Online Appendices for "Understanding the Puzzling Risk-Return Relationship for Housing"

Lu Han

September 10, 2012

Appendix I. A Housing-Consumption-Based Asset Pricing Model

In this appendix, I present a simple consumption-based asset pricing model that incorporates housing to illustrate the idea presented in Section 3.1. The model consists of a large number of infinitely-lived households that enjoy both housing services h_t and a nondurable, numeraire consumption good c_t. Given the discount factor β, preferences are

$$E\left(\sum_{t=0}^{\infty}\beta^t u(c_t, h_{t-1})\right) \tag{1}$$

where $E(\cdot)$ indicates the expectation conditional on the information at time 0. The period utility, $u(c_t, h_{t-1})$, is separable, increasing and concave in both c_t and h_{t-1}. There are two assets in the economy. One is a bond s_t, which represents a risk-free financial asset. The riskless return at time t is denoted as r_t^f. Following the convention, I assume that the bonds are in positive net supply. The other is housing capital h_t, which represents a real asset. One unit of housing stock a household purchased at the end of $t-1$ produces one unit of housing services at the beginning of t. Thus households consume h_{t-1} at time t.[1] Housing assets trade at the price P_t and depreciate at the rate δ. Let H_t denote the aggregate supply of housing at time t. For convenience, the model assumes away transaction costs.

There are N_t identical households. All households enter period 0 with initial bond endowment s_0 and housing stock h_0. At time t, each household is endowed with labor income y_t, which is defined in units of perishable non-durable consumption. The budget constraint is given by

$$c_t + P_t h_t + s_t = y_t + s_{t-1}(1 + r_t^f) + P_t h_{t-1}(1-\delta) \tag{2}$$

The stochastic house price and labor income are the key sources of uncertainty in this model, with properties to be specified later. Markets are incomplete in the sense that we close down other financial asset markets, including insurance markets. To finance future durable and non-durable

[1] The timing convention, although non-standard, is innocuous (see Piazzesi, Schneider and Tuzel, 2007). It is made to allow an analytical solution that characterizes the risk-return relationship.

consumption and insure against house price risk and income risk, households can transfer wealth over time through bond and housing holdings. The low correlation between housing and existing financial assets suggests that housing markets are highly incomplete and that adding other assets does not help lower the risk associated with house prices.

An equilibrium is a collection of processes (c_t, s_t, h_t) and price process $\{r_t^f, P_t\}$ such that (i) (c_t, h_t, s_t) solve the household's problem of maximizing utility (1) subject to the budget constraint (2) and (ii) markets clear for bonds, housing and non-durable consumption.

Following the standard capital asset pricing models with fixed asset supply, I first consider an endowment economy where housing supply, H_t, is exogenous and fixed. The Lagrangian for the household's problem is

$$L = E\left(\sum_{t=0}^{\infty}\left(\beta^t u(c_t, h_{t-1}) - \mu_t\left[c_t + s_t + P_t \cdot h_t - (1+r_t^f)s_{t-1} - P_t h_{t-1}(1-\delta)\right]\right)\right) \tag{3}$$

where μ_t is the multiplier on the budget constraint. The first order conditions are

$$\begin{aligned} \frac{\partial L}{\partial s_t} &= -\mu_t + \beta E_t \mu_{t+1}(1+r_{t+1}^f) = 0 \\ \frac{\partial L}{\partial c_t} &= \beta^t\left(u_1(c_t, h_{t-1}) - \mu_t\right) = 0 \\ \frac{\partial L}{\partial h_t} &= \beta^t\left\{-\mu_t P_t + \beta E_t\left[\mu_{t+1}\left((1-\delta)P_{t+1} + u_2(c_{t+1}, h_t)\right)\right]\right\} = 0 \end{aligned}$$

For an equilibrium to exist, the absence of arbitrage is necessary. This in turn implies the existence of a strictly positive pricing kernel, which represents the present value of a unit of numeraire one period ahead:

$$M_{t+1} = \beta\frac{u_1(c_{t+1}, h_t)}{u_1(c_t, h_{t-1})} \tag{4}$$

After substituting the pricing kernel into first order conditions, we arrive at expressions for the risk-free interest rate and house price:

$$1 = E_t\left[(1+r_{t+1}^f)M_{t+1}\right] \tag{5}$$

$$P_t = E_t\left[M_{t+1}\left(P_{t+1}(1-\delta) + \frac{u_2(c_{t+1}, h_t)}{u_1(c_{t+1}, h_t)}\right)\right] \tag{6}$$

Equation (5) defines the interest rate in equilibrium. Equation (6) is the central housing pricing formula. The price of housing asset, P_t, is the expected present value of future price, plus the marginal utility derived from housing consumption in the following period, discounted back by M_{t+1}. While the first term in the bracket represents the financial gains associated with the current home ownership, the second term can be interpreted as the rent savings from holding one unit of housing at time $t+1$. To show this, assume there is a rental market in which the time-$(t+1)$ rent for one unit of housing service is Q_{t+1}. Hypothetically, at time $t+1$ the household rents housing service h_t for just one period. In this case, the household chooses the optimal level of

h_t to minimize expenditure $c_{t+1} + Q_{t+1} h_t$ to achieve a desired utility level, $u(c_{t+1}, h_t) = \bar{u}$. This yields the following first order condition:

$$Q_{t+1} = \frac{u_2(c_{t+1}, h_t)}{u_1(c_{t+1}, h_t)} \tag{7}$$

Thus, the pricing relation can be written as

$$P_t = E_t \left[M_{t+1} \left(P_{t+1}(1-\delta) + Q_{t+1} \right) \right] \tag{8}$$

Hence, the price P_t of the housing asset is the expected present value of its future sale price, adjusted for depreciation, plus one-period rent at time $t+1$. The housing pricing relation (8) represents a close parallel to the basic asset pricing relation in the standard CCAPM of Lucas (1978), in which the role of future dividends associated with financial securities is replaced by future rents associated with home ownership in our context. Defining the total housing return as $R_{t+1} = \frac{P_{t+1}(1-\delta)+Q_{t+1}}{P_t}$, Equation (8) can now be rewritten as

$$E_t(R_{t+1} M_{t+1}) = 1 \tag{9}$$

The model presented so far has been fairly unrestrictive. To solve the model analytically, I now make assumptions about the utility function and the labor income process. Following Chambers, Garriga and Schlagenhauf (2009), I assume that the preferences are given by

$$u(c_t, h_{t-1}) = \gamma \frac{c_t^{1-\tau_1}}{1-\tau_1} (1-\gamma) \frac{h_{t-1}^{1-\tau_2}}{1-\tau_2} \tag{10}$$

For convenience, I assume that the income growth follows a stochastic martingale process:

$$\ln \frac{y_{t+1}}{y_t} = \frac{1}{\tau_1}(\ln \beta + r_t^f) + \frac{1}{2\tau_1}\sigma_{y,t}^2 + \frac{1}{\tau_1}\epsilon_{y,t+1} \tag{11}$$

with $\epsilon_{y,t+1} | I_t \sim N(0, \sigma_{y,t}^2)$.[2] Solving the household maximization problem yields $M_{t+1} = \beta \left(\frac{y_{t+1}}{y_t} \right)^{-\tau_1}$. The separability of the utility function implies that $\sigma_{M,t}^2$ is driven purely by uncertainties in per capita labor income, that is, $\sigma_{M,t}^2 = \sigma_{y,t}^2$. It then follows that

$$M_{t+1} = exp\left(-r_t^f - \frac{1}{2}\sigma_{M,t}^2 + \epsilon_{M,t+1} \right) \tag{12}$$

where $\epsilon_{M,t+1} | I_t \sim N(0, \sigma_{M,t}^2)$.

Having obtained M_{t+1} as an equilibrium outcome, I now turn to the process of house prices and rents. Let $g_{t+1}^P = \ln \frac{P_{t+1}}{P_t}$ be the growth rate of the current house's price. Let $g_{t+1}^Q = \ln \frac{Q_{t+1}}{Q_t}$ be the growth rate of the future house's rent. Both follow a stochastic process:

$$g_{t+1}^j = \alpha_0^j + \alpha_1^j g_t^j + \epsilon_{j,t+1} \tag{13}$$

$$\epsilon_{j,t+1} \sim N(0, \sigma_{j,t}^2) \tag{14}$$

[2] The coefficients in the labor income process are restricted in a way so that the equilibrium can be solved analytically. This restriction can be relaxed without affecting the model's key implications.

where $0 < \alpha_1^j < 1$ and $j = \{P, Q\}$. It is well known that housing returns exhibit positive autocorrelation in many markets; see Case and Shiller (1990) for U.S. cities and Englund and Ioannides (1997) for international comparative data. This provides empirical justification for the AR(1) process specified in Equation (13).

To pin down the risk-return relationship, I make further assumptions about the joint distribution of house price, rent and pricing kernel. First, the shocks to the price growth and pricing kernel are negatively correlated. Specifically, $cov(\epsilon_{P,t+1}, \epsilon_{M,t+1}) = \rho_{PM}\sigma^2_{P,t}$, where $\rho_{PM} < 0$. Second, the shocks to the price growth and rental growth are positively correlated: $cov(\epsilon_{P,t+1}, \epsilon_{Q,t+1}) = \rho_{PQ}\sigma^2_{P,t}$, where $\rho_{PQ} > 0$. Section 3.1 provides justification for each of these two key assumptions. Finally, for simplicity, I assume $corr(\epsilon_{Q,t+1}, \epsilon_{M,t+1}) = 0$.

Proposition 1 *Define $p_t = \ln P_t$, $q_t = \ln Q_t$ and $r_{t+1} = \ln R_{t+1}$. Given the model specified by Equations (1), (2), (10)-(12), there exists a linear solution to the log price rent ratio*

$$p_t - q_t = c_0 + c_1 g_t^P + c_2 g_t^Q + c_3\sigma^2_{P,t} + c_4\sigma^2_{Q,t} \tag{15}$$

where the parameters are

$$\begin{aligned}
c_0 &= \frac{-r_t^f + k + \gamma\ln(1-\delta) + \gamma\alpha_0^P + (1-\gamma)\alpha_0^Q}{1-\gamma} \\
c_1 &= \frac{\gamma\alpha_1^P}{1-\gamma} \\
c_2 &= \alpha_1^Q \\
c_3 &= \frac{\frac{1}{2}\gamma^2 + \gamma\rho_{PM} + \gamma(1-\gamma)\rho_{PQ}}{1-\gamma} \\
c_4 &= \frac{1}{2}(1-\gamma)
\end{aligned}$$

Furthermore, there exists a linear solution to the log return process

$$\begin{aligned}
r_{t+1} &= r_t^f + \gamma\epsilon_{P,t+1} + (1-\gamma)\epsilon_{Q,t+1} - \left(\frac{1}{2}\gamma^2 + \gamma\rho_{PM} + \gamma(1-\gamma)\rho_{PQ}\right)\sigma^2_{P,t} \\
&\quad -\frac{1}{2}(1-\gamma)^2\sigma^2_{Q,t}
\end{aligned} \tag{16}$$

Proof: To solve the model, I apply Campbell and Shiller (1988)'s approximation. Let $p_t = \ln P_t$ and $q_t = \ln Q_t$.

$$\begin{aligned}
r_{t+1} &= \ln R_{t+1} \\
&= k + \gamma\ln(1-\delta) + \gamma g_{t+1}^P + (1-\gamma)g_{t+1}^Q - (1-\gamma)(p_t - q_t)
\end{aligned} \tag{17}$$

where parameter γ is the average ratio of the house price to the sum of the house price and the rent, a number slightly smaller than one, and k is a constant related to γ. Substituting equation (4) and equation (17) into equation (9), we obtain

$$E_t exp\left[-r_t^f - \frac{1}{2}\sigma_{M,t}^2 + \epsilon_{M,t+1} + k + \gamma\ln(1-\delta) + \gamma g_{t+1}^P + (1-\gamma)g_{t+1}^Q - (1-\gamma)(p_t - q_t)\right] = 1 \tag{18}$$

I first postulate a solution to the log price rent ratio in terms of the state variables. I then verify this solution and solve for the parameters of the solution. The linear solution takes the following form.

$$p_t - q_t = c_0 + c_1 g_t^P + c_2 g_t^Q + c_3\sigma_{P,t}^2 + c_4\sigma_{Q,t}^2 \tag{19}$$

Substituting this solution to equation (18), we get

$$1 = E_t exp[A(\cdot)] \tag{20}$$

where

$$\begin{aligned}
A(\cdot) &= -r_t^f - \frac{1}{2}\sigma_{M,t}^2 + \epsilon_{M,t+1} + k + \gamma\ln(1-\delta) + \gamma g_{t+1}^P + (1-\gamma)g_{t+1}^Q \\
&\quad -(1-\gamma)(c_0 + c_1 g_t^P + c_2 g_t^Q + c_3\sigma_{P,t}^2 + c_4\sigma_{Q,t}^2) \\
&= \left(-r_t^f + k + \gamma\ln(1-\delta) - (1-\gamma)c_o + \gamma\alpha_0^P + (1-\gamma)\alpha_0^Q\right) - \frac{1}{2}\sigma_{M,t}^2 + \epsilon_{M,t+1} \\
&\quad + \left(\gamma\alpha_1^P + (1-\gamma)c_1\right) g_t^P + \left((1-\gamma)(\alpha_1^Q - c_2)\right) g_t^Q - (1-\gamma)c_3\sigma_{P,t}^2 - (1-\gamma)c_4\sigma_{Q,t}^2 \\
&\quad + \gamma\epsilon_{P,t+1} + (1-\gamma)\epsilon_{Q,t+1}
\end{aligned} \tag{21}$$

This leads to

$$\begin{aligned}
E_t A(\cdot) &= const. - \frac{1}{2}\sigma_{M,t}^2 + \left(\gamma\alpha_1^P + (1-\gamma)c_1\right) g_t^P + \left((1-\gamma)(\alpha_1^Q - c_2)\right) g_t^Q - (1-\gamma)c_3\sigma_{P,t}^2 - (1-\gamma)c_4\sigma_{Q,t}^2 \\
Var_t A(\cdot) &= \sigma_{m,t}^2 + \gamma^2\sigma_{P,t}^2 + (1-\gamma)^2\sigma_{Q,t}^2 + 2\gamma\rho_{PM}\sigma_{P,t}^2 + 2\gamma(1-\gamma)\rho_{PQ}\sigma_{P,t}^2
\end{aligned} \tag{22}$$

where $const. = -r_t^f + k + \gamma\ln(1-\delta) - (1-\gamma)c_o + \gamma\alpha_0^P + (1-\gamma)\alpha_0^Q$. Since $A(\cdot)$ is a normal random variable, we must have

$$E_t A(\cdot) + \frac{1}{2}Var_t A(\cdot) = 0 \tag{23}$$

Substituting equations (22) into equation (23), we obtain

$$\begin{aligned}
0 &= (const.) + \left(\gamma\alpha_1^P - (1-\gamma)c_1\right) g_t^P + \left((1-\gamma)(\alpha_1^Q - c_2\right) g_t^Q \\
&\quad + \left(-(1-\gamma)c_3 + \frac{1}{2}\gamma^2 + \gamma\rho_{PM} + (1-\gamma)\gamma\rho_{PQ}\right)\sigma_{P,t}^2 + \left(\frac{1}{2}(1-\gamma)^2 - (1-\gamma)c_4\right)\sigma_{Q,t}^2
\end{aligned} \tag{24}$$

Note that for this equation to hold, the terms in the five brackets must be equal to zero, since g_t^j and $\sigma_{j,t}^2$ are random variables. Solving the resulting equations produces the solutions below:

$$p_t - q_t = c_0 + c_1 g_t^P + c_2 g_t^Q + c_3\sigma_{P,t}^2 + c_4\sigma_{Q,t}^2 \tag{25}$$

where

$$\begin{aligned}
c_0 &= \frac{-r_t^f + k + \gamma\ln(1-\delta) + \gamma\alpha_0^P + (1-\gamma)\alpha_0^Q}{1-\gamma} \\
c_1 &= \frac{\gamma\alpha_1^P}{1-\gamma} \\
c_2 &= \alpha_1^Q \\
c_3 &= \frac{\frac{1}{2}\gamma^2 + \gamma\rho_{PM} + \gamma(1-\gamma)\rho_{PQ}}{1-\gamma} \\
c_4 &= \frac{1}{2}(1-\gamma)\sigma_{Q,t}^2
\end{aligned}$$

Substituting equation (25) into equation (17), we then obtain

$$r_{t+1} = r_t^f + \gamma\epsilon_{P,t+1} + (1-\gamma)\epsilon_{Q,t+1} - \left(\frac{1}{2}\gamma^2 + \gamma\rho_{PM} + \gamma(1-\gamma)\rho_{PQ}\right)\sigma_{P,t}^2 - \frac{1}{2}(1-\gamma)^2\sigma_{Q,t}^2$$

■

The solution method is consistent with Campbell and Shiller (1988) and Wu (2001). The first part of Proposition 1 shows that the log price-rent ratio ($p_t - q_t$) depends on the risk-free rate, expected house price growth and rent growth, and their conditional volatilities. The results are fairly intuitive. For example, a higher risk-free rate represents a higher opportunity cost of housing investment, which makes holding current housing less attractive. In contrast, higher price growth and rent growth increase the value of housing, leading to a higher price-rent ratio. Finally, the parameter c_3 characterizes the net effect of price risk on the price-rent ratio. The economic implications of this term are the focus of the next proposition.

The second part of Proposition 1 shows that the log realized housing return r_{t+1} also follows a simple and intuitive pattern. It depends on the risk-free rate, shocks to price growth and rent growth, and their conditional volatilities. The positive dependence of r_{t+1} on $\epsilon_{P,t+1}$ and $\epsilon_{Q,t+1}$ confirms the conventional wisdom that housing is valuable not only as a financial asset but also as a consumption good (e.g., Henderson and Ioannides, 1983; Brueckner, 1997).

Proposition 2 *Assuming* $-1 < \alpha_1^P < 1$ *and* $-1 < \alpha_1^Q < 1$.

(i) In an economy with an exogenous housing supply, the equilibrium housing risk-return relationship for the log total return process is characterized by

$$\frac{\partial \mathrm{E}_t(r_{t+1})}{\partial \sigma_{r,t}^2} = -\frac{\frac{1}{2}\gamma + \rho_{PM} + (1-\gamma)\rho_{PQ}}{\gamma + 2(1-\gamma)\rho_{PQ}} \tag{26}$$

where $\sigma_{r,t}^2 = \mathrm{VAR}_t(r_{t+1}) = \left(\gamma^2 + 2\gamma(1-\gamma)\rho_{PQ}\right)\sigma_{P,t}^2 + (1-\gamma)^2\sigma_{Q,t}^2$.

(ii)

$$\begin{aligned} \frac{\partial \mathrm{E}_t(r_{t+1})}{\partial \sigma_{r,t}^2} &< 0 \quad \textit{if} \ \rho_{PQ} > \frac{-\rho_{PM} - \frac{1}{2}\gamma}{1-\gamma} \\ \frac{\partial \mathrm{E}_t(r_{t+1})}{\partial \sigma_{r,t}^2} &> 0 \quad \textit{if} \ \rho_{PQ} < \frac{-\rho_{PM} - \frac{1}{2}\gamma}{1-\gamma} \end{aligned} \tag{27}$$

(iii)

$$\frac{\partial^2 \mathrm{E}_t(r_{t+1})}{\partial \sigma_{r,t}^2 \partial \rho_{PQ}} < 0 \tag{28}$$

Proof: (i) It follows from Equation (16) that

$$E_t(r_{t+1}) = r_t^f - \left(\frac{1}{2}\gamma^2 + \gamma\rho_{PM} + \gamma(1-\gamma)\rho_{PQ}\right)\sigma_{P,t}^2 - \frac{1}{2}(1-\gamma)^2\sigma_{Q,t}^2 \tag{29}$$

$$\sigma_{r,t}^2 = \left(\gamma^2 + 2\gamma(1-\gamma)\rho_{PQ}\right)\sigma_{P,t}^2 + (1-\gamma)^2\sigma_{Q,t}^2 \tag{30}$$

$$
\begin{aligned}
\frac{\partial E_t(r_{t+1})}{\partial \sigma^2_{r,t}} &= \frac{\partial E_t(r_{t+1})}{\partial \sigma^2_{P,t}} \frac{\partial \sigma^2_{P,t}}{\partial \sigma^2_{r,t}} \\
&= -\left(\frac{1}{2}\gamma^2 + \gamma\rho_{PM} + \gamma(1-\gamma)\rho_{PQ}\right) \frac{1}{\gamma^2 + 2\gamma(1-\gamma)\rho_{PQ}} \\
&= -\frac{\frac{1}{2}\gamma + \rho_{PM} + (1-\gamma)\rho_{PQ}}{\gamma + 2(1-\gamma)\rho_{PQ}} \qquad (31)
\end{aligned}
$$

where the last step comes from Equations (29) and (30). It follows from Equation (31) that

$$
\begin{aligned}
\frac{\partial E_t(r_{t+1})}{\partial \sigma^2_{r,t}} &< 0 \text{ if } \rho_{PQ} > \frac{-\rho_{PM} - \frac{1}{2}\gamma}{1-\gamma} \\
\frac{\partial E_t(r_{t+1})}{\partial \sigma^2_{r,t}} &> 0 \text{ if } \rho_{PQ} < \frac{-\rho_{PM} - \frac{1}{2}\gamma}{1-\gamma} \qquad \blacksquare
\end{aligned}
$$

(ii) Proof: Following Equation (31), we have

$$
\begin{aligned}
\frac{\partial^2 E_t(r_{t+1})}{\partial \sigma^2_{r,t}\partial \rho_{PQ}} &= -\frac{1}{(\gamma + 2(1-\gamma)\rho_{PQ})}(1-\gamma) + \frac{\frac{1}{2}\gamma + \rho_{PM} + (1-\gamma)\rho_{PQ}}{(\gamma + 2(1-\gamma)\rho_{PQ})^2} 2(1-\gamma) \\
&= \frac{2(1-\gamma)\rho_{PM}}{(\gamma + 2(1-\gamma)\rho_{PQ})^2} < 0
\end{aligned}
$$

where the inequality follows from the assumption that $\rho_{PM} < 0$. $\blacksquare$

Proposition 2 (i) provides a clean characterization of the net risk-return relationship. Following this, Proposition 2 (ii) shows that the consumption hedge effect dominates in markets with sufficiently large ρ_{PQ}, whereas the financial risk effect dominates in other markets. In this sense, the model delivers cross-sectional variation in the *sign* of the risk-return relationship through the relative strength of ρ_{PQ} and ρ_{PM}. Proposition 2 (iii) further delivers cross-sectional variation in the *magnitude* of the risk-return relationship. That is, across markets, the marginal expected return required to compensate for the risk decreases with ρ_{PQ}. This provides the first testable implication for our empirical work.

The analysis above focuses on the log total return process, r_t, which includes the dividend yield. In principle, one can proxy the dividends with rents and combine rents with house prices to impute the total returns. However, doing so implicitly assumes the direct comparability of owned units to rental units and of owners to renters. In a recent work by Glaeser and Gyourko (2010), they document that rental units tend to be very different from owner-occupied units and that owners are different from owners in economically meaningful ways. They also argue that, while it is possible to construct comparable rent and house price data in one market, it is infeasible to do so for large scale statistical work that involves most of the key markets in the country. For this reason, the empirical analysis in this paper examines the relationship between the expected log capital return process $E_t g^P_{t+1} = E_t\left(ln(P_{t+1}) - ln(P_t)\right)$ and the expected price risk $\sigma^2_{P,t}$. In what follows, I shall show that the implications for the log total return process in Proposition 2 essentially carry over to the log capital return process, because the capital gain return tends to dominate the total return.

Proposition 3 *Assuming* $-1 < \alpha_1^P < 1$ *and* $-1 < \alpha_1^Q < 1$.

(i) In an economy with an exogenous housing supply, the equilibrium housing risk-return relationship for the log capital return process is characterized by

$$\frac{\partial \mathrm{E}_t(g_{t+1}^P)}{\partial \sigma_{P,t}^2} = -\frac{\frac{1}{2}\gamma + \rho_{PM} + (1-\gamma)\rho_{PQ}}{1-\gamma} \tag{32}$$

. *(ii)*

$$\begin{aligned} \frac{\partial \mathrm{E}_t(g_{t+1}^P)}{\partial \sigma_{P,t}^2} &< 0 \;\; \textit{if} \;\; \rho_{PQ} > \frac{-\rho_{PM} - \frac{1}{2}\gamma}{1-\gamma} \\ \frac{\partial \mathrm{E}_t(g_{t+1}^P)}{\partial \sigma_{P,t}^2} &> 0 \;\; \textit{if} \;\; \rho_{PQ} < \frac{-\rho_{PM} - \frac{1}{2}\gamma}{1-\gamma} \end{aligned} \tag{33}$$

(iii)

$$\frac{\partial^2 \mathrm{E}_t(g_{t+1}^P)}{\partial \sigma_{P,t}^2 \partial \rho_{PQ}} < 0 \tag{34}$$

Proof: (i) Combining Equation (16) and Equation (17), we obtain

$$\begin{aligned} \frac{\partial E_t(g_{t+1}^P)}{\partial \sigma_{P,t}^2} &= \frac{\partial E_t(g_{t+1}^P)}{\partial E_t r_{t+1}} \frac{\partial E_t(r_{t+1})}{\partial \sigma_{P,t}^2} + \frac{\partial E_t(g_{t+1}^P)}{\partial (p_t - q_t)} \frac{\partial (p_t - q_t)}{\partial \sigma_{P,t}^2} \\ &= \frac{1}{\gamma}(-c_3) + \frac{1-\gamma}{\gamma} c_3 \\ &= -\frac{\frac{1}{2}\gamma + \rho_{PM} + (1-\gamma)\rho_{PQ}}{1-\gamma} \end{aligned} \tag{35}$$

(ii) and (iii) from Equation (35) naturally. ■

To complete the solution of our model economy, we need to find housing and nondurable consumption in terms of exogenous forces. The results will of course depend on what the rest of the economy looks like. Since households are identical with respect to their preferences and endowments, it follows from the market clearing conditions that in equilibrium $h_t^* = \frac{H_t}{N_t}$ and $c_t^* = y_t$. Since housing endowment is exogenous, the house price adjusts so that an exogenous shock in price risk is fully translated into changes in house prices and returns. The equilibrium risk-return relationship is therefore fully characterized by Equation (32).

Appendix II. Population-Based Urban Growth Measure

The key parameter of interest in equation (4) in the main paper is β_1, the coefficient on the interaction term $RISK_{it} \times HEDGE_{it} \times FAST_{it}$. A negative β_1 indicates that the consumption hedge effect is stronger in fast-growing markets. One econometric issue confronting this estimate is the possibility that the underlying urban growth indicator, $FAST_{it}$, is mis-measured, making it hard to interpret the resulting estimate. In measuring urban growth, Section 4.3 compares the overall population (and hence the number of households) with existing housing units. The comparison is intuitively appealing, as population is an obvious indicator of urban growth (Helsley, 2003, chapter 8). Given the extremely tight link between population and housing stock (Glaeser and Gyourko, 2005), the difference between the two would naturally reflect the imbalance between the aggregate housing demand and supply. However, what happens when we take into account the difference between the aggregate housing demand and the demand for owner-occupied housing, as well as changes in mortgage rates and expectations over future house price appreciation? Following DiPasquale and Wheaton (2000), I now present a simple model that captures these additional considerations, and show that the estimated urban growth effect remains robust after accounting for these factors.

Assume the demand for owner-occupied housing is proportional to the number of households (H_t), namely, $H_t(\alpha_0 - \alpha_1 UC_t)$. The parameter α_0 can be considered as the fraction of households who would own homes if the cost of owning a house is zero; α_1 as the responsiveness of this fraction to changes in the annual cost of owning, UC_t. A simple way to express UC_t is $P_t(M_t - I_t)$, where P_t indicates house price, M_t indicates the after-tax mortgage rate, and I_t indicates the expected rate of future house price appreciation. This specification is useful because it incorporates two important considerations. First, the proportionality factor, $\alpha_0 - \alpha_1 P_t(M_t - I_t)$, allows us to focus on the demand for owner-occupied housing, rather than the aggregate housing demand. Second, the current housing demand will be higher, all else being equal, when the after-tax mortgage rate is lower, or when expectations about future price appreciation are more optimistic.

In equilibrium, house price today adjusts so that the existing stock of housing units, S_t, equals the *ex ante* demand for owner-occupied housing units:

$$S_t = H_t(\alpha_0 - \alpha_1 P_t(M_t - I_t)) \tag{36}$$

Thus, a city exhibits *relatively* faster population growth and hence higher *owner-occupied* housing demand relative to supply if

$$\frac{H_t}{S_t} \geq \frac{1}{\alpha_0 - \alpha_1 P_t(M_t - I_t)} \tag{37}$$

Empirically, I do not observe α_0, α_1, M_t, and I_t. I therefore replace the cutoff point on the right-hand-side of equation (37) with 1. Given that $\frac{1}{\alpha_0 - \alpha_1 P_t(M_t - I_t)} \geq 1$, the fraction of fast-growing cities based on the owner-occupied housing demand (indicated by equation 37) is smaller than the fraction of fast-growing cities based on the aggregate demand – the latter being used to form the measure of urban growth in the main analysis. However, this should not affect the interpretation of our results for two reasons.

First, as discussed in Section 3.3, the urban growth measure is considered as a comparative static shift in hedging demand. In this sense, what we need here is a measure that can rank different cities based on their relative urban growth rates, rather than a measure of housing demand. I measure urban growth by comparing population with housing stock and by comparing average house price with costruct costs. This is consistent with the spirit of the urban literature (Helsley, 2003; Glaeser, Gyourko and Saks, 2006) and hence more suitable for the purpose here.

Second, to the extent that the owner-occupied housing demand matters for hedging demand in some unspecified way, equation (37) and the subsequent discussion indicate that we have a classical measurement error problem (e.g. Wooldridge, 2001, Chapter 4), which produces an attenuation bias on the parameter β_1 (i.e., a bias towards zero). In this case, a significantly negative estimate of β_1 should be considered as conservative evidence for the urban growth effect, providing reassuring support for the prediction that the consumption hedge effect is stronger in fast-growing cities.

Appendix III. Data Description

This section describes how the FHFA HPI series is collected and structured for estimation, and and how the fraction of 'staying-within-the-same-MSA' population (one of the two conditions that define within-market hedging incentives) is constructed.

FHFA HPI Series

Note that the HPI series that is currently posted on the FHFA website uses 5-digit MSA codes, which are based on the 2000 Census definition of Metropolitan Statistical Areas (MSAs) (this definition was revised again in later years). Other variables in this paper, such as hedging incentives, supply constraints and urban growth, are based on 4-digit MSA codes, which follow the 1990 Census definition of MSAs. Obviously, these two definitions of MSAs are incompatible with each other. Fortunately, for the pre-2004q2 period, the FHFA all-transaction house price indices were initially collected based on the 1990 definition of the MSAs. I therefore requested these original house price indices from the FHFA. The original HPI series, based on 4-digit MSA codes, discontinued after the second quarter of 2004. To extend the sample to the more recent years, I downloaded post-2004q2 house price indices from the FHFA website and converted 255 of the 5-digit MSA codes, used in the later series, to 4-digit MSA codes based on each MSA's county composition.[3] Putting these two series together, I have a panel dataset that contains 331 4-digit-code MSAs, among which 255 MSAs have the price series that can be extended until 2007 based on the same definition of the MSAs and the remaining 76 MSAs' series ended in 2004 due to the difficulty of matching the old and new MSA codes for these areas. This combined price series can then be merged with other data sources (e.g., for hedging incentives, supply constraints, etc.) used in this paper, allowing me to construct a panel dataset in which all key variables are defined based on the same definition of MSA.

The main empirical model is conducted based on the GARCH estimates of expected housing

[3] Unlike the MSA codes, county codes are comparable over years. This allows one to convert a 5-digit MSA code to a 4-digit MSA code if there is no change in the composition of the counties. In matching these MSAs, I drop the MSAs which contain areas that are "contaminated" either by sitting across different county boundaries or by having different county compositions over years. In the end, 255 "clean" MSAs are identified. The computer code for cross-checking the two different types of MSA codes is available upon request.

returns and risks. To conduct a sensible time series analysis, I drop the MSAs for which the number of non-missing quarterly observations is less than 70. This reduces the total sample from 331 MSAs to 313 MSAs. Among them, maximum likelihood estimates of the AR(1)-GARCH(1,1) model are obtained for 295 MSAs.[4] Restricting the sample period to 1980-2007 reduces the sample to 6659 observations for which both GARCH estimates of the expected returns and risk (including the lagged return term) are available. Given the lack of data on the population mobility fraction for some MSAs (as discussed in the next subsection), the resulting risk and return estimates cannot be perfectly merged with the within-market hedging incentive measure. As a result, the final estimation sample, as used for columns (1) and (2) of Table 3, contains 5421 observations.

Measuring "Staying-Within-the-Same-MSA" Population

As noted in Section 4.2, the within-market hedging incentive measure is jointly defined by the age distribution of the local population and the fraction of the population that stayed within the same MSA in the past 5 years. While age population data are available for all the MSAs in the FHFA sample, this is not so for the fraction of the population who have stayed within the same MSA for the past 5 years. To compute the latter, I used the household-level data from the 5 percent sample of the 1990 and 2000 Integrated Public Use Microdata Series (commonly referred to as the PUMS 1990 and the PUMS 2000), and computed the fraction of population who have stayed within the same MSA for the past 5 years.

In the PUMS 2000, households report the MSA they currently reside in and the MSA they resided in 5 years ago. Using their responses, I impute the fraction of the 'staying-within-the-same-MSA' population for 281 unique MSAs (based on 4-digit codes). Specifically, for each MSA, the fraction of recent emigrants who moved from a specific MSA to other areas in the last five years is considered as the MSA's out-migration rate. The difference between one and the imputed out-migration rate is then taken as the fraction of population who have stayed within the same MSA.

However, in the PUMS 1990, households do not report the MSA they resided in 5 years ago; instead they report the CONSPUMA (a geographic unit used by the PUMS) in which they stayed five years ago. I therefore assign the 4-digit MSA code to each CONSPUMA based on the MSA's composition of CONSPUMAs in 1990. In this process, I need to drop the CONSPUMAs that sit across different MSAs. As a result, the imputed 'staying-within-the-same-MSA' population fraction is available for only 242 MSAs.

Putting these two years' mobility data available, I can identify 236 MSAs for which the imputed fractions are available both in 2000 and in 1990. For these MSAs, I assign the 1990 estimates to years up until 1995 and the 2000 estimates to years after 1995, and merge them with the age

[4] There are 18 MSAs for which the AR(1)-GARCH(1,1) model does not converge. However, this does not necessarily imply that the conditional variance is constant in these markets. An alternative explanation could be that the series of real housing returns in these MSAs are so noisy that a systematic pattern of conditional heteroscedasticity does not hold given the relatively short time-horizon observed for some MSAs.

population data to construct within-market hedging incentives. Clearly, not all the risk-return estimates can be perfectly merged with the within-market hedging incentives. This reduces the sample size from 6659 to 5421 observations, as shown in columns (1) and (2) of Table 3.

References

Brueckner, Jan, "Consumption and Investment Motives and the Portfolio Choices of Homeowners," *Journal of Real Estate Finance and Economics*, 1997, *15*, pp. 159-180.

Campbell, John and Robert Shiller, "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors," *Review of Financial Studies*, 1988, *1*, pp. 195-228..

Case, Karl and Robert Shiller, "Forecasting Prices and Excess Returns in the Housing Markets," *AREUEA Journal*, 1990, pp. 253-273.

Chambers, Matthew, Carlos Garriga, and Don Schlagenhauf, "Accounting For Changes in the Homeownership Rate," *International Economic Review*, 2009, pp. 677-726.

DiPasquale, Denise and William Wheaton, "Urban Economics and Real Estate Markets," New Jersey: Princeton Hall, 1996.

Englund, Peter and Yannis Ioannides, "House Price Dynamics: An International Empirical Perspective," *Journal of Housing Economics*, 1997, *6*, pp. 119-136.

Glaeser, Edward and Joseph Gyourko, "Urban Decline and Durable Housing," *Journey of Political Economy*, 2005, *113* pp. 345-375.

Glaeser, Edward and Joseph Gyourko, and Raven Saks, "Urban Growth and Housing Supply," *Journal of Economic Geography*, 2006, *6* pp. 71-89.

Glaeser, Edward, Joseph Gyourko, Eduardo Morales and Charles Nathanson "Housing Dynamics," 2010. NBER Working Paper No. 12787.

Helsley, Robert "Urban and Real Estate Economics," UBC Real Estate Division, 2003.

Henderson, Vernon and Yannis Ioannides, "A Model of Housing Tenure Choice," *American Economic Review*, 1983, *73* pp. 98-113.

Piazzesi, Monika, Martin Schneider, and Selale Tuzel, "Housing, Consumption, and Asset Pricing," *Journal of Financial Economics*, 2007. *83* pp. 531-569.

Stein, Jeremy, "Prices and Trading Volume in the Housing Market: A Model with Downpayment Constraints," *Quarterly Journal of Economics,* 1995. *110,* pp. 379-406.

Wooldridge, Jeffery, "Econometric Analysis of Cross Section and Panel Data," Cambridge, MA: MIT Press, 2001.

Wu, Guojun, "The Determinants of Asymmetric Volatility," *Review of Financial Studies,* 2001. *14* pp. 837-859.

Table A1: Metropolitan Areas by Constraints and by Growth

Fast-Growing	Slow-Growing	Declining
	Less Constrained Markets	
Albuquerque, NM	Akron, OH	Detroit, MI
Allentown-Bethlehem-Easton, PA	Beaumont-Port Arthur, TX	New Orleans, LA
Atlanta, GA	Buffalo-Niagara Falls, NY	Wichita, KS
Ann Arbor, MI	Cleveland-Lorain-Elyria, OH	
Austin-San Marcos, TX	Columbus, OH	
Biloxi-Gulfport-Pascagoula, MS	Davenport-Moline-Rock Island, IA-IL	
Birmingham, AL	Dayton-Springfield, OH	
Charlotte-Gastonia-Rock Hill, NC-SC	Des Moines, IA	
Chattanooga, TN-GA	Duluth-Superior, MN-WI	
Cincinnati, OH-KY-IN	Erie, PA	
Columbia, SC	Evansville-Henderson, IN-KY	
Corpus Christi, TX	Flint, MI	
Dallas, TX	Fort Wayne, IN	
Denver, CO	Indianapolis, IN	
El Paso, TX	Gary, IN	
Fort Worth-Arlington, TX	Jacksonville, FL	
Grand Rapids-Muskegon-Holland, MI	Kalamazoo-Battle Creek, MI	
Greensboro-Winston-Salem-High Point, NC	Kansas City, MO-KS	
Houston, TX	Macon, GA	
Huntsville, AL	Omaha, NE-IA	
Jersey City, NJ	Peoria-Pekin, IL	
Knoxville, TN	Pittsburgh, PA	
Las Vegas, NV-AZ	Lansing-East Lansing, MI	
Little Rock-North Little Rock, AR	Rockford, IL	
Louisville, KY-IN	Scranton-Wilkes-Barre-Hazleton, PA	
Lubbock, TX	South Bend, IN	
Madison, WI	Springfield, MO	
Memphis, TN-AR-MS	St. Louis, MO-IL	
Minneapolis-St. Paul, MN-WI	Toledo, OH	
Mobile, AL	Topeka, KS	
Nashville, TN	Utica-Rome, NY	
Norfolk-Virginia Beach-Newport News, VA	Youngstown-Warren, OH	
Oklahoma City, OK		
Phoenix-Mesa, AZ		
Providence-Fall River-Warwick, RI-MA		
Raleigh-Durham-Chapel Hill, NC		

Table A1: Metropolitan Areas by Constraints and by Growth (cont'd)

Fast-Growing	Slow-Growing	Declining
Reno, NV		
Richmond-Petersburg, VA		
Roanoke, VA		
Rochester, NY		
Salt Lake City-Ogden, UT		
San Antonio, TX		
Springfield, MO		
Syracuse, NY		
Tampa-St. Petersburg-Clearwater, FL		
Tulsa, OK		
Washington, DC-MD-VA-WV		
More Constrained Markets		
Albany-Schenectady-Troy, NY	Milwaukee-Waukesha, WI	Philadelphia, PA-NJ
Allentown-Bethlehem-Easton, PA	Spokane, WA	
Anchorage, AK		
Baton Rouge, LA		
Bakersfield, CA		
Baltimore, MD		
Boston, MA-NH		
Bridgeport, CT		
Charleston-North Charleston, SC		
Chicago, IL		
Colorado Springs, CO		
Eugene-Springfield, OR		
Fort Lauderdale, FL		
Fresno, CA		
Hartford, CT		
Honolulu, HI		
Lawrence, MA-NH		
Lexington, KY		
Los Angeles-Long Beach, CA		
Lowell, MA-NH		
Miami, FL		
New York, NY		
Newark, NJ		
Orlando, FL		

Table A1: Metropolitan Areas by Constraints and by Growth (cont'd)

Fast-Growing	Slow-Growing	Declining
More Constrained Markets		
Phoenix-Mesa, AZ		
Portland-Vancouver, OR-WA		
Riverside-San Bernardino, CA		
Sacramento, CA		
San Diego, CA		
San Francisco, CA		
Santa Barbara-Santa Maria-Lompoc, CA		
Seattle-Bellevue-Everett, WA		
Springfield, MA		
Stockton-Lodi, CA		
Tacoma, WA		
Tucson, AZ		
Vallejo-Fairfield-Napa, CA		

Table A2: Summary Statistics of Fraction of Population Aged 20-45[a]

year	mean (%)	s.d. (%)	p50 (%)	min. (%)	max. (%)
1980	38.07	3.88	37.54	20.10	51.71
1981	38.67	3.88	38.02	21.20	52.44
1982	39.18	3.87	38.62	22.10	52.96
1983	39.61	3.85	39.05	22.90	53.37
1984	40.00	3.83	39.49	23.58	53.70
1985	40.25	3.80	39.78	24.10	54.26
1986	40.38	3.75	39.85	24.53	54.41
1987	40.39	3.71	39.91	24.87	54.59
1988	40.22	3.69	39.73	24.93	54.55
1989	40.14	3.68	39.68	25.08	54.64
1990	40.14	3.66	39.73	25.04	54.70
1991	40.28	3.66	39.97	25.30	55.55
1992	39.86	3.64	39.53	24.72	53.84
1993	39.47	3.59	39.06	24.31	53.59
1994	39.07	3.52	38.66	23.98	52.09
1995	38.71	3.49	38.27	23.75	51.35
1996	38.31	3.45	37.91	23.29	50.62
1997	37.51	3.43	37.52	22.78	49.55
1998	37.49	3.41	37.16	22.30	49.19
1999	37.10	3.40	36.76	22.16	48.79
2000	36.73	3.41	36.41	21.72	48.90
2001	36.51	3.39	36.11	22.16	49.10
2002	36.23	3.32	35.84	22.64	49.08
2003	35.87	3.22	35.49	23.01	48.83
2004	35.58	3.18	35.19	23.68	48.50
2005	35.23	3.12	34.85	23.72	48.14
2006	34.90	3.10	34.61	24.35	47.78
2007	34.57	3.05	34.18	25.13	47.35

[a]Data Source: U.S. Census. One unit of observation is an MSA in a given year.

Table A3: Hedging and Supply Effects on the Risk-Return Relationship (1980-2010)[a]

Variable	1	2	3	4	5	6
risk	0.10^{**}	0.15^{***}	0.17^{*}	0.12^{*}	0.14	0.13
	(2.01)	(3.04)	(1.68)	(1.72)	(0.89)	(0.67)
risk × hedge	-0.08^{***}	0.14	-0.09	-0.41^{*}	0.25	-0.28
	(-2.78)	(0.75)	(-1.08)	(-1.89)	(1.39)	(-1.36)
risk × hedge × constraint		-0.47^{***}	-0.36^{***}		-1.62^{**}	-0.54^{**}
		(-3.06)	(-4.12)		(-2.37)	(-1.98)
N	6091	5320	5320	1147	1122	1122

[a]This table repeats the estimation of hedging and supply effects on the risk-return relationship in Table 10 using the FHFA all-transaction house price indices between 1980-2010. The dependent variable is the expected return. Hedging incentives are based on the within-market measure in columns 1-3, and the cross-market measure in columns 4-6. Supply constraints are measured by the undevelopable land share in columns 2 and 5, and the WRLRUI in columns 4 and 6. The t-statistics are adjusted for the intra-MSA correlation and reported in parentheses. $***$, $**$, and $*$ denote statistically significant at the 1%, 5%, and 10% levels, respectively.

Table A4: Hedging and Supply Effects on the Risk-Return Relationship[a]
(Robustness Checks based on AR(2)-GARCH(1,1) Measures of Risks and Returns)

Variable	1	2	3	4	5	6
risk	0.10^{**}	0.12^{***}	0.09^{*}	0.15^{**}	0.11^{*}	0.12^{*}
	(1.98)	(2.75)	(1.91)	(1.96)	(1.79)	(1.78)
risk × hedge	-0.27^{***}	0.20	-0.05^{*}	-0.28^{**}	0.13	-0.11
	(-3.12)	(0.56)	(-1.80)	(-2.19)	(0.52)	(-0.78)
risk × hedge × constraint		-1.08^{**}	-0.47^{***}		-1.43^{**}	-0.70^{***}
		(-3.17)	(-2.82)		(-2.01)	(-2.59)
N	4350	3728	3728	920	815	815

[a]This table repeats the estimation of hedging and supply effects on the risk-return relationship in Table 10 using AR(2)-GARCH(1,1) measures of expected returns and risk based on the FHFA all-transaction house price indices between 1980-2007. The dependent variable is the expected return. Hedging incentives are based on the within-market measure in columns 1-3, and the cross-market measure in columns 4-6. Supply constraints are measured by the undevelopable land share in columns 2 and 5, and the WRLRUI in columns 3 and 6. The t-statistics are adjusted for the intra-MSA correlation and reported in parentheses. $***$, $**$, and $*$ denote statistically significant at the 1%, 5%, and 10% levels, respectively.