HOUSING DYNAMICS

by

Edward L. Glaeser
Harvard University and NBER

and

Joseph Gyourko*
University of Pennsylvania and NBER

December 15, 2006, Preliminary Draft

Abstract

The key stylized facts of the housing market are positive serial correlation of price changes at one year frequencies and mean reversion over longer periods, strong persistence in construction, and highly volatile prices and construction levels within markets. We calibrate a dynamic model of housing in the spatial equilibrium tradition of Rosen and Roback to see whether such a model can generate these facts. With reasonable parameter values, this model readily explains the mean reversion of prices over five year periods, but cannot explain the observed positive serial correlation at higher frequencies. The model predicts the positive serial correlation of new construction that we see in the data and the volatility of both prices and quantities in the typical market, but not the volatility of the nation's more extreme markets. The strong serial correlation in annual house price changes and the high volatility of prices in coastal markets are the two biggest housing market puzzles. More research is needed to determine whether measurement error-related data smoothing or market inefficiency can best account for the persistence of high frequency price changes. The best rational explanations of the volatility in high cost markets are shocks to interest rates and unobserved income shocks.

* Glaeser thanks the Taubman Center for State and Local Government at Harvard University, and Gyourko thanks the Research Sponsors Program of the Zell/Lurie Real Estate Center at The Wharton School, University of Pennsylvania for financial support. We appreciate the comments of seminar participants at the University of Chicago, the NBER Summer Institute, and the Federal Home Loan Mortgage Corporation on previous versions of the paper. Graham Elliot and James Stock provided helpful guidance. Andy Moore, Charles Nathanson, and Jon Steinnsen provided superb research assistance.

I. Introduction

Housing constitutes nearly two-thirds of the typical household's portfolio, and more than $18 trillion worth of real estate is owned within the household sector.[1] Despite the enormous size of this sector, economists' understanding of many features of the housing market remains incomplete.[2] For example, in the sample of 115 metropolitan areas from 1980 to 2005 for which we have Office of Federal Housing Enterprise Oversight (OFHEO) constant quality house price series, a $1 increase in real house prices in one year is associated with a 71 cent increase the next year. A $1 increase in local market prices over the past five years is associated with a 32 cent decrease over the next five year period. This predictability of price changes seems to pose a challenge for an efficient markets view (Case and Shiller, 1989; Cutler, Poterba, and Summers, 1991).

The large amount of inter-temporal volatility in prices within markets is also puzzling. The standard deviation of three-year real changes in our sample of metropolitan area average house prices is $26,354 (in 2000 dollars throughout the paper), which is about one-fifth of the median price level. Over one, three, and five year periods, the standard deviation of house price changes is at least three times the mean price change. Can this volatility be the result of real shocks to housing market or must it reflect bubbles and animal spirits?

---

[1] The portfolio share is from Tracy, Schneider, and Chan (1999). The dollar value figure is for the fourth quarter of 2005 and is from Table B.100 Balance Sheet of Households and Nonprofit Organizations which may be downloaded at http://www.federalreserve.gov/RELEASES/z1/Current/data.htm. The Federal Reserve's data includes market value estimates for second homes, vacant homes for sale, and vacant land owned by the household sector.

[2] The debate over whether the recent boom was a bubble is only the latest example. See McCarthy and Peach (2004), Himmelberg, Mayer and Sinai (2005), and Smith and Smith (2006) for recent analyses that conclude there is no large-scale bubble in housing prices. Shiller (2005, 2006) and Baker (2006) argue to the contrary that the bubble is both real and very large.

Another more subtle puzzle is that house price appreciation in the 1990s was negatively correlated with that in the 1980s (as shown in Figure 1), while housing unit growth was positively serially correlated over the same time periods (see Figure 2). Basic demand-driven housing models predict that prices and quantities should move symmetrically. The mismatch of quantity and price movements seems to suggest that models of housing prices need to more firmly embed supply as well as demand.

Many housing models also put great stock in macroeconomic variables such as interest rates and national income, but most variation in housing price changes is local, not national. Less than eight percent of the variation in price levels and barely more than one-quarter of the variation in price changes across cities can be accounted for by national year-specific fixed effects. The large amount of local variation and its relationship with macroeconomic variables is another challenge for a consistent economic explanation of housing market dynamics.

In this paper, we present a dynamic, rational expectations model of house price formation to see whether such a framework can explain the salient moments of housing price and quantity changes. The model follows the urban tradition of Alonso (1962), Rosen (1979) and Roback (1982) in which housing prices reflect the willingness to pay for one location versus another. In this approach, housing prices are determined endogenously by local wages and amenities, so that local heterogeneity is natural. Our model then extends the Alonso-Rosen-Roback framework by focusing on high frequency price dynamics and by incorporating endogenous housing supply.

In Section II of this paper, we present the model and four propositions regarding its implications. The model shows that the predictability of housing price changes is compatible with a no-arbitrage rational expectations equilibrium. Slow construction responses and mean reverting wage shocks imply that prices will mean revert. And, positive serial correlation of labor demand shocks at high frequencies can generate positive serial correlation of housing prices.

The model can also explain the apparent puzzle of mean reverting prices and persistent quantity changes shown in Figures 1 and 2. Proposition 4 shows that long-term trends to city productivity or local amenities will create persistence in population and housing supply changes, but will have a much smaller impact on prices, since those trends are anticipated and incorporated into initial prices. Price changes are driven by unexpected high frequency shocks, which mean revert, while quantity changes are driven by anticipated low frequency trends that persist.

The model also serves as the basis for the calibrations discussed in Sections III and IV of the paper. Section III presents our estimates of the model's key parameters: the real rate of interest, the degree to which construction responds to higher prices, and the variance and serial correlation of local demand shocks. We assume constant interest rates for most of the paper, but do turn to time-varying interest rates in Section V. We estimate construction cost parameters using data on construction and price variance. The literature on housing demand provides our estimates of the heterogeneity in preferences for particular locales. And, we use Bureau of Economic Analysis (BEA) income data to infer the time series properties of local income shocks.

In Section IV, we compare the moments of the real data with the moments predicted by the model based on the parameter estimates from Section III. We first investigate the serial correlation properties of prices and quantities. The parameter values described in Section III predict that housing prices will mean revert over five year periods at almost exactly the same rate that we see in the data. This mean reversion is the result of new construction satisfying demand and the observed mean reversion of economic shocks to local productivity. We fit the modest mean reversion of construction quantities less perfectly, but the patterns in the real data are quite compatible with reasonable parameter values.

Over one year periods, we predict strong serial correlation of new construction, but in the data serial correlation of new permits is even greater than the level that our model predicts. The model does not predict the strong serial correlation of price changes at one and three year intervals. This serial correlation could be due to the artificial smoothing of the underlying data or less rational factors. Persistence itself is not enough to reject a rational expectations model, but the mismatch between data and model at annual frequencies indicates that Case and Shiller's (1989) conclusion regarding inefficiency could be right. Future work needs to deal with the data smoothing problem to see whether the actual serial correlation still is far too high relative to the model.

Reasonable parameter values predict variances of new construction and price changes that are quite close to the variances seen in the median metropolitan area in our sample. We do overestimate the volatility of price changes at annual frequencies, but that could be the result of data smoothing. The model does not predict 'too much' variation for three and five year changes, where smoothing should be less of an issue.

While the model can fit the median market, it cannot explain the volatility of either prices or construction in the nation's more extreme markets. The model does not fit the price volatility in California which has huge price changes and it does not fit the construction volatility in sunbelt cities such as Atlanta and Houston. The high construction volatility in the sunbelt areas is most plausibly the result of lumpiness in the construction process and the possibility that the sunbelt areas have much lower construction cost parameters than we estimate. However, these explanations cannot help us explain the high price volatility areas, and we try to understand them in the penultimate section of the paper.

We consider three added sources of volatility: amenity fluctuation, unmeasured income volatility, and volatile interest rates. The one high frequency amenity variable that we have—crime rates—shows little ability to increase predicted demand and price variability. Using data from New York City, we examine whether the volatility of incomes for recent home buyers is higher than the volatility for average income, and find that it is. The variance of income in areas with big price change areas is higher than the variance of incomes for the average market. These factors may explain the high variation of prices in the most volatile markets on the east coast, but do little to help us understand California outside of the bay region, which has less income volatility.

Volatile interest rates will not increase the volatility of prices or construction in markets with prices close to construction costs (or to the national median price in our model), but they can increase the predicted variance for places with permanently high amenities or productivity. For interest rates to generate high levels of volatility, shocks to interest rates must be extremely high and areas must be innately extremely attractive, but these conditions may be true for California over the last two decades.

II. A Dynamic Model of Housing Prices

Our dynamic model of housing prices is based on three equilibrium conditions. Following Rosen (1979) and Roback (1982) we require consumers to be indifferent across space at all points in time, which requires utility U(W, A, R) to be equal across space, where W refers to wages, A to amenities, and R to the flow cost of housing. Our simplifying assumption that this spatial equilibrium must hold in all periods is the housing equivalent of assuming no financial transaction costs (as in Hansen and Jagannathan, 1991). Our second equilibrium condition is in the housing markets: we require the expected returns from making a house (its expected price) to be equal to the cost of construction. If the city is not growing, this equilibrium condition need not hold (as in Glaeser and Gyourko, 2005), but we make the simplifying assumption that the city is always adding new units. Our final equilibrium condition concerns wages which must equal the marginal product of labor to firms in the city.

We implement the spatial equilibrium condition by assuming that there is a "reservation locale" that delivers utility of $\underline{U}(t)$ in each period "t" and that the cost of building a home there always equals "C," which reflects the physical costs of construction. Since housing can be built in the reservation locale freely at cost C, we assume that the price of a house there always equals C.[3] The reservation locale represents the many metropolitan areas in the American hinterland with steady growth and where prices stay close to the physical costs of construction (Glaeser, Gyourko and Saks, 2005).[4] The annual cost of living in the reservation locale equals the difference

[3] While it is possible that prices will deviate around this value because of temporary over- or under-building, we simplify and assume that the price of a house always equals C.
[4] Van Neiuwerburgh and Weill (2006) present a similar model in their exploration of long run changes in the distribution of income (also studied by Gyourko, Mayer and Sinai, 2006). Our paper was produced

between the price of the house at time t and the discounted value of the house at time t+1, or C-C/(1+r) = rC/(1+r), where r is the assumed fixed rate of interest.[5] We abstract from taxes, maintenance costs and allow time-varying interest rates only in Section 5.[6]

The spatial equilibrium requires all cities at all times to deliver to the marginal resident the same utility that always is available in the reservation locale. We focus on the dynamics in a single representative city (which is different from the reservation city). The utility flow for person i living in that city during period t is $W(i,t)+A(i,t)$, or wages plus amenities. We assume that there are a fixed number of firms each of which has output that is quadratic in labor. This assumption ensures that the marginal product at each firm is linearly decreasing with the number of workers and that wages in the city are linearly decreasing with the number of workers. These labor demand schedules generated by firm optimization underpin our assumption that wages at the city level include an exogenous component and is linearly decreasing in total city population.

We assume that the time-specific and individual-specific effects that make up the net utility flow from the city are separable so $W(i,t)+A(i,t)-\underline{U}(t)$ can be written as $D(t)+\theta(i)$. The composite variable D(t) reflects wages and amenities, which in turn reflect exogenous shocks and city size. We let N(t) denote the housing stock in the city and assume that the city's population and labor force equal a constant times the amount

independently of theirs, and our focus on high frequency variation in prices and quantities is quite different from their focus on changes in the long run distribution of housing prices. More generally, the approach taken here differs from most research into housing prices, which employs the user cost approach introduced by Hendershott and Slemrod (1983) and Poterba (1984). That branch of the literature is too voluminous to describe in detail. The first three papers referenced in footnote 2 employ a user cost framework to examine the recent housing boom.

[5] This difference would also be the rent that a landlord earning zero profits would charge a tenant.

[6] If maintenance costs are independent of housing values and constant over space, they will not change the analysis. If maintenance costs scale with housing and if there are property taxes, then the cost of owning a house would be higher than the after-tax interest rate. For this reason, we will assume a relatively high real rate in our simulations. See below for more on that.

of housing.[7] We further assume that D(t) moves linearly with city population to allow for the fact that wages and amenities may fall due to congestion or rise because of agglomeration economies as city size increases. We assume that $\theta(i)$ is a uniformly distributed taste for living in this particular locale, so that the value of $\theta(i)$ for the marginal resident at time t (denoted $\theta(i^*(t))$) is also linearly decreasing in locale size.

The exogenous components of city amenities and wages include a city-specific component (denoted $\overline{D}$), a city-specific time trend (denoted qt) and a mean zero stochastic component (denoted x(t)). Thus, the flow of utility for the city's marginal resident at time t with index i*(t) relative to the reservation locale, $D(t)+\theta(i^*(t))$, can be written $\overline{D}+qt+x(t)-\alpha N(t)$, where α captures the assumption that wages, amenities and the taste of the marginal resident for living in the locale can fall linearly with city size. We further assume that x(t) follows an auto regressive moving average (ARMA) (1, 1) process so that $x(t)=\delta x(t-1)+\varepsilon(t)+\theta\varepsilon(t-1)$, where $0<\delta<1$, and the $\varepsilon(t)$ shocks are independently and identically distributed with mean zero.

The expected cost of housing in the representative locale equals $H(t)$ minus $E_t(H(t+1))/(1+r)$, where $E_t(.)$ denotes the time t expectation operator. The difference between the cost of housing in the representative city and housing costs in the reservation locale, rC/(1+r), should be understood as the cost of receiving the extra utility flow associated with locating in the city. If extra housing costs in the city equals extra utility delivered by the city then:

(1) $$H(t)-\frac{E_t(H(t+1))}{1+r}-\frac{rC}{1+r}=\overline{D}+qt+x(t)-\alpha N(t).$$

[7] Glaeser, Gyourko and Saks (2006) provide evidence showing that population is essentially proportional to the size of the housing stock.

Equation (1) represents a dynamic version of the Rosen-Roback spatial indifference equation where differences in housing costs equal differences in wages plus differences in amenities. We assume a transversality condition on housing prices such that $\lim_{j\to\infty}\left(\frac{H(t+j)}{(1+r)^j}\right)=0$.[8] If housing supply was fixed, so N(t)=N (as might be the case in a declining city as analyzed in Glaeser and Gyourko, 2005) then:

$$(2)\ H(t)=C+\frac{(1+r)(\overline{D}-\alpha N+qt)}{r}+\frac{(1+r)q}{r^2}+\frac{(1+r)x(t)+\theta\varepsilon(t)}{1+r-\delta}\ .$$

Housing prices are a function of exogenous population and exogenous shocks to wages and amenities, and the derivative of housing prices with respect to a one dollar permanent increase in wages will be (1+r)/r. Note that house price changes are predictable in this framework as long as there are predictable components to changes in urban wages and amenities. The ARMA(1,1) structure of the shocks makes it possible to have the positive correlation of changes at high frequencies and the negative correlation at low frequencies that we see in the data.

The city can grow with new construction so that $N(t)$ equals $N(t-1)+I(t)$, where I(t) is the amount of construction in time *t*.[9] The physical, administrative and land costs of producing a house are $C+c_0t+c_1I(t)+c_2N(t-1)$, where $c_1>c_2$ because current housing production should have a bigger impact on current construction costs than housing production many years ago.[10] Investment decisions for time *t* are made based on time *t-1* information, and there is free entry of risk neutral builders. Thus, if there is any

[8] This assumption limits the possible role of housing bubbles. While our focus here is on a purely rational model, we expect that future work will consider dropping this assumption.

[9] For simplicity, we do not allow depreciation which may be reasonable for shorter term housing dynamics, but would not be appropriate for a very long term analysis of city population changes.

[10] We deviate from the investment cost assumptions of Topel and Rosen (1988) by assuming that costs are increasing with the total level of development and not with changes in the level of investment.

building, construction costs will equal the time t expected housing price as described in equation (3):

(3) $E_{t-1}(H(t)) = C + c_0 t + c_1 I(t) + c_2 N(t-1)$.

As mentioned above, we assume that demand for the city is sufficiently robust so that there is always a positive quantity of new construction and this equation always holds.[11] Equations (1) and (3) then together describe housing supply and demand.

These equations give us the steady state values of housing prices, investment and housing stock:

$$\hat{H}(t) = \frac{\begin{pmatrix} c_2^2(q(1+r) + r(rC + \overline{D}(1+r))) + \\ \alpha(1+r)(C + c_1(q(1+r) - rc_0)) + \\ \alpha c_2(1+r)((c_0 + \overline{D} - q)(1+r) + 2rC) \end{pmatrix}}{(rc_2 + \alpha(1+r))^2} + \frac{(1+r)(\alpha c_0 + qc_2)}{rc_2 + \alpha(1+r)} t,$$

$$\hat{I}(t) = \hat{I} = \frac{q(1+r) - rc_0}{rc_2 + \alpha(1+r)} \text{ and } \hat{N}(t) = \frac{\begin{pmatrix} rc_1(rc_0 - q(1+r)) + \alpha(1+r)(c_0 + \overline{D}(1+r)) \\ + c_2(q(1+r)^2 + r(\overline{D}(1+r) - rc_0)) \end{pmatrix}}{(rc_2 + \alpha(1+r))^2} + \hat{I}t.$$

If x(t)=0 for all t, and $\hat{N}(t) = N(t)$ for some initial period, then these quantities would fully describe this representative city.[12] Secular trends in housing prices can come from trend in housing demand as long as $c_2 > 0$, or the trend in construction costs as long as $\alpha > 0$. If $c_2 = 0$ so that construction costs don't increase with total city size, then trends in wages or amenities will impact city size but not housing prices. If $\alpha = 0$ and city size doesn't decrease wages or amenities, then trends in construction costs will impact city size but not prices.

[11] The model can be extended to allow for the possibility that, in some states of the world, new construction will be zero. This adds much complication and only a modest amount of insight into our questions.

[12] In this case, the assumption that there is always some construction requires that $q(1+r) > rc_0$.

Proposition 1 describes housing prices and investment when there are shocks to demand and when $\hat{N}(t) \neq N(t)$. All proofs are in the appendix.

Proposition 1: At time *t*, housing prices equal

$$H(t) = \hat{H}(t) + \frac{\bar{\phi}}{\bar{\phi} - \delta} x(t) + \frac{\theta}{\bar{\phi} - \delta} \varepsilon(t) - \frac{\alpha(1+r)}{1+r-\phi}\left(N(t) - \hat{N}(t)\right)$$

and investment equals

$$I(t+1) = \hat{I} + \frac{(1+r)}{c_1(\bar{\phi} - \delta)}\left(\delta x(t) + \theta\varepsilon(t)\right) - (1-\phi)\left(N(t) - \hat{N}(t)\right),$$

where $\bar{\phi}$ and ϕ are the two roots of

$c_1 y^2 - ((2+r)c_1 + (1+r)\alpha - c_2)y + (1+r)(c_1 - c_2) = 0$ and $\bar{\phi} \geq 1 + r \geq 1 > \phi \geq 0$.

This proposition describes the movement of housing prices and construction around their steady state levels. A temporary shock, ε, will increase housing prices by $\frac{\bar{\phi} + \theta}{\bar{\phi} - \delta}$ and increase construction by $\frac{(1+r)(\delta + \theta)}{c_1(\bar{\phi} - \delta)}$. Higher values of δ (i.e., more permanent shocks) will make both of these effects stronger. Higher values of c_1 mute the construction response to shocks and increase the price response to a temporary shock (by reducing the quantity response). These comparative statics provide the intuition that places which are quantity constrained should have less construction volatility and more price volatility.

The next proposition provides implications about expected housing price changes.

Proposition 2: At time *t*, the expected home price change between time *t* and $t + j$ is

$$\hat{H}(t+j) - \hat{H}(t) + \left(\frac{\alpha(1+r)}{1+r-\phi} - \phi^{j-1}\left((1-\phi)c_1 - c_2\right)\right)\left(N(t) - \hat{N}(t)\right)$$

$$-x(t) + \frac{1}{\bar{\phi} - \delta}\left(\frac{1+r}{c_1}\frac{\delta^{j-1}\left((1-\delta)c_1 - c_2\right) - \phi^{j-1}\left((1-\phi)c_1 - c_2\right)}{\phi - \delta} - 1\right)E_t(x(t+1)),$$

the expected change in the city housing stock between time *t* and $t + j$ is

$$j\hat{I} + \frac{1+r}{c_1(\bar{\phi}-\delta)}\frac{\phi^j - \delta^j}{\phi - \delta}E_t(x(t+1)) - (1-\phi^j)\left(N(t) - \hat{N}(t)\right),$$

and expected time $t+j$ construction is

$$\hat{I} + \frac{1+r}{c_1(\bar{\phi}-\delta)}\left(\frac{\delta^{j-1}(1-\delta) - \phi^{j-1}(1-\phi)}{\phi - \delta}\right)E_t(x(t+1)) - \phi^{j-1}(1-\phi)\left(N(t) - \hat{N}(t)\right).$$

Proposition 2 delivers the implication that a rational expectations model of housing prices is fully compatible with predictability in housing prices. If utility flows in a city are high today and expected to be low in the future, then housing prices will also be expected to decline over time. Any predictability of wages and construction means that predictability in housing price changes will result in a rational expectations model.

The predictability of construction and prices comes in part from the convergence to steady state values. If $x(t) = \varepsilon(t) = 0$ and initial population is above its steady state levels, then prices and investment are expected converge on their steady state levels from above. If initial population is below its steady state level and $x(t) = \varepsilon(t) = 0$, then price and population is expected to converge on their steady state levels from below. The rate of convergence is determined by r and the ratios $\frac{c_1}{\alpha}$ and $\frac{c_2}{\alpha}$. Higher levels of these ratios will cause the rate of convergence to slow by reducing the extent that new construction will respond to changes in demand.

The impact of a shock, x(t), is explored in the next proposition.

Proposition 3: If $N(t) = \hat{N}(t)$, $x(t-1) = \varepsilon(t-1) = 0$, $\theta > 0$, $c_2 = 0$, and $\varepsilon(t) > 0$, then investment and housing prices will initially be higher than steady state levels, but there exists a value j^* such that for all $j > j^*$, time t expected values of time $t+j$ construction and housing prices will lie below steady state levels. The situation is symmetric when $\varepsilon(t) < 0$.

Proposition 3 highlights that this model not only delivers mean reversion, but overshooting. Figure 3 shows the response of population, construction and prices relative to their steady state levels in response to a one time shock. Construction and prices immediately shoot up, but both start to decline from that point. At first, population rises slowly over time, but as the shock wears off, the heightened construction means that the city is too large relative to its steady state level. Eventually, both construction and prices end up below their steady state levels because there is too much housing in the city relative to its wages and amenities. Places with positive shocks will experience mean reversion, with a quick boom in prices and construction, followed by a bust.

Finally, we turn to the puzzling empirical fact that, across the 1980s and 1990s, there was strong mean reversion of prices and strong positive serial correlation in population levels. We address this by looking at the one period covariance of price and population changes. We focus on one-period for simplicity, but think of this proposition as relating to longer time periods. Since mean reversion dominates over long time periods, we assume $\theta = 0$ to avoid the effects of serial correlation:

Proposition 4: If $N(0) = \hat{N}(0)$, $\theta = 0$, $x(0) = \varepsilon(0)$, cities differ only in their demand trends *q* and their shock terms $\varepsilon(0)$, $\varepsilon(1)$ and $\varepsilon(2)$, and the demand trends are uncorrelated with the demand shocks, then the coefficient estimated when regressing second period population growth on first period population growth will be positive if and only if $\frac{Var(q)}{Var(\varepsilon)} > (1-\delta-\phi)\left(\frac{\delta(rc_2+\alpha(1+r))}{c_1(\bar{\phi}-\delta)}\right)^2$ while the coefficient estimated when regressing second period price growth on first period price growth will be negative if and only if $\Omega\left(\frac{rc_2+\alpha(1+r)}{(1+r)c_1c_2(\bar{\phi}-\delta)}\right)^2 > \frac{Var(q)}{Var(\varepsilon)}$, where

$$\Omega = \left(\frac{\alpha(1+r)^2\delta}{1+r-\phi} + c_1(1-\delta)\bar{\phi}\right)\left((1-\delta-\phi)\frac{\alpha\delta(1+r)^2}{1+r-\phi} + c_1(1-\delta+\delta^2)\bar{\phi}\right).$$

Proposition 4 tells us that positive correlation of quantities and negative correlation of prices are quite compatible in the model. The positive correlation of quantities is driven by the heterogeneous trends in demand across urban areas. As long as the variance of these trends is high enough relative to the variance of temporary shocks, then there will be positive serial correlation in quantities as in Figure 2.

The mean reversion of prices is driven by the shocks, and as long as c_2 is sufficiently low, prices will mean revert. As discussed above, when c_2 is low, trends will have little impact on steady state price growth. The positive trends show up mainly in the level of prices. However, regardless of the value of c_2, unexpected shocks impact prices and, if these shocks mean revert, then so will prices.

This suggests two requirements for the observed positive correlation of quantities and negative correlation of prices: city-specific trends must differ significantly and the impact of city size on construction costs must be small. The extensive heterogeneity in city-specific trends is much commented on, with the recent papers by Gyourko, Mayer, and Sinai (2006) and Van Nieuwerburgh and Weill (2006) attempting to explain the phenomenon. The literature on housing investment suggests that the impact of city size on construction costs is small (Topel and Rosen, 1988; Gyourko and Saiz, 2006). Thus, we shouldn't be surprised to see positive serial correlation in quantity changes and negative serial correlation in price changes.

III. Key Parameter Values for the Calibration Exercises

We now use the model as a calibration tool to see what moments of the data, including its serial correlation properties and variances, can and cannot be explained by our framework. We focus on the movements in prices and construction intensity around steady state levels. The appendix contains the formulae for the predicted values of these moments.[13] The model's predictions about variances and serial correlations depend on seven parameters: the real interest rate (r), the degree to which demand declines with city population (α), the degree to which construction responds to higher costs (c_1 and c_2), the time series pattern of local economic shocks (δ and θ), and the variation of those shocks (σ_ε^2). Table 1 reports the value of these parameters which are used in the calibration exercise, with the remainder of this section discussing how we estimate or impute them.

The Real Interest Rate (r)

The first row of Table 1 shows that we use a real interest rate (r) of 4 percent in all calibrations. This value is higher than standard estimates of the real rate because it is also meant to reflect other facets of the cost of owning, such as taxes or maintenance expenses, that might scale with housing. Experimentation shows that the simulation results are robust to a wide range of alternative values of *r* (e.g., from 2.5-5 percent).

Supply Side Parameters: c_1 and c_2

The housing cost parameters are both particularly important for our model and relatively understudied by the literature. The parameter c_1 reflects the extent that construction costs, including land assembly, permitting and physical construction costs,

[13] We do not use the high frequency correlations of prices with other variables to pin down parameter values. Changes in house prices and changes in income accompany each other at longer horizons (e.g., over the past twenty years, the correlation of the two changes is over 50 percent), but the correlation is much weaker at higher frequencies. Higher frequency correlations are difficult to interpret because the real world information structure may not match that presumed in our model. For example, if income shocks are known a period earlier, this will not matter much for predicted variances and serial correlations, but it will dramatically alter the predicted relationship between income and price changes.

rise with the level of current construction activity. The c_2 parameter measures the sensitivity of costs to the level of overall development, or market size.

Topel and Rosen (1988) use national data and estimate a supply elasticity ranging from 1.4 and 2.2. This supply elasticity is the relationship between the logarithm of investment and the logarithm of price, which in our model would equal $H(t)/c_1 I(t)$. Using the mean values of investment and housing prices across our cities and an elasticity of 1.8, this generates a range of c_1 from 1 to 151. The median value is 18 which seems too high to us. The range between 25th and 75th percentiles of the distribution is 5 to 28, which provides us with one range of parameter estimates.

Alternatively, we can estimate c_1 from the variance of prices and quantities within the data. While this violates the typical calibration rule of using parameter estimates from outside the data to be explained, we think that this exercise is useful in generating a range of possible construction cost values. We estimate them for each of our 115 metropolitan areas and use this broad range in our simulations.[14] Our hope is that by showing results for a wide range of estimates, we can diffuse worries that naturally arise from the fact that our estimates were made using the data that we are trying to explain.

We start with the basic supply-side equation $E_{t-1}(H(t)) = C + c_0 t + c_1 I(t) + c_2 N(t-1)$ and let $H(t) = E_{t-1}(H(t)) + \mu$, where μ is the prediction error. We further assume that the variance of μ equals

[14] Physical construction costs vary the least across U.S. markets, but Gyourko and Saiz (2006) report a 20 percent gap across the interquartile range of major metropolitan areas. The differences are driven largely by the degree of union penetration in the local construction trades and select other local factors. The nature of local land use regulation varies much more dramatically by market, with little binding constraint on new development in markets such as Atlanta and Las Vegas, while the Boston and Bay Area markets have very stringent and expensive regulation that makes it very hard to build even though market prices of homes are well above physical construction costs in those places. [See Glaeser and Gyourko (2003), Glaeser, Gyourko, and Saks (2006), and Saks (2006)) for studies on the impacts of differing land use regulation.]

$\kappa Var(H(t) - C - c_0 t - c_1 \hat{I} - c_2 (N(0) + t\hat{I})$, where the parameter κ is one minus the R^2 from the best possible prediction of next period's housing price. To determine this unexplained share of the deviation from actual house prices, we begin by regressing house prices on year and metropolitan area dummies. The R^2 from that fixed effects specification is 0.935. When we add two lags of house prices and housing permits to the specification, the R^2 increases to 0.993. Thus, κ = 10% or (0.1 ~ 0.07/0.7) based on these values. We then let $c_2 = \omega c_1$ and using the equation

$$(4)\quad \begin{array}{l} Var(H(t) - C - c_0 t - c_1 \hat{I} - c_2 (N(0) + t\hat{I}) = Var(\mu) + \\ c_1^2 Var\Big(((I(t) - \hat{I}) + \omega\big(N(t-1) - N(0) - t\hat{I}\big)\Big) \end{array},$$

it follows that,

$$(5)\quad c_1^2 = \frac{Var(H(t) - C - c_0 t - c_1 \hat{I} - c_2 (N(0) + t\hat{I})(1-\kappa)}{Var\Big(((I(t) - \hat{I}) + \omega\big(N(t-1) - N(0) - (t-1)\hat{I}\big)\Big)}.$$

The numerator of this ratio is the variance of the house price prediction error, weighted by the share of unexplained variation. The denominator contains two components. The first reflects the variation associated with yearly new construction deviating from its average level. The second term in the denominator, which is weighted by ω, captures the variation associated with the market's housing stock being off its trend amount.

To empirically use equation (5), we must impute the housing stock (the N(t-1) term) because the census provides actual counts of the stock only once each decade. For each metropolitan area, we know the housing stock at the beginning and end of each decade and the permits issued each year in between. Our estimate of the housing stock at time t+j, is $N(t) + \frac{\sum_{i=0}^{j-1} Permits_{t+i}}{\sum_{i=0}^{9} Permits_{t+i}} (N(t+10) - N(t))$, where N(t) and N(t+10) are the housing stocks measured during the two closest censuses. The change in housing stock is portioned across years based on the observed permitting activity. To measure how many units the market should have had each year (the " $N(0)+(t-1)\hat{I}$ " term), we use the actual count of the housing stock from the decennial census in our initial year of 1980 and assume that I-hat is the average change in the actual housing stock between 1980 and 2000 (again using census data to measure the stock).

Since costs do not appear to vary much based on city development, we conclude that $c_1 > c_2$, but beyond that, the literature yields little to help us pin down the relationship between c_1 and c_2. We consider a range of values for ω, including 0, 0.25, and 0.50.[15]

The bottom two panels of Table 1 report the distribution of estimated values of c_1 and c_2 for the three different values of omega. If ω=0.25, there is a wide range of values for c_1 from 1.5 for the 10th percentile metropolitan area to 28.1 for the 90th percentile area. The mean value of c_1 is 13.6 which is twice the median market's c_1 value of 6.4,

[15] An alternative method of estimating these parameters suggests a value of 0.25 for ω. That approach to estimating the construction cost parameters follows Rosen and Topel (1988) in inverting the construction cost equation to obtain $I(t) = (1/c_1)(E_{t-1}[H(t) - C) - (c_2/c_1)(N_{t-1})$. In empirically implementing this equation, we used total housing permits in period t to proxy for new construction in period $t+1$, and actual house prices to measure expected values. Obviously, the use of actual prices in lieu of expected prices introduces some bias, but it should be small since the annual time period over which price is measured is relatively short. We also imputed the housing stock (N(t)) each year as described above. A simple regression of each market's resulting c_2 value on its c_1 value (with no intercept, as suggested by our assumed functional form) yielded a coefficient of 0.25. The estimated coefficient is 0.21 if we allow for an intercept. The simple correlation between c_1 and c_2 values estimated this way is quite high at 0.92. Finally, it is noteworthy that the absolute level of these c_1 values is higher than those reported in Table 1 based on the approach described in equations (4) and (5). This was to be expected given that this alternative strategy effectively assumes that price only reflects demand, not supply, shocks. To the extent that prices incorporate productivity-enhancing changes, the relationship between new construction and values is magnified. More problematic for this alternative approach is the issue of the potential endogeneity of housing prices with respect to new construction. We also estimated a singe nationwide regression using the interaction of initial industrial characteristics and national economic success of the industries as instruments (as in Bartik, 1989). When a linear specification is estimated, the coefficients are quite imprecise. When we follow Rosen and Topel (1988) and use a log-log specification, we precisely estimate an elasticity of 2, which implies a range of estimates of c_1 and c_2 similar to those presented in Table 1.

which reflects the skewness of the distribution of costs. A handful of large and expanding markets such as Atlanta, Charlotte, Houston, and Dallas have c_1 values below one.[16] The top ten percent of markets in terms of c_1 values are all in Hawaii, along the coast of California, or in the suburbs of New York City. These high values of c_1 appear to reflect both high labor costs and regulations that constrain construction.[17] The interquartile range of c_2 runs from 0.7 to 3.3. The median value is 1.6

Increases in Population and the Marginal Valuation of an Area: α

The value of α reflects the impact that an increase in the housing stock will have on the willingness to pay to live in a locale. If population was fixed, equation (2) tells us that the derivative of housing prices with respect to the housing stock equals $-(1+r)\alpha/r$, which can be seen as the slope of the housing demand curve. Typically, housing demand relationships are estimated as elasticities. Consequently, we must transform estimated demand elasticities into a levels estimate by multiplying by r/(1+r).

While many housing demand elasticity estimates are around one (or slightly below-in absolute value), there is a wide range in the literature, so we experiment with a range from 0 to 2. We begin the transformation from an elasticity to a level by multiplying by the ratio of price to population, which produces a range of estimates for $(1+r)\alpha/r$ of from 0 to 3. Multiplying this span by r/(1+r) yields a range from 0 to 0.15. We will use a parameter estimate of 0.1 in our simulations which implies that for every 10,000 extra homes sold, the marginal purchaser likes living in the area $1,000 less per

[16] The Atlanta area has the lowest value at 0.39.
[17] The values for the metropolitan areas of Honolulu, Salinas, Santa Cruz, and Napa each are above 70. While we cannot tell for sure, these magnitudes probably are associated with regulatory costs, as they seem too high to solely reflect labor.

year (see row 5 of Table 1). This estimate seems high to us, but lower estimated values of α do not significantly change the simulations.

Time Series Properties and Variance of Shocks: δ, θ, and σ_ε^2

The model does not separately address wages and amenities. There is little evidence on high frequency changes in amenities, except for crime rates which we will discuss in Section V. Consequently, we assume here that the high frequency movement in demand is driven by changes in labor demand, not changes in the valuation of amenities. More specifically, observed wages $W(t)$ are presumed to equal $w_0 - \gamma N(0) + w_1 t + x(t) - \gamma(N(t) - N(0))$, where $w_0 - \gamma N(0)$ is a component of $\overline{D}$, w_1 is a component of q and γ is a component of α. Controlling for a city-specific fixed effect and trend will eliminate the term $w_0 - \gamma N(0) + w_1 t$, and the residual component of wages equals $x(t) - \gamma(N(t) - N(0))$.[18]

The most difficult part of estimating the x(t) process is our attempt to control the impact of population changes, but while our procedure is debatable, it has little impact on the estimated properties of x(t). The parameter γ represents the impact that an increase in city size will have on wages, which is proportional to the impact of labor supply on wage (or the slope of the labor demand function). Customarily, labor demand is estimated as an elasticity, $\frac{Labor\ \ Force}{Wage}\frac{\partial Wage}{\partial Labor\ \ Force}$, and most estimates of this elasticity are statistically indistinguishable from zero (e.g., Card and Butcher, 1991). Borjas (2003) finds a higher estimate of -0.3, although this is at the national level. We use this upper-

[18] The distinguished literature on regional shocks (e.g. Blanchard and Katz, 1992) does not yield the parameter estimates that we need to calibrate the model.

bound estimate, but note that it has little differential effect on our results compared to assuming an elasticity of zero.

Just as with housing demand, we must convert this elasticity into an estimate of γ. We use BEA data on personal income per capita as our measure of wages, and for our sample of metropolitan areas, the mean of this variable in 1990 (the middle of our sample period) was $26,965 (in $2,000). Mean employment in 1990 across these metropolitan areas was 539,215, so our ratio of wage to the labor force is about 0.05 (~26,965/539,215). Based on these numbers, an elasticity of -0.3 suggests that each worker is associated with 1.5 cents less annual income in the city. In our sample, there are on average 1.26 workers per home, so each extra home is associated with 1.9 cents per year less annual income, which serves as our estimate of γ.

The per capita income series is converted into household income by multiplying by 2.63 (the average ratio of people per household in our sample in 1990). We adjust this income variable for N(t) using our estimate of 0.019 for γ and by using

$N(t)+\frac{\sum_{i=0}^{j-1} Permits_{t+i}}{\sum_{i=0}^{9} Permits_{t+i}}(N(t+j)-N(t))$ (as above) as our estimate of total housing stock between census years.

With this corrected income series, we can estimate the time series properties of income shocks at the local level, by fitting an ARMA(1,1) to the wage series that is first demeaned with city and year fixed effects and then corrected for city size changes as discussed above. As shown in Table 1, this estimation procedure yields estimates of δ=0.87, θ=0.17, and σ_ε^2= \$3,603,463.[19]

IV. Calibrating the Model and Matching the Data

In this section, we calibrate the model using the parameters values discussed above. We then compare this calibration to the moments of the real data. We first discuss the time series coefficients of prices and construction, and then discuss the volatility of these series. Our "real data" sample is a set of 115 metropolitan areas for which we have continuously defined price data from 1980-2005. As discussed above, our simulations assume that the real interest rate (r) equals 0.04, the variance of the local economic shock (σ_ε^2) equals \$3,603,463, the parameters of the ARMA process describing that shock are 0.87 (δ) and 0.17 (θ), and the valuation of the metropolitan area by the marginal entrant (α) is 0.1. We let c_1 and c_2 equal the fifteen different pairwise combinations associated with the three different values of omega and report simulation results for the full set of those values.

Short-Term Momentum and Longer-Term Mean Reversion in Prices, Rents and Permits

The top row of Table 2 provides evidence on momentum and mean reversion in OFHEO house prices within market over time. We use absolute price changes rather than changes in the logarithm of prices in order to be compatible with the model, but our empirical results are not sensitive to such changes in functional form. Since the OFHEO index only provides price increases relative to a base year, we convert this into an implied

[19] Largely because γ is so small throughout its relevant range, this adjustment to wages does not have a material impact on our results. If we use a value of 0 for γ, we estimate a value of δ = 0.86, an estimate of θ = 0.18, and an estimate of σ_ε^2=\$3,408,250 In addition, we attempted joint maximum likelihood estimation of δ, θ, and σ_ε^2 for given trend effects and metropolitan area fixed effects, but the program would not converge because the panel was too short relative to the number of markets.

price series by using the median housing value in the metropolitan area in 1980 as a base price in the metropolitan area and then scaling that value by the appreciation in the OFHEO index each year.[20]

The results are estimates from a regression of the current change in prices on the lag change in prices

$$(6) \quad Price_{t+j} - Price_t = \alpha_{MSA} + \gamma_{Year} + \beta\left(Price_t - Price_{t-j}\right),$$

for j equal to one, three and five years. Because fixed effects estimates such as these which remove market-specific averages can be biased (with spurious mean reversion produced especially when the number of time periods is relatively low), in the first row of Table 2, we report Arellano-Bond estimates which use lagged values of the dependent variable (price changes) as instruments.[21]

Our one year estimate of price change serial correlation is 0.71, so a $1 increase in housing prices between time t and t+1 is associated with a 71 cent increase between time t+1 and t+2. Our estimate is larger than that reported by the pioneering work of Case and Shiller (1989). It is now well understood that smoothing of the underlying data series can bias one towards finding short-run momentum. Case and Shiller (1989) were able to address this problem by splitting their sample, which consisted of extensive micro data on sales transactions in four markets (Atlanta, Chicago, Dallas, and San Francisco). They report coefficients ranging from 0.2-0.5, although they use the logarithm, not the

[20] This procedure essentially provides the real price for a constant quality house with the quality being that associated with the median house in 1980. We have experimented with using values from the 1990 and 2000 censuses as the base. All the results reported below are robust to such changes.
[21] See Arellano and Bond (1991) for more detail on this estimation procedure. More specifically, we use the "xtabond" Stata command with year and area fixed effects.

level, of prices so the results are not exactly comparable. Because we cannot perform any comparable procedure with the OFHEO data, our estimate is surely biased upwards.[22]

Over three years, there is still momentum. The estimate of 0.27 means that a $1 increase in housing prices between time t and t+3 is associated with a 27 cent increase between time t+3 and t+6. Over five year periods, we estimate a mean reversion coefficient of -0.32, so a $1 increase between times t and t+5 is associated with a 32 cent decline between time t+5 and t+10.[23] These estimates are not an artifact of the Arellano-Bond procedure. The analogous ordinary least squares estimates over 1, 3, and 5 year horizons are 0.74, 0.18 and -0.39, respectively.[24]

The mean reversion in prices that we estimate over five-year horizons is quite similar in magnitude to that observed for financial assets by Fama and French (1988). Unfortunately, the short time period for which we have constant quality data at less than decadal frequencies makes it difficult to know whether this mean reversion is a permanent feature of urban life or whether it represents the impact of shocks that are specific to the post-1980 time period. Cutler, Poterba and Summers (1991) also find this

[22] The OFHEO index includes data on repeat sales or refinancings of the same house. The latter typically rely on an appraisal, not a market sale price. Undoubtedly, this results in smoothing of the series and biases upward our estimate of short-run momentum. Even the Case and Shiller (1989) estimates, which rely only on actual sales, could be upward biased. Working with a split sample, bias can result if, randomly, some fraction of homes on which a buyer and seller agree on a price have delayed closings that move their reported sales dates into the next reported period (quarter, year, etc.). Whatever shock there was in period *t* that influenced the agreed upon price, some of its measured impact will spill over into period *t+1*. Obviously this is potentially more of a problem the shorter the measurement period.

[23] As noted in the Introduction, decadal changes also find significant mean reversion across the 1980s and 1990s.

[24] We also addressed concerns about spurious mean reversion by estimating specifications without metropolitan area fixed effects. If we estimate the following equation, $Price_{t+5} - Price_t = \alpha + \gamma_{Year} + \beta(Price_t - Price_{t-5})$, the mean reversion coefficient drops to -0.11 and becomes only marginally significant. However, as soon as we include percent of adults with college degrees as a control, the coefficient becomes -0.18 with a t-statistic of three. If we estimate the same change regression using the logarithm of prices instead of the levels, the coefficient is -0.20 (-0.22 with the college graduate control) and has a t-statistic of four.

pattern of short run momentum and longer- run mean reversion for housing and a number of other asset markets.

Table 3 reports the comparable results for our simulations using the different c_1 and c_2 values discussed above. All other parameter values are fixed at the values listed in Table 1. The first three columns show results for annual serial correlation in prices, the next three columns present the analogous findings over three year periods, with the final three columns being for five year periods. Within each time horizon, the fifteen cells correspond to the fifteen c_1, c_2 pairs reported in Table 1.

The first three columns document the model's failure to match the positive serial correlation observed in the annual data. In fact, our parameter estimates suggest a mild amount of mean reversion even at such a high frequency. Similarly, the results for three year horizons reported in the middle columns of Table 3 find a mismatch with the data. Assuming the middle case for omega (ω=0.25, column 4), we predict mean reversion coefficients from -0.18 to -0.28, not the positive persistence we see in the data.

Our model does a much better job of fitting the -0.32 mean reversion seen at five year intervals (columns 6-9). At five year horizons, if c_1 and c_2 are at their medians for the case of ω=0.25, we come within ten percent of matching the data (see row 3, column 8). And, if c_1=2.7 and c_2=0.7, the prediction literally is -0.32 (row 2, column 8). Thus, the predictable mean reversion of prices at five year intervals cannot be seen as a challenge for a rational expectations model of housing price movements.

In the model, this mean reversion reflects both the tendency of shocks to mean revert and of new construction to cause future declines in prices. To decompose the impact of the two forces, we also looked at the cases where there is no new construction impact (i.e., α=0) and where there is no mean reversion in the x(t) parameter. When mean reversion in wages is turned off (i.e., δ=1), our simulations predict very low levels of overall mean reversion in house prices. However, we predict higher levels of mean reversion close to those found in the data when there is no effect allowed from construction increasing market size. Thus, our analysis suggests that the majority of mean reversion is coming from the mean reversion of local demand shocks.

One way to check whether short run momentum reflects the dynamics of euphoria in an asset market is to see if the same phenomenon appears in rents. In the second row of Table 2, we report the results from rental regressions of the form:

$$(7) \quad \mathrm{Rent}_{t+j} - \mathrm{Rent}_{t} = \alpha_{MSA} + \gamma_{Year} + \beta\left(\mathrm{Rent}_{t} - \mathrm{Rent}_{t-j}\right).$$

Rental data on apartments is collected by an industry consultant and data provider, REIS Inc. Their data covers only a limited number of metropolitan areas (46 in our sample), and in general, rental units are not comparable to owner-occupied housing.[25]

Over one- and three-year horizons, there is strong evidence of persistence, with the Arellano-Bond estimates being 0.27 in both cases. Over five year time horizons, we estimate a mean reversion parameter of -0.64. The presence of momentum and mean reversion in rents suggests that these features do not reflect something unique to housing asset markets, but rather something about the changing demand for cities.[26]

Table 4 then reports the predicted values of serial correlation from the simulations of the model. At annual frequencies, we tend to predict very modest persistence, with the

[25] Rental units are overwhelmingly in multi-unit buildings, while owner-occupied housing is overwhelmingly single-family detached housing. These differences in housing types and the problem of accurately measuring maintenance costs are two reasons why it is extremely difficult to tell whether housing prices are high or low relative to rents.

[26] The ordinary least squares estimates of these coefficients are 0.28, 0.08 and -0.51 for one, three and five year horizons, respectively.

results ranging from -0.05 to 0.08 when ω=0.25 (column 2). These estimates are well below the 0.27 seen in the data over one year horizons. Over three year horizons, we consistently predict mean reversion, while there is still a positive correlation of rents in the data. For five year intervals, we predict that rent changes should have a mean reversion coefficient of about -0.30 if we use the median value of 6.4 and 1.6, respectively, for c_1 and c_2 when ω=0.25 (row 3, column 8). Slightly higher mean reversion is predicted if c_1 is lower, but our estimates still are only about one-half of the observed mean reversion in that case.

We are again unable to explain the strong positive serial correlations at shorter time horizons. Since there are many reasons to be suspicious about the properties of the rental data, especially because of artificial smoothing, we do not attach much importance to the quantitative mismatch with the data here.[27] However, the short run momentum and long run mean reversion of rents, which are predicted by the model, suggest that these features could reflect something other than irrationality in an asset market.

To examine the dynamics of housing quantity, we look at housing permit data from the *Census of Construction*. The final set of results reported in Table 2 use housing permits estimated in the following regression: $Permits_t^{t+j} = \alpha_{MSA} + \gamma_{Year} + \beta Permits_{t-j}^t$, where $Permits_{t-j}^t$ refers to the number of permits issued between time t-j and time t. The one-three and five year Arellano-Bond coefficient estimates are 0.84, 0.43, and -0.07,

[27] For example, smoothing is a greater problem in the rental data. The industry consultant that provides the rent data does not survey actual renters, but the landlord owners of apartment buildings. Undoubtedly, averages are being reported.

respectively. Thus, construction also displays high frequency momentum, but little or no persistence or mean reversion at longer horizons.[28]

The calibration results for this variable are provided in Table 5. For the case where c_1 and c_2 are the median values when ω=0.25, the predicted coefficients are 0.60, 0.29 and 0.06, for one, three, and five year horizons, respectively. These are reasonably close to the actual parameters, and minor changes in the values of one or both of the supply side parameters enable us to fit the data more exactly. While the predictions about the serial correlation of construction are not as accurate as the predictions about the mean reversion in prices, the moments of the real data cannot be said to reject the model.

Thus, the model does a reasonable job at fitting the time series properties of new building and an excellent job at fitting the long term mean reversion of rents. It does a poor job of fitting the high frequency positive serial correlation of price changes. This failure may be the result of data smoothing causing us to empirically overestimate momentum, or as Case and Shiller (1989) suggest, some sort of irrationality in the housing market.

House Price Change and Construction Variances Across Markets

Table 6 reports the variance of price changes and of new construction in our sample.[29] The volatility of both prices and construction varies enormously across cities and is quite skewed, with the mean variance much higher than the median variance. Consequently, our approach to documenting both housing price change and construction variance is to first run a pooled market regression (separately for each variable, of course)

[28] As is the case with the other data, this pattern is not an artifact of our estimation procedure. The analogous ordinary least squares coefficients are 0.82, 0.37, and 0.07, respectively.
[29] Since the rent data are smoothed, we don't put much weight on the variance of rents and exclude them from this part of the analysis.

controlling for year effects, and then to compute the variance of the residuals from this regression by metropolitan area. This variance gives us the volatility of prices and construction, respectively, within a metropolitan area controlling for nationwide effects.

The top panel of Table 6 shows that that the variance of one year price changes equals $14 million in the tenth percentile metropolitan area and $209 million in the 90th percentile market. The median market has a one year price change variance of $34 million, which is much smaller that the sample mean of $83 million. This skewness is driven primarily by California markets and Honolulu. The variance of one-year price changes in Honolulu is $763 million, which is the largest in our sample. Five other markets—San Jose, San Francisco, Santa Barbara, Santa Ana and Salinas--had variances that were ten times greater than the sample median.

The second and third columns of this table report the distribution of variances of three and five year price changes. The distribution of longer horizon price changes is again quite skewed, with the mean price change substantially exceeding the change for the median area. The volatility of price changes is very high at longer horizons. The variance in five-year price changes is $625 million for the median market, with one quarter of the metropolitan areas having variances of at least $1.1 billion.

Table 7 reports predicted price change variances from our simulations with the results arrayed in the same manner as in the serial correlation tables above. At annual frequencies (columns 1-3), we predict a range of price change variances from a low of $50 million to a high of $181 million. Not surprisingly, price change volatility is lower the smaller are c_1 and c_2—markets in which quantity changes a lot with changes in costs. However, our variance prediction for the lowest (c_1, c_2) combination still is well above the $34 million variance found in the median market. Data smoothing would bias price volatility downward over short time periods, and if so, we would expect this problem to be less severe for longer time periods.

Our ability to match the volatility of price changes does increase with the horizon over which those changes are measured. For example, the range of predicted variances of 3-year price changes across all 15 (c_1, c_2) combination runs from $123-$488 million (columns 3-6, Table 7). This spans the interquartile range of $124-$445 million found in the data (column 2, top panel of Table 6), but we still need to have a relatively low value for c_1 (2.7) and a c_2=0 to closely match the median. This is not the case for the 5-year price change variance predictions listed in the final three columns of Table 7. Our range of predictions runs from $163-$718 million, allowing us to capture much of the lower half of the distribution of actual price change variation reported in the third column of Table 6 (top panel).

The median (c_1, c_2) pair of 6.4 and 1.6 when ω=0.25 is associated with a predicted variance of $469 million, which still underpredicts the sample median ($625 million), but higher c_1 values allow us to come much closer. However, we are still unable to approach matching the very high price change volatility found in the top quarter of markets, and the top ten percent, especially. That is an issue to which we will turn in the next section. In sum, the general pattern of results in Table 7 shows an overestimate of price volatility at high frequencies which then disappears at lower frequencies. This trend is consistent with data smoothing, but it could also reflect a flaw in our model.

The bottom panel of Table 6 reports the variance in units permitted across our 115 metropolitan area sample. As with price changes, there is substantial heterogeneity in the volatility of construction intensity across markets, and this distribution is skewed by a few outliers. For example, the bottom quartile of markets has a new construction variance of about 2 million units per annum, while the top quartile is at least five times more volatile. Moreover, this distribution is skewed by relatively few markets in the right tail that have variances of at least 38 million units (column 1, bottom panel of Table 6). Six markets—Phoenix, Dallas, Riverside-San Bernardino, Atlanta, Los Angeles, and Houston---stand out in this regard, having construction intensity variances that are at least double the next six highest variance markets. There also is great heterogeneity in construction intensity variance over longer horizons, as the second and third columns of this part of Table 6 document.

Table 8 reports the construction intensity variance estimates from our standard set of simulations. Over annual periods, our model does a decent job of matching this data. Our range of predicted construction variances runs from 300 thousand to 42 million units. This captures the bulk of the range across markets in the actual data, which is 2 million units in the 10^{th} percentile market and 38 million units in the 90^{th} percentile market (Table 6). The median market in our sample has a one year standard deviation of 3 million units, which precisely equals the 3 million unit variance we predict if c_1=6.4, the median value in our sample when ω=0.25.

Over three year horizons, we can match the median market in the nation as well, but the fit is not as perfect. For example, the three-year construction variance in the median market is 26 million units (middle column, bottom panel of Table 6), while we predict a 21 million unit variance for our ‘median’ (c_1, c_2) combination of 6.4 and 1.6 (see the middle cell in columns 4-6 of Table 8). Our range of estimates using our ‘lower’ (but not ‘lowest’ (c_1, c_2) values as in row 2 of these columns) spans the 84 million unit variance associated with the 75th percentile metropolitan area (Table 6). Our ability to match outcomes weakens for the most volatile markets. Our top variance estimate of 189 million units for three year horizons is only 58 percent of the 328 million unit variance observed for the 90th percentile market.

Our under-prediction of construction intensity variation is greater at five year horizons. Our full range of estimates runs from 3-328 million units, versus a 10th-90th percentile range of 29-760 million in the data. Our ‘median’ estimate of 44 million units is only three-quarters of the national market median of 59 million. Just as in the case of prices, we are able to fit the volatility of the median markets with plausible parameter values, but we completely fail to predict the volatility of the more extreme cases.

The model implies that markets with elastic supplies of housing (i.e., those with low c_1 and c_2 values) should have relatively low price change variation and relatively high construction volatility. Concomitantly, those areas with high c_1 and c_2 values that have inelastic supplies will have little variation in quantities and more volatility in prices. Hence, if our estimation procedure for determining these two supply side parameters biases us from extreme values, we could be underestimating the impact of supply conditions on both price and quantity volatility.

One natural explanation for these extreme markets is that we have significantly overestimated the value of c_1 and c_2 in sunbelt region markets with elastic supplies of housing. To address the role that mismeasurement of these parameters can play, we ran four additional simulations using values for c_1 and c_2 from the Atlanta, Phoenix, San Francisco, and Salinas (CA) metropolitan areas. The Atlanta market has the lowest c_1

value in our sample, and assuming ω=0.25, its (c_1, c_2) combination is (0.4, 0.1). Phoenix also experiences new supply if prices rise even slightly, and its (c_1, c_2) combination is (1.0, 0.25). San Francisco and Salinas are at the opposite end of the spectrum, with Salinas being a particularly extreme outlier. Their (c_1, c_2) combinations are estimated to be (40, 10) and (100, 25) respectively.

Lower c_1 and c_2 values are associated with increases in predicted construction volatility. Our parameter estimates for Phoenix predict 1, 3, and 5 year construction variances of 34 million, 186 million, and 357 million units, respectively. The analogous figures for our Atlanta parameter estimates are 74 million units, 361 million units, and 651 million units, respectively. These numbers still are well below what we observe in the most highly volatile construction markets (e.g., Atlanta's five year variance is 4.52 billion units), but they do match the values for the market in the 90th percentile of the distribution (see Table 6). The most volatile construction markets still remain a puzzle for our model, but we interpret these results to imply that underestimating how elastic the supply side is could help account for a meaningful part of our underestimation of the variation in construction intensity in many sunbelt markets.

Significantly higher c_1 and c_2 values do not result in commensurate increases in predicted price change variation, so even if we are underestimating these parameters in highly volatile coastal markets, that cannot explain the dramatic price change variation we see in those place. More specifically, the 1, 3, and 5 year predictions assuming supply side conditions are those found in San Francisco are \$181 million, \$487 million, and \$714 million, respectively. The Salinas estimates yield similar predictions: \$190 million, \$515 million, and \$759 million, respectively. These estimates are barely above the maximum values reported in Table 7. Although supply elasticity can explain a significant amount of the high construction volatility in the sunbelt, supply inelasticity cannot explain much of the high price volatility in coastal markets. We now turn to other potential explanations.

V. Explaining the High Volatility of Housing Prices

To see if we can match the high volatility of housing prices observed in certain markets in the data, we first turn to omitted demand factors such as local tax rates and amenities. Following that, we examine the impact of potential mismeasurement of income shocks. Finally, we look at the role of time-varying real interest rates.

Local Tax Rates and Amenities

While we assume that income volatility is the only source of high frequency changes in demand in our simulations, changes in local tax rates and amenity flows also could affect volatility. To address volatility related to tax rates, we used data from the NBER TAXSIM website on the average tax rate on wage income earned in a given state each year to create an after-tax income measure for each market. Our analysis of after-tax income showed that controlling for this factor cannot be responsible for more than a 10 percent increase in local demand variability which would translate into a ten percent increase in price and construction volatility. Appendix II provides the details, but we conclude that changes in state level tax rates cannot be driving the high price volatility.

We think of most amenities as being relatively permanent characteristics of a place, (e.g., the weather, local architecture). The demand for these amenities may change slowly over time as a society becomes richer or more unequal or as new technologies become available, but it is hard to imagine that their value will fluctuate a lot at annual frequencies. Crime represents one of the few amenities that does change relatively rapidly and for which there is available data. Since crime represents a possible omitted amenity, we collected violent crime rates for the largest cities in each of our metropolitan areas using continuous crime data from 1985-2005.[30] We then created an adjusted income variable that subtracted the negative effect of crime from our BEA real income measure. As detailed in Appendix III, the results showed almost no impact on the variability of local demand from controlling for crime. We infer from this that we are unlikely to be find an amenity with high frequency variation that can explain much of the observed volatility in prices or construction.

Measurement of Income Shocks

There are two reasons why our estimates of income volatility might be understating the true magnitude of income shocks in high volatility markets. First, our estimate of σ_ε^2 is based on all 115 markets in the sample, and if income variability were systematically higher in the high price change variance markets, then our estimates of price volatility would be biased downward in those markets. Second, our use of BEA per capita income makes no allowance for the possibility that the volatility of the marginal home buyer's income could be relatively high.

To look at the impact of heterogeneous income variability across markets, we focus on coastal markets that have particularly volatile housing prices. We re-estimated σ_ε^2 using the same ARMA procedure described above for a subsample of 31 markets whose centroids are within 50 miles of the Atlantic or Pacific Oceans. While the AR (δ)

[30] We emphasize that this measure is for the local political jurisdiction, which we then impute to the metropolitan area.

and MA (θ) components were little changed from those reported for the 115 market national sample, the estimate of σ_ε^2 is almost 50 percent higher in the 31 coastal markets: $5.3 million versus $3.6 million. This difference in volatility of local wage shocks is large, and to our knowledge, has not been well-documented and is not well-understood. Highly productive coastal areas might specialize in highly volatile idea-intensive industries, but this is an appropriate topic for future research.

The puzzle of excess price change variance also could be at least partially explained if our measure of income variance understates the true year-to-year variation in the returns from living in the city for the marginal homebuyer. For example, if marginal buyers are young, then this might mean that their incomes are more volatile. In cities with vibrant economies, buyers on the margin might be people in the cities' fastest growing and most volatile industries. In New York City, the marginal homebuyer might be more likely to be in finance, for example, and have a more volatile income. In San Francisco, the marginal homebuyer might be more likely to work in the volatile technology sector.

To investigate this hypothesis, we turned to the *New York City Housing and Vacancy Survey* (NYCHVS) which provides extremely detailed information on New York City residents. This survey was available for the years 1978, 1981, 1984, 1987, 1991, 1993, 1996 and 2002. In each year, we calculated the mean income for the sample of people who bought a house within the last two years. This two year window is meant to capture the population of recent buyers, while also providing a decent sample size (e.g., there are between 100 and 400 observations in each year).

The incomes for this sample could be more volatile than the BEA per capita figures simply because of smaller numbers, so our approach is not to look at its variance, but at its correlation with BEA real income in New York. Specifically, we estimated the following regression:

(8) Recent Buyer Real Income=-32,451+ 1.29*BEA Per Capita Real Income,
(15,102) (0.19)

There are nine observations (one for each survey year), the R^2=0.87, and standard errors are in parentheses. This regression suggests that recent buyer income increases by $1.29 for every $1 increase in BEA-measured income.[31] With respect to our simulations, if we assume that the value of x(t), the income shock, for recent owners is 1.29 times the shock for the entire population, then the variance of shocks is 1.66 times the variance for the entire population.[32]

Including both effects implies predicted variances in high (c_1, c_2) markets are 2.4 times greater than those reported in Tables 7 and 8 (1.66*1.44~2.4). Because our evidence certainly is not strong enough to warrant confidence in such a precise result, we report ranges of values for new simulation results assuming the local demand shock (σ_ε^2) is 50 percent, 100 percent, and 150 percent greater than the $3.6 million figure used in

---

[31]While the small samples and nine observations make it hard to draw too much from this, it is noteworthy that this procedure does not generate estimates above $1 automatically. When we performed the same analysis using the sample of renters in the survey, the following resulted:

Renter Real Income= -2,885 + 0.47*BEA Per Capita Real Income
(5,436) (0.07)

The number of observations again is nine, the R^2 still is 0.87, and standard errors are in parentheses.

[32] In addition, if recent buyer incomes were more cyclical than average incomes, then the variance of prices might be substantially higher than the variance predicted by the model. A related point is that if the New York City regressions are to be believed, then the income of renters is far less volatile than the income of owners. The regression predicts that the variance of their income shocks is one quarter of the variance of income shocks for the general population. This difference may be one reason why rents are so stable relative to housing prices, and it also suggests that comparing rents and prices is a very tricky process, indeed.

the baseline simulations reported above. Because the key outstanding price-related puzzle for our model is the high volatility of lower frequency price changes in the top quartile of metropolitan areas, the top panel of Table 9 provides new estimated variances of five-year price changes for markets with the 75th and 90th percentile values of our supply-side parameters (c_1 and c_2), assuming more variable local demand shocks.

The first column reproduces our baseline estimates for these parameter values from Table 7. Recall that those simulation results are well below the $1.17 billion variance in five-year price changes found in the metropolitan area that is in the 75th percentile for all such changes, and not even close to the $3.58 billion variance observed for the 90th percentile market (see column 3, top panel of Table 6 for those values). The next three columns report predicted variances assuming the three higher estimates of local demand variability. The first row of the third column indicates that σ_ε^2 needs to be doubled in order to account for all of the mismatch between predicted and actual variance at the 75th percentile of the distribution (i.e., $1,190/$1,170~1).

While this represents a large increase in local volatility, based on the discussion above, it is not implausible. However, the results in the second row of the top panel of Table 9 indicate that our model still cannot account for the most volatile ten percent of markets. If our baseline σ_ε^2 estimate is off by 150 percent, the predicted variance of $1.71 billion is only 48 percent of the $3.58 billion variance seen in the 90th percentile metropolitan area.

Even the best of models have trouble accounting for the outliers in any distribution. Still, examining the 12 most volatile markets (the top 10 percent of our 115 metropolitan areas), in terms of price changes over five-year periods, shows that what our model just cannot explain is the volatility of coastal California (plus Honolulu).[33] Literally, there are no east coast markets in this group, with the Nassau-Suffolk and Bethesda-Gaithersburg-Frederick metropolitan areas having the 15th and 16th biggest price change volatilities.

The bottom panel of Table 9 reports new estimated variances of construction intensity assuming supply-side parameters (c_1 and c_2) consistent with values associated with the 10th and 25th percentiles of that distribution and assuming a 66 percent increase in σ_ε^2. We focus on the impact in markets with more elastic supply sides because the remaining puzzle about quantities is the very high variation observed in many of the high growth sunbelt markets. We only allow demand side volatility to increase by 66 percent because there is no evidence to suggest their intrinsic income variability is higher than the average we use in the baseline simulations[34], but we would expect the volatility of the marginal buyer's income to be greater than the average we are using in the baseline simulations. Even assuming the NYCHVS results are applicable to these markets, the results show that underestimating local demand variability cannot account for our underestimate of quantity variation for the most volatile markets in this respect.

Time-Varying Interest Rates

We have so far assumed that interest rates are fixed for reasons of tractability, but understand that many authors have claimed that the dramatic rise in house prices,

[33] The top ten percent of the most volatile metropolitan areas in terms of five-year price changes (in ascending order from #104-#115) are as follows: Oakland-Fremont-Hayward, Santa Cruz-Watsonville, San Luis Obispo-Paso Robles, San Diego-Carlsbad-San Marcos, Oxnard-Thousand Oaks-Ventura, Los Angeles-Long Beach-Glendale, San Jose-Sunnyvale-Santa Clara, Salinas, Santa Ana-Anaheim-Irvine, San Francisco-San Mateo-Redwood City, Santa Barbara-Santa Maria, and Honolulu.

[34] Mathematically, it is somewhat smaller given that coastal market income volatility is higher. We ignore this effect for simplicity. It does not change any of our conclusions.

especially in high cost markets, over the past decade is best understood as a response to declining interest rates that make housing in those areas more affordable (e.g., Himmelberg, Mayer and Sinai, 2005). A full treatment of interest rates would require an analysis of long period mortgages and prepayment that lies well beyond the scope of this paper. We can, however, adjust the model modestly to acquire some understanding of the potential impact of time-varying interest rates.

To do so, we decompose interest rates into permanent and transitory components, $\bar{r}$ and $\rho(t)$, resprectively, where $r(t) = \bar{r} + \rho(t)$, and use the approximations $\frac{1}{1+r(t)} \approx \frac{1}{1+\bar{r}}$ and $r(t)(H(t)-C) = \bar{r}(H(t)-C) + \rho(t)(\overline{H}-C)$, where $\overline{H}$ is meant to reflect the average housing price in the city.[35] If we adjust equation (2) for time-varying interest rates using these approximations, equation (2') results:

$$(2')\quad H(t) - \frac{\bar{r}C}{1+\bar{r}} - \frac{E_t(H(t+1))}{1+\bar{r}} = \overline{D} + qt + x(t) - \rho(t)(\overline{H}-C) - \alpha N(t)$$

If we then make the somewhat unrealistic assumption that $\rho(t) = \lambda\rho(t-1) + \eta(t)$ so the difference equation remains linear, the model can be solved in a relatively straightforward fashion.

This results in equation (9)'s description of prices,

$$(9)\quad \begin{aligned} H(t) = \hat{H}(t) &+ \frac{(c_1\phi + \Psi)x(t)}{c_1(\phi-\delta)+\Psi} + \frac{c_1\theta\varepsilon(t)}{c_1(\phi-\delta)+\Psi} - \frac{\alpha(1+r)}{1+r-\phi}\left(N(t) - \hat{N}(t)\right) \\ &- \frac{(c_1\phi+\Psi)(\overline{H}-C)\rho(t)}{c_1(\phi-\lambda)+\Psi} \end{aligned}$$

[35] The first approximation is minor and would have been unnecessary if we assumed that the utility flow was received at the end of the period rather than the beginning of each period. The second approximation eliminates interactions between transitory changes in value and transitory changes in the interest rate and it may be more consequential.

This differs from the price equation in Proposition 1 because of its last term which multiplies $\frac{c_1\phi+\Psi}{c_1(\phi-\delta)+\Psi}$ times the interest rate shock times $\overline{H}-C$ (the gap between average housing prices in the area and construction costs). This term reflects the fact that a decline in interest rates essentially is a positive demand shock for high amenity and productivity places. The shock makes it cheaper to live in such places, pushing up demand and prices.

Since our regressions correct for year effects, this interest rate effect can have no impact on our empirical estimates for the average market which will have prices close to construction costs. The interest rate effect does, however, have the capacity to generate increased variance in both price changes and construction levels for places that are considerably more expensive on average. We consider four different values of $(\overline{H}-C)$: \$25,000, \$50,000, \$100,000 and \$200,000, which over the past 25 years captures most of the range of American metropolitan areas.[36]

We set c_1, and c_2 equal to 6.4 and 1.6, respectively, in order to focus on interest rates effects. We assume that λ=0.90, but experimentation with values as high as 0.95 yield similar results. The variance of interest rates is far more important, and we use a range of standard deviations for $\eta(t)$ from 0.005 to 0.02 which appears to encompass most assessments of the amount of real rate variation.[37]

[36] For example, in 1980 the highest price metropolitan area had a median house value that was about \$170,000 greater than in the median market. The real value of median market's median house price is barely changed between 1980 and 2000. Except for a handful of markets in the upper tail of the metropolitan area price distributions, gaps in excess of \$200,000 with the median market do not exist. Finally, we omit runs with a value of zero because they correspond to the simulations from the previous section.

[37] Campbell's (2000) review of the asset pricing literature notes that the standard deviation on a one period riskless asset is 1.76 percent, but concludes that "… perhaps half … is due to ex post inflation shocks (p.

Simulations suggest that these changes to our baseline model make little difference to the amount of predicted mean reversion. Thus, the results in Tables 10 and 11 focus on the impact of interest rate volatility on the variance of price changes and construction, respectively. Table 10 shows that including interest rates shocks can generate significant increases in the variance of price changes if the interest rate shock is quite high and if the market has house prices much greater than the average (and, thus, much more than construction costs). For example, comparing the predicted variances in Table 10 for markets $25,000 or $50,000 above the average market with those reported in Table 7 (for the same (c_1, c_2) parameters) shows that interest rate volatility does not increase the predicted volatility of price changes much at all. To generate a predicted variance above the $209 million observed over annual periods for the 90th percentile metropolitan area (see Table 6) requires that H-C be $200,000 if the standard deviation of $\eta(t)$ is one percent.

At three year intervals, the predicted variance without interest rate shocks is $330 million (see the middle cell of Table 7), which is higher than the median market (Table 6), but less than the mean and far less than the 90th percentile metropolitan area. Including interest rate shocks with a standard deviation of 0.01 and a $100,000 gap between prices and construction costs increases the predicted variance to $418,000, which is near the volatility observed in the 75th percentile city according to the figures in Table 6. To fit the 90th percentile city's price change variation of $1.38 billion, the gap between average prices and construction costs needs to be $200,000 dollars and the

---

1519)." Thus, the lower half of this range may be more plausible. Recent asset pricing papers such as Bansal, Kiku, and Yaron (2006) assume a standard deviation of 1 percent.

standard deviation of interest rates needs to be much greater than 0.01. We think that both of these assumptions are extreme.

At five year intervals, including interest rate shocks again increases the predicted variation significantly in the most attractive or productive markets, but the predicted variation is still far less than is actually observed in the most volatile markets according to the data in Table 6. A 0.01 standard deviation interest rate shock and a $100,000 gap between housing prices in the city and construction costs increase the predicted variance by about one quarter. To get the much higher variances that we seen in the data, the gulf between prices and construction costs must be over $100,000 and the shock to interest rates much have a standard deviation greater than 0.01.

The ability of interest rate shocks to explain the high level of variation in construction is far more limited. For example, Table 11 documents that even if the standard deviation of shocks is 0.02 and $\overline{H}-C$ equals $200,000, the model still does not predict the sample mean variance of construction. At more reasonable parameter values, such as a standard deviation of interest rate shocks of 0.01 and $\overline{H}-C$ equal to $100,000, interest rate shocks predict almost no increase in the variance of construction at one year intervals, an increase in construction variance of 30 percent at three year intervals and an increase in construction variance of 25 percent at five year intervals. Thus, interest rate shocks do increase predicted construction variance, but they do so modestly.

One clear implication of the model is that if interest rate shocks are important, then the variance of price changes and construction should be higher in high price areas. Figure 4 graphs the variance of one year price changes for each metropolitan area against its average price in 1980. The graph shows a strong positive relationship, just as predicted by the role of interest rates. The most volatile places in the country are places that were most expensive in 1980. Interest rate shocks are one explanation of this phenomenon. However, another possible explanation is that these places had high costs because they restricted construction, so that Figure 4 is showing the impact of restricted construction on volatility. However, a quick comparison of Table 7 and Table 10 shows that interest rates can generate this high level of price volatility more readily than restricted construction without interest rate shocks. We suspect both phenomena are at work in high cost, high volatility areas.

In Figure 5, we graph the variance of one year construction rates on the average price in the metropolitan area in 1980. In this case, there is no visible relationship, perhaps because restrictions on construction in high cost areas ensure low levels of construction volatility. Nonetheless, we think that both Tables 10 and 11 and Figures 4 and 5 suggest that interest rates shocks can plausibly play a role in explaining some of the observed price volatility in high cost areas. Those results also suggest that interest rates are unlikely to explain high construction volatility in lower cost, sunbelt areas.

VI. Conclusion

This paper presents a dynamic rational explanations model of housing markets based on a cross-city spatial equilibrium. The model predicts that housing markets will be largely local, which they are, and that construction persistence is fully compatible with price mean reversion. The model is also consistent with price changes being predictable.

The model has successes and failures at fitting the real data. The model can explain the serial correlation of construction quantities reasonably well and can explain the five year mean reversion of prices almost perfectly. However, the model cannot explain the high frequency positive serial correlation of price changes. The model can explain the price and construction volatility of the nation's typical housing markets, but it does a poor job of explaining the most volatile markets.

Time-varying interest rates can in principle explain some of the price variation in high cost markets, but interest rates are unlikely to explain much of the extreme cases of construction variation. Our simulations showed that interest rate shocks do a better job of fitting the markets with high levels of price variance than the markets with high levels of construction variance. Across cities, price volatility is concentrated in high costs areas, which is a prediction of the model when it includes interest rates. Construction volatility is concentrated in lower cost markets, which the model suggests should have little responsiveness to interest rates.

There are two problems with concluding too much from our interest rate findings. First, on a theoretical level, we have omitted many important features of mortgage contracts such as a the prepayment option which seem crucial to us in understanding the impact the interest rates will have on price dynamics. Second, empirically, interest rates explain only a small portion of price volatility even in high price areas. We hope that future research will focus more on this important topic.

Finally, the value of this model is as much in what it cannot explain as in what it can explain. It suggests that housing economists should focus their attention on high price volatility in coastal markets and on the positive serial correlation of price changes. The average volatility and longer-term mean reversion of prices should no longer be viewed as puzzles.

References

Alonso, William (1964). *Location and Land Use.* Cambridge, MA: Harvard University Press.

Arellano, Manuel and Stephen Bond (1991). "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations", *Review of Economic Studies*, Vol. 58, no. 2 (April): 277-97.

Baker, Dean (2006). "The Menace of an Unchecked Housing Bubble", *The Economists' Voice*, Vol. 3, no. 4, Article 1.

Bansal, Ravi, Dana Kiku, and Amir Yaron (2006). "Risks for the Long Run: Estimation and Inference", working paper, October 2006.

Blanchard, Olivier and Lawrence Katz (1992) "Regional Evolutions" *Brookings Papers on Economic Activity* 1: 1-75.

Borjas, George (2003). "The Labor Demand Curve Is Downward Sloping: Reexamining the Impact of Immigration on the Labor Market", *Quarterly Journal of Economics*, Vol. 188, no. 4 (November): 1335-1374.

Campbell, John (2000). "Asset Pricing at the Millennium", *Journal of Finance*, Vol. LX, no. 4 (August): 1515-1567.

Card, David and Kristin Butcher (1991). "Immigration and Wages: Evidence from the 1980s", *American Economic Review*, Vol. 81, no. 2 (May): 292-296.

Case, Karl E. and Robert Shiller (1989). "The Efficiency of the Market for Single Family Homes", *American Economic Review*, Vol. 79, no. 1: 125-37.

Cutler, David, James Poterba, and Lawrence Summers (1991). "Speculative Dynamics", *Review of Economic Studies*, Vol. 58, no. 3 (May 1991): 529-46.

Fama, Eugene and Kenneth French (1988). "Permanent and Temporary Components of Stock Prices", *Journal of Political Economy*, Vol. 96, no.2 (April 1988): 246-73.

Glaeser, Edward and Joseph Gyourko (2005). "Urban Decline and Durable Housing", *Journal of Political Economy*, Vol. 113, no. 2 (April): 345-375.

Glaeser, Edward, Joseph Gyourko and Raven Saks (2005). "Why Have House Prices Gone Up?", *American Economic Review*, Vol. 95, no. 2 (May): 329-333.

Glaeser, Edward, Joseph Gyourko and Raven Saks (2006). "Urban Growth and Housing Supply", *Journal of Economic Geography*, Vol. 6, no. 1 (January): 71-89.

Glaeser, Edward and Jesse Shapiro (2003). "Urban Growth in the 1990s" Is City Living Back?', *Journal of Regional Science*, Vol. 43 (February): 139-165.

Gyourko, Joseph, Christopher Mayer, and Todd Sinai (2006). "Superstar Cities", NBER Working Paper 12355, July 2006.

Gyourko, Joseph and Albert Saiz (2006). "Construction Costs and the Supply of Housing Structure", *Journal of Regional Science*, Vol. 46, no. 4 (November 2006): forthcoming.

Hansen, Lars and Ravi Jagannathan (1991). "Implications of Security Market Data for Models of Dynamic Economies", *Journal of Political Economy*, Vol. 99, no. 2 (April): 225-262.

Hendershott, Patric and Joel Slemrod (1983). "Taxes and the User Cost of Capital for Owner-Occupied Housing", *Journal of the American Real Estate and Urban Economics Association*, Vol. 10, no. 4 (Winter): 375-93.

Himmelberg, Charles, Christopher Mayer and Todd Sinai (2005). "Assessing High House Prices: Bubbles, Fundamentals, and Misperceptions", *Journal of Economic Perspectives*, Vol. 19, no. 4 (Fall 2005): 67-92.

McCarthy, Jonathan and Richard Peach (2004). "Are Home Prices the Next Bubble?", *Economic Policy Review*, Vol. 10, no. 3: 1-17.

Poterba, James (1984). "Tax Subsidies to Owner-Occupied Housing: An Asset Market Approach", *Quarterly Journal of Economics*, Vol. 99, no. 4: 729-52.

Roback, Jennifer (1982). "Wages, Rents, and the Quality of Life", *Journal of Political Economy*, Vol. 90, no. 4 (December 1982): 1257-78.

Rosen, Sherwin (1979). "Wage-Based Indexes of Urban Quality of Life". In *Current Issues in Urban Economics*, edited by Peter Mieszkowski and Mahlon Straszheim. Baltimore: Johns Hopkins Univerity Press, 1979.

Schwartz, Amy Ellen, Scott Susin and Ioan Voicu (2003). "Has Falling Crime Driven New York's Real Estate Boom?", *Journal of Housing Research*, Vol. 14, no. 1: 101-135.

Shiller, Robert (2005). *Irrational Exuberance* (2nd edition). Princeton University Press.

____________ (2006). "Long-Term Perspectives on the Current Boom in Home Prices", *The Economists' Voice*, Vol. 3, no. 4, Article 4.

Smith, Margaret Hwang and Gary Smith (2006). "Bubble, Bubble, Where's the Housing Bubble", Pomona College mimeo, March 2006.

Thaler, Richard (1978). "A Note on Crime Control: Evidence from the Property Market", *Journal of Urban Economics*, Vol 5, no. 1 (January): 137-145.

Topel, Robert and Sherwin Rosen (1988). "Housing Investment in the United States", *Journal of Political Economy*, Vol 96, no. 4: 718-740.

Tracy, Joseph, Henry Schneider and Sewin Chan (1999). "Are Stocks Overtaking Real Estate in Household Portfolios?", *Current Issues in Economics and Finance*, Vol. 5, no. 5 (April): 1-6

Van Nieuwerburgh, Stijn and Pierre-Olivier Weill (2006). "Why Has House Price Dispersion Gone Up?", NBER Working Paper #12538, September.

Table 1: Model Parameters				
R	0.04			
Δ	0.87			
Θ	0.17			
σ_ε^2	$3,603,463			
α	0.1			
C_1		ω=0.0	ω=0.25	ω=0.50
	10th percentile value:	1.7	1.5	1.0
	25th percentile value:	3.5	2.7	2.0
	50th percentile value:	7.0	6.4	4.8
	75th percentile value:	14.4	13.3	10.3
	90th percentile value:	28.8	28.1	22.9
C_2		ω=0.0	ω=0.25	ω=0.50
	10th percentile value:	0.0	0.4	0.5
	25th percentile value:	0.0	0.7	1.0
	50th percentile value:	0.0	1.6	2.4
	75th percentile value:	0.0	3.3	5.2
	90th percentile value:	0.0	7.0	11.5

Table 2: Variation in Prices and Quantities Within-Market Over Time Arellano-Bond Estimates of Coefficients on Lagged Dependent Variable 1, 3, & 5 year horizons			
Dependent Variable	*1-year changes*	*3-year changes*	*5-year changes*
House Price Change	0.71 (0.01) N=2,819	0.27 (0.04) N=690	-0.32 (0.07) N=345
Rent Change	0.27 (0.03) N=1,007	0.27 (0.08) N=274	-0.64 (0.17) N=91
New Permits	0.84 (0.01) N=2,645	0.43 (0.04) N=690	-0.07 (0.06) N=460

Notes:

1. Sample for house price, employment, and permit specifications is 115 metropolitan area sample described in text.
2. Sample for rent specification is 46 metropolitan areas tracked by REIS.

Table 3: Predicted Serial Correlation in House Prices									
	One-Year			Three-Year			Five-Year		
	c_2 if ω = 0	c_2 if ω= 0.25	c_2 if ω = 0.5	c_2 if ω = 0	c_2 if ω= 0.25	c_2 if ω = 0.5	c_2 if ω = 0	c_2 if ω= 0.25	c_2 if ω = 0.5
lowest c_1	-0.09	-0.11	-0.13	-0.29	-0.28	-0.25	-0.41	-0.35	-0.31
lower c_1	-0.07	-0.09	-0.09	-0.25	-0.25	-0.22	-0.36	-0.32	-0.28
medium c_1	-0.05	-0.07	-0.07	-0.21	-0.21	-0.19	-0.32	-0.29	-0.26
higher c_1	-0.05	-0.05	-0.05	-0.19	-0.19	-0.17	-0.29	-0.27	-0.25
highest c_1	-0.04	-0.05	-0.05	-0.18	-0.18	-0.17	-0.27	-0.26	-0.25

Notes: Within each time horizon over which serial correlation is estimated, 15 simulations were run, corresponding to the 15 (c_1, c_2) pairs reported in Table 1. Thus, the (c_1, c_2) combinations used in the first, fourth, and seventh columns for which ω=0 are (1.7, 0.0), (3.5, 0.0), (7.0, 0.0), (14.4, 0.0), and (28.8, 0.0), going from the top to the bottom of the table. The (c_1, c_2) combinations used in the second, fifth, and eighth columns for which ω=0.25 are (1.5, 0.4), (2.7, 0.7), (6.4, 1.6), (13.3, 3.3), and (28.1, 7.0). The (c_1, c_2) combinations used in the third, sixth, and ninth columns for which ω=0.50 are (1.0, 0.5), (2.0, 1.0), (4.8, 2.4), (10.3, 5.2), and (22.9, 11.5).

Table 4: Predicted Serial Correlation in Apartment Rents									
	One-Year			Three-Year			Five-Year		
	c_2 if $\omega = 0$	c_2 if $\omega = 0.25$	c_2 if $\omega = 0.5$	c_2 if $\omega = 0$	c_2 if $\omega = 0.25$	c_2 if $\omega = 0.5$	c_2 if $\omega = 0$	c_2 if $\omega = 0.25$	c_2 if $\omega = 0.5$
lowest c_1	0.02	-0.05	-0.14	-0.25	-0.31	-0.34	-0.37	-0.39	-0.38
lower c_1	0.05	-0.00	-0.07	-0.21	-0.26	-0.27	-0.32	-0.35	-0.33
medium c_1	0.07	0.04	0.01	-0.17	-0.21	-0.20	-0.29	-0.30	-0.28
higher c_1	0.08	0.07	0.05	-0.15	-0.17	-0.17	-0.26	-0.27	-0.25
highest c_1	0.09	0.08	0.07	-0.14	-0.15	-0.15	-0.24	-0.25	-0.24

Notes: Within each time horizon over which serial correlation is estimated, 15 simulations were run, corresponding to the 15 (c_1, c_2) pairs reported in Table 1. Thus, the (c_1, c_2) combinations used in the first, fourth, and seventh columns for which ω=0 are (1.7, 0.0), (3.5, 0.0), (7.0, 0.0), (14.4, 0.0), and (28.8, 0.0), going from the top to the bottom of the table. The (c_1, c_2) combinations used in the second, fifth, and eighth columns for which ω=0.25 are (1.5, 0.4), (2.7, 0.7), (6.4, 1.6), (13.3, 3.3), and (28.1, 7.0). The (c_1, c_2) combinations used in the third, sixth, and ninth columns for which ω=0.50 are (1.0, 0.5), (2.0, 1.0), (4.8, 2.4), (10.3, 5.2), and (22.9, 11.5).

Table 5: Predicted Serial Correlation in Construction									
	One-Year			Three-Year			Five-Year		
	c_2 if $\omega = 0$	c_2 if $\omega = 0.25$	c_2 if $\omega = 0.5$	c_2 if $\omega = 0$	c_2 if $\omega = 0.25$	c_2 if $\omega = 0.5$	c_2 if $\omega = 0$	c_2 if $\omega = 0.25$	c_2 if $\omega = 0.5$
lowest c_1	0.68	0.52	0.34	0.40	0.19	0.02	0.17	-0.03	-0.15
lower c_1	0.74	0.56	0.36	0.50	0.24	0.04	0.28	0.01	-0.14
medium c_1	0.78	0.60	0.39	0.57	0.29	0.06	0.37	0.06	-0.13
higher c_1	0.81	0.62	0.37	0.63	0.32	0.05	0.45	0.08	-0.13
highest c_1	0.83	0.63	0.40	0.67	0.33	0.07	0.51	0.09	-0.12

Notes: Within each time horizon over which serial correlation is estimated, 15 simulations were run, corresponding to the 15 (c_1, c_2) pairs reported in Table 1. Thus, the (c_1, c_2) combinations used in the first, fourth, and seventh columns for which ω=0 are (1.7, 0.0), (3.5, 0.0), (7.0, 0.0), (14.4, 0.0), and (28.8, 0.0), going from the top to the bottom of the table. The (c_1, c_2) combinations used in the second, fifth, and eighth columns for which ω=0.25 are (1.5, 0.4), (2.7, 0.7), (6.4, 1.6), (13.3, 3.3), and (28.1, 7.0). The (c_1, c_2) combinations used in the third, sixth, and ninth columns for which ω=0.50 are (1.0, 0.5), (2.0, 1.0), (4.8, 2.4), (10.3, 5.2), and (22.9, 11.5).

Table 6: Variance in House Price Changes and Construction Intensity 1, 3, and 5 Year Horizons			
	House Price Change Variance (millions of $2000)		
	1 year	3 years	5 years
10th percentile market	$14	$69	$183
25th percentile market	$26	$124	$452
50th percentile market	$34	$185	$625
75th percentile market	$70	$445	$1,170
90th percentile market	$209	$1,380	$3,580
Sample mean	$83	$484	$1,310
	Construction Intensity Variance (millions of units)		
	1 year	3 years	5 years
10th percentile market	2	13	29
25th percentile market	2	19	41
50th percentile market	3	26	59
75th percentile market	11	84	212
90th percentile market	38	328	760
Sample mean	21	160	417

Table 7: Predicted Variance of Price Changes ($millions) 1, 3, and 5 Year Horizons									
	One-Year			Three-Year			Five-Year		
	c_2 if $\omega = 0$	c_2 if $\omega = 0.25$	c_2 if $\omega = 0.5$	c_2 if $\omega = 0$	c_2 if $\omega = 0.25$	c_2 if $\omega = 0.5$	c_2 if $\omega = 0$	c_2 if $\omega = 0.25$	c_2 if $\omega = 0.5$
lowest c_1	50	66	64	123	154	148	163	203	200
lower c_1	70	90	97	178	221	239	246	303	337
medium c_1	91	128	138	239	330	360	340	469	520
higher c_1	114	155	166	306	411	442	443	595	647
highest c_1	136	175	181	367	468	488	538	685	718

Notes: Within each time horizon over which serial correlation is estimated, 15 simulations were run, corresponding to the 15 (c_1, c_2) pairs reported in Table 1. Thus, the (c_1, c_2) combinations used in the first, fourth, and seventh columns for which ω=0 are (1.7, 0.0), (3.5, 0.0), (7.0, 0.0), (14.4, 0.0), and (28.8, 0.0), going from the top to the bottom of the table. The (c_1, c_2) combinations used in the second, fifth, and eighth columns for which ω=0.25 are (1.5, 0.4), (2.7, 0.7), (6.4, 1.6), (13.3, 3.3), and (28.1, 7.0). The (c_1, c_2) combinations used in the third, sixth, and ninth columns for which ω=0.50 are (1.0, 0.5), (2.0, 1.0), (4.8, 2.4), (10.3, 5.2), and (22.9, 11.5).

Table 8: Predicted Variance of Construction Intensity (millions of units) 1, 3, and 5 Year Horizons									
	One-Year			Three-Year			Five-Year		
	c_2 if $\omega = 0$	c_2 if $\omega = 0.25$	c_2 if $\omega = 0.5$	c_2 if $\omega = 0$	c_2 if $\omega = 0.25$	c_2 if $\omega = 0.5$	c_2 if $\omega = 0$	c_2 if $\omega = 0.25$	c_2 if $\omega = 0.5$
lowest c_1	15	23	42	101	128	189	221	250	328
lower c_1	7	12	19	48	68	87	110	135	153
medium c_1	3	3	5	21	21	25	49	44	44
higher c_1	1	1	1	8	7	7	19	14	12
highest c_1	0.4	0.3	0.3	3	2	2	7	4	3

Notes: Within each time horizon over which serial correlation is estimated, 15 simulations were run, corresponding to the 15 (c_1, c_2) pairs reported in Table 1. Thus, the (c_1, c_2) combinations used in the first, fourth, and seventh columns for which ω=0 are (1.7, 0.0), (3.5, 0.0), (7.0, 0.0), (14.4, 0.0), and (28.8, 0.0), going from the top to the bottom of the table. The (c_1, c_2) combinations used in the second, fifth, and eighth columns for which ω=0.25 are (1.5, 0.4), (2.7, 0.7), (6.4, 1.6), (13.3, 3.3), and (28.1, 7.0). The (c_1, c_2) combinations used in the third, sixth, and ninth columns for which ω=0.50 are (1.0, 0.5), (2.0, 1.0), (4.8, 2.4), (10.3, 5.2), and (22.9, 11.5).

Table 9: The Impact of Greater Local Demand Variability				
	Five-year Price Change Variance ($millions)			
	Baseline σ_ε^2 (=$3.6)	Baseline σ_ε^2 * 1.5	Baseline σ_ε^2 * 2.0	Baseline σ_ε^2 * 2.5
Market with 75th percentile (c_1, c_2) values (c_1=13.3, c_2=3.3)	595	893	1,190	1,488
Market with 90th percentile (c_1, c_2) values (c_1=28.8, c_2=7.0)	685	1,028	1,370	1,713
	Five-Year Quantity Change Variance (millions of units)			
	Baseline σ_ε^2 (=$3.6)	Baseline σ_ε^2 * 1.66		
Market with 25th percentile (c_1, c_2) values (c_1=2.7, c_2=0.7)	135	224		
Market with 10th percentile (c_1, c_2) values (c_1=1.5, c_2=0.4)	250	415		

Notes: Unless noted explicitly in the table, each simulation reported above using common parameter values for the national sample as reported in Table 1. In each case, ω=0.25.

Table 10: Predicted Variance of Price Changes ($millions): Interest Rate Volatility 1, 3, and 5 Year Horizons									
	One-Year			Three-Year			Five-Year		
$\overline{H} - C$	**Stand. Dev. (η) = 0.005**	**Stand. Dev. (η) = 0.01**	**Stand. Dev. (η) = 0.02**	**Stand. Dev. (η) = 0.005**	**Stand. Dev. (η) = 0.01**	**Stand. Dev. (η) = 0.02**	**Stand. Dev. (η) = 0.005**	**Stand. Dev. (η) = 0.01**	**Stand. Dev. (η) = 0.02**
$25,000	128	130	136	331	335	352	471	477	501
$50,000	130	136	162	335	352	418	477	501	597
$100,000	136	162	266	352	418	683	501	597	980
$200,000	162	266	681	418	683	1743	597	980	2513

Table 11: Predicted Variance of Construction (millions of units): Interest Rate Volatility 1, 3, and 5 Year Horizons									
	One-Year			Three-Year			Five-Year		
$\overline{H} - C$	**Stand. Dev. (η) = 0.005**	**Stand. Dev. (η) = 0.01**	**Stand. Dev. (η) = 0.02**	**Stand. Dev. (η) = 0.005**	**Stand. Dev. (η) = 0.01**	**Stand. Dev. (η) = 0.02**	**Stand. Dev. (η) = 0.005**	**Stand. Dev. (η) = 0.01**	**Stand. Dev. (η) = 0.02**
$25,000	4	4	4	21	22	23	44	45	47
$50,000	4	4	4	22	23	28	45	47	57
$100,000	4	4	7	23	28	46	47	57	97
$200,000	5	7	19	28	46	121	57	97	255

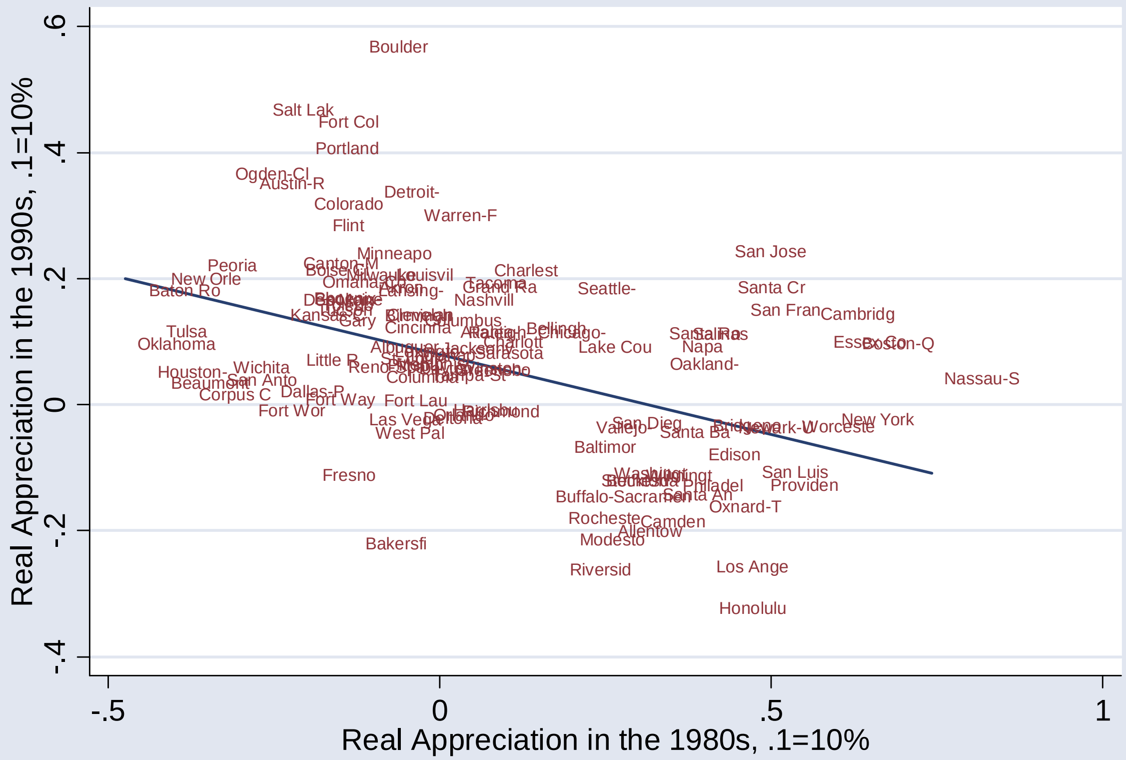
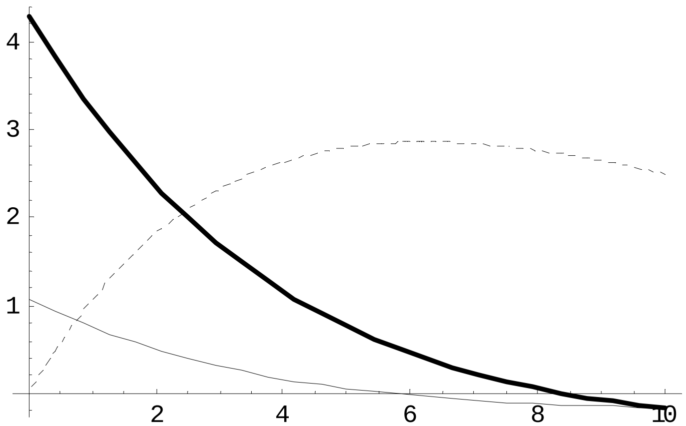


Figure 1: Real House Price Appreciation in the 1980s and 1990s

Figure 2: Housing Unit Growth in the 1980s and 1990s

Figure 3: One-Time Shock



Population: ------------
Construction: ————
Price: ━━━━

($\alpha = 0.1$, $c_1 = 3$, $c_2 = 0.1$)

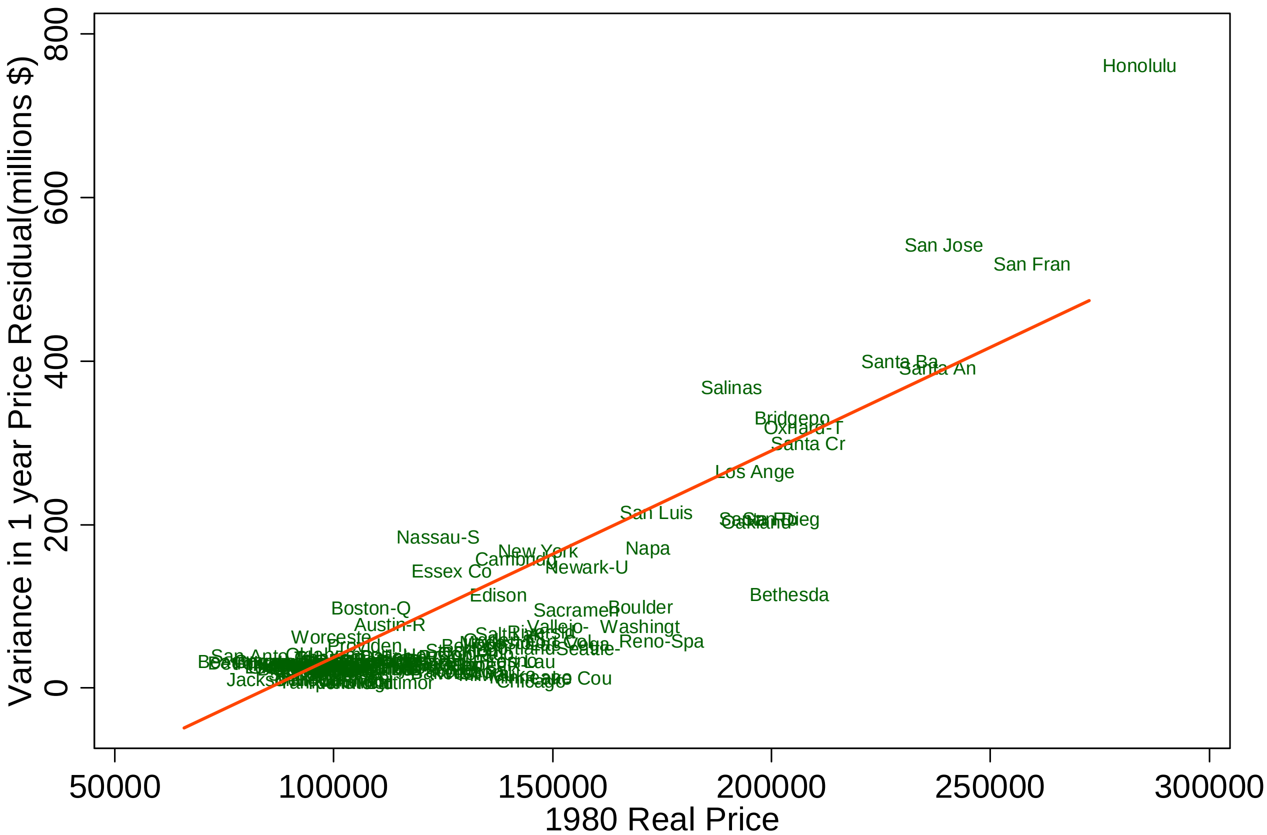


Figure 4: Variance in 1 yr Price Residual Vs. 1980 Price

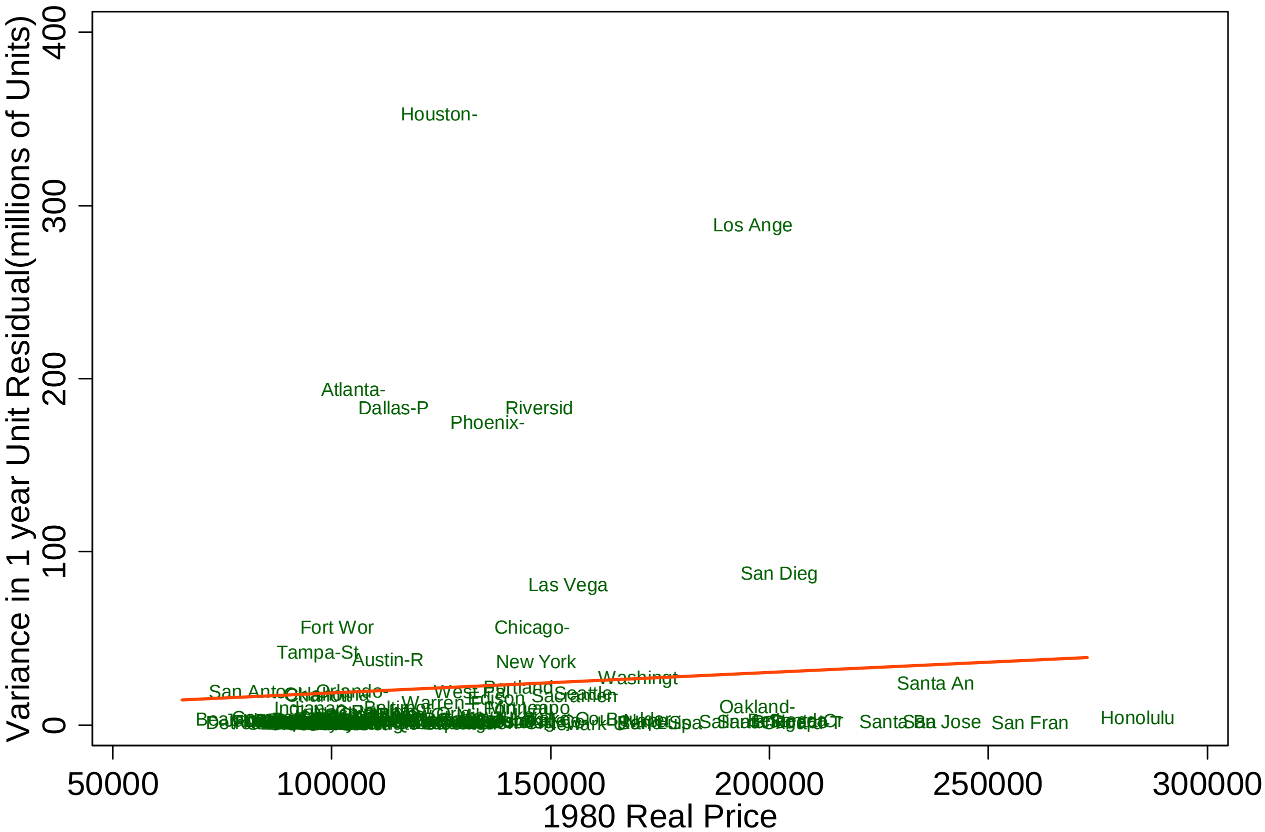


Figure 5: Variance in 1 yr Unit Residual Vs. 1980 Price

Appendix 1: Proofs of Propositions

Proof of Proposition 1: We use the change of variables $I(t) = m(t) + \hat{I}(t)$, $N(t) = n(t) + \hat{N}(t)$, and $H(t) = z(t) + \hat{H}(t)$. Substituting in our definitions of $\hat{I}$, $\hat{N}$, and $\hat{H}$, we reduce the core pricing equation

$H(t) = \overline{D} + qt + x(t) - \alpha N(t) + \frac{rC}{1+r} + \frac{E_t(H(t+1))}{1+r}$ to

$$z(t) = x(t) - \alpha n(t) + \frac{E_t(z(t+1))}{1+r}, \text{ (*)}$$

the optimality condition for production $C + c_0(t+1) + c_1 I(t+1) + c_2 N(t) = E_t(H(t+1))$ to

$$c_1 m(t+1) + c_2 n(t) = E_t(z(t+1)), \text{ (**)}$$

and the defining equation $I(t+1) = N(t+1) - N(t)$ to

$$m(t+1) = n(t+1) - n(t). \text{ (***)}$$

We seek functions *n*, *z*, and *m* that satisfy the starred equations.

Define $u \equiv \frac{\alpha}{c_1}$ and $v \equiv \frac{c_2}{c_1}$; $0 \le u$ and $0 \le v < 1$ by the conventions in force. Then

$$\phi = \frac{1}{2}\left(2 + r + (1+r)u - v - \sqrt{r^2 + v^2 + 2(1+r)(2+r)u + (1+r)^2 u^2 + 2v(r-(1+r)u)}\right)$$

and

$$\overline{\phi} = \frac{1}{2}\left(2 + r + (1+r)u - v + \sqrt{r^2 + v^2 + 2(1+r)(2+r)u + (1+r)^2 u^2 + 2v(r-(1+r)u)}\right).$$

Because $0 \le v < 1$, the expression under the radical is positive. Note that $\phi + \overline{\phi} = 2 + r + (1+r)u - v > (1+r)(1+u) > 0$ and $\phi\overline{\phi} = (1+r)(1-v) > 0$, so $\phi, \overline{\phi} > 0$.

Also, $\sqrt{r^2 + v^2 + 2(1+r)(2+r)u + (1+r)^2 u^2 + 2v(r-(1+r)u)} >$

$$\sqrt{2(1+r)(2+r)u + (1+r)^2 u^2 - 2(1+r)u} > \sqrt{2(1+r)(1+r)u + (1+r)^2 u^2} > 1 + r,$$

so

$$\overline{\phi} > \frac{1}{2}\left(1 + r + (1+r)u + 1 + r\right) \ge 1 + r > 1,$$

which in turn gives

$$0 \le \phi = \frac{(1+r)(1-v)}{\overline{\phi}} < 1 - v \le 1.$$

Now define *n* by the difference equation

$$n(t) - \phi n(t-1) = \frac{1+r}{c_1(\overline{\phi} - \delta)} E_{t-1}(x(t)). \text{ (1)}$$

A unique solution *n* exists because $\overline{\phi} > 1 > \delta$ ensures $\overline{\phi} - \delta \ne 0$ and because $|\phi| < 1$ allows us to solve for *n* explicitly as

$$n(t) = \frac{1+r}{c_1(\bar{\phi} - \delta)} \sum_{i=0}^{\infty} \phi^i L^i E_{t-1}(x(t)) \quad (2),$$

where L denotes the lag operator. Now that we have defined n, we set

$$z(t) \equiv x(t) + \frac{1}{\bar{\phi} - \delta} E_t(x(t+1)) - \frac{\alpha(1+r)}{1+r-\phi} n(t) \quad (3)$$

and

$$m(t+1) \equiv \frac{1+r}{c_1(\bar{\phi} - \delta)} E_t(x(t+1)) - (1-\phi)n(t). \quad (4)$$

With these choices for z and m, (*) reduces to

$$\frac{\alpha}{1+r-\phi}\left(n(t+1) - \phi n(t)\right) = \frac{\bar{\phi} - 1 - r}{(\bar{\phi} - \delta)(1+r)} E_t(x(t+1)), \quad (5)$$

which by (1) is equivalent to

$$\frac{u(1+r)}{(1+r-\phi)(\bar{\phi} - \delta)} E_t(x(t+1)) = \frac{\bar{\phi} - 1 - r}{(\bar{\phi} - \delta)(1+r)} E_t(x(t+1)),$$

which is true, as one sees from cross-multiplying the coefficients and using the previously established formulas for the product and sum of ϕ and $\bar{\phi}$. (**) reduces to

$$\frac{\alpha(1+r)}{1+r-\phi}\left(n(t+1) - \frac{1+r-\phi}{\alpha(1+r)}\left(c_1(1-\phi) - c_2\right)n(t) \right) = \frac{\bar{\phi} - 1 - r}{\bar{\phi} - \delta} E_t(x(t+1)),$$

which is equivalent to (5), and thus true, because

$$\phi = \frac{1+r-\phi}{\alpha(1+r)}\left(c_1(1-\phi) - c_2\right) = \frac{(1+r-\phi)(1-\phi-v)}{u(1+r)},$$

itself evident from cross-multiplying and using the fact that ϕ satisfies the quadratic equation $y^2 - (2 + r + (1+r)u - v)y + (1+r)(1-v) = 0$. Finally, (***) reduces to (1). This shows that our choices for n, z, and m solve the starred equations. To recover Proposition 1, we use $I(t) = m(t) + \hat{I}(t)$, $N(t) = n(t) + \hat{N}(t)$, and $H(t) = z(t) + \hat{H}(t)$. [The result then follows from $E_t(x(t+1)) = \delta x(t) + \theta\varepsilon(t)$.]

Proof of Proposition 2: First note that by induction on $i \geq 1$, $E_t(x(t+i)) = E_t(\delta x(t+i-1) + \theta\varepsilon(t+i-1)) = \delta E_t(x(t+i-1)) = \delta^{i-1} E_t(x(t+1))$, so from (2),

$$\begin{aligned}
E_t(n(t+j)) &= \frac{1+r}{c_1(\bar{\phi} - \delta)} E_t \sum_{i=0}^{\infty} \phi^i L^i E_{t+j-1}(x(t+j)) \\
&= \frac{1+r}{c_1(\bar{\phi} - \delta)} \left(\sum_{i=j}^{\infty} \phi^i L^i E_{t+j-1}(x(t+j)) + \sum_{i=0}^{j-1} \phi^i E_t(x(t+j-i)) \right) \\
&= \frac{1+r}{c_1(\bar{\phi} - \delta)} \left(\phi^j \sum_{i=0}^{\infty} \phi^i L^i E_{t-1}(x(t)) + \sum_{i=0}^{j-1} \phi^i \delta^{j-1-i} E_t(x(t+1)) \right) \\
&= \phi^j n(t) + \frac{1+r}{c_1(\bar{\phi} - \delta)} \frac{\phi^j - \delta^j}{\phi - \delta} E_t(x(t+1)). \quad (6)
\end{aligned}$$

Using (4) and (6), we next find that

$$E_t(m(t+j)) = \frac{1+r}{c_1(\bar{\phi}-\delta)} E_t(x(t+j)) - (1-\phi)E_t(n(t+j-1))$$

$$= \frac{1+r}{c_1(\bar{\phi}-\delta)}\left(\delta^{j-1} - (1-\phi)\frac{\phi^{j-1}-\delta^{j-1}}{\phi-\delta}\right)E_t(x(t+1)) - \phi^{j-1}(1-\phi)n(t)$$

$$= \frac{1+r}{c_1(\bar{\phi}-\delta)}\left(\frac{\delta^{j-1}(1-\delta)-\phi^{j-1}(1-\phi)}{\phi-\delta}\right)E_t(x(t+1)) - \phi^{j-1}(1-\phi)n(t)\,. \quad (7)$$

Finally, using (**), (6), and (7), we get

$$E_t(z(t+j)) = c_1E_t(m(t+j)) + c_2E_t(n(t+j-1))$$

$$= \frac{1+r}{c_1(\bar{\phi}-\delta)}\left(c_1\frac{\delta^{j-1}(1-\delta)-\phi^{j-1}(1-\phi)}{\phi-\delta} + c_2\frac{\phi^{j-1}-\delta^{j-1}}{\phi-\delta}\right)E_t(x(t+1))$$

$$- \phi^{j-1}\left(c_1(1-\phi)-c_2\right)n(t)$$

$$= \frac{1+r}{\bar{\phi}-\delta}\left(\frac{\delta^{j-1}(1-v-\delta)-\phi^{j-1}(1-v-\phi)}{\phi-\delta}\right)E_t(x(t+1)) - c_1\phi^{j-1}(1-v-\phi)n(t)\,. \quad (8)$$

To recover Proposition 2, we use $I(t) = m(t) + \hat{I}(t)$, $N(t) = n(t) + \hat{N}(t)$, and $H(t) = z(t) + \hat{H}(t)$ with equations (3), (8), (4), (7), and (6).

Proof of Proposition 3: Given the hypotheses, we have $x(t) = \delta x(t-1) + \theta\varepsilon(t-1) + \varepsilon(t) = \varepsilon(t) > 0$, $E_t x(t+1) = \delta x(t) + \theta\varepsilon(t) = (\delta+\theta)\varepsilon(t) > 0$, and $n(t) = 0$, so from (3) and (4) we deduce that $z(t) > 0$ and $m(t+1) > 0$: prices and investment will initially be higher than steady state levels. By assumption, $v = 0$, so by (7) and (8), each of expected time $t+j$ construction, $E_t(m(t+j))$, and expected time $t+j$ price, $E_t(z(t+j))$, is negative if and only if

$$\frac{\delta^{j-1}(1-\delta)-\phi^{j-1}(1-\phi)}{\phi-\delta} < 0\,. \quad (9)$$

If $\phi > \delta$, then (9) holds if and only if

$$\frac{1-\delta}{1-\phi} < \left(\frac{\phi}{\delta}\right)^{j-1},$$

which holds for sufficiently large j because $\phi/\delta > 1$. If $\phi < \delta$, then (9) holds if and only if

$$\frac{1-\phi}{1-\delta} < \left(\frac{\delta}{\phi}\right)^{j-1},$$

which holds for sufficiently large j because $\delta/\phi > 1$. If $\phi = \delta$, then we reduce (9) to

$$0 < \frac{\phi^j - \delta^j - (\phi^{j-1}-\delta^{j-1})}{\phi-\delta} = \frac{(\phi-\delta)\left(\sum_{i=0}^{j-1}\phi^i\delta^{j-1-i} - \sum_{i=0}^{j-2}\phi^i\delta^{j-2-i}\right)}{\phi-\delta} = \phi^{j-2}(j\phi - j + 1),$$

which holds for sufficiently large j because $0 \le \phi < 1$. This shows that there exists j^* such that for all $j > j^*$, time t expected values of time $t+j$ construction and housing prices will lie below steady state levels. When $\varepsilon(t) < 0$, we swap $>$ and $<$ to recover the symmetric case.

Proof of Proposition 4: By assumption, $n(0) = 0$, so from (1), we have

$$n(1) = \frac{(1+r)\delta}{c_1(\bar{\phi} - \delta)}\varepsilon(0). \quad (10)$$

From (***), $m(1) = n(1)$, and from (4) and (10),

$$m(2) = \frac{(1+r)\delta}{c_1(\bar{\phi} - \delta)}\big((\delta + \phi - 1)\varepsilon(0) + \varepsilon(1)\big).$$

By definition, $I(t) = \dfrac{q(1+r) - rc_0}{rc_2 + \alpha(1+r)} + m(t)$, so

$$Cov\big(I(2), I(1)\big) = \left(\frac{1+r}{rc_2 + \alpha(1+r)}\right)^2 Var(q) + \left(\frac{(1+r)\delta}{c_1(\bar{\phi} - \delta)}\right)^2 (\delta + \phi - 1)Var(\varepsilon),$$

which is positive if and only if

$$\frac{Var(q)}{Var(\varepsilon)} > (1 - \delta - \phi)\left(\frac{\delta\big(rc_2 + \alpha(1+r)\big)}{c_1(\bar{\phi} - \delta)}\right)^2.$$

From (3) and (10),

$$z(0) = \frac{\bar{\phi}}{\bar{\phi} - \delta}\varepsilon(0) \quad (11)$$

and

$$z(1) = \frac{1}{\bar{\phi} - \delta}\left(\left(\bar{\phi} - \frac{u(1+r)^2}{1+r-\phi}\right)\delta\varepsilon(0) + \bar{\phi}\varepsilon(1)\right). \quad (12)$$

To compute $z(2)$, we use (1) and (10) to get

$$n(2) = \frac{(1+r)\delta}{c_1(\bar{\phi} - \delta)}\big((\phi + \delta)\varepsilon(0) + \varepsilon(1)\big),$$

which yields via (3) that

$$z(2) = \frac{1}{\bar{\phi} - \delta}\left(\left(\bar{\phi}\delta - \frac{u(1+r)^2(\phi + \delta)}{1+r-\phi}\right)\delta\varepsilon(0) + \left(\bar{\phi} - \frac{u(1+r)^2}{1+r-\phi}\right)\delta\varepsilon(1) + \bar{\phi}\varepsilon(2)\right). \quad (13)$$

By definition, $H(t) = \hat{H}(0) + \dfrac{(1+r)(\alpha c_0 + qc_2)}{rc_2 + \alpha(1+r)}t + z(t)$, so from (11), (12), and (13),

$$Cov\big(H(2) - H(1), H(1) - H(0)\big) = \left(\frac{(1+r)c_2}{rc_2 + \alpha(1+r)}\right)^2 Var(q)$$
$$- \frac{\left(\dfrac{u(1+r)^2\delta}{1+r-\phi} + (1-\delta)\bar{\phi}\right)\left(\dfrac{u(1+r)^2\delta}{1+r-\phi}(1 - \delta - \phi) + (1 - \delta + \delta^2)\bar{\phi}\right)}{(\bar{\phi} - \delta)^2} Var(\varepsilon),$$

which is negative if and only if

$$\Omega\left(\frac{rc_2+\alpha(1+r)}{(1+r)c_1c_2(\bar{\phi}-\delta)}\right)^2 > \frac{Var(q)}{Var(\varepsilon)}.$$

Appendix II: The Contribution of Taxes Local Demand Variance

Data on the average tax rate paid each year in each state was matched to our metropolitan areas using files from the NBER's TaxSim web page. We then multiplied our income numbers by one minus the average tax rate, and calculated new values of δ, θ and σ_ε^2 for this adjusted after-tax income measure. The new "after-tax" values of the three parameters are very similar to those used in our simulations: δ=0.87, θ=0.18, and σ_ε^2=\$3.3 million. The latter is 92 percent of the \$3.6 million figure obtained without any adjustment for taxes. Hence, correcting for taxes creates an eight percent reduction in the variance and almost no change in the other parameters. Consequently, we conclude that including state level tax rates does not offer any hope of explaining the particularly high volatilities.

Appendix III: The Contribution of Crime to Local Demand Variance

We began by drawing on the hedonic literature on the costs of crime. The range of estimates of the elasticity of property value with respect to the violent crime rate ran from 0.05 to 0.15.[38] To turn these housing price elasticities into estimates of the impact of crime on the flow of utility measured in dollar units, we multiply the elasticity by the average housing price per crime to obtain a relationship between the price of housing and the level of crime. We then followed our model and multiplied this figure by r/(1+r) to generate an estimate of the impact of crime on the flow of utility measured in dollars.

Using this method, our elasticity range from 0.05 to 0.15 implies that the impact of violent crime on the flow of well-being ranges from \$35 to \$105. The upper bound estimate of \$105 dollars implies that, if the violent crime rate in a city increases from 12 violent crimes per 1,000 inhabitants (the national mean) to 24 violent crimes per 1,000 inhabitants, then this is equivalent to an income loss of about \$1,260 dollars, which we believe is a reasonable result.

We then used this upper bound impact to adjust the underlying BEA real income variable and δ, θ and σ_ε^2. As with taxes, crime had little impact on the volatility of the local income shock. Specifically, there is only a 1.4 percent greater shock variance when controlling for crime.[39] While the crime data is far from perfect for our purposes, this exercise leads us to conclude that variation in local amenities will explain little of the high variance price change or construction markets.

---

[38] See Thaler (1978) for the lower bound estimate and Schwartz, Susin, and Voicu (2003) for the upper bound number.

[39] We were able to obtain crime data for the major cities of 105 of our 115 metropolitan areas. The ARMA estimates of δ and θ are virtually unchanged depending upon whether income is adjusted for crime in these 105 markets. As noted, the variability of the 'after-crime' income shock is marginally higher.