Journal of Econometrics 158 (2010) 160–173

Contents lists available at ScienceDirect

Journal of Econometrics

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

A spatio-temporal model of house prices in the USA[☆]

Sean Holly[a], M. Hashem Pesaran[a,b,*], Takashi Yamagata[c]

[a] Faculty of Economics and CIMF, University of Cambridge, United Kingdom

[b] University of Southern California, United States

[c] Department of Economics and Related Studies, University of York, United Kingdom

ARTICLE INFO

Article history:
Available online 17 March 2010

JEL classification:
C21
C23

Keywords:
House prices
Cointegration
Cross-sectional dependence
Spatial dependence

ABSTRACT

This paper provides an empirical analysis of changes in real house prices in the USA using State level data. It examines the extent to which real house prices at the State level are driven by fundamentals such as real per capita disposable income, as well as by common shocks, and determines the speed of adjustment of real house prices to macroeconomic and local disturbances. We take explicit account of both cross-sectional dependence and heterogeneity.This allows us to find a cointegrating relationship between real house prices and real per capita incomes with coefficients $(1, -1)$, as predicted by the theory. We are also able to identify a significant negative effect for a net borrowing cost variable, and a significant positive effect for the State level population growth on changes in real house prices. Using this model we then examine the role of spatial factors, in particular, the effect of contiguous states by use of a weighting matrix. We are able to identify a significant spatial effect, even after controlling for State specific real incomes, and allowing for a number of unobserved common factors. We do, however, find evidence of departures from long run equilibrium in the housing markets in a number of States notably California, New York, Massachusetts, and to a lesser extent Connecticut, Rhode Island, Oregon and Washington State.

1. Introduction

Recent developments in the housing markets in the USA and elsewhere have once again highlighted the importance of large changes in house prices for the functioning of credit and money markets (International Monetary Fund, 2004). Changes in housing wealth also play an important role in household behaviour with real implications for output and employment. There is also the possibility of bubbles in house prices with prices moving well away from their fundamental drivers, such as household disposable income (Case and Shiller, 2003; McCarthy and Peach, 2004). This in turn raises the issue of whether there is cointegration between real house prices and real per capita disposable incomes. The evidence on this is mixed.

Using US national-level data, Meen (2002) and Gallin (2006) do not find strong evidence of a cointegrating relationship, possibly because of the short time span of the data they consider. To cope with this problem, other studies use panel data. Malpezzi (1999) uses panel data on 133 metropolitan areas in the USA over 18 years from 1979 to 1996 and he is able to reject the null of a unit root in the residuals of the regressions of real house prices on real per capita incomes, using the panel unit root test of Levin et al. (2002, LLC). However, the LLC test does not take account of possible cross-sectional dependence of house prices and this could bias the test results. Capozza et al. (2002) recognize this problem and try to control for cross-sectional dependence by adding time dummies to their error correction specifications. However, as Gallin (2006) points out, local housing market shocks are likely to be correlated in ways that are not captured by simple time effects. To allow for more general error cross-sectional dependence, Gallin (2006) adopts a bootstrap version of Pedroni's 1999 residual-based cointegration test procedure, originally advanced in Maddala and Wu (1999), but fails to reject the hypothesis of no cointegration. However, his bootstrap approach is likely to be biased when the cross section dimension (N) is much larger than the time series dimension (T), as in Gallin's application.

In this paper, using recently developed econometric techniques for the analysis of heterogeneous dynamic panels subject to cross-sectional dependence, we study the determination of real house prices in a panel of 49 US States over 29 years. We examine the extent to which real house prices at the State level are driven by fundamentals such as real per capita disposable income, as well as by common shocks, and determine the speed of adjustment of real house prices to macroeconomic and local disturbances.

[☆] We would like to thank Jörg Breitung, Peter Burridge, Bernie Fingleton, David Hendry, Andreas Pick, Ron Smith, Elisa Tosetti and Joakim Westerlund for helpful comments. We have also benefited from constructive comments and suggestions from two anonymous referees and the editors. Financial support from the ESRC under Grant No. RES-000-23-0135 is gratefully acknowledged.

* Corresponding address: Faculty of Economics, University of Cambridge, Cambridge CB3 9DD, United Kingdom.

E-mail address: mhp1@econ.cam.ac.uk (M.H. Pesaran).

0304-4076/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.jeconom.2010.03.040

There are considerable differences across US States in both the level and rates of growth of real incomes.[1] This heterogeneity should in turn be reflected in real house prices. The importance of heterogeneity in spatially distributed housing markets has been highlighted by Fratantoni and Schuh (2003). They quantify the importance of spatial heterogeneity in US housing markets for the efficacy of monetary policy. Depending on local conditions monetary policy can have differing effects on particular US regions (Carlino and DeFina, 1998). However, there are significant dependences in house prices and real incomes across States that cannot be captured by spatial effects alone. For example, Pollakowski and Ray (1997), using vector autoregressive (VAR) models, show that at the national level (dividing the USA into nine regions) there are significant non-spatial diffusion patterns.

Real house prices can vary between States because real incomes differ, but they can also differ because of scarcity of land or other idiosyncratic factors. The effects of common shocks on house prices whether observed, such as changes in interest rates and oil prices, or unobserved, such as technological change, could also differ across States. We take account of these influences by making use of the common correlated effects (CCE) estimator of Pesaran (2006) which is consistent under heterogeneity and cross-sectional dependence. The CCE estimator can be computed by ordinary least squares (OLS), applied to an auxiliary regression where the observed regressors are augmented by (weighted) cross-sectional averages of the dependent variable and the individual specific regressors. Notably, CCE estimation allows for unobserved common factors to be possibly correlated with exogenously given State-specific regressors and it is invariant to the (unknown but fixed) number of unobserved common factors as N and T tend to infinity (jointly). These features are not shared with other approaches in Banerjee and Carrion-i-Silvestre (2006), Groen and Kleibergen (2003), Nelson et al. (2005), Westerlund (2005), Pedroni and Vogelsang (2005), Chang (2005), and Bai and Kao (2006).

The CCE procedure also copes with the presence of spatial effects (Pesaran and Tosetti, 2010). This is because spatial dependence is dominated by the common factor error structure that underlies the CCE estimators. But once the model parameters are estimated, the importance of spatial effects can be ascertained by fitting a spatial model to the residuals from the panel. It is also worth noting that the CCE procedure is robust to the choice of the spatial model, although for estimation of spatial effects some parametric formulation would be needed. In this paper we estimate a spatial autoregressive (SAR) error model and show that spatial effects are indeed statistically highly significant in the analysis of house prices in the USA.

Finally, to test for cointegration between real house prices and real disposable incomes, we apply the panel unit root tests of Moon and Perron (2004) and Pesaran (2007) to the log price–income ratio which allow for cross-sectional dependence.

The remainder of the paper is organized as follows. Section 2 discusses the theory underlying house price determination. Section 3 provides a review of the panel data model and estimation methods. Section 4 provides a preliminary data analysis. Section 5 reports the estimation results. Section 6 provides some concluding remarks.

2. Modelling house prices

It is now standard to see the determination of house prices as the outcome of a market for the services of the housing stock and as an asset. Demand for housing can be met either through rental of a residential property or by owner occupation. The expected net benefit from owner occupation needs to be set against the rental cost of the same property. Denote the real house price at the start of period t by P_t, and the real rental cost of the same house over the period t by R_t. Then the net benefit of owning the house over the period t to $t+1$ is given by $P_{t+1} - P_t(1+r_t) + S_t$, where $r_t > 0$ is the real rate of interest, and S_t is the real value of housing services. This expression abstracts from transaction costs, depreciation, and other costs of home ownership. These can be readily incorporated into the analysis without affecting the long run relationship between real house prices and incomes that is the focus of our empirical analysis.

For a risk neutral household the one period arbitrage condition for the asset market equilibrium in real house prices is given by[2]

$$E\left(P_{t+1}|\mathcal{F}_t\right) - P_t(1+r_t) + S_t = R_t,$$

or

$$P_t = \left(\frac{1}{1+r_t}\right)\left[E\left(P_{t+1}|\mathcal{F}_t\right) + S_t - R_t\right],$$

where $\mathcal{F}_t$ is the information set available at time t.[3] To complete the model we shall assume that R_t cannot exceed household's real disposable income, Y_t, and represent this relationship by

$$R_t = \alpha_t Y_t, \quad 0 < \alpha_t < 1,$$

where α_t is assumed to follow a stationary process. We shall also assume that

$$S_t = \beta_t^{-1} R_t, \quad 0 < \beta_t < 1,$$

which ensures positive real house prices in all periods.[4] Under these assumptions

$$P_t = \left(\frac{1}{1+r_t}\right)\left[E\left(P_{t+1}|\mathcal{F}_t\right) + \theta_t Y_t\right], \tag{2.1}$$

where

$$\theta_t = \frac{\alpha_t(1-\beta_t)}{\beta_t} > 0.$$

It is also reasonable to assume that θ_t, the fraction of income allocated to net housing services, $S_t - R_t$, is stationary.

Accordingly, under rational expectations and assuming that r_t is sufficiently large relative to the growth of real disposable income, $g_t = \Delta \ln(Y_t)$, bubble-free real house prices will be given as the discounted stream of future net housing services, $S_t - R_t = \theta_t Y_t$. The solution simplifies considerably under $r_t = r$,

$$P_t = \sum_{j=0}^{\infty}\left(\frac{1}{1+r}\right)^j E\left(\theta_{t+j} Y_{t+j}|\mathcal{F}_t\right),$$

which can be written equivalently as

$$\frac{P_t}{Y_t} = \sum_{j=0}^{\infty} E\left(\theta_{t+j}\prod_{s=1}^{j}\left(\frac{1+g_{t+s}}{1+r}\right)\middle|\,\mathcal{F}_t\right).$$

Therefore, under fairly general assumptions regarding the processes generating g_t and θ_t, the price–income ratio, P_t/Y_t, would also be stationary. In particular, $p_t = \ln(P_t)$ will be cointegrated with $y_t = \ln(Y_t)$ with the cointegrating vector given by $(1, -1)$, if y_t is an integrated variable of order 1. For example, if θ_t and g_t are

[1] For a recent review of the USA housing market see Green and Malpezzi (2003).

[2] Feldstein et al. (1978), Hendershott and Hu (1981) and Buckley and Ermisch (1982).

[3] Alternatively, expectations can be taken under the risk-neutral measure which does not necessarily require households to be individually risk neutral, although it does imply that at the aggregate households are treated as if they are risk neutral.

[4] The condition $0 < \beta_t < 1$ is sufficient but not necessary for $P_t > 0$ for all t.