization.¹³ Over the full sample period, the nominal log returns on French stock investments equal 6.5% per year, which is similar to the 6.3% we find on housing. Given the larger volatility of stock returns of 12.8%, Sharpe ratios on housing (0.25) are still higher than those for equities

¹³ The index of Le Bris and Hautcoeur (2010) is an improvement over that of Arbulu (1998), since the Arbulu (1998) index weights returns by the market capitalization of each industry, but averages returns across firms within an industry, which overestimates returns at the end of the sample period. Le Bris and Hautcoeur (2010) instead use weights at the firm level to compute total returns. Before 1854, the bias in Arbulu (1998) is likely negligible given the low number of stocks per industry. Between 1854 and 1890, returns on both indexes are very similar.

(0.17). We find similar differences when we restrict our sample to the 1870–1943 period, when Jordà et al. (2019) use the index of Le Bris and Hautcoeur (2010), with a Sharpe ratio for housing of 0.31 relative to 0.19 for equities.

For the Netherlands, Jordà et al. (2019) use the series of Eichholtz, Koedijk, and Otten (2000) to estimate stock returns, with an average nominal log return of 6.3% and a volatility of 18.8% between 1900–1979. Using bond yields, this results in a low Sharpe ratio of 0.10 relative to our estimates for Amsterdam rental housing of about 0.35. We should note that the Dutch equity returns and Sharpe ratios—based on the data of Eichholtz et al. (2000)—are low compared to those in other countries in this period (Jordà et al. 2019). Given data limitations, Eichholtz, Koedijk, and Otten (2000) estimate dividend yields in the first half of the sample rather than actually measuring them. Some evidence indicates that this approach underestimates equity returns. For the 1901–1928 period, Derks (1933) reports total returns to investments in companies with the largest market capitalizations across industries, including all dividends, approximately equally weighting sampled stocks, and including firms that eventually went bankrupt. The average log total return for these stocks in 1901–1928 is 6.4% relative to 4.3% in the data in Eichholtz, Koedijk, and Otten (2000). This is similar to the return on housing for this period.

Although our improved estimates close about two-thirds of the gap in Sharpe ratios between housing and equities that Jordà et al. (2019) report for Paris, and about half of the gap in Amsterdam, our findings still point to higher risk-adjusted returns for housing relative to equities.¹⁴ Beyond potential limitations in data, we also see more fundamental explanations for this finding.

First, there were substantial transfer taxes on Parisian housing during the period that we study, while such taxes were absent or negligible for stocks. The transfer tax rate in Paris varied between 4% and 15% during our sample period, averaging 6.8%.¹⁵

If we would compute yields on the basis of property prices plus transfer taxes (so assuming an infinite holding period), property returns would fall by about 0.25% per year, reducing the Sharpe ratio on Paris rental housing to 0.21. This is a lower bound on the impact of transaction taxes on returns. If we compute taxes based on the median holding period in our repeat-sales sample of 9.57 years, total returns would fall by 0.7% per year. For our Parisian sample, this would bring the Sharpe ratio on housing down to 0.17, equal to the Sharpe ratio on equities.

For Amsterdam, transaction taxes were lower, amounting to 2.5% before 1970 and 6% after that. As a result, computing yields on the basis of sales

¹⁴ Note that Jordà et al. (2019) use arithmetic returns to compute Sharpe ratios, whereas we report here on geometric average returns. Using arithmetic average returns reduces the relative gap in Sharpe ratios, given the higher volatility of equity returns relative to housing returns.

¹⁵ From the Revolution to 1816, the transfer tax was 4%, after which it rose to 5.5% and remained unchanged until 1905. In 1905, the tax was raised to 7%, then to 10% in 1920 before rising to 15% in 1926. It fell to 12% in 1929 before rising again to 13.5% around 1935.

prices plus transaction taxes reduces yields only by 0.15%. Assuming median holding periods similar to Paris, total returns would fall by 0.3% per year. This would reduce Sharpe ratios by a small amount to 0.32.

A second key reason for this gap is that diversification is more costly for real estate than for equities, and for nearly all investors it is impossible to own more than just a few properties. As we noted when estimating capital gains, estimated return volatilities increase once we base them on smaller samples, so reducing Sharpe ratios. We will analyze the role of idiosyncratic risk in the next section, but we will first discuss the limitations of our own index estimations, and do some analyses of their robustness.

3.3 Limitations

Although our indexes offer a substantial improvement over existing ones, both because of the much larger sample and because of the application of a consistent methodology, some limitations apply to our series of capital gains and yields. We briefly discuss these in this subsection. We provide a more extensive discussion and the results from our robustness checks in Internet Appendix E.

Regarding the measurement of capital gains, the repeat-sales methodology we employ alleviates most concerns regarding unobserved quality differences, but a few measurement issues remain. First, we do not observe changes in the quality of a given property over time. Our estimation method tracks the same parcels, and given the advances in housing quality over the last two centuries, quality changes are likely to be salient. For Paris, we are able to use the tax registers to assess the importance of this issue and find an unobserved quality improvement of about 0.4% per year, which implies that we may overestimate capital gains a bit. Second, we do not observe the universe of transactions in a city, but rely on sampling so our series possibly still overestimate the volatility in aggregate capital gains. Our robustness tests for Paris and Amsterdam regarding this issue suggest this leads to a slight overestimation of the capital gains volatility. A third possible concern relating to the capital component of the total return is that our Amsterdam index not only relies on transaction prices but also relies on appraisals. If these appraisals deviate from actual transaction prices in systematic ways, our results will be biased, most likely in the volatility of the capital component. However, when we exclude the observations stemming from appraisals, the resultant change in volatility can be attributed to the reduced observation numbers rather than the appraisal effect.

There also may be some measurement issues in our gross yield series. To assess the role of outliers, we compare the average gross yields to annual median yields and find that the median gross yield levels are similar to the average yields for Paris and slightly higher for Amsterdam. For Amsterdam, the use of appraisals may also affect the yield series, but when we compare the series with and without the appraisals, the yields we obtain are nearly identical.

For Paris, our gross yield estimates are based on rent observations that are typically not from the same year as the price observation to which we match it. We use a 30-year matching period in the main specification, but when we reduce this difference to 10 years, our yield estimates do not change significantly. Another concern for the Paris data is that most yields are based on self-reported values of rental leases in declarations of donations or inheritances, only verified ex post by tax registrars. We assess whether our estimates deviate from actual rents by comparing our Parisian yield series to gross yields based on the subset of rent data coming from actual rental contracts and from gross yields we obtain from advertisements in *Le Figaro*, a French newspaper. The resultant estimates are almost identical to our main series, although the yield series based on *Le Figaro* advertisements appears somewhat smoother than our succession-based estimates.

4. Idiosyncratic Risks

The indivisibility of assets, high transaction costs, and the capital intensity of real estate investments hamper the construction of well-diversified direct property portfolios.¹⁶ For markets in which investors cannot fully diversify, theoretical studies (for instance, Levy 1978; Merton 1987) and evidence (e.g., Fu 2009; Eiling et al. 2019) suggest a link between idiosyncratic risk and expected returns in the cross-section.

How relevant is idiosyncratic risk at the property level? Giacoletti (2021) shows that idiosyncratic risk accounts for the majority of housing risk in California and that the idiosyncratic variance is constant across holding periods. Sagi (2020) models the price process for U.S. commercial real estate and also finds a dominant role of idiosyncratic risk. Because of data limitations, these and other papers (e.g., Peng and Thibodeau 2017; Eisfeldt and Demers 2018) ignore the contribution of income to total return and risk and concentrate on the capital return only.¹⁷ The novelty of our study is to investigate asset-level systematic and idiosyncratic risk based on total gross returns for different investment horizons of up to 20 years and to highlight the role of yield risk in total return risk. Also, we broaden the geographic and temporal scope of this discourse by using European data and by covering large parts of the 19th and 20th centuries.

Ignoring costs, the total gross log return r for a portfolio of properties of any size i with the holding period n going from year $t=0$ to year $t=n$ is defined as follows, with $y_{i,t}$ denoting $\log(\text{Yield} + 1)$ at time t , and g_{i,t_n} the log capital

¹⁶ In 1832, half of Amsterdam's property investors owned one building only. About 90% of investors owned fewer than five properties (Fryske Akademy 2014).

¹⁷ For commercial real estate, Peng (2016) combines capital returns and net operating income (NOI) figures to arrive at asset-level total return estimates but does not analyze idiosyncratic risk.

gain from period 0 to period n .¹⁸

$$r_{i,t_n} = \sum_{j=0}^{n-1} y_{i,t=j} + g_{i,t_n}. \quad (4)$$

Correspondingly, the variance of any individual gross property return can be written as follows:

$$\begin{aligned} Var(r_{i,t_n}) = & \sum_{j=0}^{n-1} Var(y_{i,t=j}) + 2 \times \sum_{0 \leq j < k \leq n-1} Cov(y_{i,t=j}, y_{i,t=k}) + 2 \\ & \times \sum_{j=0}^{n-1} Cov(y_{i,t=j}, g_{i,t_n}) + Var(g_{i,t_n}). \end{aligned} \quad (5)$$

Equation (5) shows that the variance of a gross housing return is a function of the variance of the yield and the capital gain, as well as the covariance in yields over time, and the covariance of the yield with the capital gain. In the remainder of this section, we aim to assess each of these quantities both at the property level and across space, with a focus on the first three terms of Equation (5).

Because we do not have dwelling-level observations on costs, we only study gross returns. We both look at the variance of total property-level returns and the variance of property-level idiosyncratic returns. We obtain the latter by deflating yields and capital gains with their average market values. To address spatial variation in yields, we will use both the Amsterdam and Paris samples.

Most of this section will be based on a subset of properties from our Amsterdam database. For this subset, we observe transaction prices and yields repeatedly on the same properties, resulting in 5,582 pairs of yields. Because this information is only available for a subset of data, is difficult to estimate each of the quantities in Equation (5) precisely, even when pooling data from different time periods. We attempt to quantify the degree of noise in Internet Appendix F, and also perform robustness checks.

4.1 Dispersion of yields

Starting with the first term of Equation (5), we look at dispersion in yields across different properties. For Paris and Amsterdam, the standard deviations of all log gross property yields are 4.2% and 4.8%, respectively. Only a minor part of these variations in yields is due to changes in aggregate housing yields: the volatility of residual yields after controlling for changes in market yields is 4.0% for Paris and 4.1% for Amsterdam.

Still, citywide yield estimates might not capture all systematic variation in local yields and total returns. For instance, Giacoletti (2021) studies

¹⁸ Note that we assume that properties transact at the start of each year and that annual rents are paid directly afterward, so that yields are earned from $t=0$ until $t=n-1$.

idiosyncratic capital gains risk at the ZIP code level and finds pronounced differences across low- and high-income areas. While the historic cores of Paris and Amsterdam are very compact and at least one order of magnitude smaller than today's cities, yields and returns might differ even at more granular geographic levels.

Figure 5 displays the spatial dispersion of gross yields at different periods for neighborhoods in both cities.¹⁹ In both cities, we find higher yields in poor areas, in line with evidence for modern cities from smaller databases (Bracke 2015; Desmond and Wilmers 2019).

Investigating the degree of spatial structure in yields more formally, we estimate Moran's I statistics (Moran 1950) to tests for correlations in yields among nearby houses. For Amsterdam, gross yields in excess of the market *without* neighborhood controls are found to be strongly correlated.²⁰ The null hypothesis of yields being randomly distributed across Amsterdam can be rejected firmly (p -value: $< .001$). This implies that local factors introduce nonrandom deviations from the citywide trends for Amsterdam and that investors cannot expect, on average, the same level of yields across the city.

When calculating excess returns at the neighborhood level, the Moran's I test statistics cease to be significant (p -value: .29). This implies that the time-invariant neighborhood fixed effects capture most of the spatial heterogeneity in yields and more detailed demarcation of submarkets or more granular neighborhood-level indexes will not improve the empirical fit. For Paris, the Moran's I statistic is positive, but not statistically different from zero (p -value: .20), even before accounting for neighborhood differences.

When calculating systematic risk, we, therefore, assume persistent differences in yield levels but similar time trends across neighborhoods.

4.2 Covariance of yields and capital gains

We now present some stylized facts about the second and third terms of Equation (5), the covariance of yields over time, and the covariance of yields with capital gains. To do so, we use a subset of 5,852 transaction pairs from Amsterdam where we observe the yield at purchase, the yield at the time of sale, and the capital gain.

Aggregating across all holding periods, we find that gross yields at purchase have a low but statistically significant correlation of 0.13 with subsequent capital gains, while gross yields at sale have a negative correlation of -0.21 with realized capital gains. If we compute these correlations on the basis of residual yields, controlling both for yield differences across neighborhoods and

¹⁹ These differences are not driven by taxation. The distribution of net-of-tax yields in Amsterdam (see Internet Appendix Figure 21) is similar to that of gross yields. The lower number of observations for net-of-tax yields, however, does not allow for a breakdown into subperiods.

²⁰ We find Moran's I, which measures spatial autocorrelation, to be 0.010, with an expectation of -0.00006 and a variance of 0.0000016 .

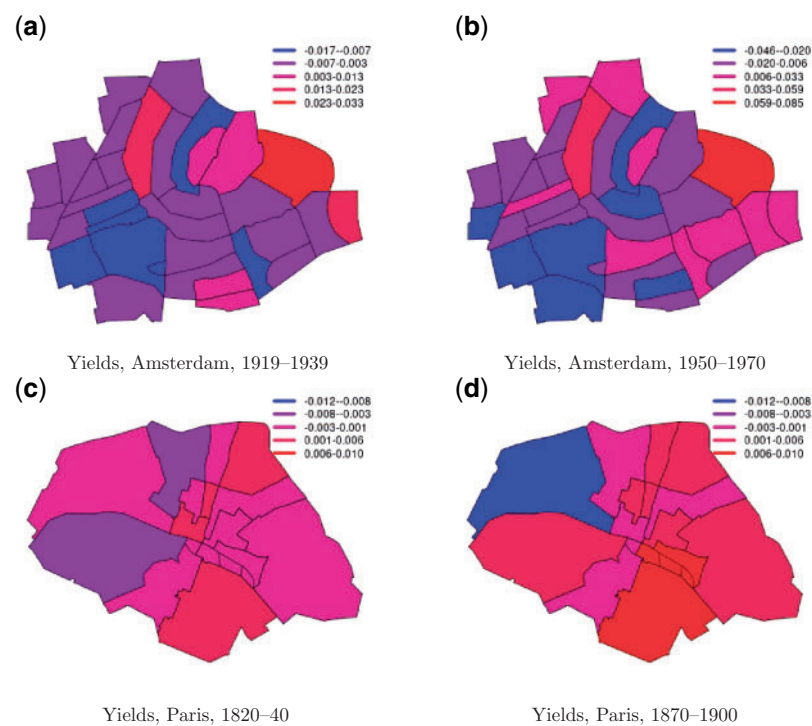


Figure 5
Median excess gross yields, per neighborhood

Excess yields are calculated as the difference between property-level gross yields and citywide median yields for a given year. Excess yields are not homogeneously distributed in space as clusters of low yields are found next to high-yield areas. For Amsterdam, the deviations from the citywide averages are large to begin with, getting more pronounced in time but tend to keep the relative ranking: the upmarket canal belt and museum areas, for instance, persistently deliver the lowest yields right next to the high-yield, working-class Jordaan. Differences in yields are only partially offset by higher subsequent capital gains. The correlation of asset-level yields at time of purchase and the subsequent capital gains are estimated to be only 0.22 for Amsterdam. In Paris, yields are more densely distributed in space but again mostly persistent in time: only the northwest experienced a distinct yield shift. Boundaries are based on Vasserot arrondissements²¹ for Paris and contemporary neighborhoods for Amsterdam.

market yields, we find stronger correlations of 0.20 and -0.46 , respectively. Thus, investors purchasing a high-yield or selling a low-yield property typically realize higher capital gains, regardless of the holding period.

This result is partially mechanical when transaction prices are distributed around fundamental asset values. However, the positive correlations between initial yields and subsequent capital gains are not strong enough to flatten out spatial differences in yields that persist in time: Figure 5 illustrates the persistence of yield differences at the neighborhood level.²²

²² At the market level, Ambrose, Eichholtz, and Lindenthal (2013) show that, while rents and prices are cointegrated, deviations from long-run rent-to-price ratios often take decades to correct. Price adjustments account for most

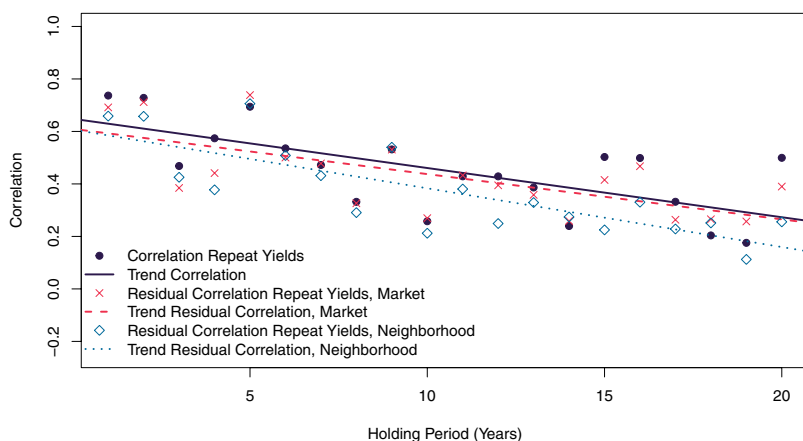


Figure 6

Correlation repeated yields

Based on 4,371 pairs of repeat sales in Amsterdam with holding periods of 20 years or less, the correlation of yields at purchase and sale is estimated for various holding periods (in years). In general, both yields are positively correlated. The effect fades away for longer holding periods, with an estimated trend of -0.018 per year ($t = -3.92$). The residual correlations are estimated after subtracting the median yield for the entire market at the time of each transaction from the respective yields. Correlations of neighborhood-level yield residuals fall more quickly over time than correlations for marketwide residuals (-0.022 per year, $t = -5.32$).

The question is to what extent this yield persistence also applies to individual properties, which brings us to the second term in Equation (5), the covariance in yields over time. If yields are very persistent over time, the yield at purchase will have a lasting consequence on returns over the entire holding period.

Figure 6 plots the correlation in yields across holding periods of up to 20 years, based on the same set of repeated yields. Because we aggregate data by holding period, we only have a limited number of observations for each point in the plots, starting at 587 repeated yields for 1-year holding periods and dropping to below 100 repeated yields for holding periods longer than 15 years. This implies that the correlations for individual holding periods are difficult to estimate precisely, particularly when considering that data points within a given holding period can come from the entire 1900–1979 period. We, therefore, focus on interpreting their trends across holding periods. Second, we also report the residual correlation after controlling for market and neighborhood yields, by subtracting the median yield per year and neighborhood across the sample at the time of each transaction from the property-level yields.

Importantly, the correlation across yields over time is substantial, and only gradually decays over time. Thus, properties with higher or lower gross yields will continue to earn above- or below-market gross yields for decades after purchase. Since the variation in median gross yields over time is small relative

of the eventual reversion to the mean, which is in line with the positive correlations between yields and capital gains found in this study.

to the total variation in yields, we do not find large differences in this pattern when controlling for market prices. However, given persistent differences in yields across neighborhoods, we find a stronger decrease in correlations across holding periods after controlling for neighborhood differences.

4.3 Total return risk and its components

We now proceed to estimate the contribution of each component of Equation (5) to the variance of total returns to housing investments for different holding periods. We can compute these numbers both based on total returns and based on residuals returns, the latter adjusted for average neighborhood yields and changes in market prices. To estimate these, we pool all yield pairs by holding period and estimate each of the individual variance terms.

We can estimate the variance of yields and capital gains for each holding period directly, but we need to make two assumptions to estimate the covariance terms since these contain terms that we do not observe in our data.

First, we assume that the covariance in yields on a property during a holding period is independent of the length of that holding period, so that we can use the observed covariance between the yield at purchase and the yield at sale to estimate covariances in yields during holding periods. For example, we assume that we can use the observed yield covariance for 1-year holding periods to estimate the unobserved yield covariance between the yield in year t and year $t+1$ for investments that have holding periods longer than 1 year (for which 1-year covariances are not observed). Formally, this implies that for all observations where $k-j=m-l$, $Cov(y_{i,t=j}, y_{i,t=k}) = Cov(y_{i,t=l}, y_{i,t=m})$.

Second, we assume that the covariance between yields during the holding period and the total capital gain is a weighted average of the observed covariance between the yield at purchase and the capital gain, and the observed covariance of the yield at sale and the capital gain. This implies that the third term of Equation (5) can be rewritten as $2 \times \sum_{j=0}^{n-1} Cov(y_{i,t=j}, g_{i,t_n}) = n \times Cov(y_{i,t=0}, g_{i,t_n}) + n \times Cov(y_{i,t=n}, g_{i,t_n})$.

The main drawback of this assumption is that it magnifies any noise in our estimated covariances between yields and capital gains by a factor n . We noted already that we have relatively few repeated-transactions per holding period, implying that we cannot estimate the covariance between yields and capital gains very precisely for each holding period. This noise increases for longer holding periods when the number of observations per holding period drops.

Figure 7 plots for each holding period between 1 and 20 years the estimated variance of the total return and its components. Figure 7a differentiates between idiosyncratic and systematic risk, and Figure 7b shows the composition of total property-level risk.

A clear pattern emerges. For investments with short holding periods, nearly all risk is idiosyncratic and unsystematic risk hardly seems to play a role. But the importance of idiosyncratic risk gradually reduces once we consider longer holding periods.

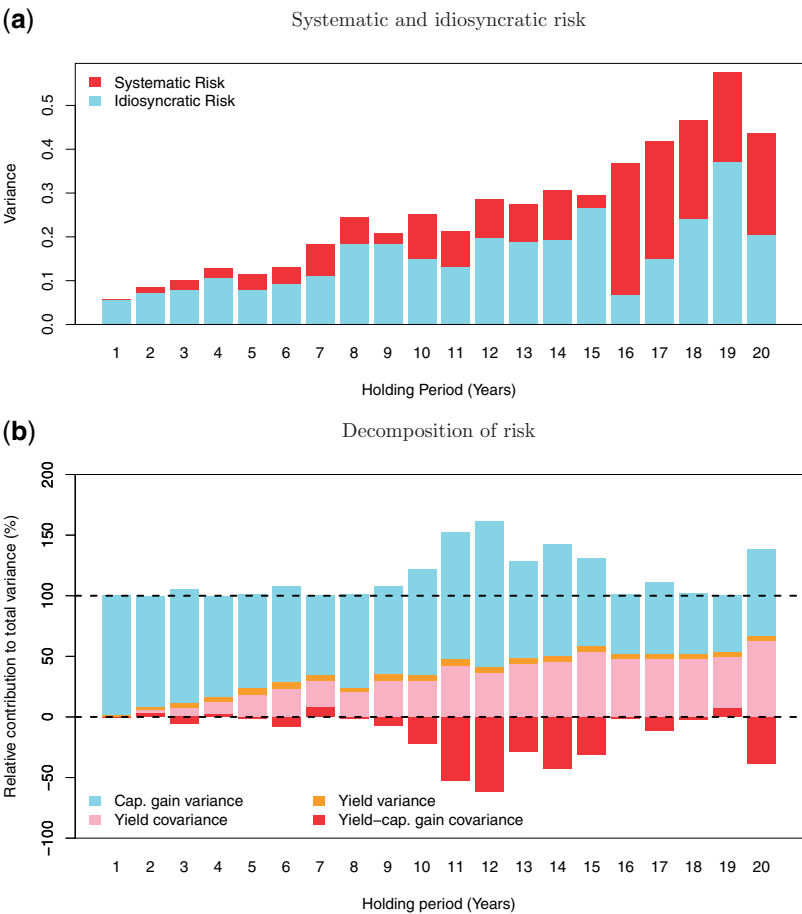


Figure 7
Decomposition of total risk

Panel A compares total systematic and idiosyncratic risk based on variance, and panel B depicts the components of total property-level risk. The residual variance is calculated in excess of citywide yield and price trends. Most variance in the short term can be attributed to capital gains risk because the variance of yields is small relative to the variance of capital gains. However, the importance of capital gains variance decays over time.

Importantly, the contribution of each component of total return volatility, depicted in figure 7b, changes over time. Unsurprisingly, most variance in the short term comes from capital gains risk, because the variance of yields is small relative to the variance of capital gains. However, the importance of capital gains variance decays over time, and yield covariance becomes an increasingly important component of total return risk.²³ Because the correlation

²³ This pattern also persists when plotting residual yields and adjusting for neighborhood yields and market price changes in Internet Appendix Figure 22.

between yields and capital gains is positive for the yield at purchase but negative for the yield at sale, we do not find large aggregate impacts of the covariance between yields and capital gains. Because the negative correlation outweighs the positive correlation, this covariance is negative on average but moves over time due to estimation noise.

4.4 Implications

Taken together, these results have two important implications for our understanding of the risk and return of individual properties. First, nearly all short-term investment risk at the property level is idiosyncratic, but the fraction of idiosyncratic risk decreases with the holding period, as changes in marketwide trends of capital gains and yields become more important in the long term.

This also implies that the impact of idiosyncratic risk on total return volatility changes across holding periods. For the average holding period of 10 years, we find property-level gross volatility of about 50%, including idiosyncratic risk, relative to systematic return volatility of 32%.²⁴ Thus, if we would compute Sharpe ratios on the basis of property-level gross return volatilities, they would drop by about 35% when including idiosyncratic risk. However, the impact of idiosyncratic risk on Sharpe ratios is even larger for short holding periods. For example, the 3-year total gross property-level return volatility in our sample is about 32%, whereas the systematic 3-year volatility in the sample equals about 15%.

If idiosyncratic risk followed a random walk, the idiosyncratic variance of total returns would scale exactly by the holding period (and volatility by the square root). Sagi (2020) and Giacoletti (2021) soundly reject this assumption for the idiosyncratic risk of capital gains, and our sample of capital gains confirms this finding.

However, due to the persistent positive covariance in yields, we document that this effect is gradually compensated over time by the increasing importance of yield risk in total idiosyncratic risk. Although idiosyncratic return risk increases more steeply over holding periods when also including yield risk, the idiosyncratic risk still does not scale fully with the holding period. For example, the 5-year idiosyncratic variance of total returns is about 0.1 (volatility of around 30%), while this increases to about 0.25 for 20-year periods (volatility around 50%).

Second, and most importantly, we document that persistence in property-level yields is a crucial risk component for investors, especially in the long run. This implies that for a long-term investor the initial yield is a much more important source of risk than for a short-term investor, who primarily bets on capital gains. Basically, rents are sticky and will not quickly revert back to their

²⁴ Note that the number of 32% slightly differs from the 37.8% reported in Table 4. This is because the figure of 37.8% is based on all data covering the entire 1900–1979 period, whereas the 32% is only based on the small set of repeated observations with a 10-year holding period.

economic fundamentals in case of deviations at purchase. A small part of this persistence can be explained by differences in yields across neighborhoods. Another factor that could be causing these yield differences are structural differences in nontax costs across properties. However, representative long-term data on these costs are difficult to obtain. With short-term data from Milwaukee, Wisconsin, Desmond and Wilmers (2019) show that gross yields are higher in poor areas and that less expensive rental units face higher costs, in line with our evidence for Amsterdam. However, they find that landlords still earn higher net rental returns in poor areas.

We again want to note that our estimates of the variance components and yield correlations contain noise when estimated for individual holding periods, because of the more limited samples we have available. In Internet Appendix F, we discuss this issue in more detail, quantify some of the uncertainty in our individual holding period estimates, and present robustness checks, considering a sample excluding appraisal data and splitting the sample in the pre- and post-WWII period. The trends we document across holding periods are robust to considering these alternative specifications. If many more observations were available, one could estimate capital gains indexes at the neighborhood level and fully separate marketwide risk, neighborhood risk, and idiosyncratic risk. Some of the variation in capital gains that we still account for as idiosyncratic risk would then likely be captured by a neighborhood risk factor.

5. Conclusion

This paper creates new indexes describing the net total returns of rental housing for extended periods and compares these to returns reported in recent work (i.e., Jordà et al. 2019; Chambers, Spaenjers, and Steiner (2021)). We create total return indexes for Paris and Amsterdam, for the periods 1809–1943 and 1900–1979, respectively. These indexes are based on previously unexplored archival data that we hand-collected and digitized for this study. Our unique contribution lies in the fact that we observe rental yields and values for the same properties, and that we have enough observations to use a repeated measures approach to reliably control for changes in asset quality.

The first main finding is that the geometric average net total return to rental housing is 6.3% in Paris and 8.0% in Amsterdam. These returns come with considerable volatility of 8.6% and 10.3%, respectively. We show that using actual rental yields and capital gains for the same set of properties is essential to obtain reliable estimates of housing return and risk. Relative to Jordà et al. (2019), who use secondary series, we find substantially lower risk-adjusted returns to housing for both cities and a low correlation with their total return series. This confirms the conclusion of Chambers, Spaenjers, and Steiner (2021).

We show that most of the real long-term total return to rental housing stems from the net yield, and that capital returns are small and even negative in real

terms for Amsterdam. This is in contrast to Eisfeldt and Demers (2018), who find a more important role for the capital return.

Besides these findings at the index level, our study also makes important contributions regarding property-level investment performance. Our findings regarding the geographic dispersion of rental housing performance and the importance of the holding period show that the yield at purchase is a key determinant of the holding period return at the individual asset level. We find that yields are persistent, even over holding periods as long as 20 years.

Regarding the composition of total return risk, we find that the idiosyncratic risk is the dominant part of total risk in the short term, but that the importance of market risk increases over the holding period. Moreover, we show that variation in the capital gain is the dominant factor in total asset risk only in the short term. For holding periods going up to 20 years, yield covariance becomes as important.

References

- Ambrose, B., P. Eichholtz, and T. Lindenthal. 2013. House prices and fundamentals: 355 years of evidence. *Journal of Money, Credit and Banking* 45:477–91.
- Arbulu, P. 1998. La bourse de Paris au XIX^{ème} siècle: l'exemple d'un marché émergent devenu efficient. *Revue d'économie financière* 49:213–49.
- 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* 58:933–42.
- Bracke, P. 2015. House prices and rents: Microevidence from a matched data set in Central London. *Real Estate Economics* 43:403–31.
- . 2021. How much do investors pay for houses? *Real Estate Economics* 49:41–73.
- Bracke, P., E. W. Pinchbeck, and J. Wyatt. 2018. The time value of housing: Historical evidence on discount rates. *Economic Journal* 128:1820–43.
- Brounen, D., P. Eichholtz, S. Straetmans, and M. A. J. Theebe. 2014. Inflation protection from homeownership: Long-run evidence, 1814–2008. *Real Estate Economics* 42:662–89.
- Campbell, S. D., M. A. Davis, J. Gallin, and R. F. Martin. 2009. What moves housing markets: A variance decomposition of the rent–price ratio. *Journal of Urban Economics* 66:90–102.
- Chambers, D., C. Spaenjers, and E. Steiner. 2021. The rate of return on real estate: Long-run micro-level evidence. *Review of Financial Studies* 34:3572–607.
- Derks, J. F. A. 1933. *De waarde van aandelen en obligaties als object van belegging*. Uitgeverij Lumax - Utrecht.
- Desmond, M. and N. Wilmers. 2019. Do the poor pay more for housing? Exploitation, profit, and risk in rental markets. *American Journal of Sociology* 124:1090–124.
- Dimson, E., P. R. Marsh, and M. Staunton. 2018. *Credit Suisse Global Investment Returns Yearbook 2018*. Credit Suisse Research Institute, Zürich, Switzerland.
- Duon, G. 1943. Évolution de la valeur vénale des immeubles parisiens. *Journal de la société française de statistique* 84:169–92.
- . 1946. Documents sur le problème du logement à Paris. Paris: Ministère de l'économie nationale. Service national des statistiques. Direction de la statistique générale.

- Eichholtz, P. 1997. A long run house price index: The Herengracht index, 1628-1973. *Real Estate Economics* 25:175–92.
- Eichholtz, P., K. Koedijk, and R. Otten. 2000. De eeuw van het aandeel. *Economisch-Statistische Berichten* 85:24–24.
- Eichholtz, P., M. Korevaar, and T. Lindenthal. 2019. 500 years of urban rents, housing quality and affordability. Working Paper, Erasmus University Rotterdam.
- Eiling, E., E. Giambona, R. Lopez Aliouchkin, and P. Tuijp. 2019. Homeowners' risk premia: Evidence from zip code housing returns. Working Paper, University of Amsterdam.
- Eisfeldt, A., and A. Demers. 2018. Total returns to single family rentals. Working Paper, University of California, Los Angeles.
- Faure, A. and C. Lévy-Vroelant. 2007. *Une chambre en ville. Hôtels meublés à Paris, 1860-1990*. Grane: Creaphis.
- Favilukis, J., S. C. Ludvigson, and S. Van Nieuwerburgh. 2017. The macroeconomic effects of housing wealth, housing finance, and limited risk sharing in general equilibrium. *Journal of Political Economy* 125:140–223.
- Fishback, P. and T. Kollmann. 2014. New multicity estimates of the changes in home values, 1920-1940. In *Housing and mortgage markets in historical perspective*, 203–44. Chicago: University of Chicago Press.
- Flavin, M., and T. Yamashita. 2002. Owner-occupied housing and the composition of the household portfolio. *American Economic Review* 92:345–62.
- Fryske Akademy. 2014. HISGIS-project Amsterdam.
- Fu, F. 2009. Idiosyncratic risk and the cross-section of expected stock returns. *Journal of Financial Economics* 91:24–37.
- Gatzlaff, D. H., and D. C. Ling. 1994. Measuring changes in local house prices: An empirical investigation of alternative methodologies. *Journal of Urban Economics* 35:221–44.
- Gemeente Amsterdam. 2018. Statistical yearbooks Amsterdam, 1895–2017.
- Giaconelli, M. 2021. Idiosyncratic risk in housing markets. *Review of Financial Studies* 34:3695–741.
- Giglio, S., M. Maggiori, and J. Stroebe. 2015. Very long-run discount rates. *Quarterly Journal of Economics* 130:1–53.
- Giglio, S., M. Maggiori, K. Rao, J. Stroebe, and A. Weber. 2018. Climate Change and Long-Run Discount Rates: Evidence from Real Estate. Chicago Booth Research Paper No. 17–22.
- Himmelberg, C., C. Mayer, and T. Sinai. 2005. Assessing high house prices: Bubbles, fundamentals and misperceptions. *Journal of Economic Perspectives* 19:67–92.
- Jordà, O., K. Knoll, D. Kuvshinov, M. Schularick, and A. M. Taylor. 2019. The rate of return on everything, 1870–2015. *Quarterly Journal of Economics* 134:1225–98.
- Jordà, O., M. Schularick, and A. M. Taylor. 2019. The total risk premium puzzle. Working Paper, Federal Reserve Bank of San Francisco.
- Knoll, K., M. Schularick, and T. Steger. 2017. No price like home: Global house prices, 1870–2012. *American Economic Review* 107:331–53.
- Le Bris, D., and P.-C. Hautcoeur. 2010. A challenge to triumphant optimists? A blue chips index for the Paris stock exchange, 1854-2007. *Financial History Review* 17:141–83.
- Levy, H. 1978. Equilibrium in an imperfect market: A constraint on the number of securities in the portfolio. *American Economic Review* 68:643–58.
- Merton, R. C. 1987. A simple model of capital market equilibrium with incomplete information. *Journal of Finance* 42:483–510.

- Mills, J., R. Molloy, and R. Zarutskie. 2019. Large-scale buy-to-rent investors in the single-family housing market: The emergence of a new asset class. *Real Estate Economics* 47:399–430.
- Moran, P. A. P. 1950. Notes on continuous stochastic phenomena. *Biometrika* 37:17–23.
- Peng, L. 2016. The risk and return of commercial real estate: A property level analysis. *Real Estate Economics* 44:555–83.
- Peng, L., and T. G. Thibodeau. 2017. Idiosyncratic risk of house prices: Evidence from 26 million home sales. *Real Estate Economics* 45:340–75.
- Piazzesi, M., M. Schneider, and S. Tuzel. 2007. Housing, consumption and asset pricing. *Journal of Financial Economics* 83:531–69.
- Rosen, S. 1974. Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy* 82:34–55.
- Sagi, J. S. 2020. Asset-level risk and return in real estate investments. *Review of Financial Studies* 34:3647–94.
- Shiller, R. J. 2005. *Irrational exuberance*. Princeton, NJ: Princeton University Press.
- Sinai, T., and N. S. Souleles. 2005. Owner-occupied housing as a hedge against rent risk. *Quarterly Journal of Economics* 120:763–89.
- Smid, R. 2019. *Speculanten en revolutiebouwers: Projectontwikkeling in Amsterdam 1877-1940*. Nijmegen: Uitgeverij van Tilt.
- Sommer, K., P. Sullivan, and R. Verbrugge. 2013. The equilibrium effect of fundamentals on house prices and rents. *Journal of Monetary Economics* 60:854–70.
- Verwey, G. 1943. Prijsverloop van 426 huizen te Amsterdam, welke voor 1880 reeds van eigenaar wisselden. *Economisch Historisch Jaarboek* 22:251–303.