Journal of Housing Economics 61 (2023) 101946

Contents lists available at ScienceDirect

Journal of Housing Economics

journal homepage: www.elsevier.com/locate/jhec

Revisiting metropolitan house price-income relationships

Elias Oikarinen [a,b,c,*], Steven C. Bourassa [d], Martin Hoesli [e,f], Janne Engblom [c]

[a] Oulu Business School, University of Oulu, Finland
[b] Department of Built Environment, Aalto University, Espoo, Finland
[c] School of Economics, University of Turku, Turku, Finland
[d] Runstad Department of Real Estate, University of Washington, Seattle, WA, United States of America
[e] Geneva School of Economics and Management and Swiss Finance Institute, University of Geneva, Geneva, Switzerland
[f] School of Business, University of Aberdeen, Aberdeen, United Kingdom

ARTICLE INFO

JEL codes:
C33
R10
R31

Keywords:
House prices
Personal income
Spatial equilibrium
Regional heterogeneity
Supply elasticity

ABSTRACT

We explore long-term patterns of the house price-income relationship across the 70 largest U.S. metropolitan areas. In line with a standard spatial equilibrium model, our empirical findings indicate that regional house price-income ratios are typically not stable, even over the long run. In contrast, panel regression models that relate house prices to aggregate personal income and allow for regional heterogeneity yield stationary long-term relationships in most areas. The house price-income relationship varies significantly across locations, underscoring the importance of using estimation techniques that allow for spatial heterogeneity. The substantial regional differences are closely related to the elasticity of housing supply.

1. Introduction

The relationship between house prices and income is important for urban and regional development and the overall economy for several reasons. The income elasticity of house prices affects housing affordability, and spatial differences in elasticity influence regional growth: the larger the long-term elasticity of prices with respect to income, the greater the counterforce posed by rising house prices on regional growth. Moreover, the relationship between house prices and income is expected to influence household consumption levels, consumption inequality, savings, and perceived financial well-being (Cooper, 2013; Berger et al., 2018; Etheridge, 2019; Atalay and Edwards, 2022). If regional house prices and incomes have stable long-term relationships, then deviations of prices from this relationship can be used to assess whether prices are under or over their long-term equilibrium levels. As house price dynamics influence credit and macroeconomic cycles, the house price-income relationship and its regional heterogeneity are of importance not only for regional policy makers, but also for central authorities aiming to stabilize macroeconomic cycles (Piazzesi and Schneider, 2016).

We explore the long-term relationship between house prices and income by considering the implications of a standard spatial equilibrium model and testing the predictions of this model empirically using data for the 70 largest U.S. metropolitan statistical areas (MSAs). Following numerous earlier studies of house price dynamics, our focus is on metropolitan-level relationships, as it is well established that house price dynamics vary across markets (Glaeser et al., 2008; Oikarinen et al., 2018; Ma, 2020, among others) and these dynamics affect the development paths of urban economies (Glaeser and Gyourko, 2018). For numerous agents, including households, construction companies, investors, and credit institutions, local developments are of great importance.

Further understanding of the house price-income relationship is desirable as the assumptions and implications of extant research on the relationship vary considerably across studies. Also, the house price-income ratio is an often-used metric to identify house price overvaluations in academic studies (e.g., Case and Shiller, 2003; Himmelberg et al., 2005; Vogiazas and Alexiou, 2017; Bourassa et al., 2019) and many institutions, such as central banks, use the ratio as a bubble indicator. In order to apply the price-income ratio as a reliable bubble indicator, stationarity of the ratio needs to be assumed. A stationary price-income ratio would imply that the long-term coefficient on income

* Corresponding author.
E-mail address: elias.oikarinen@oulu.fi (E. Oikarinen).

https://doi.org/10.1016/j.jhe.2023.101946
Received 17 October 2022; Received in revised form 27 March 2023; Accepted 21 May 2023
Available online 23 May 2023
1051-1377/© 2023 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

is one in a regression model explaining house prices. There are reasons to question the validity of the assumption, including variations in supply elasticities across locations (Saiz, 2010) that are expected to affect regional house price-income relations.

Previous empirical findings on the issue are mixed. The results of Gregoriou et al. (2014) are overall not supportive of stationary price-income ratios in U.K. regions. Oikarinen and Engblom (2016), in turn, observe cointegration, i.e., a stationary relationship, between apartment prices and income in Finnish cities, but the coefficient on income generally differs significantly from one. While these two studies conduct pairwise city- or region-specific tests, it is common in the literature to apply tests using panel data. In a panel study of U.S. metropolitan areas, Malpezzi (1999) cannot reject the hypothesis of non-stationary price-income ratios. In contrast, based on more advanced panel techniques that control for spatial dependence, Holly et al. (2010) conclude that the hypothesis of non-stationary ratios can be rejected in U.S. state level data.

There are some panel studies that address regional heterogeneity in the price-income relationship, allowing the coefficient on income to vary across regions and thereby also to differ from one. Harter-Dreiman (2004) finds a stationary vector between prices and aggregate personal income in U.S. metropolitan areas separately in panels of supply-constrained and unconstrained cities using conventional pooled panel estimators. Gallin (2006), using panel tests that allow for cross-correlations across cities, concludes that U.S. metropolitan area house prices are not cointegrated with per capita income and population. While Gallin (2006) is the first to control for spatial dependence in the context of unit root tests for the house price-income relationship, Holly et al. (2010) control for spatial dependence in their regression models in addition to unit root tests. They reject the hypothesis of a unit root in a regression between U.S. state-level house prices and per capita income, thereby accepting the alternative hypothesis of stationary relationship between the variables. However, rejection in such unit root tests does not necessarily imply that the price-income relationship is stationary in all or even most regions, as we will discuss in Section 4. Overall, the results reported in earlier studies are mixed not only regarding the stationarity of the house price-income ratio but more generally concerning the question of whether there is a stable long-term relationship between house prices and per capita or aggregate income at all (allowing the coefficient on income to differ from one and across locations).

To our knowledge, ours is the first study in which the implications of a standard spatial equilibrium framework for the house price-income relationship are considered. While Leung (2014) derives a dynamic stochastic general equilibrium model that informs the understanding of the house price-income relationship at the country level,[1] the spatial equilibrium model provides important insights regarding the city- or region-level relationships. The framework predicts that local house price-income ratios are generally not stable over the long run. Interdependence among urban economies plays a notable role in this result. In the empirical analysis, we apply panel econometric tools – including estimators and tests that have not been applied to housing price-income relation analyses before – that allow us to explore the implications of the spatial equilibrium model.

A further contribution is that, in addition to the price-income ratio, we study cointegration of MSA-level house prices and income using both per capita and aggregate personal income and test formally for a one-to-one relation between house prices and these income measures. We also permit income elasticity to be heterogeneous across MSAs, and control for spatial dependence in the unit root and cointegration tests. As additional contributions, we test formally for spatial heterogeneity in the price-income relations and investigate how variations in the price-income relationships across MSAs are related to supply elasticities and demand side factors. In addition to housing affordability and regional growth, our findings are relevant to the analysis of house price bubbles because a stable (i.e., stationary) long-term relationship between income and prices is needed to measure deviations from long-term equilibrium.

Consistent with the predictions of the spatial equilibrium model, our empirical results indicate that long-term stability of local house price-income ratios is the exception rather than the rule and that the income elasticity of house prices varies considerably across MSAs. As expected, MSA-specific price-income ratio trends and income elasticities are strongly correlated with supply elasticities in particular, but demand side factors also play a role. In line with the theoretical predictions, a panel regression model that allows for regional variation in the coefficient on aggregate personal income demonstrates a stationary long-term relationship for house prices in many more metropolitan areas than does the ratio of price to per capita income. We also show that the overall panel unit root test statistics may be misleading because they do not say anything about the proportion of regions, or cities, with stable relationships between prices and incomes – an issue not considered previously in the related literature.

Similar to several earlier studies, we are interested in whether MSA-level *income variables alone* form a long-term trend for local house prices, and hence do not include other possible fundamentals as control variables in the cointegration analysis. If we added other variables, we would no longer be studying the stationarity of pure price-income relationships. In other words, a stationary relation is not an indication of a stable price-income relationship if the stationarity necessitates the inclusion of some other (non-stationary) variable(s) in the model. Although our focus is on long-term relations rather than short-term dynamics, we provide a brief investigation of the potential causes of shorter-term deviations of house prices from their long-term trends. This analysis implies that cycles of metropolitan house prices around their long-term relationship with local aggregate incomes are associated with developments in local unemployment rates and in the national mortgage market. Also, a stationary relationship between house prices and aggregate income is more often observed for MSAs with relatively inelastic housing supply.

The next section of the paper considers the key predictions of a standard spatial equilibrium framework for the house price-income relationship. Section 3 discusses our data, including preliminary analysis of the variables' time series properties. Section 4 contains the empirical analysis. A final section concludes.

2. Spatial equilibrium and the house price-income relationship

Understanding the factors affecting the house price-income relationship and its development over time in a given city or region requires a theoretical framework that considers the whole system of cities or regions.[2] Partial equilibrium models (i.e., models that consider a single city in isolation, such as the closed city model that assumes no migration and takes local population and income as exogenous) miss important effects because housing costs, wages, city populations, and their growth rates are jointly determined and, therefore, population and income are endogenous to house prices (Glaeser and Gottlieb, 2009; Moretti, 2011).

A suitable framework for such analysis is provided by a general spatial equilibrium model with the typical assumption that welfare is equalized across space (Glaeser and Gottlieb, 2009) and is assumed to be determined by three factors: wages, housing costs, and the quality of amenities. Our theoretical predictions are based on a derivation of the standard Rosen (1979) and Roback (1982) model with spatial

[1] The model suggests that the price-income ratio can be stationary in some countries and nonstationary in others, but that nonstationarity of the ratio does not necessary imply lack of cointegration between house prices and income.

[2] While we refer to 'cities' in the theoretical discussion, the same logic applies to wider metropolitan areas or regions as well.

equilibrium and are largely grounded on the model presented in Moretti (2011). Carlino and Saiz (2019), for instance, provide empirical evidence consistent with such a model. We consider the *long-term* developments in the price-income relationship because, in the short run, there are frictions that can restrain labor and firm mobility and the adjustment of house prices and supply toward equilibrium (Anenberg, 2016; Moretti, 2011). Our considerations also provide insights into the complex relationship between labor and housing markets, which plays a key role in the model. We are not aware of earlier studies that would have considered the implications of spatial equilibrium for the house price-income ratio.

Here, we present the principal idea of the standard model we apply to investigate the house price-income ratio. Appendix A describes the model and its derivations in more detail and provides a simple numerical illustration.

Following Glaeser and Gottlieb (2009), among others, we assume that the utility of workers in city i (U_i) is given by the Cobb-Douglas utility function

$$U_i = M_i C_{H,i}^{\gamma} C_{O,i}^{1-\gamma},\ 0 < \gamma < 1. \tag{1}$$

In (1), M_i is the quality of amenities[3] in city i, C_H and C_O represent the consumption of housing and other goods, respectively, and γ is the share of expenditure on housing, which is assumed to be similar over time and across cities. Piazzesi et al. (2007), Davis and Ortalo-Magné (2011), and Piazzesi and Schneider (2016) provide support for this assumption, which is common in spatial equilibrium models. Similar to Glaeser and Gottlieb (2009) and Hsieh and Moretti (2015), the indirect utility (V_i) then equals[4]

$$V_i = M_i W_i (P_i)^{-\gamma}, \tag{2}$$

where W_i denotes the nominal wage level and P_i is the cost (or price) of housing in city i. In log form

$$v_i = m_i + w_i - \gamma p_i, \tag{3}$$

where the lower-case letters denote natural logs. Utility is positively related to wage level and the quality of amenities, and negatively affected by higher housing costs. In spatial equilibrium, the utility levels are the same across cities; i.e., workers are indifferent between locations. Hence, in spatial equilibrium

$$w_i - \gamma p_i + m_i = w_j - \gamma p_j + m_j \tag{4}$$

holds for every city i and j.

In our model, the function for *long-term* equilibrium house prices at the city level follows conventional assumptions in the literature and is given by [for derivations, see Appendix A, Eqs. (A6)-(A9)]

$$p_i = \alpha_i + \beta_{1i} w_i + \beta_{2i} n_i;\ \beta_{1i} = \beta_{2i} = \frac{1}{\omega_i + 1} > 0, \tag{5}$$

where n_i and ω_i denote (log of) population and the price elasticity of housing supply in city i, respectively, while β_{1i} and β_{2i} are the city-specific coefficients on income and population to be estimated in the empirical analysis. Higher wages, greater population, and lower supply elasticity due to topographic or regulatory constraints cause higher housing costs. The considerable spatial variation in ω_i (Saiz, 2010) is expected to yield notable variation in β_{1i} and β_{2i} across cities. In the model, the quality of amenities affects housing demand and thereby house prices indirectly through city populations.

The interdependence among cities affects local house prices through population movements, in particular. For instance, an increase in productivity (and thereby wages) or the quality of amenities in city j leads to lower house price levels in city i due to some households moving to j. On the other hand, house prices increase in city j as the population inflow increases housing demand, and the spatial equilibrium condition in (4) is maintained. As the local wage level is determined by productivity in the city, the house price-income ratio ($p_i - w_i$) and its time path in city i is dependent on developments in other cities j.

In summary, the conventional spatial equilibrium model predicts that :[5]

1) The equilibrium house price-income ratio is not necessarily stable over the long run – in fact, long-term stability of the ratio is expected to be a special case rather than the rule.
2) The price-income ratio can be altered by various shocks, such as a shock in productivity or in perceived quality of amenities, in the city itself or in other cities.
3) The elasticity of housing supply is a key determinant of the influence of various shocks on the house price-income ratio, and the elasticities in other cities, too, affect the outcomes in a given city. Greater elasticity of supply is related to smaller growth trends in the price-income ratio.

Other implications of the framework for the relationships between house prices, incomes, and population are more familiar from the literature (although assumptions and results regarding point 5 vary, as discussed above):

4) House prices, wages, and population are jointly determined.
5) The income elasticity of house prices is expected to vary across cities.

We empirically investigate whether the predictions of the theory model hold true in practice. In the empirical analysis, we focus on points 1 and 5, but we additionally relate income elasticities and trends in price-income ratios to supply elasticities (point 3) and examine correlations between wages, populations, and house prices across MSAs (point 4).

3. Data

Our empirical analysis is based on quarterly data for the 70 largest (as of 2018) U.S. Metropolitan Statistical Areas or Divisions (referred to below as MSAs) for the period 1979Q3 through 2018Q2. This period includes one or more prominent house price cycles in all the MSAs. These cycles take place especially during the 2000s but also in the late 1980s through early 1990s. For house prices, we use the quarterly Federal Housing Finance Agency (FHFA) all transactions house price indexes (p).[6] The MSA per capita personal income (w) and aggregate personal income (wa) series are from the Bureau of Economic Analysis (BEA). As the income series are annual, we interpolate quarterly values

[3] Urban amenities are defined as local-specific characteristics that positively influence household utility and hence increase local housing demand and prices. Carlino and Saiz (2019) provide a review of literature supporting the role of amenities in households' location choices. Spatial variation in amenities can affect house values within cities as well. However, our focus is solely on the variations of local amenities *across* cities.

[4] As is typical, we abstract from the constant term that is assumed to be the same across cities.

[5] While spatial equilibrium models conventionally assume perfect labor mobility, these predictions hold even in the case of frictions to mobility, such as transaction costs and other relocation costs.

[6] We limit the sample to the 70 largest MSAs since smaller MSAs tend to exhibit too much implausible volatility in the FHFA house price indexes and many lack complete data. Due to extreme volatility (likely due to measurement error) in the early years of the price index, we also exclude Honolulu, which would have been ranked 69th with respect to population. In terms of short- and long-run house price dynamics, the FHFA data are similar to CoreLogic data (Oikarinen et al., 2018).

Table 1
Descriptive statistics.

Variable			Mean across all MSAs (annualized)	Standard deviation of MSA-specific means (annualized)	Lowest mean across MSAs (annualized)	Highest mean across MSAs (annualized)
Real house price growth (Δp)			0.011	0.049	–0.004	0.036
Real per capita income growth (Δw)			0.016	0.021	0.008	0.028
Real aggregate income growth (Δwa)			0.028	0.023	0.001	0.053
Population growth (Δn)			0.012	0.008	–0.007	0.041
Correlations			p	y	ya	pop
Real house price (p)			1.000			
Real per capita income (w)			0.630**	1.000		
Real aggregate income (wa)			0.392**	0.749**	1.000	
Population (n)			0.090**	0.323**	0.869**	1.000
			Δp	Δy	Δya	Δpop
Δp			1.000			
Δw			0.403**	1.000		
Δwa			0.393**	0.945**	1.000	
Δn			0.073**	0.087**	0.407**	1.000
			p	y	ya	pop
Mean of cross-sectional correlations			0.581	0.956	0.963	0.693
CIPS unit root test statistics						
Variable	p_{it}	w_{it}	wa_{it}	Δp_{it}	Δw_{it}	Δwa_{it}
Test value	–2.477	–2.115	–2.300	–3.937**	–4.906**	–4.742**

Note: The sample period is 1979Q3–2018Q2. For the correlations and unit root tests, * and ** denote statistical significance at the 5%, and 1% level, respectively. Correlations are reported for all MSA quarterly data stacked together. The mean of cross-sectional correlations is the average of cross-sectional correlations between all MSA pairs. The CIPS test values are based on city-specific CADF regressions. An intercept is included in all the CADF and ADF regressions. The regressions in the tests for levels additionally include a linear trend following Holly et al. (2010). The number of lags in the CADF regressions is allowed to vary across cities. For each MSA, the lag length is based on the general-to-specific method, using a threshold significance level of 5% and a maximum lag length of four. All the average residual cross-correlations of ADF regressions are statistically significant.

of per capita income based on changes in the national GDP, which is also from the BEA. While the quarter-to-quarter variations in the income variables do not affect the long-term estimates in cointegrating equations, the use of quarterly data provides us with a much greater number of observations and thereby more powerful tests. All variables are in real terms and in natural log form.[7] Table 1 provides descriptive statistics of the data. Although not separately included in the regression models, we report statistics for MSA-level populations as well (also from the BEA).

As expected, there are considerable regional variations in the mean growth rates of house prices, incomes, and population. The mean real house price growth was negative between 1979 and 2018 in six MSAs, all of which are inland. The highest price growth (annualized rate of 3.6%) was observed in San Francisco. In San Jose and Nassau-Suffolk, too, the figure was over 3%, while in Tulsa it was –0.4%. Population growth was very rapid in Las Vegas, 4.1% per year on average, and the growth rate reached 3% in two other MSAs as well. The highest house price growth rates were not in any of the MSAs with the highest population growth rates. There were five MSAs with contracting population, four of them in the Great Lakes region and the other being Philadelphia; of these, Detroit had the largest rate of population loss (0.7% per year).

The mean real per capita and aggregate income growth rates were positive in all 70 MSAs. Across all the MSAs, the annual mean growth rates were 1.6% and 2.8%, respectively. In San Francisco, per capita income growth was 2.8% per year, while the growth rate was only 0.8% in Detroit and Riverside. Real aggregate income growth was highest in Austin (5.3%) and lowest in Detroit (0.1%).

Portland, OR offers an interesting illustration of how city-specific developments of the three variables, *p, w*, and *n*, can differ relative to the average developments across cities. While Portland's annual real house price and population growth rates were relatively large during the sample period, 1.9% (the mean across the MSAs is 1.1%) and 1.6% (1.2%), respectively, in terms of real per capita income growth Portland was ranked only 50th (1.4%). Based on the spatial equilibrium framework, these patterns could be explained by growth in the perceived quality of amenities in the city: higher quality of amenities leads to lower required income net of housing costs, inducing greater population and thereby higher prices relative to income. Indeed, Portland is perceived as a city in which the quality of amenities has substantially increased, thereby increasing the supply of labor (population) in the city (Miller, 2014).

In accordance with the theoretical model, Table 1 shows that all correlations between variables are positive both in levels and in differences, and the mean of cross-sectional correlations across MSAs is large in all cases. That is, as expected, house prices are higher in larger cities with higher income levels.

As a preliminary check, we conducted panel unit root tests to examine the stationarity of each variable used in the regression analysis. Since the residual series from conventional augmented Dickey-Fuller (ADF) regressions include significant cross-sectional correlation and hence the conventional panel ADF test statistics could be biased, we follow Holly et al. (2010) and report the cross-sectional augmented IPS (CIPS) panel unit root test (Pesaran, 2007). The CIPS test is based on ADF regressions that are augmented with cross-sectional averages of the variables (CADF) and is thereby not biased by spatial dependence in the data. As shown by Gueye (2021), it is important to take account of cross-sectional dependence when conducting unit root tests for house prices and economic fundamentals, and for residuals from regressions between house prices and fundamentals (i.e., in cointegration tests, conducted in Section 4 of this paper). The test also allows for regional heterogeneity, as CADF regressions are estimated separately for each MSA. The results reported in the lower part of Table 1 indicate that the variables should be treated as non-stationary in levels. For all the differenced variables, the test statistics indicate stationarity.

4. Empirical analysis

In this section, we test some of the key implications of the spatial equilibrium model. In particular, we study the stationarity of the house price-income ratio and report regression results and cointegration tests based on several alternative estimators and model specifications. We also investigate the extent of heterogeneity across MSAs and relate this heterogeneity to the supply elasticity of housing and demand side factors. Appendix B provides more information on the various estimators and tests used in the empirical analysis.

The spatial equilibrium model predicts that the house price-income ratio is generally not stationary at the city level and therefore also points to complications with using the ratio to identify house price misalignments. In the regression models, we relax the restrictive assumption – a coefficient of one on income – imposed implicitly by the ratio. The regression models are based on the house price Eq. (5).

$$p_i = \alpha_i + \beta_{1i} w_i + \beta_{2i} n_i;\ \beta_{1i} = \beta_{2i} = \frac{1}{\omega_i + 1} > 0,$$

The estimated models are:

$$\text{Model 1}:\ p_{i,t} = \alpha_{i1} + \beta_{y,i} w_{i,t} + \varepsilon_{i1,t} \tag{6a}$$

$$\text{Model 2}:\ p_{i,t} = \alpha_{i2} + \beta_{ya,i} wa_{i,t} + \varepsilon_{i2,t}, \tag{6b}$$

where wa_i is the natural log of aggregate income in city i (i.e., equals $w_i + n_i$), α_i are the MSA-specific fixed-effects, $\beta_{y,i}$ and $\beta_{ya,i}$ are MSA-specific slope coefficients, and ε_{i1} and ε_{i2} are MSA-specific error terms. That is, we let the coefficients on w and wa vary across MSAs. Model 1 allows the coefficient on income per capita to differ from one but ignores the effects of population growth on housing demand. Model 2 takes account of population developments by including aggregate instead of per capita income. Note that Model 2 corresponds to the price Eq. (5): since $\beta_{1i} = \beta_{2i}$, there is no need to include income and population in the regression model separately. This also circumvents the collinearity complication that is typically present in house price regressions when incomes and populations are separately included as explanatory variables.[8]

The price-income ratio can be presented in the same form as Model 1: $p_{i,t} - w_{i,t} = \alpha_i + \varepsilon_{i,t} \rightarrow p_{i,t} = \alpha_i + \beta_{w,i} w_{i,t} + \varepsilon_{i,t}$, where $\beta_{w,i} = 1$ for all cities i.

If $\varepsilon_{i,t}$ is stationary, then the (estimated) relationship can be regarded as a stable long-run relation, implying that factors other than ω_i and either $w_{i,t}$ or $wa_{i,t}$ have only temporary effects on house prices. In the case of the price-income ratio, stationarity of $\varepsilon_{i,t}$ would additionally suggest that $\beta_{w,i} = 1$. In contrast, if $\varepsilon_{i,t}$ is non-stationary, the respective model does not imply a stable long-term relationship. The theoretical model predicts that Model 2 should outperform both the price-income ratio and Model 1 in terms of producing stable long-run relationships, as it corresponds to (5). As noted previously, if we added variables other

[7] All variables, including income, are deflated by the national urban CPI less shelter costs. If we were to deflate income by the CPI for all items (including shelter), house price growth would affect the deflated income series. Although house prices are not included in the CPI, rents are included, and house prices and rents are essentially measuring the same thing in the long run. For example, if housing demand grows substantially, inducing greater house prices and rents, while other prices stay constant, then the all-items CPI would increase. This would lower our real income measure even though income and all other (non-housing) components of the CPI have remained constant. Therefore, housing demand growth would not only cause higher real house prices but also lower real incomes, meaning that house price growth would have a disproportionate impact on the relationship between prices and incomes.

[8] Apparently due to collinearity complications, some studies of house price dynamics report negative coefficients on population when income and population enter the model separately.

than $w_{i,t}$ or $wa_{i,t}$ in the estimated long-term equations, we would no longer be studying the stationarity of pure price-income relationships.

The stationarity of residuals – which indicates that the model is cointegrated, i.e., that house prices and (aggregate) income have a cointegrating relationship – also has some other practical implications. First, since a stationary (i.e., cointegrating) equation for house prices can be interpreted as a long-term fundamental price level towards which house prices adjust, any short-term deviations of prices from this equation need to correct over time – that is, house prices being notably over their long-term level indicated by such equation would reflect substantial overpricing of housing that should correct in the future. Therefore, such relationship may be reasonably used as a bubble indicator. In contrast, a nonstationary regression equation can exhibit spurious regression bias and cannot be reliably interpreted as a long-term price equation, as there is no tendency of prices to revert to such a relationship. Given the importance of house price cycles for the economy and the fact that house price-income relationships are used as bubble indicators by many prominent institutions, this issue is highly relevant to policy. Second, as cointegration of an estimated model indicates that house prices react to deviations from the equation, observing stationary residuals has important implications regarding the price dynamics and thus concerning the predictability of future house price movements. Neglecting cointegrating relations would result in excluding information contained in the non-stationary levels variables (Engle and Granger, 1987).

The fully-modified OLS (FMOLS) estimator of Pedroni (2000, 2001) is a good starting point for estimating our regressions given that at least the population component of aggregate income is likely to be endogenous. While the estimators generally used in previous studies, such as conventional fixed-effects or random-effects OLS estimators, can exhibit endogeneity bias, the FMOLS estimator is consistent in the presence of endogenous regressors and endogeneity due to possible omitted variables (Pedroni, 2001, 2007). We report results for both the pooled FMOLS (PFMOLS) estimator that allows regional heterogeneity only through city-specific fixed-effects and the FMOLS mean-group (FMOLS-MG) estimator that allows regional heterogeneity in all parameter estimates. The FMOLS estimators are also super-consistent in the presence of non-stationary but cointegrated data, which is not the case for the fixed-effects OLS estimator. For comparison purposes, we also report results from the basic pooled fixed-effects OLS (POLS) estimator.

A potential complication with the aforementioned estimators is that they do not control for spatial dependence. Hence, we also report results based on the Pesaran (2006) common correlated effects mean group (CCEMG) estimator and the Chudik and Pesaran (2015) dynamic CCEMG (DCCEMG) estimator. Although these two estimators aim to remove the potential biasing impact of spatial dependence by including the cross-sectional averages of the dependent and independent variables as additional regressors (while allowing for regional heterogeneity), they can exhibit bias due to endogeneity. Moreover, due to the several additional variables that aim to remove cross-sectional dependence, the slope coefficient estimation may no longer be super-consistent. Hence, some of the attractive robustness features associated with super-consistent estimation under cointegration are potentially lost (Pedroni, 2007). Indeed, it turns out that the (D)CCEMG estimators do not work well with our data, which could be due to these complications. Based on the properties of the estimator, FMOLS-MG is preferred.

4.1. Baseline results

Consistent with Holly et al. (2010), the CIPS unit root tests reported in Table 2 reject the hypothesis of a unit root in the log price-income ratio, $p_{i,t} - w_{i,t}$. The CIPS tests also reject the hypothesis of no-cointegration (i.e., of a unit root in $\varepsilon_{i,t}$) in all the regression models except for those based on the (D)CCEMG estimator.

Stationarity of $p - w$ would indicate that the long-run coefficient on w is one and homogenous across MSAs, which is in contrast with the theoretical predictions. However, based on the size-adjusted F-test for the FMOLS-MG model (Pedroni, 2007) and the Swamy test of slope homogeneity for the (D)CCEMG models (Pesaran and Yamagata, 2008), the hypothesis of homogeneous coefficients on w is clearly rejected. Moreover, Wald F-test statistics reject the hypothesis that the group mean or pooled coefficient on w equals one for all models. Hence, the

Table 2
Cointegration tests and regression results.

	FMOLS-MG	POLS	PFMOLS	CCEMG	DCCEMG
CIPS cointegration test statistics					
$p - w$: –1.774**					
Model (1)	–2.156**	–2.081**	–2.052**	–1.329	–1.332
Model (2)	–2.549**	–2.076**	–2.007**	–.674	–0.714
Coefficient estimates and test statistics					
	Model 1				
y_{it}	0.745**	0.844**	0.771**	1.776**	1.744**
	(0.042)	(0.008)	(0.044)	(0.177)	(0.190)
Average residual cross-correlation	0.340	0.347	0.341	–0.008	–0.009
F-test of homogeneity (p-value)	0.000**				
Swamy test (p-value)				0.000**	0.000**
Wald F-test on $\beta_y = 1$ (p-value)	0.000**	0.000**	0.000**	0.000**	0.001**
	Model 2				
ya_{it}	.479**	.412**	.351**	1.650**	1.642**
	(0.023)	(0.005)	(0.028)	(0.142)	(0.146)
Average residual cross-correlation	0.338	0.332	0.343	–0.002	–0.003
F-test of homogeneity (p-value)	0.000**				
Swamy test (p-value)				0.000**	0.000**
Wald F-test on $\beta_{ya} = 1$ (p-value)	0.000**	0.000**	0.000**	0.000**	0.000**

Note: The sample period is 1979Q3–2018Q2. $p - w$ is the log house price-income ratio. Dependent variable $= p_{i,t}$. The intercepts are not reported. * and ** denote statistical significance at the 5%, and 1% level, respectively. Except for POLS and PFMOLS, the reported regression coefficients represent the mean group estimates, i.e., the mean estimates across all MSAs. The standard errors for the mean group estimates are computed following Pesaran and Smith (1995). The models include MSA-specific intercepts (fixed-effects). The null hypothesis in the Swamy test and F-test on homogeneity is that of homogeneous slope coefficients across MSAs. The CIPS statistics are based on CADF regressions that do not include intercepts, as these are cointegration tests on model residuals. The number of lags in the CADF regressions is allowed to vary across cities. For each MSA, the lag length is based on the general-to-specific method, using a threshold significance level of 5% and a maximum lag length of four. Critical values in the CIPS test are –1.53 –1.65 at the 5% and 1% level of significance, respectively. The lag length in the Bartlett (Newey-West) window width in the FMOLS estimations is four. The lag length choice does not notably affect the results.

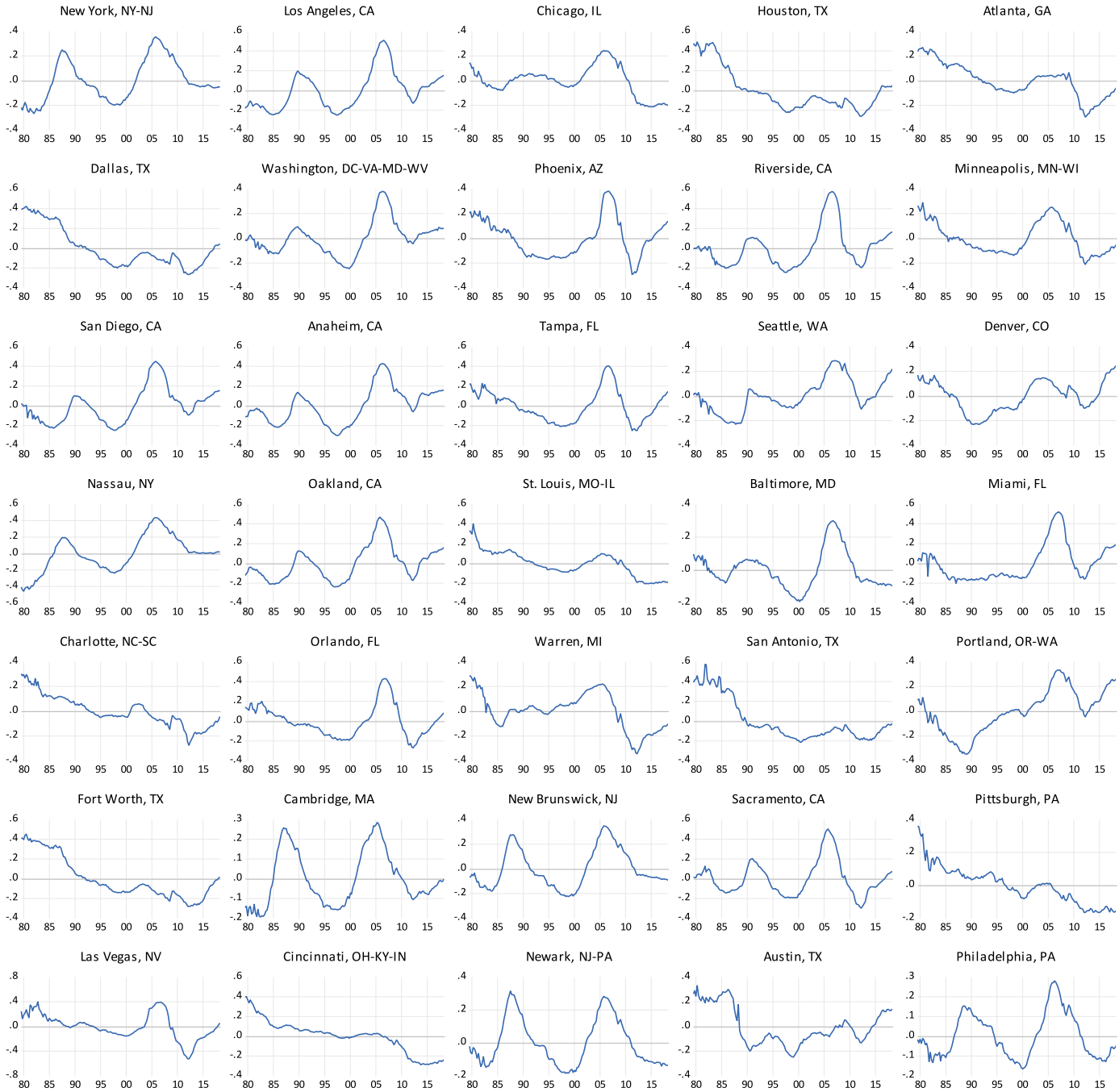


Fig. 1. Residuals from (log of) house price-income ratios (demeaned).

regression results for Model 1 are in stark contrast with the concept of a stationary house price-income ratio. The test statistics for Model 2 (with aggregate personal income) also reject the hypothesis of $\beta = 1$ and indicate significant variations in the coefficient estimates across MSAs.

The two estimators that aim to control for cross-sectional dependence, CCEMG and DCCEMG, remove practically all cross-sectional correlation from the model residuals (the remaining correlation is less than 0.01). However, in these models the residual unit root hypothesis cannot be rejected, and the city-specific residual series are clearly trending for most MSAs. This implies that the (D)CCEMG estimators do not work well for our data: The FMOLS results clearly indicate that there are stationary relationships between house prices and income in these MSAs, which is also supported by the lack of obvious trends in most of the model residuals (shown in Fig. 4 for Model 2) – if one estimator yields stationary relationships but another does not, instead producing clearly trending residuals, the former is obviously preferred as it provides an equation with non-trending residuals and in effect finds existing cointegrating (i.e., long-term equilibrium) relationships that the other is unable to detect.[9] These complications are not unexpected given the properties of the (D)CCEMG estimators discussed above.

The preferred FMOLS-MG estimator yields mean group estimates of 0.75 on *w* and 0.48 on *wa*, and the POLS and PFMOLS estimates differ somewhat from the FMOLS-MG ones. The mean-group estimate of 0.48 in FMOLS-MG Model 2 is close to what Eq. (3) would predict

[9] Despite these complications, we also report the results based on the CCEMG and DCCEMG estimators because they are more recent innovations aimed at data with cross-section correlation, and also are methods that have been used in previous related literature (in Holly et al., 2010, in particular).

$[1/(\omega_i+1) = 0.41]$ based on the median supply elasticity of 1.44 across the MSAs estimated by Saiz (2010).

4.2. A closer look at the results

The interpretation of the unit root test results from panel level analysis is complicated due to the nature of the alternative hypothesis. While the null hypothesis is that of a unit root in each series, the alternative hypothesis is more complex, especially in heterogeneous panels: rejecting the null does not necessarily mean that all or even most individual series are stationary; this point has not been considered in the related literature. The null hypothesis (H_0) and the alternative hypothesis (H_1) in our panel cointegration tests are:

H0. *Each* of the residual series is non-stationary (i.e., none of the MSA-specific equations is cointegrated).

H1. *One or more* residual series are stationary (i.e., one or more MSA-specific equations are cointegrated).

Pesaran (2012) suggests that the rejection of the panel unit root null hypothesis should be interpreted as evidence that a statistically significant fraction of the individual series is stationary. That is, a rejection of the null hypothesis does not necessarily mean that the respective relationship is stationary for all or even most cities: a relatively small group of MSAs with stationary relations can cause the panel unit root test to reject the null hypothesis.

In accordance with theory, Fig. 1 shows that many of the (demeaned) price-income ratios have notable trends, implying that in many MSAs the ratio is not stable even over the long run. In line with the visual inspection, a unit root in the residuals from $p - w$ can be rejected in only 11 of the 70 MSAs at the 5% level of significance based on individual CADF statistics. Given the well-known power problems with individual ADF-type tests, the 10% level of significance may be a more reasonable threshold, but even at the 10% level the unit root is rejected in only 17 MSAs. Hence, the fact that the CIPS test rejects the null hypothesis of a unit root in $p - w$ cannot be used as evidence of stationarity of the price-income ratio in all, or even most of, the MSAs.

If we regress the panel of price-income ratios on an intercept and a

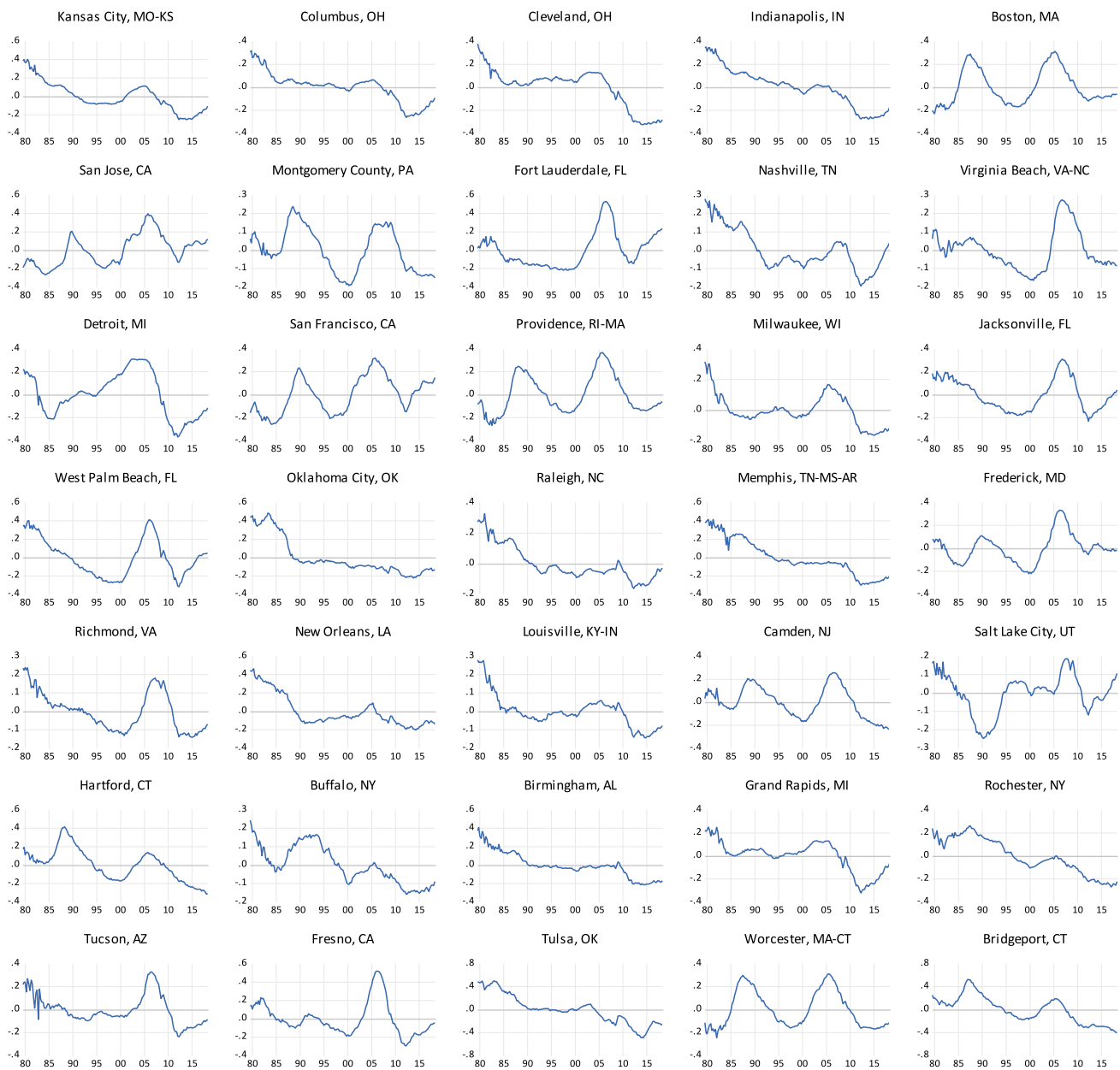


Fig. 1. (*continued*).

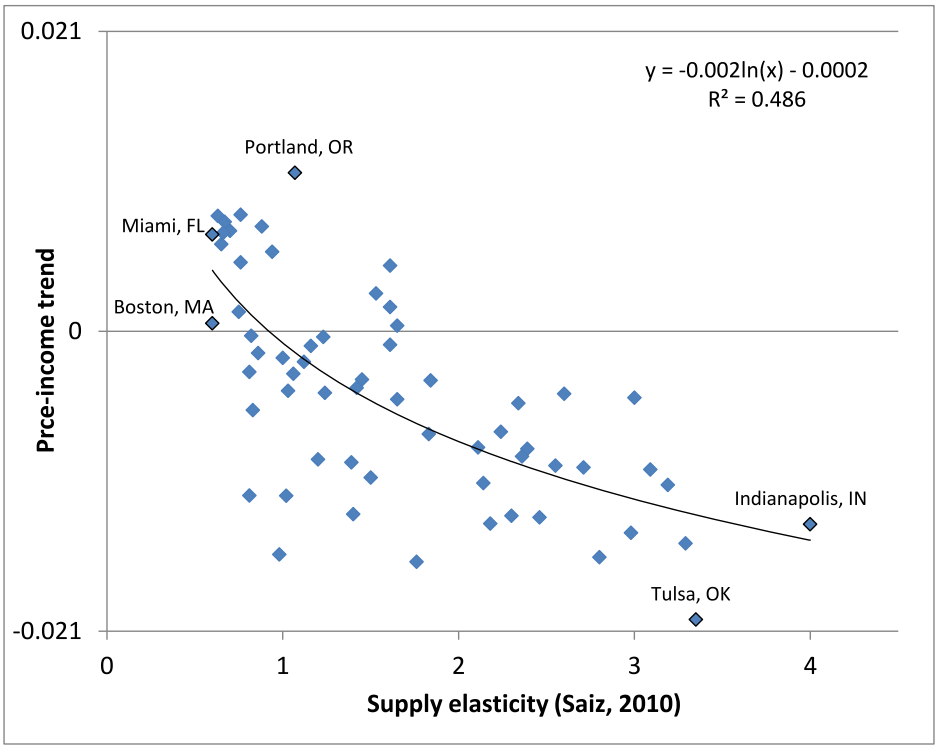


Fig. 2. Annualized trends in house price-income ratio and the supply elasticity of housing.

time trend using the Pesaran et al. (1999) mean group estimator that allows for regional heterogeneity in the coefficients, we find statistically significant trends in 57 MSAs (approximately 80%). The MSA-specific trends are significantly associated with the price elasticities of housing supply reported in Saiz (2010). Fig. 2 illustrates that, generally, the slope of the trend in the observed $p - w$ relationship is larger, i.e., house prices have increased more relative to income, in cities with relatively inelastic supply. For example, the two MSAs with the least elastic supply – Boston and Miami – both have positive price-income trends. In contrast, Indianapolis has the highest supply elasticity and one of the lowest price-income slopes. Consistent with its perceived increase in the quality of amenities, Portland is the MSA with the highest price-income slope; it has a relatively low (although slightly greater than one) supply elasticity. Tulsa, OK, has the lowest price-income slope and the second highest supply elasticity.

In fact, the most common trend in $p - w$ is negative, suggesting that housing affordability has increased in a majority of the MSAs.[10] Fig. 3 (panel A) shows that the $p - w$ trends tend to be positive, and therefore housing affordability tends to get worse, on the east and west coasts and negative elsewhere. In line with the price-income trends and theoretical considerations, the MSA-specific FMOLS-MG estimates on w and wa (the income elasticities of house prices) are highly negatively correlated with supply elasticities: the correlations are −0.63 (Model 1) and −0.44 (Model 2).

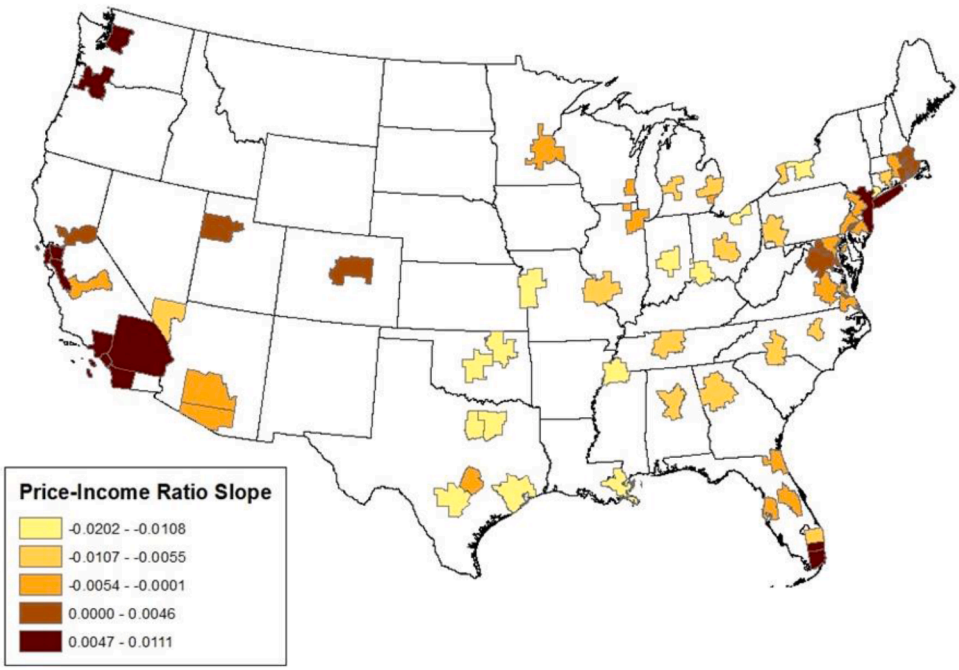
The developments in house price-income ratios are relevant to trends in the wealth-income relationship. The findings of Piketty and Zucman (2014) suggest that capital gains on housing explain a large part of the rise of wealth-income ratios in several countries, including the U.S., since 1970, and Knoll et al. (2017) report a substantial rise in house prices relative to GDP across a number of developed countries. However, the price-income developments at the country level can hide heterogeneous developments across regions within a country. Indeed, our data provide evidence of downward trending price-income ratios in a large number of MSAs, suggesting that increases in the wealth-income ratios due to house price trends have not occurred in these cities since 1979 and are not inevitable in the future.

The price-income trends also are in line with Glaeser and Gottlieb (2009), who argue that the rise of Sunbelt cities is related to abundant housing supply rather than rising amenity values. If amenity values drove the growth of Sunbelt cities, then we would expect the price-income trends to be increasing in these cities. However, with the exception of most California MSAs and Miami and Fort Lauderdale in Florida, all of which are supply constrained, the price-income ratio has trended downwards in the Sunbelt metropolitan areas (in 15 out of 17 such areas outside California) as shown by Fig. 3A. Moreover, the price-income trends are not significantly correlated with the MSA-specific average January temperatures.

Table 3 summarizes the MSA-level unit root statistics for the price-income ratio and both FMOLS-MG models. If the assumption of a coefficient of one on per capita income (imposed by the price-income ratio) is relaxed, and the coefficient is allowed to vary across cities (Model 1), the number of MSAs for which the unit root can be rejected at the 10% level in individual CADF tests increases from 17 to 35. The model with aggregate income (Model 2) works even better, with stationary relationships in 43 cities. Thus, the relationship is stationary in more and more cities when the restrictive assumptions – that are not consistent with the theoretical considerations – are progressively relaxed. The results therefore indicate that population should be included in regressions (by using aggregate instead of per capita income) to better capture regional price dynamics and to reach more reliable conclusions regarding possible disequilibria in regional house price levels.

[10] While decreasing house price-income ratios imply improved affordability at the metropolitan area level, affordability can still deteriorate for some households as pointed out by Gan and Hill (2009).

Panel A: Annualized price-income ratio slopes



Panel B: Aggregate income coefficient estimates (Model 2)

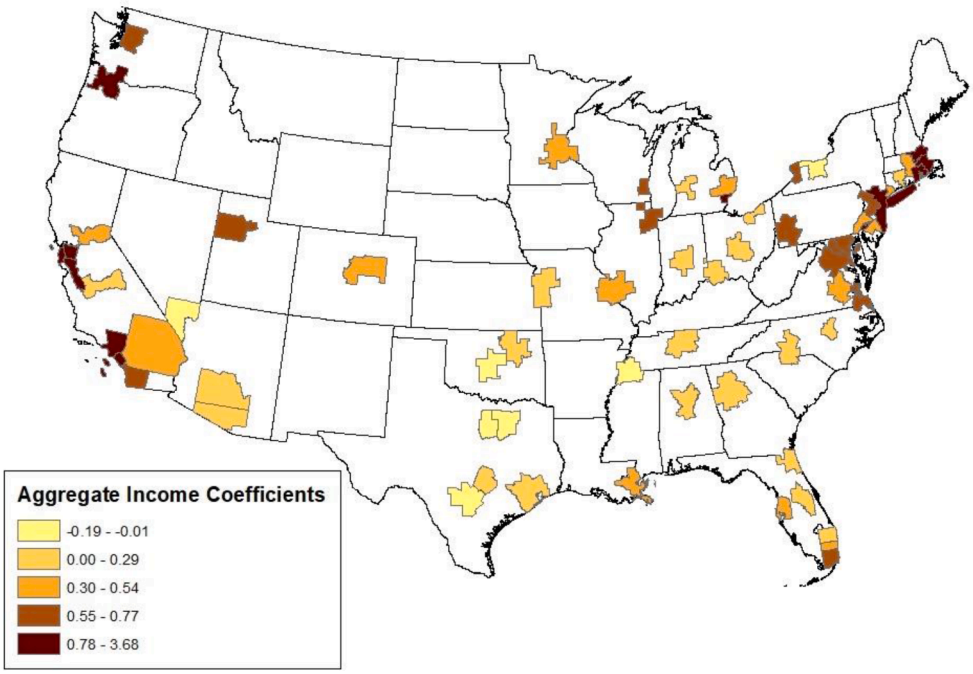


Fig. 3. Geographic distribution of house price-income relationships.

Table 3
MSA-specific CADF unit root test statistics for house price-income ratio and FMOLS-MG models (MSAs ordered by 2018 population).

			Regression model					Regression model	
		p–w	1	2			*p–w*	1	2
1	New York, NY-NJ (MSAD)	**	**	***	36	Kansas City, MO-KS			
2	Los Angeles, CA (MSAD)		**	**	37	Columbus, OH			
3	Chicago, IL (MSAD)				38	Cleveland, OH			
4	Houston, TX				39	Indianapolis, IN			
5	Atlanta, GA		**	***	40	Boston, MA (MSAD)	***	**	**
6	Dallas, TX (MSAD)				41	San Jose, CA		**	**
7	Washington, DC-VA-MD-WV (MSAD)			***	42	Montgomery, PA (MSAD)			
8	Phoenix, AZ			**	43	Fort Lauderdale, FL (MSAD)		*	**
9	Riverside, CA	*		***	44	Nashville, TN	*		
10	Minneapolis, MN-WI				45	Virginia Beach, VA-NC		**	*
11	San Diego, CA		*	***	46	Detroit, MI (MSAD)		*	***
12	Anaheim, CA (MSAD)			**	47	San Francisco, CA (MSAD)		*	*
13	Tampa, FL				48	Providence, RI-MA		*	*
14	Seattle, WA (MSAD)		**	**	49	Milwaukee, WI	*	*	**
15	Denver, CO			*	50	Jacksonville, FL			*
16	Nassau, NY (MSAD)	***	***	**	51	West Palm Beach, FL (MSAD)		*	*
17	Oakland, CA (MSAD)		**	***	52	Oklahoma City, OK	**		
18	St. Louis, MO-IL		**	***	53	Raleigh, NC	*		
19	Baltimore, MD				54	Memphis, TN-MS-AR		*	*
20	Miami, FL (MSAD)		*		55	Frederick, MD (MSAD)		*	**
21	Charlotte, NC-SC	*		**	56	Richmond, VA	***	**	
22	Orlando, FL		**	***	57	New Orleans, LA			*
23	Warren, MI (MSAD)				58	Louisville, KY-IN	**		*
24	San Antonio, TX				59	Camden, NJ (MSAD)			
25	Portland, OR-WA			*	60	Salt Lake City, UT		*	*
26	Fort Worth, TX (MSAD)				61	Hartford, CT		***	***
27	Cambridge, MA (MSAD)	***	***	***	62	Buffalo, NY			
28	New Brunswick, NJ (MSAD)		**	**	63	Birmingham, AL	**	*	**
29	Sacramento, CA	**	*	***	64	Grand Rapids, MI			
30	Pittsburgh, PA	*	**	***	65	Rochester, NY		*	*
31	Las Vegas, NV		**	**	66	Tucson, AZ			
32	Cincinnati, OH-KY-IN				67	Fresno, CA	***	**	***
33	Newark, NJ-PA (MSAD)		**	**	68	Tulsa, OK			**
34	Austin, TX				69	Worcester, MA-CT	***	**	**
35	Philadelphia, PA (MSAD)				70	Bridgeport, CT		*	

Note: *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively. Critical values at the 10%, 5% and 1% level of significance are: -2.26, -2.60, and -3.30. $p - w$ is the log house price-income ratio. MSAD refers to areas that are metropolitan divisions.

Hence, Model 2 is the most useful for examining regional house price cycles relative to long-term fundamental levels.[11] Importantly, the residuals from Model 2 do not exhibit evident trends in any of the MSAs. This is in stark contrast with the simple price-income ratio, as shown in Fig. 4. However, the inability to detect cointegration in Model 2 in over one third of the MSA-specific equations may indicate that other fundamentals should be included in models aiming to capture long-term trends in house prices in some MSAs or that there have been structural changes in the price elasticities over time.[12]

Fig. 3B displays the geographic distribution of the aggregate income coefficients from Model 2. These tend to be higher in relatively supply inelastic coastal locations on the west coast and in the North, Mid-Atlantic, and Great Lakes regions. Overall, Fig. 3 highlights that it is particularly important to consider policy actions that can improve housing supply elasticity in the coastal regions where water bodies and other topographic constraints restrict housing supply.

In contrast with the MSA-specific residual series based on the FMOLS-MG equations, the POLS and PFMOLS equations – that assume homogenous slope coefficients across MSAs – yield clearly trending residuals in many MSAs. This reinforces our conclusion that the homogeneity assumption is too restrictive and that heterogeneity across MSAs should be allowed to get more reliable assessments of house price elasticities and misalignments.

4.3. Analysis of residuals and trends

Our Model 2 is a success in terms of such a simple model – only one explanatory variable – being able to capture the long-term developments of house prices in most of the MSAs. Even for the remaining 27 MSAs the model residuals do not trend up or down, suggesting that the model captures, to a large extent, the house price developments during the sample period in those MSAs as well. Also, in a set of 70 MSAs, the CADF test power complications may lead to acceptance of the null of no cointegration in some MSAs where the model actually is cointegrated. Nevertheless, we provide a simple further analysis of the potential reasons for the lack of cointegration for some MSAs and of potential factors associated with the deviations of house prices from estimated long-run equilibrium levels and with the price-income ratio trends, as such considerations may entail additional policy implications.

For this purpose, we include in the analysis national-level variables that may influence the relationship between house prices and income: the loan-to-GDP ratio and the real mortgage interest rate. Metropolitan level growth rates of income and population, housing supply elasticity, unemployment rate, college employment ratio, and an amenity index

[11] Although Model 2 corresponds to our theoretical considerations, we also estimated a model where income and population are included separately thus allowing for different coefficients on them. This model is highly problematic, since the coefficient on population takes the wrong (negative) sign in 47 MSAs and the mean-group estimate also is negative. Moreover, the model is cointegrated in fewer MSAs than Model 2. These complications are likely due to the strong time series collinearity between income and population in many of the MSAs (correlation of up to 0.98 in several MSAs). Gallin (2006), for instance, faces a similar complication with negative coefficients on population.

[12] Bourassa et al. (2019) report that a parsimonious regression model with only aggregate income on the right-hand side works as a better indicator for house price bubbles than a model that also includes other explanatory variables.

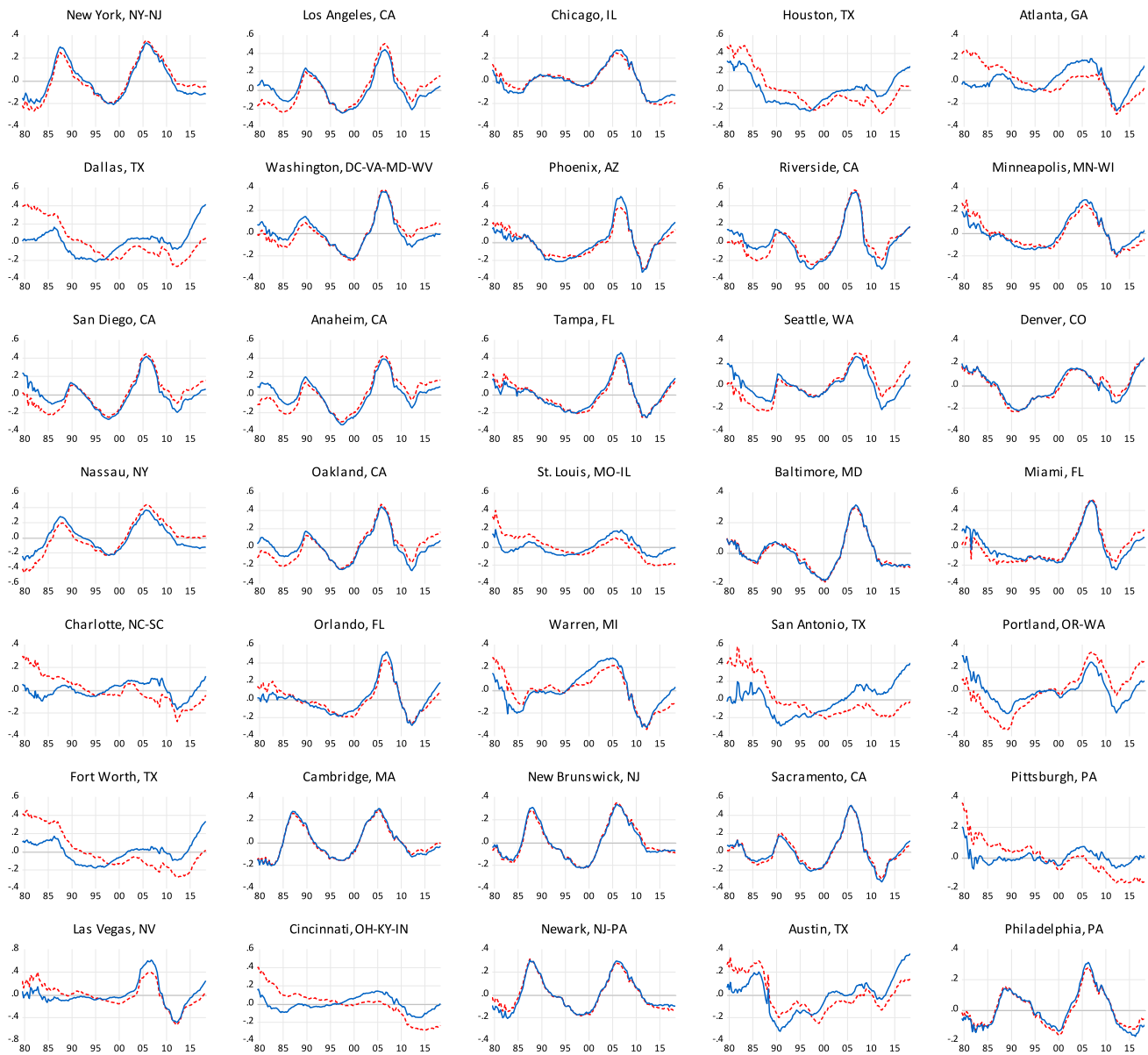


Fig. 4. Residuals from regression Model 2 (continuous blue) and from price-income ratio (dashed red).

are also included.[13] The loan-to-GDP ratio is aimed at capturing developments in housing loan constraints over time (Oikarinen, 2009; Karpestam and Johansson, 2019) that likely affect housing demand. The mortgage rate reflects time-variations in the discount rate and in liquidity constraints.

Diamond (2016) shows that, in addition to increased concentration of college graduates in cities with high wages, there have been endogenous increases in amenities within these higher skill cities. The differing time trends in the MSA shares of college workers and in the quality of amenities across MSAs may have affected the house price-income ratio trends. The MSA-specific change in the college employment ratio from 1980 to 2000, used by Diamond (2016), is added in the analysis to investigate these potential effects on the price-income relationship. Moreover, we use the MSA-level amenity index over 1990–2015 of Broxterman and Kuang (2019), aiming to capture time trends in the quality of amenities.

The spatial equilibrium model indicates that amenity variations over time can cause price-income ratio trends. In our Model 2, the influence of amenities and worker skill distributions are in principle captured indirectly through the aggregate income variable, i.e., through local populations and income levels: higher quality of amenities leads to greater population, and greater share of highly educated workers should cause higher productivity and thereby higher incomes and greater population. Nevertheless, we also investigate whether the college ratio and amenity variables are associated with the stationarity vs.

[13] The loan-to-GDP ratio and mortgage interest rate data were downloaded from the Federal Reserve Bank of St. Louis economic database: https://fred.stlouisfed.org; the unemployment data were sourced from the Bureau of Labor Statistics: https://www.bls.gov; the college employment ratio data used by Diamond (2016) were downloaded from https://www.openicpsr.org/openicpsr/project/112969/version/V1/view?path=/openicpsr/112969/fcr:versions/V1/replication&type=folder; the amenity indices are described in detail in Broxterman and Kuang (2019), we are thankful to those authors for providing these data. This is a revealed preference amenity index that ranks cities by amenity level using travel demand as a proxy for amenity quality; this index is apparently the only one providing a time series dimension for as long as 25 years (1990-2015).

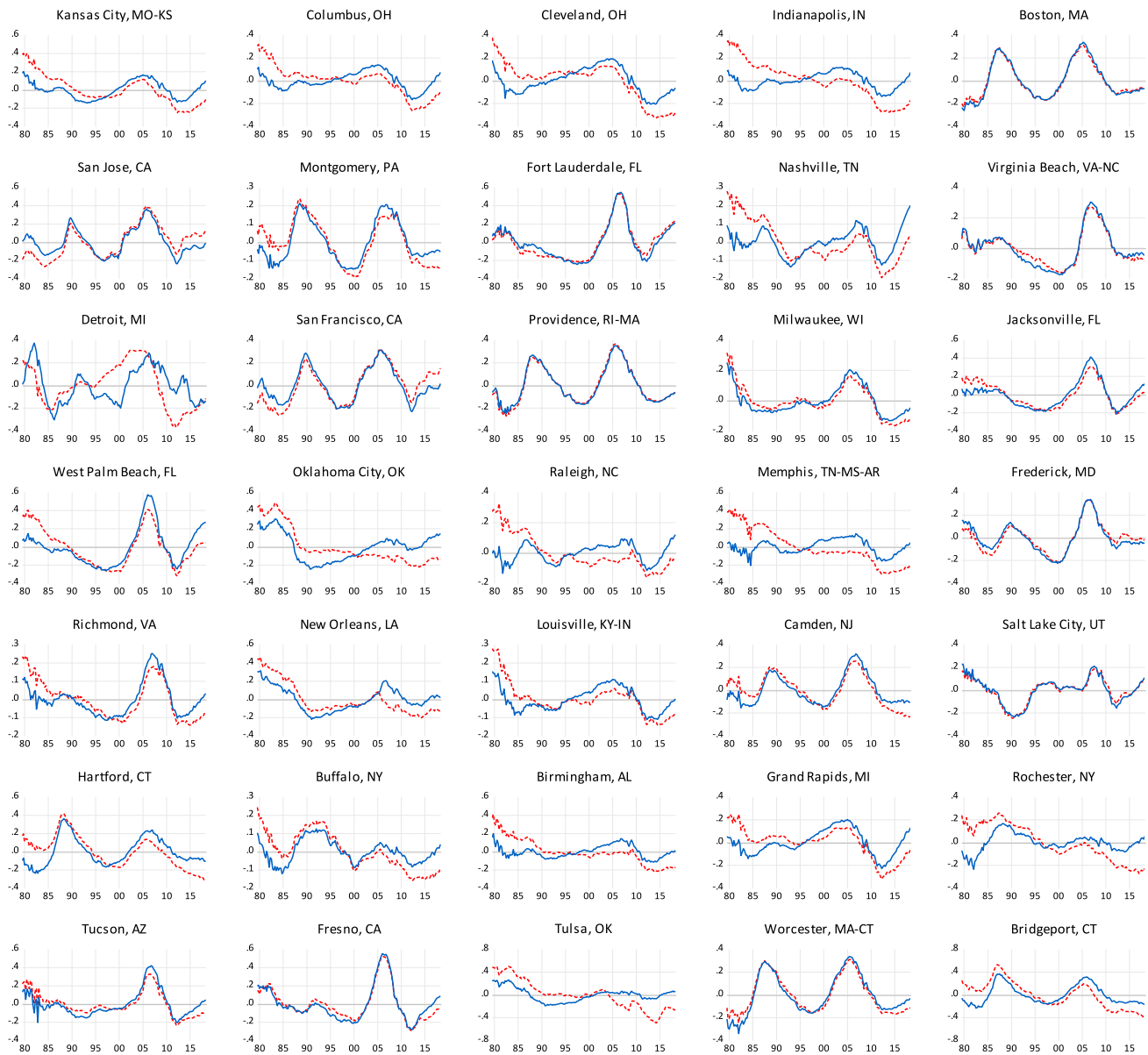


Fig. 4. (*continued*).

nonstationarity of Model 2 across the MSAs. We further add the local seasonally adjusted unemployment rates that likely capture, at least to some extent, the differences in MSA-specific education levels. The unemployment data has the benefit of being available at the quarterly frequency from 1990 onwards, whereas the college ratio data of Diamond (2016) are available only for census years 1980, 1990, and 2000.

We summarize the key observations and implications of this analysis below. Table 4 presents some of the most interesting estimated regressions.

On average, nonstationary MSAs (referring to Model 2) have much lower quality of amenities and higher supply elasticity than the MSAs for which we observe a cointegrating relationship. However, there is no difference between the stationary and nonstationary MSAs in terms of the change in amenity index values or college shares, and trends in neither amenities nor college shares exhibit notable correlation with the house price-income ratio trends.

The loan-to-GDP, mortgage rate, and unemployment data allow for panel regression analysis. The loan-to-GDP ratio is positively associated with the equilibrium errors, i.e., deviations from estimated long-run relations of Model 2. This is in accordance with prior expectations, since higher loan-to-GDP ratios are associated with looser loan constraints that stimulate housing demand. While the ratio appears to some extent to explain cycles of house prices around their long-run relationship with aggregate income, the inclusion of the ratio helps to make only one of the 27 nonstationary MSAs stationary (Nashville), and controlling for the mortgage rate does not help either. The equilibrium errors and unemployment developments, in turn, are negatively related: a lower price level relative to aggregate income is associated with a higher unemployment rate.[14] This may relate to both income distribution and uncertainty effects on housing demand. These panel analysis results apply regardless of whether we use the FMOLS-MG or the basic MG

[14] We prefer to use the whole sample period (1979-2018) to investigate long-run stationarity of the price-income relations, since having more years increases the power of cointegration tests. As expected, with a shorter sample period (1990-2018 with the unemployment data), several MSAs for which we observe long-run equilibrium relationships with aggregate income over 1979-2018 would falsely seem to have no such relationship.

Table 4
Regressions for Model 2 residuals and house price-income ratio trends.

Explanatory variable	PANEL MODEL Dependent: Residuals from Model 2 Sample period: 1990Q1–2018Q2	CROSS-SECTION (1) Dependent: price-income ratio slope 1979–2018	CROSS-SECTION (2)	CROSS-SECTION (3) Dependent: price-income ratio slope 1980–2000	CROSS-SECTION (4)
Loan-to-GDP ratio	0.742**				
	(0.021)				
Unemployment rate	–2.769*				
	(0.137)				
Constant		0.005**	0.000	–0.003**	–0.001
		(0.002)	(0.003)	(0.001)	(0.001)
Supply elasticity		–0.006**	–0.006**	–0.002**	–0.002**
		(0.001)	(0.001)	(0.000)	(0.000)
Population growth		0.112	0.164*		–0.249*
		(0.068)	(0.082)		(0.098)
College share growth			0.047*	0.025**	0.018*
			(0.023)	(0.009)	(0.008)
Average residual cross-correlation	.308				
F-test of homogeneity (p-value)	.000**				
Breusch-Pagan-Godfrey homoscedasticity test (p-value)		0.438	0.096	0.249	0.235
Jarque-Bera test on residual normality (p-value)	.000**	0.483	0.533	0.250	0.286
RESET test (p-value)		0.185	0.448	0.627	0.969
R^2	.116	0.525	0.540	0.351	0.421
adj. R^2		0.508	0.514	0.327	0.389

Note: * and ** denote statistical significance at the 5%, and 1% level, respectively. Standard errors are reported in the parenthesis. The panel model is estimated with FMOLS-MG estimator, and the reported regression coefficients represent the mean group estimates, i.e., the mean estimates across all MSAs. The standard errors for the mean group estimates are computed following Pesaran and Smith (1995). The panel model includes MSA-specific intercepts. The lag length in the Bartlett (Newey-West) window width in the FMOLS estimation is four. The lag length choice does not notably affect the results. R^2 for the panel FMOLS-MG model is computed as the squared correlation between predicted and actual Model 2 residuals. Price-income ratio slopes are quarter-level growth trends estimated with the MG estimator. Population growth is the mean quarterly growth rate over 1979–2018 (cross-section models 1–2) or 1980–2000 (cross-section models 3–4). College share growth is the (%-point/100) increase in the share of college workers from 1980 to 2000. In the Breusch-Pagan-Godfrey F-test, the null hypothesis is that of homoscedastic residuals. Ramsay RESET test includes squared fitted values, and the null hypothesis is no model misspecification.

estimator and level variables or differenced (stationary) variables. Table 4 reports the FMOLS-MG model in levels including both loan-to-GDP ratio and unemployment rate.

Cross-section probit regressions are conducted with the dependent binary variable being the Model 2 residual stationarity: 1 = stationary, 0 = nonstationary. These regressions include as explanatory variables the college employment ratio and amenity index mean values and their changes, supply elasticity, and income and population growth rates. Supply elasticity is the dominant explanatory variable, and is highly significant regardless of the other explanatory variables included in the regression. Without the elasticity, the probit models do not have an in-sample prediction power better than the naïve model of always guessing 1, whereas the models with elasticity perform substantially better than the naïve guess.[15] The relationship between elasticity and stationarity is negative as indicated by the observed negative correlation. None of the other variables are statistically significant when elasticity is controlled for. It should also be noted that the nonstationarity in some MSAs may be due to structural changes in the supply side. We cannot study this, unfortunately, as we are not aware of any time series of MSA-level supply elasticities.

Finally, we estimate cross-section regressions for the price-income ratio slopes. Regardless of the other explanatory variables included and whether we consider the price-income slope of the whole sample period or during 1980–2000 (when including the college employment share growth), or 1990–2015 (corresponding to the amenity data), the supply elasticity is always significant at the 0.1% level. While the elasticity alone can explain up to 50% of cross-sectional variation in the slope, the other variables together have much less explanatory power. This underlines the strong connection between supply elasticity and housing affordability trends over time. The population growth rate also has significant explanatory power with a positive sign.[16] This result is intuitive: a larger population means higher prices at a given income level and can likely capture amenity quality growth too, to some extent at least. Interestingly, the college employment ratio growth rate is significant with a positive sign even when controlling for supply elasticity and population growth. The population growth rate should capture amenity quality increases, at least to some extent, but this finding suggests that growth in the share of highly-educated population has had independent explanatory power that may relate to the influence of amenities and income distribution on housing demand. The pure amenity variable does not possess explanatory power. Table 4 presents two different cross-section regression specifications for both the full sample period price-income slope and for that corresponding to the college share data (1980–2000).

Overall, the results highlight the role of supply elasticity for the house price-income relationship and its developments. In any case, the data in this simple analysis are far from perfect and allow for suggestive evidence only. The potential influences of amenity developments and the increased geographic sorting of high-educated workers across the U. S. is an interesting topic for further research in this area.

5. Conclusions

This study contributes to the analysis of the relationship between house prices and income and regional heterogeneity in this relationship

[15] For instance, in a model with the elasticity as the only explanatory variable the gain in correct predictions over naïve model is 13 percentage points. In models with additional variables the gain is somewhat smaller. Naïve model refers to always predicting the outcome that takes place more often, i.e., in our case always predicting stationarity.

[16] This explanatory variable is potentially endogenous, as a higher price-income ratio can restrict population growth. The estimated coefficient may therefore be biased downwards.

in several ways. We consider a standard spatial equilibrium model and conduct an empirical analysis that examines whether results using panel data from the 70 largest U.S. MSAs are in line with that model's predictions – which they are. Our primary conclusion is that, at the MSA-level, an assumption of a constant house price-income ratio over the long run is in line with neither theory nor empirical facts. Instead, long-term stability of the price-income ratio in a given area or region is expected to be a special case rather than the rule, and house price predictions as well as evaluations of house price deviations from their long-term fundamental levels should be based on less restrictive assumptions, allowing income elasticities of house prices to differ from one and vary across regions. In addition, population growth should be considered when assessing local house price levels and dynamics by using aggregate income measures.

Our analysis leads to several additional conclusions: (1) It supports the argument that local supply constraints are related to greater increases in regional house prices relative to incomes, thus generating a counterforce for regional growth through adverse effects on the affordability of housing (while on the other hand supporting wealth accumulation). (2) Panel level cointegration, or unit root, tests can lead to misleading conclusions regarding the nature of the regional house price-income relationships. (3) Consistent with variations in supply elasticities across locations, our results underscore the importance of allowing for spatial heterogeneity when modeling regional house price dynamics.

CRediT authorship contribution statement

Elias Oikarinen: Conceptualization, Methodology, Formal analysis, Investigation, Writing – original draft. **Steven C. Bourassa:** Conceptualization, Investigation, Writing – original draft. **Martin Hoesli:** Conceptualization, Investigation, Writing – original draft. **Janne Engblom:** Methodology, Formal analysis, Writing – original draft.

Declaration of Competing Interest

None.

Data availability

The authors do not have permission to share data.

Acknowledgments

For helpful comments, we thank the editor Jeff Zabel and two anonymous reviewers, as well as Christian Hilber, Simon Lapointe and Oskari Vähämaa, and participants at the International AREUEA conference in Amsterdam, ARES conference in San Diego, ERSA conference in Cork, ERES conference in Delft, Annual Meeting of the Finnish Economic Association, and research seminar at the University of Cambridge. We also thank Andrew VanValin for preparing the maps. This work was supported by the Strategic Research Council at the Academy of Finland (first author; decision numbers 352450 and 352451) and the West-Finland Housing Association of Public Utility.

Appendix A. Spatial equilibrium model

The spatial equilibrium model

Our theoretical predictions are based on a derivation of the standard Rosen (1979) and Roback (1982) model with spatial equilibrium and is largely grounded on the model presented in Moretti (2011). This framework considers the whole system of cities and thus the impacts of interdependence among urban economies on local house price-income relations and their dynamics. We consider the *long-term* developments in the price-income relationship because, in the short run, there are frictions that can restrain labor and firm mobility and the adjustment of housing prices and supply toward equilibrium (Anenberg, 2016; Moretti, 2011).[17]

We start by assuming that each city is a competitive economy in a system of cities and produces a single output good Y. This good is traded in the 'international' market so that its price is the same in all cities. The price of one unit of Y is set to be 1. Similar to, e.g., Moretti (2011) and Kline and Moretti (2014), the production function in city i takes the Cobb-Douglas form with constant returns to scale:

$$Y_i = X_i N_i^h K_i^{1-h},\ 0 < h < 1. \tag{A1}$$

Here N_i represents the number of workers, K_i is the amount of capital in city i, and X_i is a city-specific productivity shifter. Firms and workers are mobile and locate where their profits and utility are maximized. It is assumed that the number of workers determines the number of households and is perfectly correlated with population in each city.

For simplicity, we assume homogeneous labor and that each worker provides one unit of labor. Hence, local labor supply is determined solely by the location decisions of workers. Following Glaeser and Gottlieb (2009), among others, we assume that the utility of workers in city i (U_i) is given by the Cobb-Douglas utility function

$$U_i = M_i C_{H,i}^{\gamma} C_{O,i}^{1-\gamma},\ 0 < \gamma < 1. \tag{A2}$$

In (2), M_i is the quality of amenities in city i, C_H and C_O represent the consumption of housing and other goods, respectively, and γ is the share of expenditure on housing, which is assumed to be similar over time and across cities.[18] Piazzesi et al. (2007), Davis and Ortalo-Magné (2011), and Piazzesi and Schneider (2016) provide support for this assumption, which is common in spatial equilibrium models. Similar to Glaeser and Gottlieb (2009) and Hsieh and Moretti (2015), the indirect utility (V_i) then equals[19]

[17] While spatial equilibrium models, such as ours, conventionally assume perfect labor mobility, the key predictions hold even in the case of frictions to mobility such as transaction costs and other relocation costs.

[18] Carlino and Saiz (2019) provide a review of literature supporting the role of amenities in households' location choices. Spatial variation in amenities can affect house values within cities as well (e.g., Letdin and Shim, 2019). However, our focus is solely on the variations of local amenities *across* cities.

[19] As is typical, we abstract from the constant term that is assumed to be the same across cities.

$$V_i = M_i W_i (P_i)^{-\gamma}, \tag{A3}$$

where W_i denotes the nominal wage level and P_i is the cost (or price) of housing in city i. In log form

$$v_i = m_i + w_i - \gamma p_i, \tag{A4}$$

where the lower-case letters denote natural logs. Utility is positively related to wage level and the quality of amenities, and negatively affected by higher housing costs. In spatial equilibrium, the utility levels are the same across cities; i.e., workers are indifferent between locations. Hence, in spatial equilibrium

$$w_i - \gamma p_i + m_i = w_j - \gamma p_j + m_j \tag{A5}$$

holds for every city i and j.

Given the utility function in (A2), the Marshallian demand for housing of a household located in i (D_i^{hh}) is

$$D_i^{hh} = \gamma W_i / P_i;\ d_i^{hh} = \ln\gamma + w_i - p_i. \tag{A6}$$

The market level demand in city i (d_i) then equals (in logs)

$$d_i = c_1 + w_i + n_i - p_i, \tag{A7}$$

where c_1 ($= \ln\gamma$) is a constant term.

Local housing supply (S_i), in turn, is provided by absentee landlords, and is positively related to the level of housing costs (which reflect the return on housing investment), with ω_i (> 0 for every i) denoting the price elasticity of housing supply :[20]

$$S_i = C_{2i} P_i^{\omega_i};\ s_i = c_{2i} + \omega_i p_i;\ \omega_i > 0. \tag{A8}$$

In the short and medium run, the elasticity of housing supply can vary depending on whether prices are decreasing or increasing (Glaeser and Gyourko, 2005). This model focuses on *long-term* trends in the price-income relationship, however. Moreover, the demand for housing, measured as real aggregate income, trended upwards during the sample period (1979–2018) in all the metropolitan areas included in our empirical analysis. Therefore, we do not distinguish between upwards and downwards adjustment of housing supply.

To keep the framework tractable, we assume that housing production does not involve the use of locally varying inputs. In equilibrium, housing supply equals housing demand; hence, the equilibrium price level is given by

$$p_i = \alpha_i + \beta_{1i} w_i + \beta_{2i} n_i;\ \alpha_i = \frac{c_1 - c_{2i}}{\omega_i + 1},\ \beta_{1i} = \beta_{2i} = \frac{1}{\omega_i + 1} > 0. \tag{A9}$$

Higher wages, greater population, and lower supply elasticity (smaller ω) due to topographic or regulatory constraints cause higher housing costs. If the number of households increases in a city but wages do not (i.e., population growth is induced by relative improvement in the quality of amenities), housing space per person must decrease in the city. The considerable spatial variation in ω_i (Saiz, 2010) is expected to yield notable variation in β_{1i} and β_{2i} across cities. For simplicity, we assume that the constant term (α) in the price equation is the same across cities.

We follow Moretti (2011) and Kline and Moretti (2014) by assuming that there are two cities, a and b. This allows us to keep the model simple while still being able to illustrate the key implications of the spatial equilibrium condition for the house price-income ratio. Given the spatial equilibrium condition, the inverse labor supply function in city a is:

$$w_a = w_b + \gamma(p_a - p_b) + (m_b - m_a). \tag{A10}$$

Using Eq. (A9) for p_a and p_b yields

$$w_a = [(1 - \gamma\beta_{1b}) w_b + \gamma(\beta_{2a} n_a - \beta_{2b} n_b) + (m_b - m_a)] / (1 - \gamma\beta_{1a}). \tag{A11}$$

A higher quality of amenities in city a induces larger local labor supply, i.e., a greater number of households and therefore higher house prices. In other words, the utility gain from higher amenities makes workers willing to live in a city even if their net wages after housing costs are lower. Given the upward sloping housing supply curve, the labor supply curve is also upward sloping: since greater n_i causes higher p_i, wages need to be higher to attract more workers in the city.

The total number of workers, N, is exogenous and divided between the two cities ($N = N_a + N_b$) so that the spatial equilibrium condition is fulfilled. The impact of a greater number of workers on local housing costs restricts city growth when wages increase (due to a positive productivity shock, for instance) or the quality of amenities improves relative to the other city.

Finally, the model is closed by the labor demand equation. We assume that firms are perfectly mobile and price takers, and labor is paid its marginal product. Hence, the (inverse) labor demand is[21]

$$W_i = h X_i N_i^{h-1} K_i^{1-h};\ w_i = x_i + (h-1) n_i + (1-h) k_i + \ln h. \tag{A12}$$

Labor market equilibrium is obtained by equating (A11) and (A12) for each city.

[20] Although supply elasticity could be endogenous to city size (S), it is conventional in spatial equilibrium models to assume that it is exogenous (Hsieh and Moretti, 2015; Kline and Moretti, 2014; Moretti, 2011). This assumption does not have any bearing on the conclusions we derive from the model and allowing supply elasticity to be endogenous would greatly diminish the model's tractability.

[21] It is assumed that there is an 'international' capital market where capital is infinitely supplied at a given price, so that firms in each city can rent as much capital as is optimal at this price.

Productivity shock and the house price-income relationship

Next, we use this standard spatial equilibrium model to consider the influence of a local labor demand shock on house price-income ratios. Following Moretti (2011), we assume that the two cities are identical initially, after which total factor productivity increases in city a due to a shock in the local productivity shifter. That is, there is a small shock in x_a, causing a wage increase $w_{a2} - w_{a1} = \Delta$ (> 0) in city a, where subscripts 1 and 2 indicate time periods before and after the shock, respectively, and Δ equals the productivity increase. Using (A11), we can write :[22]

$$\Delta = [\gamma\beta_{2a}(n_{a2} - n_{a1}) - \gamma\beta_{2b}(n_{b2} - n_{b1})]/(1-\gamma\beta_{1a}). \tag{A13}$$

Eq. (13) cannot readily be used to compute the effect of the shock on the number of workers in a, as the number of workers in b is dependent on that in a. To circumvent this complication, we utilize the fact that $N_b = N - N_a$. Assuming that the cities are identical before the shock, $n_{b2} - n_{b1} = -(n_{a2} - n_{a1})$. However, if $n_{a1} \neq n_{b1}$, for a *small* change in N_b (corresponding to a *small* change in x_a and thereby a *small* Δ): $n_{b2} - n_{b1} \approx -N_{a1}/(N - N_{a1}) \times (n_{a2} - n_{a1})$. Using this approximation to achieve greater generality, we get the population change in a due to the shock:

$$n_{a2} - n_{a1} \approx (1-\gamma\beta_{1a})/[\gamma(\beta_{2a} + \delta_{a1}\beta_{2b})] \times \Delta, \tag{A14}$$

where δ_{a1} ($= 1$ in the case of identical cities) is the initial number of workers located in city a relative to workers located in b $[N_{a1}/(N - N_{a1})]$. The growth of city a after the productivity shock is moderated by more inelastic housing supply in a (greater β_{1a} and β_{2a}) and in b (greater β_{2b}). The elasticity of housing supply in b affects the growth rate of a, because less elastic supply in city b leads to a greater drop in housing costs in the city as workers move to city a.[23] This greater housing cost decline yields greater growth in income net of housing costs in b, which lessens the movement of workers from b to a.

Taking advantage of Eqs. (A9) and (A10), the change in the house price-income ratio in city a due to the productivity shock is

$$d(p_a - w_a) = (1-\gamma)\beta_{2a}(n_{a2} - n_{a1}) + \gamma\beta_{2b}(n_{b2} - n_{b1}) + (1-\gamma)\beta_{1a}\Delta, \tag{A15}$$

where $d(p_a - w_a) = (p_{a2} - w_{a2}) - (p_{a1} - w_{a1})$. Using the above approximation for $n_{b2} - n_{b1}$ and Eq. (A14), we can express $d(p_a - w_a)$ solely in terms of model parameters and the productivity increase:

$$d(p_a - w_a) = [\theta(1-\gamma\beta_{1a}) + (1-\gamma)\beta_{1a}] \times \Delta, \tag{A16}$$

where $\theta = \frac{(1-\gamma)\beta_{2a} - \delta_{a1}\gamma\beta_{2b}}{\gamma(\beta_{2a}+\delta_{a1}\beta_{2b})} = \frac{\beta_{2a}}{\gamma(\beta_{2a}+\delta_{a1}\beta_{2b})} - 1$. Based on (A16), the price-income ratio remains constant after a productivity shock only in the special case where $\theta(1 - \gamma\beta_{1a}) + (1 - \gamma)\beta_{1a} = 0$.

Comparative static predictions can be formulated with the following partial derivatives:

$$\frac{\partial[d(p_a - w_a)]}{\partial\beta_{1a}} = \frac{\delta_{a1}\beta_{2b}}{\beta_{2a} + \delta_{a1}\beta_{2b}} > 0$$

$$\frac{\partial[d(p_a - w_a)]}{\partial\beta_{2a}} = (1-\gamma\beta_{1a})\frac{\delta_{a1}\beta_{2b}}{\gamma(\beta_{2a} + \delta_{a1}\beta_{2b})^2} > 0$$

$$\frac{\partial[d(p_a - w_a)]}{\partial\beta_{2b}} = -(1-\gamma\beta_{1a})\frac{\delta_{a1}\beta_{2a}}{\gamma(\beta_{2a} + \delta_{a1}\beta_{2b})^2} < 0$$

Since higher income and population elasticities of house prices (β_{1a} and β_{2a}, respectively) in city a yield greater $d(p_a - w_a)$ as productivity in a increases, and β_1 and β_2 are positively dependent on supply elasticity, the model predicts that productivity increases lead to higher price-income ratios and thus greater growth in the ratio in more supply restricted cities. The decreasing influence of greater population elasticity (i.e., smaller supply elasticity) in city b on $d(p_a - w_a)$ is that fewer people will move from b to a after the productivity shock, since house prices drop more in b in response to declining population. Because of the stronger downwards reaction of house prices in city b, fewer people move from b to a and house prices increase less in city a.

A Simple numerical illustration

Consider two cities that are identical initially: $N = 1$ so that N_a and N_b reflect the population shares in the cities; the supply elasticity is 1.5 in both cities[24] so that the elasticity of house prices with respect to income (β_1) and population (β_2) is 0.40; and the parameters in the production function are $X_{a1} = X_{b1} = 10$ and $h = 0.5$. Since $N_{a1} = N_{b1} = 0.5$, we get $W_a = W_b = 5$. Finally, we set the share of expenditure on housing (γ) to 0.25 based on the findings reported in Davis and Ortalo-Magné (2011) for U.S. cities.

Suppose there is a total factor productivity shock in city a so that w_a increases by 5%. As shown in Table A1 (column I), in the new equilibrium population is 22.5% greater and the house price level is 11% higher in city a than before the shock. Consequently, the house price-income ratio is 6% higher. In city b, in turn, the 22.5% population decline causes a 9% decrease in the price-income ratio; i.e., housing becomes more affordable.

To illustrate the role of supply elasticity, consider the same shock but with a lower supply elasticity (0.6) in city a.[25] Column II in Table A1 shows

[22] Note that the wage level in city b does not change, since the amount of capital used by firms in b offsets the effect of the change in n_b (Moretti, 2011).

[23] In the short and medium term, when housing supply tends to be more inelastic downwards than upwards, the price level in city b would drop even more and population would decrease (increase) somewhat less in b (a). This would not affect the key conclusions of our model and, in any case, we are interested in the long-term dynamics.

[24] The assumed supply elasticity is close to the median for the 70 metropolitan areas investigated in this study.

[25] The 0.6 supply elasticity is the lowest among the 70 metropolitan areas that we consider.

Table A1
Numerical illustrations.

Assumed supply elasticities and impacts of shocks	5% productivity increase in city *a*		5% productivity increase in both cities		5% increase in the value of amenities in city *a*	
City a	I	II	III	IV	V	VI
Assumed supply elasticity	1.500	0.600	1.500	0.600	1.500	0.600
Change in price-income ratio	0.060	0.084	–0.030	–0.026	0.080	0.122
Population change	0.225	0.165	0.000	–0.011	0.200	0.195
Price change	0.110	0.134	0.020	0.024	0.080	0.122
City b						
Assumed supply elasticity	1.500	1.500	1.500	1.500	1.500	1.500
Change in price-income ratio	–0.090	–0.066	–0.030	–0.026	–0.080	–0.078
Population change	–0.225	–0.165	0.000	0.011	–0.200	–0.195
Price change	–0.090	–0.066	0.020	0.024	–0.080	–0.078

Note: All examples are based on the following parameters before the shock: $N = 1$; $N_a = N_b = 0.5$; $\gamma = 0.25$; $X_a = X_b = 10$; $h = 0.5$. An exception is that, in the cases with 0.6 supply elasticity in city *a* (columns II, IV and VI), the productivity shifter in *a*, X_a, is initially 10.63.

that the price-income ratio increase in *a* is greater and the decrease in *b* is milder because of the more inelastic supply in *a*; housing is more expensive in *both cities* than in the baseline case. Given the assumption of equal city sizes before the shock, the lower supply elasticity in *a* also means that initially the productivity shifter in city *a*, X_a, is 10.63: higher wages are needed to compensate for the more expensive housing (which is an outcome of the supply inelasticity).

Now assume that a similar productivity shock takes place in both cities, with both w_a and w_b rising by 5%. As there is a similar wage increase in both cities, there is no flow of workers between *a* and *b* in the baseline case (column III). The income increase induces house price growth of 2%. Thus, the price-income ratio decreases by 3%. Column IV reports the effects assuming an elasticity of 0.6 in *a*: because $\beta_{1a} > \beta_{1b}$, some households need to move from *a* to *b* so that the spatial equilibrium condition is maintained. Due to the inelastic supply in *a*, housing costs increase more in *both cities* than in the baseline case.

Finally, suppose that, instead of a productivity shock, there is a positive shock in the value of amenities in city *a* (column V). This shock could take place due to a change in workers' preferences for various amenities (e.g., quality of public transportation or climate) or a change in the amenities themselves (e.g., better services, less crime, or cleaner environment). The wage levels in the two cities are unaltered as there is no change in productivity. Hence, the spatial equilibrium condition requires that some workers move from *b* to *a*, causing housing costs to adjust so that the equilibrium condition is maintained: the price level increases in *a* and decreases in *b* thereby causing a higher house price-income ratio in *a* and a lower ratio in *b*.[26] A lower supply elasticity in *a* (column VI) would yield greater price-income changes in *both cities* and less movement from *b* to *a*.

Appendix B. Panel data estimators and tests

Panel OLS (POLS)

In a panel dataset, y_{it} is the dependent variable (house price in our empirical analysis) and x_{it} is a vector of explanatory variables (one-dimensional vector including income), which are observed for cross-sections (MSAs) $i = 1, \ldots, N$ and time periods $t = 1, \ldots, T$. OLS estimator of slope (β) of equation

$$y_{it} = \alpha + \beta x_{it} + \varepsilon_{it} \tag{B1}$$

is given by

$$\widehat{\beta}_{OLS} = \left(\sum_{i=1}^{N} \sum_{t=1}^{T} (x_{it} - \bar{x}_i)^2 \right)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} (x_{it} - \bar{x}_i)(y_{it} - \bar{y}_i) \tag{B2}$$

where $\bar{x}_i$ and $\bar{y}_i$ refer to the MSA-specific means. If the slope coefficients are heterogeneous across cross-sections, the estimator is asymptotically biased and its asymptotic distribution will be dependent on nuisance parameters associated with dynamics of the underlying process. Only for the special case in which the regressors are homogenous across members of the panel can valid inferences be made from the standardized distribution of $\widehat{\beta}_{OLS}$ or its associated t-statistic (Pedroni, 2000).

Mean group estimator (MG)

Unlike the POLS estimator, the Mean Group Estimator (MG) of Pesaran et al. (1999) allows for heterogeneity of slope coefficients across cross-sections. One can estimate separate equations for each group (MSA) and examine the distribution of the estimated coefficients across groups. Of particular interest is the mean of the estimates, i.e., MG estimator, which is defined by means of individual model parameters over panel subjects. For example, intercept and slope estimates are here defined by

[26] Regulatory restrictiveness – and thus supply elasticity – could be correlated with amenities (Hilber and Robert-Nicoud, 2013): greater value of amenities can give rise to lower supply elasticity (through more regulation). This could add another channel from an amenity shock to the price-income ratios. Additionally, the quality of amenities may rise with income (Diamond, 2016), which could increase the influence of income growth on house prices. These potential effects do not alter the key conclusions of the model.

$$\widehat{\alpha}_{MG} = \bar{\alpha} = N^{-1}\sum_{i=1}^{N}\widehat{\alpha}_{i,OLS} \tag{B3}$$

$$\widehat{\beta}_{MG} = \bar{\beta} = N^{-1}\sum_{i=1}^{N}\widehat{\beta}_{i,OLS}. \tag{B4}$$

In earlier work, Pesaran and Smith (1995) showed that in heterogeneous panels the MG estimator will produce consistent estimates of the average of the parameters and thus is preferred over POLS.

The fully-modified panel OLS (FMOLS)

While the POLS and MG estimators are vulnerable to endogeneity bias, the fully-modified OLS (FMOLS) estimator of Pedroni (2000, 2001) that is aimed at cointegrating regressions is consistent in the presence of endogenous regressors and endogeneity due to possible omitted variables. The FMOLS estimators are also asymptotically unbiased and super-consistent in the presence of non-stationary but cointegrated data, which is not the case for the conventional OLS and MG estimators. (Pedroni, 2001, 2007.)

For Pooled Fully-modified Panel OLS estimator (PFMOLS), consider the following system for a panel of $i = 1, \ldots, N$ members:

$$y_{it} = \alpha + \beta x_{it} + \mu_{it} \tag{B5}$$

$$x_{it} = x_{it-1} + \varepsilon_{it}. \tag{B6}$$

The slope estimator is given as

$$\widehat{\beta}_{PFMOLS} = \sum_{i=1}^{N}\widehat{L}_{22i}^{-2}\left(\sum_{t=1}^{T}(x_{it}-\bar{x}_i)^2\right)^{-1}\sum_{i=i}^{N}\widehat{L}_{11i}^{-1}\widehat{L}_{22i}^{-1}\left(\sum_{t=1}^{T}(x_{it}-\bar{x}_i)y_{it}^{*} - T\gamma_i\right) \tag{B7}$$

where $\widehat{L}$ and γ_i refer to components of long-run variance-covariance matrix of μ_{it} and ε_{it} (see Pedroni, 2000, for further details). The dependent values y_{it} are transformed using variance-covariance matrix of $(\mu_{it},\ \varepsilon_{it})$ to y_{it}^{*}. FMOLS mean-group (FMOLS-MG) estimator is given by Pedroni (2001) as

$$\widehat{\beta}_{FMOLS-MG} = N^{-1}\sum_{i=1}^{N}\widehat{\beta}_{i,FMOLS} \tag{B8}$$

where $\widehat{\beta}_{i,FMOLS}$ are the individual (MSA-specific) FMOLS estimates. An advantage of the group mean estimator over the pooled panel FMOLS estimator is that by allowing heterogenous coefficients this estimator allows for a more flexible hypothesis testing in the presence of heterogeneity in cointegrating vectors. This is because the group mean estimator is based on the so called 'between dimension' of the panel, while the pooled estimators are based on the 'within dimension' of the panel.

An additional advantage of the between-dimension estimators is the more useful interpretation of the point estimates in the event that the true cointegrating vectors (slope coefficients) are heterogeneous: Point estimates for the between-dimension estimator can be interpreted as the mean value for the cointegrating vectors, which is not the case for the within-dimension estimators. (Pedroni, 2001.) Furthermore, as Pesaran and Smith (1995) argue in the context of OLS regressions, when the true slope coefficients are heterogeneous, group mean estimators provide consistent point estimates of the sample mean of the heterogeneous cointegrating vectors, while pooled within dimension estimators do not.

Common correlated effects estimator (CCEMG)

The Common Correlated Effects Estimator (CCE) of Pesaran (2006) includes the cross-sectional means of y_{it} and x_{it} as additional regressors to account for cross-section dependence in the data. The CCE method is robust to different types of error cross-sectional dependence.

The CCEMG estimator is, similar to MG estimator, a simple average of the individual estimates of the model. Estimates are given by

$$y_{it} = \widehat{\alpha}_{CCEMG} + \widehat{\beta}_{CCEMG}x_{it} + \widehat{\delta}_i\bar{z}_t + \varepsilon_{it} \tag{B9}$$

$$\widehat{\beta}_{CCEMG} = N^{-1}\sum_{i=1}^{N}\widehat{\beta}_{i,CCEMG} \tag{B10}$$

where $\bar{z}$ refers to cross-sectional averages of individual-specific variables (both y and x) and $\widehat{\delta}_i$ to a vector of coefficients on $\bar{z}$. The basic idea is to filter the individual-specific regressors by means of cross-section aggregates so that asymptotically the differential effects of unobserved common factors are eliminated. The estimation can be conducted with OLS. (Pesaran, 2006.)

The Dynamic Common Correlated Effects estimator of Chudik and Pesaran (2015) additionally includes lags of $\bar{z}$ and therefore performs well even in the case of dynamic models with weakly exogenous regressors.

Panel unit root / cointegration testing: the CIPS test

Pesaran (2007) proposes a CIPS panel unit root test in the presence of cross-section dependence. The test is based on individual augmented Dickey-Fuller (ADF) regressions augmented with cross-sectional means of lagged levels and first-differences of the series y_{it}. The cross-sectionally augmented ADF (CADF) regression thus is:

$$\Delta y_{it} = \widehat{\alpha}_i + \widehat{\varnothing}_i y_{it-1} + \widehat{\rho}_i \bar{y}_{t-1} + \widehat{\theta}_i \bar{\Delta} y_t + \widehat{\kappa}_i \Delta y_{it-1} + \widehat{\psi}_i \bar{\Delta} y_{t-1} + \varepsilon_{it} \tag{B11}$$

This regression can be applied to individual time series in the usual manner, t-value for coefficient $\widehat{\varnothing}_i$ being the CADF test statistic. Specification (B11) includes one lag of the dependent variable $\Delta\bar{y}_t$, but there can be more lags and the number of lags can vary across cross-sections.

The cross-sectionally augmented IPS panel unit root test then is based on the CIPS statistic:

$$CIPS(N,T) = N^{-1} \sum_{i=1}^{N} t_i(N,T), \tag{B12}$$

where $t_i(N,T)$ is the t-ratio of the coefficient of $\widehat{\varnothing}_i$ in the CADF regression. Pesaran (2007) provides critical values for both CADF and CIPS test statistic under various cases.

The null hypothesis (H_0) and the alternative hypothesis (H_A) in the CIPS test are:

H_0: Each of the individual-specific series is non-stationary (when testing for cointegration: none of the individual-specific equations is cointegrated, i.e., each of the residual series is non-stationary).

H_A: One or more individual-specific series are stationary (when testing for cointegration: one or more individual-specific equations are cointegrated, i.e., one or more individual-specific residual series are stationary).

Tests of the null hypothesis of homogeneity

The null hypothesis of homogeneity of slope coefficients across cross-sections in FMOLS-MG can be tested with F-test (Pedroni, 2007). The F-test is based on the residuals of the individual and group-mean FMOLS estimated regressions. Specifically, the test is constructed for the restrictions implied by the case in which the slope coefficients are assumed to be common across the cross-sections. The corresponding Wald statistic compares the sum of squared errors across all periods and cross-sections for the restricted case when $\beta_i = \beta$ for all i versus the case with unrestricted heterogeneous β values.

Swamy (1970) bases his test of slope homogeneity on the dispersion of individual slope estimates from a suitable pooled estimator. Similar to the F-test, the null hypothesis in Swamy's test is that of homogeneous slope coefficients. The test statistic S can be written as

$$S = \sum_{i=1}^{N} (\widehat{\beta}_i - \widehat{\beta}_{WFE})' \frac{X_i' M_i X_i}{\sigma_i^2} (\widehat{\beta}_i - \widehat{\beta}_{WFE}) \tag{B13}$$

where

$$\sigma_i^2 = \frac{(y_i - X_i\widehat{\beta}_i)' M_t (y_i - X_i\widehat{\beta}_i)}{T-k-1}. \tag{B14}$$

In (13), X refers to vector of exogenous regressors, M is a construction matrix of the panel data, σ_i^2 denotes error variance, k is the number of regressors, and $\widehat{\beta}_{WFE}$ is the vector of weighted Fixed-Effects (WFE) pooled estimators of slope coefficients.

Pesaran and Yamagata (2008) propose a standardized version of Swamy's test. The test statistic is given by

$$S_{PY} = \sqrt{N}\left(\frac{N^{-1}S - k}{\sqrt{2k}}\right). \tag{B15}$$

S_{PY} approaches standard normal distribution under certain assumptions.

References

Anenberg, E., 2016. Information frictions and housing market dynamics. Int. Econ. Rev. (Philadelphia) 57 (4), 1449–1479.

Atalay, K., Edwards, R., 2022. House prices, housing wealth and financial well-being. J. Urban Econ. 129, 103438.

Berger, D., Guerrieri, V., Lorenzoni, Vavra, J., 2018. House prices and consumer spending. Rev. Econ. Stud. 85 (3), 1502–1542.

Bourassa, S.C., Hoesli, M., Oikarinen, E., 2019. Measuring house price bubbles. Real Estate Econ. 47 (2), 534–563.

Broxterman, D.A., Kuang, C., 2019. A revealed preference index of urban amenities: using travel demand as a proxy. J. Reg. Sci. 59 (3), 508–537.

Carlino, G.A., Saiz, A., 2019. Beautiful city: leisure amenities and urban growth. J. Reg. Sci. 59 (3), 369–408.

Case, K., Shiller, R., 2003. Is there a bubble in the housing market? Brookings Pap Econ Act 2, 299–342.

Chudik, A., Pesaran, M.H., 2015. Common correlated effects estimation of heterogeneous dynamic panel data models with weakly exogenous regressors. J. Econom. 188 (2), 393–420.

Cooper, D., 2013. House price fluctuations: the role of housing wealth as borrowing collateral. Rev. Econ. Statistics 95 (4), 1183–1197.

Davis, M., Ortalo-Magné, F., 2011. Household expenditures, wages, rents. Rev. Econ. Dyn. 14 (2), 248–261.

Diamond, R., 2016. The determinants and welfare implications of US workers' diverging location choices by skill: 1980-2000. Am. Econ. Rev. 106 (3), 479–524.

Engle, R.F., Granger, C.W.J., 1987. Co-integration and error-correction: representation, estimation and testing. Econometrica 55 (2), 251–276.

Etheridge, B., 2019. House prices and consumption inequality. Int. Econ. Rev. (Philadelphia) 60 (4), 1781–1822.

Gallin, J., 2006. The long-run relationship between house prices and income: evidence from local housing markets. Real Estate Econ. 34 (3), 417–438.

Gan, Q., Hill, R.J., 2009. Measuring housing affordability: looking beyond the median. J. Hous. Econ. 18 (2), 115–125.

Glaeser, E., Gyourko, J., 2018. The economic implications of housing supply. J. Econ. Perspectives 32 (1), 3–30.

Glaeser, E.L., Gottlieb, J.D., 2009. The wealth of cities: agglomeration economies and spatial equilibrium in the United States. J. Econ. Lit. 47 (4), 983–1028.

Glaeser, E.L., Gyourko, J., 2005. Urban decline and durable housing. J. Political Econ. 115 (2), 345–375.

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

Gregoriou, A., Kontonikas, A., Montagnoli, A., 2014. Aggregate and regional house price to earnings ratio dynamics in the UK. Urban Stud. 51 (13), 2916–2927.

Gueye, G.N., 2021. Pitfalls in the cointegration analysis of housing prices with the macroeconomy: evidence from OECD countries. J. Hous. Econ. 51, 101748.

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

Hilber, C.A.L., Robert-Nicoud, F., 2013. On the origins of land use regulations: theory and evidence from US metro areas. J. Urban Econ. 75, 29–43.

Himmelberg, C., Mayer, C., Sinai, T., 2005. Assessing high house prices: bubbles, fundamentals and misperceptions. J. Econ. Perspectives 19 (4), 67–92.

Holly, S., Pesaran, M.H., Yamagata, T., 2010. A spatio-temporal model of house prices in the USA. J. Econom. 158 (1), 160–173.
Hsieh C.-.T. and Moretti E. (2015) Why do cities matter? Local growth and aggregate growth. NBER Working Paper No. 21154.
Karpestam, P., Johansson, S., 2019. Interest-only-mortgages and housing market fluctuations in Denmark. J. Hous. Econ. 46, 101627.
Kline, P., Moretti, E., 2014. People, places, and public policy: some simple welfare economics of local economic development programs. Annu. Rev. Econom. 6, 629–662.
Knoll, K., Schularik, M., Steger, T., 2017. No price like home: global house prices, 1870-2012. Am. Econ. Rev. 107 (2), 331–353.
Letdin, M., Shim, H.S., 2019. Location choice, life cycle and amenities. J. Reg. Sci. 59 (3), 567–585.
Leung, C.K.Y., 2014. Error correction dynamics of house prices: an equilibrium benchmark. J. Hous. Econ. 25, 75–95.
Ma, C., 2020. Momentum and reversion to fundamentals: are they captured by subjective expectations of house prices? J. Hous. Econ. 49, 101687.
Malpezzi, S., 1999. A simple error correction model of house prices. J. Hous. Econ. 8 (1), 27–62.
Miller, C.C., 2014. Will Portland always Be a Retirement Community For the young? New York Times. September 21. https://www.nytimes.com/2014/09/21/magazine/will-portland-always-be-a-retirement-community-for-the-young.html?emc=eta1&_r=0.
Moretti E. (2011) Local labor markets. In: D Card and O Ashenfelter (Eds.), *Handbook of Labor Economics*, Vol. 4B Amsterdam: Elsevier, pp. 1237–1313.
Oikarinen, E., 2009. Household borrowing and metropolitan housing price dynamics – Empirical evidence from Helsinki. J. Hous. Econ. 18 (2), 126–139.
Oikarinen, E., Bourassa, S.C., Hoesli, M., Engblom, J., 2018. U.S. metropolitan house price dynamics. J. Urban Econ. 105, 54–69.
Oikarinen, E., Engblom, J., 2016. Differences in housing price dynamics across cities: a comparison of different panel model specifications. Urban Studies 53 (11), 2312–2329.
Pedroni, P., 2000. Fully modified OLS for heterogeneous cointegrated panels. In: Baltagi, B.H. (Ed.), Nonstationary Panels, Cointegration in Panels and Dynamic panels. Elsevier, Amsterdam, pp. 93–130.
Pedroni, P., 2001. Purchasing power parity tests in cointegrated panels. Rev. Econ. Statistics 83 (4), 727–731.
Pedroni, P., 2007. Social capital, barriers to production and capital shares: implications for the importance of parameter heterogeneity from a nonstationary panel approach. J. Appl. Econom. 22 (2), 429–451.
Pesaran, M.H., 2006. Estimation and inference in large heterogeneous panels with a multifactor error structure. Econometrica 74 (4), 967–1012.
Pesaran, M.H., 2007. A simple panel unit root test in the presence of cross-section dependence. J. Appl. Econom. 22 (2), 265–312.
Pesaran, M.H., 2012. On the interpretation of panel unit root tests. Econ. Lett. 116 (3), 545–546.
Pesaran, M.H., Shin, Y., Smith, R.P., 1999. Pooled mean group estimation of dynamic heterogeneous panels. J. Am. Stat. Assoc. 94 (446), 621–634.
Pesaran, M.H., Smith, R.P., 1995. Estimating long-run relationships from dynamic heterogeneous panels. J. Econom. 68 (1), 79–113.
Pesaran, M.H., Yamagata, T., 2008. Testing slope homogeneity in large panels. J. Econom. 142 (1), 50–93.
Piazzesi M. and Schneider M. (2016) Housing and macroeconomics. In: JB Taylor and H Uhlig (Eds.), *Handbook of Macroeconomics*, Vol. 2B Amsterdam: Elsevier, pp. 1547–1640.
Piazzesi, M., Schneider, M., Tuzel, S., 2007. Housing, consumption and asset pricing. J. Financ. Econ. 83 (3), 531–569.
Piketty, T., Zucman, G., 2014. Capital is back: wealth-income ratios in rich countries 1700-2010. Q. J. Econ. 129 (3), 1255–1310.
Roback, J., 1982. Wages, rents, and the quality of life. J. Political Econ. 90 (6), 1257–1278.
Rosen, S., 1979. Wage-based indexes of urban quality of life. In: Miezkowski, PN, Straszheim, MR (Eds.), Current Issues in Urban Economics. Johns Hopkins University Press, Baltimore (MD), pp. 74–104.
Saiz, A., 2010. The geographic determinants of housing supply. Q. J. Econ. 125 (3), 1253–1296.
Swamy, P.A.V.B., 1970. Efficient inference in a random coefficient regression model. Econometrica 38 (2), 311–323.
Vogiazas, S., Alexiou, C., 2017. Determinants of housing prices and bubble detection: evidence from seven advanced economies. Atlantic Econ. J. 45 (1), 119–131.