Monetary Policy, Housing, and Heterogeneous Regional Markets

Michael Fratantoni*
Fannie Mae

Scott Schuh† ‡
Federal Reserve Bank of Boston

This version: September 2001
Original version: January 2000

*3900 Wisconsin Avenue, N.W., Mail Stop 1H-2E-01 Washington, D.C. 20016-2892; *Michael_Fratantoni@fanniemae.com*

†P.O. Box 2076, Boston, MA 02106-2076; *Scott.Schuh@bos.frb.org*

‡We thank Seth Carpenter, Gordon Crawford, Jeff Fuhrer, Mark Gertler, Simon Gilchrist, Kathryn Holliday, Richard Kopcke, Len Mills, Jim Stock, Randy Verbrugge, Charles Weise, and two anonymous referees for helpful comments. We are grateful to Bob Tannenwald for assistance in obtaining data. Kevin Daly, Catherine Humblet, and Jennifer Young provided excellent research assistance. The views expressed in the paper are those of the authors and do not necessarily represent the views of Fannie Mae, the Federal Reserve Bank of Boston, or the Federal Reserve System.

Abstract

This paper quantifies the importance of heterogeneity in regional housing markets for the conduct of monetary policy using a new model called a heterogeneous-agent VAR (HAVAR), which generalizes conventional macro VARs. The HAVAR model integrates a national monetary authority and financial market with regional housing markets, imposing exact aggregation. Monetary policy is transmitted to the national to regional markets via the mortgage rate. Although the HAVAR model is based on linear regional VARs, its aggregate impulse responses exhibit two nonlinearities: (1) time variation, stemming from aggregation over heterogeneous regions; and (2) state dependence on initial economic conditions in regions. Thus, the effects of monetary policy on the economy depend on the extent and nature of regional heterogeneity, which vary over time. Using longitudinal data for a subsample of detailed U.S. regions, we estimate the effects of time variation and state dependence on the dynamic responses of the HAVAR model. The estimated model provides plausible and tangible explanations for "long and variable" lags in monetary policy. To provide a policy-relevant illustration, we show how coastal housing booms influence the efficacy of monetary policy.

JEL Codes: E22, E52, R21, R31
Keywords: monetary policy, housing, aggregation, heterogeneity, regional, VAR

1 Introduction

A widely held conventional view is that monetary policy should focus only on aggregate economic conditions because it cannot control or target the conditions of particular geographic regions. This paper examines the largely overlooked flip side of this conventional view. Regional economic conditions can (and do) significantly influence aggregate responses to monetary policy actions, for two simple and intuitive reasons. First, economic sensitivity to monetary policy varies across regions, as shown recently by Carlino and DeFina [7], [8]. Second, economic conditions prevailing at the time of monetary policy actions vary across regions, as we show later. For both reasons, aggregate dynamic responses to monetary policy actions are nonlinear so the magnitude and duration of the responses vary over time.

In other words, although monetary policy cannot target regional economic performance, regional heterogeneity may matter for the efficacy of monetary policy. For example, the extent to which the economy slows in response to a monetary tightening will depend on issues such as which regions are growing fastest, and whether the most rapidly expanding regions are the most interest sensitive. More generally, the aggregate effects of monetary policy depend on the distribution of regional sensitivities to monetary policy and on the initial distribution of regional economic conditions at the time of monetary tightening. Both distributions vary over time, so small changes in the configuration of heterogeneity can produce economically significant changes in aggregate responses.

To evaluate the importance of this idea, we develop a macroeconometric framework called a heterogeneous-agent vector autoregression (HAVAR) model.[1] The HAVAR model is a blend of VAR models from two literatures. At the aggregate level, the HAVAR model is analogous to the class of monetary VAR models, described in Christiano, Eichenbaum, and Evans [12], which approximate the reduced-forms of small-scale structural macro models.[2] At the disaggregate level, the HAVAR model is analogous to VAR models of international, regional and industrial hetero-

[1]This name replaces the aggregation vector autoregression (AVAR) terminology used in previous versions of the paper. We thank John Keating for pointing out the acronym conflict with asymmetric VAR models.

[2]Indeed, the HAVAR model framework is fully applicable to structural models of monetary policy such as Fuhrer and Moore [22]; Chari, Christiano, and Kehoe [10]; King and Wolman [34]; Goodfriend and King [24]; Rotemberg and Woodford [53]; McCallum and Nelson [44]; Clarida, Gali, and Gertler [13]; and the November 1995 issue of the *Journal of Money, Credit and Banking*, 27(4). These models can be solved to obtain reduced-forms that are essentially restricted VARs, then incorporated into the HAVAR framework.

geneity.[3] The latter typically include only one or two variables, such as employment or income, and focus on the relative importance of aggregate versus disaggregate shocks or the extent and nature of heterogeneity.

Our HAVAR model combines the strengths of both literatures – the relative breadth of monetary VARs and the heterogeneity of disaggregate VARs. The aggregate and disaggregate sectors are linked in an econometrically tractable manner by exploiting theory-based simplifying assumptions about the economic relationships among them. Because it imposes exact aggregation conditions and focuses primarily on aggregate behavior, the HAVAR model quantifies the macroeconomic importance of microeconomic heterogeneity in a way that previous studies have not. A recent paper by Abadir and Talmain [1] contains a closely related analysis of aggregation of heterogeneous firms in a real business cycle model.

HAVAR models also offer a more general framework for policy analysis than conventional macro VAR models. By relaxing implicit restrictions on micro heterogeneity and aggregating, the HAVAR model becomes a simple form of nonlinear macro VAR model – even though the underlying micro models are linear. Specifically, the macro parameters vary over time for two reasons: 1) region size varies over time; and 2) initial economic conditions in regions vary over time. Consequently, HAVAR models can quantify the extent to which the efficacy of monetary policy depends on the distributions of regional conditions and the nature of regional shocks. Conventional VARs are simply special cases of HAVAR models with fixed macro parameters and without state dependence.[4]

In our application of a HAVAR model, we quantify the importance of regional heterogeneity in housing markets for the efficacy of monetary policy. Although housing is not the only potential source of regional heterogeneity, it is an ideal candidate. Housing is a volatile leading indicator of the business cycle and a critical channel of monetary transmission. But housing is determined in local markets because supply and demand depend heavily on idiosyncratic and regional factors. Other sources of regional heterogeneity, such as industrial or demographic composition, state and

[3] Clark and Shin [15] contains a nice survey of this literature. Other recent related examples include the Carlino and DeFina studies, Blanchard and Katz [6], Quah [52], Sill [55], Clark [14], Davis and Haltiwanger [18], Coulson [17], and Carlino, DeFina, and Sill [9].

[4] Although the particular form of nonlinearity is quite simple, the HAVAR model is a specific example of the nonlinear time series model described in Granger and Newbold [27]. Thus, it is part of the literature on nonlinear impulse response functions that includes Gallant, Rossi, and Tauchen [23], Potter [49], [50], and [51], Koop, Pesaran, and Potter [35], Aoki [2], Weise [62], and Balke [3].

local fiscal policies, and financial market conditions, would also be interesting applications.

Our housing-monetary HAVAR model assumes the following decomposition of a benchmark monetary VAR model. At the national level, inflation and interest rates (nominal variables) are determined by the monetary authority and financial markets. At the regional level, income, housing investment, and housing prices (real variables) are determined by households and firms. National and regional markets are linked by the mortgage rate, which serves as the central channel for monetary transmission. As an example, the housing-monetary HAVAR model enables comparison of the effects of monetary policy during the late 1980s coastal housing boom with policy during periods when regional housing activity was more balanced.

Using a longitudinal panel of detailed regional data, we find economically and statistically significant differences between the dynamic responses of the housing-monetary HAVAR model and the conventional VAR for up to two years. Aggregate variables exhibit much less persistence in the HAVAR model, with mean lags of dynamic responses to monetary policy shocks about one year shorter than in the conventional VAR.[5] Even larger differences arise between the dynamic responses of the HAVAR model to variation in regions' initial economic conditions. Peak responses (magnitude) can vary by more than one percentage point and mean lags (duration) can vary by more than one year, depending on the configuration of regional conditions.

The HAVAR framework and empirical results offer one practical and tangible interpretation of the "long and variables lags" to monetary policy. We interpret the lesser persistence in HAVAR dynamic responses as evidence of aggregation bias in conventional VAR responses, and thus a partial explanation for the perceived "long" lags. We interpret variation in the magnitude and duration of HAVAR dynamic responses across regional initial conditions as a partial explanation for "variable" lags.

Overall, our findings point to the importance of addressing the role of microeconomic heterogeneity in assessing the effects of monetary policy on macroeconomic conditions. Incorporating other sources of heterogeneity along with housing may yield even greater effects. Furthermore, our methodology has obvious extensions to fiscal policy and other types of shocks, and to international settings such as the European Union.

[5] This result contributes to the literature on aggregation and persistence (long-memory) in time series models. See Abadir and Talmain [1] and the references therein.

2 Context and Motivation

2.1 Housing and Macro Models

Housing investment long has been acknowledged as an important factor in macroeconomic models. One reason is that housing is one of the most volatile sectors of the economy. Housing investment growth is 28 times more variable than GDP growth, whereas nonresidential investment growth is only seven times more variable. However, housing investment only accounts for about 5 percent of GDP and 7 percent of the variance of GDP growth, so more than striking volatility is required to justify a role for housing.

Housing investment also is closely connected with the business cycle and monetary policy. Green [28] provides evidence that housing investment Granger causes GDP, whereas nonresidential investment does not, and the Conference Board includes housing permits among its leading economic indicators. Moreover, the data show a strong inverse lead-lag relationship between monetary policy, as measured by the federal funds rate or the spread between mortgage and fed funds rates, and housing and other real activity.[6] Simple cross-correlations reveal that the monetary policy variables lead housing investment growth by one quarter, which in turn leads nonhousing GDP growth by one quarter. We also find strong cross-correlations between inflation and housing appreciation (growth of real housing prices) over the business cycle with much longer lags. These correlations are not widely known and suggest an additional link between housing and monetary policy.[7]

Although these correlations are suggestive, multivariate dynamic analyses provide more convincing evidence of the independent importance of housing. Consequently, we estimated a benchmark monetary VAR model that is representative of the literature except that it disaggregates output and price into housing and nonhousing components. The variables are: P^n = log of the implicit deflator for nonhousing GDP; P^h = log of the implicit deflator for housing investment; Y = log of

[6] See Bernanke and Blinder [4] and Friedman and Kuttner [21] regarding these measures of monetary policy. The latter use the commercial paper-Treasury bill spread, but mortgage-fed funds spread is qualitatively similar.

[7] From a macroeconomic perspective, it is easy to understand why inflation and interest rates have occupied center stage. Contemporaneously, inflation and interest rates are much more closely aligned with the business cycle than housing appreciation, and inflation and interest rates vary over time by an order of magnitude more than aggregate housing appreciation. We do not claim that monetary policymakers set policy based on rules that incorporate housing appreciation, and we are unaware of evidence that the empirical correlation is widely known, much less exploited by policymakers.

real per-capita nonhousing GDP; H = log of real per-capita housing investment; M = the federal funds rate; and R = the nominal interest rate on conventional 30-year mortgages. Conventional tests indicate that all six variables clearly belong in the system. The VAR is estimated over the period 1966:Q1 to 1998:Q2 and is identified by orthogonalization based on the common ordering of prices, quantities, and interest rates: P^n, P^h, Y, H, M, and R.

Figure 1 plots impulse responses from the benchmark housing-monetary VAR with two standard error bands. The fifth column demonstrates the main point: monetary policy shocks have significantly larger and more rapid effects on housing than nonhousing. After a shock to M, Y declines slowly and steadily for 10 quarters before beginning to rise – a well-known, robust feature of VAR and other small-scale macro models.[8] In contrast, H declines much more sharply and quickly. The peak decline is about seven times that of Y and it only takes six quarters, a year less than Y. Likewise, P^h declines much more and more quickly than P^n; the so-called price puzzle appears only for P^n. This difference between housing and nonhousing dynamics reflects heterogeneity among VAR coefficients that motivates disaggregation of these components.

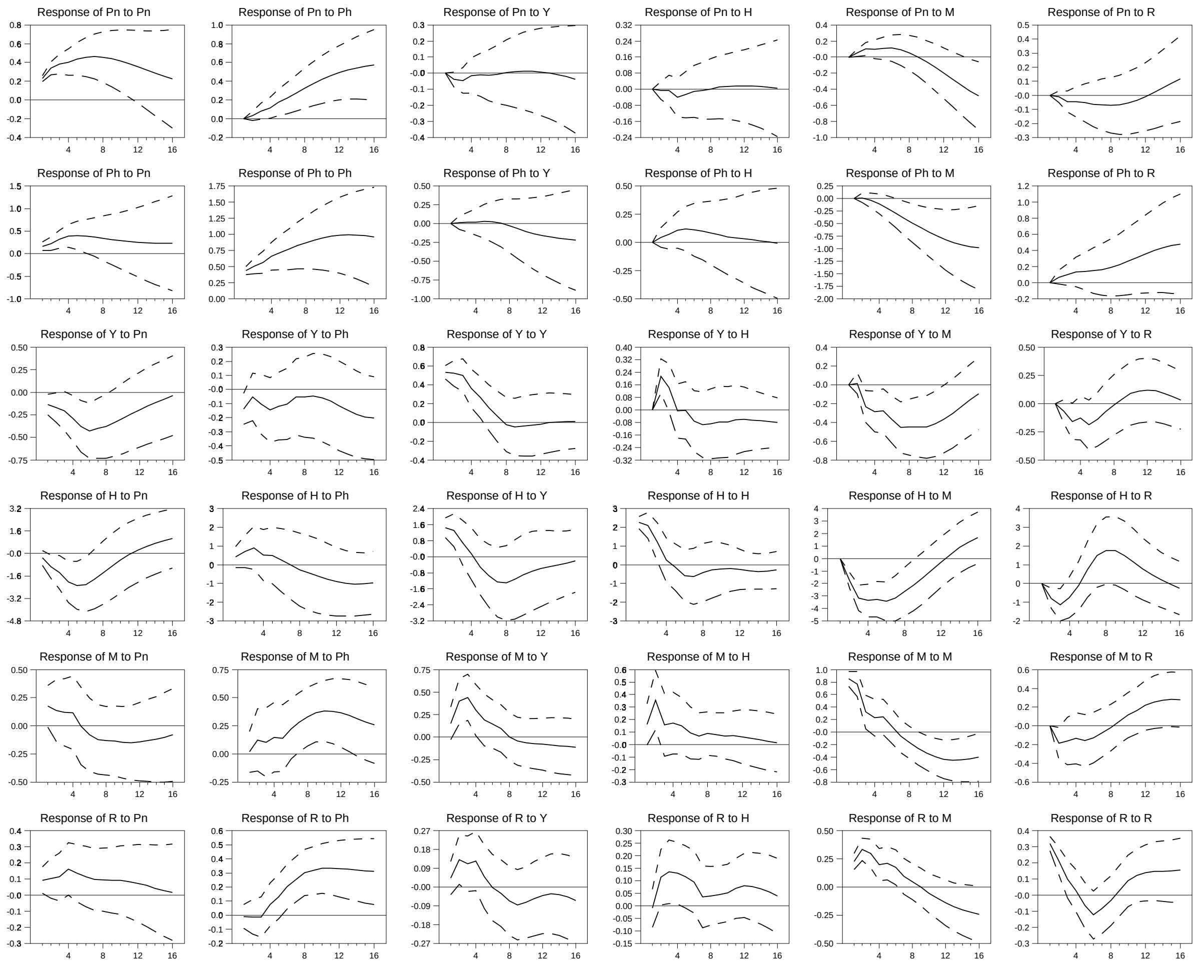
Table 1 shows the importance of housing in variance decompositions. Monetary policy accounts for much larger fractions of variation in housing than nonhousing, especially in the first year where M explains five times more variation in H than Y. Although the contributions of both quantity variables is relatively small, the contribution of H is disproportionately large compared to Y, considering its small GDP share. Finally, P^h plays a surprisingly large role in explaining the variance of system variables, especially for two years or more and for prices and interest rates.

2.2 Regional Disaggregation

Perhaps more than any other component of GDP, housing warrants a disaggregated approach because of its special characteristics. According to Smith, Rosen, and Fallis [56], "Foremost among the special characteristics are the durability, spatial fixity, and heterogeneity of housing and the extensive involvement of governments in housing and related input markets." For these reasons and more, housing supply and demand are determined in distinct local markets. This reasoning

[8] For VAR evidence, see Christiano, Eichenbaum, and Evans [12] and Leeper, Sims, and Zha [37]. The output response to a monetary policy shock, our main focus, is quite robust to variable orderings, data sources, handling of nonstationarity, and other specification issues. Walsh [61] points out that this response is typical in traditional and contemporary large-scale macro models.

Figure 1: Monetary VAR with Housing



motivates our regional disaggregation.[9]

Another factor motivating disaggregation is the volatility of regional housing markets relative to macroeconomic fluctuations. The federal funds rate is almost three times more variable than *aggregate* housing appreciation (standard deviations of 3.4 versus 1.3 percent), but interest rates tend to be relatively homogeneous across regions whereas time series variation in regional appreciation

[9]Because we take a macroeconomic approach, the paper differs from the traditional focus of the housing literature in two important ways. Housing research centers on supply or demand elasticities, but reduced-form VARs do not identify such elasticities (although VAR coefficients embed them). Instead, our focus is on the aggregate dynamic responses of housing investment and prices to system shocks. Housing research has been concerned predominantly with price elasticities, whereas we are interested in interest rate elasticities as well.

Table 1: Variance Decompositions of Forecasted Variables

Forecast Variable	Percent of variance accounted for by: After 4 quarters P^n	P^h	Y	H	M	R	After 8 quarters P^n	P^h	Y	H	M	R	After 16 quarters P^n	P^h	Y	H	M	R
P^n	89	4	1	0	4	1	77	18	0	0	3	1	40	43	0	0	15	1
P^h	21	75	0	1	1	2	16	70	0	1	10	3	7	54	1	0	31	7
Y	12	4	66	5	10	3	27	3	37	3	26	4	24	6	25	3	38	4
H	13	3	8	22	49	5	18	2	17	10	55	7	17	5	8	8	51	10
M	3	2	19	9	63	3	4	10	18	8	55	4	4	20	10	5	51	10
R	9	1	7	8	49	26	10	22	5	7	37	19	6	45	4	4	27	15

NOTE: Rows may not add to 100 exactly due to rounding.

rates (standard deviations of 8.3 percent in the median region and 11.2 percent in the 90th percentile region) swamps that of the funds rate. Thus, while interest rates vary more at the aggregate level, housing appreciation varies more in regional markets.

Regional disaggregation, however, runs against conventional wisdom. Macroeconomists widely agree that monetary policy cannot be used to influence, control, or target the economic conditions of particular regions.[10] We subscribe to this view as well, but argue that the reverse is not true. Regional economic conditions can, in fact, influence aggregate outcomes. In particular, the magnitude and duration of aggregate responses of income, inflation, and other macroeconomic variables to monetary policy shocks vary over time and depend on regional economic conditions. Two conditions are necessary to support this hypothesis: 1) regions must respond differentially to monetary policy actions; and 2) initial economic conditions must vary across regions when monetary policy actions are taken.

Figure 2, reproduced from Carlino and DeFina [8], provides evidence on heterogeneous sensitivity of regions to monetary policy. After two years, the dynamic response of real income to a monetary policy shock (tightening) varies widely across U.S. states. In some states, such as Michigan, Indiana, and Arizona, income declines on average by about 2 percent whereas in other states, such as Texas, Wyoming, and New York, income only declines by about 1/2 percent. Both

[10]This view is illustrated by a video game in the visitors' lobby at the Federal Reserve Board called "You Are the Chairman," which asks players to select a monetary policy — tighter, looser, or no change — in response to different scenarios. Most scenarios focus on traditional policy responses to macroeconomic developments in real GDP, inflation, and interest rates. But in one scenario, unemployment begins to rise in several farm-belt states. Players are told the correct answer is no change.

estimates deviate significantly from the national average of 1.2 percent.

Figure 2: Sensitivity of Regional Income to Monetary Policy

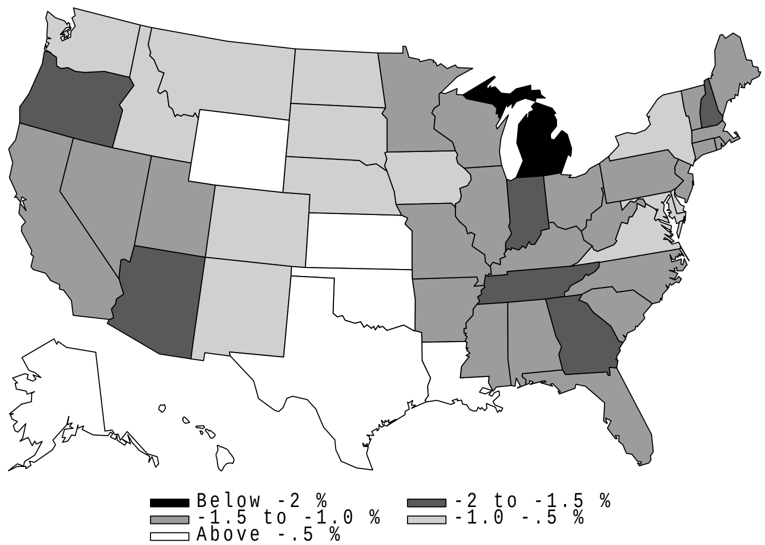


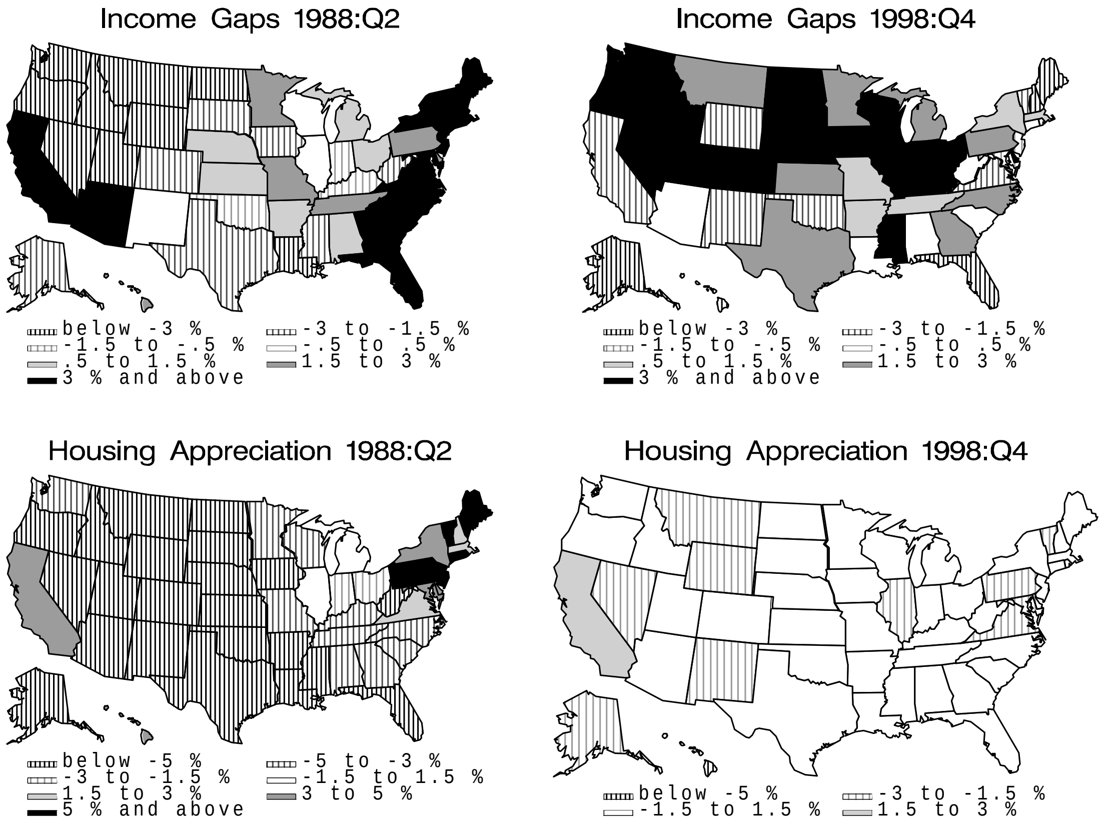
Figure 3 provides evidence on heterogeneous initial conditions in real income gaps and housing price appreciation in the late 1980s and 1990s.[11] These periods follow long economic expansions and precede monetary tightening, so the aggregate output gap was relatively high at both times and regional conditions would be expected to be homogenous. In fact, however, regional economic conditions varied quite significantly across these periods. In the late 1980s, income and housing appreciation were relatively high along the east coast and parts of the west coast whereas economic conditions were weaker among central states. In the late 1990s, income was relatively high in central states while the coastal regions were weaker, and housing appreciation was much more uniform across all regions.[12]

The data show that regions respond differentially to monetary policy and that regional economic conditions vary significantly across tightening episodes. These conditions, and the macro

[11] State-level real income gaps are actual income less trend income, where the trend is log linear with a break in 1973, expressed in percent. State-level housing price appreciation is the 2-year growth in housing prices less the total U.S. growth, expressed in percentage points.

[12] Regional conditions in the late 1970s (not shown) were also quite different in that they were very homogeneous. For example, virtually all states experienced large positive income gaps.

Figure 3: Regional Economic Conditions in the 1980s and 1990s



NOTE: See text for details.

evidence on housing, suggest that housing is a reasonable choice for introducing heterogeneity into a macroeconometric model for evaluating monetary policy. Numerous other sources of heterogeneity besides housing are relevant as well, and many of these could provide complementary justification for regional heterogeneity and disaggregation.[13]

[13] Hayo and Uhlenbrock [30] emphasize the industry mix across regions, with some industries more interest-sensitive than others. Mankiw and Weil [42] highlight significant fluctuations in national demographic composition, which likely varies across states as well. Johnes and Hyclak [33] find that variation in local labor market conditions and human capital stocks affect regional housing prices. Poterba [47], [48] emphasizes regional variation in federal, state, and local tax policies. Mayer and Somerville [43] and Malpezzi [40] show strong effects of local government land regulation and public goods on housing price and supply dynamics. Lamont and Stein [36] demonstrate that leverage varies across U.S. cities, and that cities with relatively high loan-to-value ratios react more to regional income

3 Theoretical Framework

This section sketches out the economic framework underlying our HAVAR model. To our knowledge, no structural dynamic general equilibrium model with optimizing behavior exists that satisfactorily represents our housing-monetary framework. The models of Manchester [41], Ortalo-Magne and Rady [46], Muellbauer and Murphy [45] contain many essential elements and yield vectors of endogenous variables similar to our regional VAR models, but they do not include monetary policy or heterogeneous regional housing markets. Development of the complete structural model underlying the housing-monetary HAVAR model is feasible but ambitious and left for future research.[14]

Figure 4 provides an overview of the theoretical framework underlying the HAVAR model. In this flow diagram, the boxes denote agents and markets, and arrows denote the economic connections among agents characterizing monetary transmission. The economy comprises two main sectors. At the national level, the monetary authority and financial intermediary determine inflation and interest rates. The monetary authority follows a policy rule to achieve a price or inflation goal P, the weighted average of P^h and P^n, using a monetary instrument M, the short-term federal funds interest rate.[15] The financial intermediary matches savers and investors by creating mortgages secured by the housing stock, and determines the mortgage asset prices, long-term interest rate R, by a term-structure relationship with expected future M. At the local level, households and firms take R as exogenous and determine income, y_i, housing investment, h_i, and the real housing price, $q_i = p^h/P$, for region i. The macroeconomic values of these variables (Y, H, and, Q) are simply the appropriately weighted aggregates across all regions.

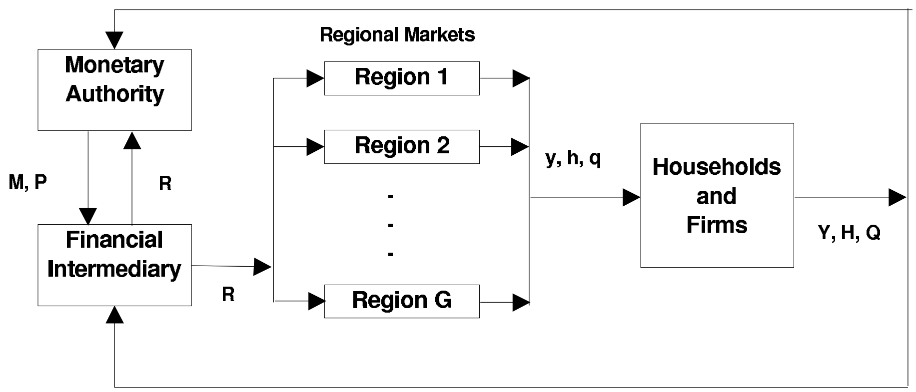
Obviously, this theoretical framework quickly becomes econometrically intractible as the number of regions increases, so we impose three simplifying assumptions. First, we adopt the traditional representative agent view for macroeconomic agents. The monetary authority and financial intermediary consider only Y, H, and Q in their decision making, and not particular regional val-

shocks. Finally, Fergus [19] notes that climate and abnormal weather conditions contribute to regional variation in depreciation and maintenance costs.

[14]Throughout the paper, the reader should remember that our HAVAR framework is not based on an explicit and structural theoretical model. Consequently, one cannot place structural interpretations on the econometric results.

[15]Bernanke and Blinder [4] recommend the funds rate, but the literature is replete with alternative monetary instruments, targets, and policy rules that could be used instead. For VAR-based examples that incorportate monetary aggregates, bank reserves, and other aspects of monetary intervention see Cochrane [16], Christiano, Eichenbuam, and Evans [11]; Leeper, Sims, and Zha [37], and Bernanke and Mihov [5].

Figure 4: The Housing-Monetary HAVAR Model Framework



ues y_i, h_i, and q_i. Second, only R influences regional activity contemporaneously. Regional agents don't particularly care about M *per se*, only its indirect effect on R. Regional agents also don't care about P *per se*, only its indirect effect on real interest rates. Because P moves sluggishly, and expectations of P even more so, real interest rates defined with lagged P – which are embedded in the system – are quite similar to real rates defined with current or expected P. Third, activity in individual regions does not affect other regions contemporaneously. The latter two assumptions only apply contemporaneously; all variables and all regions influence each other with a one-period lag.

This theoretical framework yields the following monetary transmission mechanism. The monetary authority sets policy by moving the funds rate to achieve an aggregate price target. This policy decision directly, and indirectly with a lag, affects the mortgage rate. Movements in the mortgage rate then directly affect regional income, housing investment, and housing appreciation, both contemporaneously and with lags.[16] This monetary transmission mechanism is a bit simplistic because it focuses on an interest rate spread and excludes traditional money and banking channels. However, spread-based transmission mechanisms have some empirical support, as shown in Stock and Watson [58] and Friedman and Kuttner [21].

[16]Note that because the system includes inflation, implicitly it is the *real* mortgage rate that influences regional activity.

4 The Housing-Monetary HAVAR Model

4.1 General Specification

The preceding theoretical framework suggests a housing-monetary HAVAR data vector $X_t = [P_t \quad Q_t \quad Y_t \quad H_t \quad M_t \quad R_t]'$, which should be divided into two parts: $X_t = [X_t^m \quad X_t^a]'$. One part is $X_t^m = [P_t \ \ M_t \ \ R_t]'$, a data vector of macroeconomic (superscript m) variables determined nationally. The other is $X_t^a = [Q_t \ \ Y_t \ \ H_t]'$, a data vector of aggregated (superscript a) variables determined regionally.

Using this notation, the structural representation of the conventional monetary VAR with housing is

$$\Gamma_0 X_t = K + \sum_{l=1}^{L} \Gamma_l X_{t-l} + U_t \ . \tag{1}$$

The partitioned structural parameters and innovations are

$$\begin{aligned}
\Gamma_l &= \left[\begin{array}{c|c} \Gamma_l^{mm} & \Gamma_l^{ma} \\ \hline \Gamma_l^{am} & \Gamma_l^{aa} \end{array}\right] \quad \forall \ \ l \geq 0 \\
K &= [K^m | K^a]' = [K^P \ \ K^M \ \ K^R \mid K^Q \ \ K^Y \ \ K^H]' \\
U_t &= [U_t^m \mid U_t^a]' = [U_t^P \ \ U_t^M \ \ U_t^R \mid U_t^Q \ \ U_t^Y \ \ U_t^H]' \ ,
\end{aligned}$$

where the dimension of Γ_l is (6×6) and the dimensions of K and U_t are (6×1). The reduced-form representation of this VAR is

$$X_t = \Pi_0 + \sum_{l=1}^{L} \Pi_l X_{t-l} + V_t \tag{2}$$

where

$$\Pi_0 = \Gamma_0^{-1} K \qquad \Pi_l = \Gamma_0^{-1} \Gamma_l \ \ \forall \ \ l > 0 \qquad V_t = \Gamma_0^{-1} U_t$$

By placing restrictions on Γ_0 and the structural innovation covariance matrix, $\Sigma = E(U_t U_t{}')$, one can identify and estimate the structure in the usual way (see Hamilton [29], chapter 11).

The theoretical framework underlying the HAVAR model motivates disaggregation of X_t^a into regional data vectors $x_{it}^a = [q_{it} \quad y_{it} \quad h_{it}]'$ for $i = 1, \ldots G$ regions, where lowercase letters denote regional variables. Following the methodology pioneered by Theil [59], we incorporate regional VAR models — the fundamental microeconomic relations — into the conventional monetary VAR model to obtain the structural HAVAR model

$$\Gamma_0 Z_t = K + \sum_{l=1}^{L} \Gamma_l Z_{t-l} + U_t \tag{3}$$

where now

$$\begin{aligned}
\Gamma_l &= \left[\begin{array}{c|c} \Gamma_l^{mm} & \Gamma_l^{ma} \\ \hline \Gamma_l^{am} & \gamma_l \end{array}\right] = \left[\begin{array}{c|cccc} \Gamma_l^{mm} & s_{1,t-l}\Gamma_{l,1}^{ma} & s_{2,t-l}\Gamma_{l,2}^{ma} & \dots & s_{G,t-l}\Gamma_{l,G}^{ma} \\ \hline s_{1,t-l}\Gamma_{l,1}^{am} & \gamma_{l,11} & \gamma_{l,12} & \dots & \gamma_{l,1G} \\ s_{2,t-l}\Gamma_{l,2}^{am} & \gamma_{l,21} & \gamma_{l,22} & \dots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ s_{G,t-l}\Gamma_{l,G}^{am} & \gamma_{l,G1} & \dots & \dots & \gamma_{l,GG} \end{array}\right] \quad \forall \ l \geq 0 \\
Z_t &= [X_t^m \mid x_{1t}^a \ \ x_{2t}^a \ \ \dots \ \ x_{Gt}^a]' \\
K_t &= [K^m \mid k_1 \ \ k_2 \ \ \dots \ \ k_G]' \\
U_t &= [U_t^m \mid u_{1t}^a \ \ u_{2t}^a \ \ \dots \ \ u_{Gt}^a]' \ .
\end{aligned}$$

The time-varying weights, $s_{i,t-l}$, are elements of the size-weighted linear aggregator $X_t^a = \sum_{i=1}^{G} s_{it} x_{it}^a$, where s_{it} is 1 for quantity variables (X_t^a is a sum) and a measure of region size for prices, growth rates, and per-capita variables (X_t^a is a mean).[17] Aggregating the regionally determined variables in equation (3) yields what Theil called the "true" macro model (and we call the HAVAR model) denoted by *:

$$\Gamma_t^* X_t = K_t^* + \sum_{l=1}^{L} \Gamma_{t-l}^* X_{t-l} + U_t^* \tag{4}$$

where

$$\begin{aligned}
\Gamma_{t-l}^* &= \left[\begin{array}{c|c} \Gamma_l^{mm} & \overline{\Gamma}_{t-l}^{ma} \\ \hline \overline{\Gamma}_{t-l}^{am} & \overline{\gamma}_{t-l}^{aa} \end{array}\right] = \left[\begin{array}{c|c} \Gamma_l^{mm} & \sum_{i=1}^{G} s_{i,t-l}\Gamma_{l,i}^{ma} \\ \hline \sum_{i=1}^{G} s_{i,t-l}\Gamma_{l,i}^{am} & \sum_{j=1}^{G}\sum_{i=1}^{G} \gamma_{l,ij} w_{j,t-l} \end{array}\right] \quad \forall \ l \geq 0 \\
w_{j,t-l} &= s_{j,t-l}\left(\frac{x_{j,t-l}^a}{X_{t-l}^a}\right) \\
K_t^* &= [K^m \mid \overline{k}^a]' = [K^m \mid \sum_{i=1}^{G} s_{it} k_i^a]'
\end{aligned}$$

[17] The size variable reflects the magnitude, or share, of region i relative to the aggregate and depends on the nature of x_{it}. For example, s_{it} is the population share of the region for per-capita variables and is $1/G$ for unweighted means.

$$U_t^* \;=\; [U_t^m \;\mid\; \overline{u}_t^a]' = [U_t^m \;\mid\; \sum_{i=1}^{G} s_{it} u_{it}^a]' \; .$$

Unlike standard macro models, the parameters of the HAVAR macro model, Γ_{t-l}^*, are time-varying, except for the nine parameters in Γ_l^{mm} associated with the nationally determined variables. In particular, the γ_{t-l}^{aa} matrices depend on $s_{j,t-l}$ and $(x_{i,t-l}^a / X_{t-l}^a)$, which both vary over time for secular and cyclical reasons. As a result, the reduced-form of the HAVAR model,

$$X_t = \Pi_t^* + \sum_{l=1}^{L} \Pi_{t-l}^* X_{t-l} + V_t^* \tag{5}$$

where

$$\Pi_t^* = (\Gamma_t^*)^{-1} K^* \qquad \Pi_{t-l}^* = (\Gamma_t^*)^{-1} \Gamma_{t-l}^* \;\; \forall \;\; l > 0 \qquad V_t^* = (\Gamma_t^*)^{-1} U_t^* \;\; ,$$

also exhibits time-varying parameters.[18]

By comparing the HAVAR, equation (5), and conventional VAR, equation (2), one can see the implications for dynamic analysis. The HAVAR model exhibits time-varying impulse responses because of time variation in the Π_{t-l} parameters. Furthermore, the distribution of regional structural innovations, u_{it}^a, affects aggregate dynamic responses because Γ_t^* is time-varying and has heterogeneity in the elements of γ_0. Thus, HAVAR aggregate dynamic responses are state dependent. Conventional macro VARs implicitly impose the highly restrictive, and empirically inconsistent, assumption that all regions are in equilibrium when shocks hit the system. HAVAR models relax that assumption, so aggregate dynamic responses depend on how current and cumulative shocks are distributed across regions.

Time variation means the HAVAR model falls in the class of nonlinear time-series models cited in the introduction. Despite this nonlinearity, the HAVAR model does not exhibit multiple *equilibria* because the micro models are linear and assumed stable. However, any given aggregate shock can produce a multiplicity of aggregate impulse responses, depending on the exact distribution of micro shocks. At the same time, any given aggregate impulse response is consistent with a multiplicity of distributions of micro shocks. The technical appendix explains these issues in more detail.

Differences in the dynamic responses of the HAVAR and conventional VAR models also are related to the issue of aggregation bias raised by Theil [59]. Because the regional models have

[18] Note that the constant also is time-varying, unlike in Theil's original analysis, because we are working with the underlying structural VAR.

fixed parameters, an approximation of the HAVAR model can be obtained using time-series average weights, $\overline{s}_i$ and $\overline{w}_i$, to aggregate the regional fixed parameters. The difference between the weighted sum of regional parameters and the conventional parameters reflects the extent of bias in the conventional parameter due to aggregation.[19] In time series models like HAVAR, a central effect of aggregation is to bias aggregate autoregressive parameters upward, or equivalently adjustment speeds downward. In other words, dynamic responses of aggregate models exhibit more persistence than do the responses of micro models.[20]

4.2 Simplifying Assumptions

Obviously, neither equation (4) nor (5) of the HAVAR model can be estimated like conventional fixed-parameter macro VAR models. A conventional version of the housing-monetary VAR has six variables and thus $36L+6$ reduced-form parameters to estimate, or 180 with $L = 4$. In comparison, the HAVAR model has $3G + 3$ variables and thus $(3G + 3)^2L + 3G + 3$ reduced-form parameters to estimate. With state-level data ($G = 50$ and $L = 4$), the HAVAR model has 117,045 reduced-form parameters. Clearly, it is infeasible to estimate such a model given the limited availability of disaggregated time series data.

Consequently, simplifying restrictions are required to make the model empirically tractable. Fortunately, the theoretical framework and the behavioral dichotomy between the national and regional markets suggest three economically sensible restrictions, which we impose on the HAVAR model. These restrictions, along with the standard unit normalization of the diagonal of Γ_0, dramatically reduce the number of parameters to be estimated.

ASSUMPTION 1: ***Macro Representative Agent*** *– The monetary authority and financial intermediary ignore regional heterogeneity. Thus, only the aggregate values (sum or mean) of the regional variables, X_t^a, matter for the behavior of the macro variables, X_t^m.*[21]

[19]Goodman [25] examines aggregation bias in a relatively simple *calibrated* model of housing demand and finds that the effects on aggregate behavior are modest in the long run. The HAVAR model, however, is more sharply influenced by aggregation because it is considerably more complicated, incorporates actual data, and focuses on the short run.

[20]For examples in the literature, see Granger [26], Trivedi [60], Lippi [39], Lewbel [38], Schuh [54], and Abadir and Talmain [1].

[21]Although this assumption does not necessarily imply that monetary and financial agents impose the representative agent assumption in formulating their behavior, the implications are essentially the same. It also does not suggest that the monetary authority does not care about the distribution of regional variables, or is unaware of the distribution. It

This assumption eliminates heterogeneity in parameters governing the feedback from aggregated variables, x_{it}^a, to macro variables, X_t^m. Specifically,

$$\Gamma_{l,i}^{ma} = \Gamma_l^{ma} = \begin{bmatrix} \Gamma_l^{PQ} & \Gamma_l^{PY} & \Gamma_l^{PH} \\ \Gamma_l^{MQ} & \Gamma_l^{MY} & \Gamma_l^{MH} \\ \Gamma_l^{RQ} & \Gamma_l^{RY} & \Gamma_l^{RH} \end{bmatrix} \ \forall \ i, l \ .$$

In other words, all changes in x_{it}^a have the same effect on X_t^m regardless of the region in which they occur and whether the changes are spread evenly across regions. This assumption reduces the number of parameters in these matrices from $9GL$ to $9L$.

ASSUMPTION 2: ***Monetary Transmission*** *— Contemporaneously, monetary policy is transmitted from the funds rate, M_t, to the mortgage rate, R_t, to regional variables, x_{it}^a. Other macro variables influence x_{it}^a indirectly with a lag. The responses of x_{it}^a to monetary policy are heterogeneous across regions.*

This assumption limits the *contemporaneous* feedback from macro variables, X_t^m, to aggregated variables, x_{it}^a. It implies that

$$\Gamma_{0,i}^{am} = \begin{bmatrix} 0 & 0 & \Gamma_{l,i}^{QR} \\ 0 & 0 & \Gamma_{l,i}^{YR} \\ 0 & 0 & \Gamma_{l,i}^{HR} \end{bmatrix} \ \forall \ i \qquad \Gamma_{l,i}^{am} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \ \forall \ i, \ \ l > 0 \ .$$

However, all macro variables, X_t^m, eventually affect all regional variables, x_{it}^a, with a lag through the mortgage rate.[22] Clearly, it would be preferable to include the contemporaneous federal funds and inflation rates. The former is used by rational agents to incorporate government policy and the latter is necessary for real rate calculations. With slightly more time series data, it would be feasible to drop this restriction. In addition, this assumption is only marginally helpful, reducing the number of parameters from $9GL$ to $3GL$.

ASSUMPTION 3: ***Limited Regional Interdependence*** *– Economic activity (income and housing) in a region is not directly affected by activity in other regions contemporaneously. Activity in other regions has effects with a lag, both through the mortgage rate and directly through lagged aggregate activity. But in all cases the pairwise economic interactions among regions implicitly are homogeneous.*

only means that it does not, and perhaps cannot, account for heterogeneity in setting policy.

[22]This assumption finds some empirical support from Sill [55], which reports that it takes three months before regions exhibit similar responses to a common aggregate shock.

This assumption limits the extent and heterogeneity of interregional economic relationships. It implies that

$$\gamma_{l,ij} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \forall \;\; i \neq j \;\; \text{and} \;\; \forall \;\; l \; .$$

Regions are interrelated, but only indirectly through lags of X_t^m. Furthermore, the fact that only aggregate activity affects a region means that all regions affect each other in the same way. For example, an increase in aggregate income will affect all regions the same regardless of the region(s) in which income rose. This assumption dramatically reduces the impact of heterogeneity in the system, but it also gains most of the identification power by reducing the number of parameters exponentially from $3G^2L$ to $3GL$ as $\gamma_{l,t}^{aa}$ becomes block diagonal.[23]

After imposing assumptions 1-3, each Γ_l matrix contains up to $12G + 18$ parameters to be estimated. This is still a lot — for example, with $G = 50$ and $L = 4$ there are 2,540 structural parameters plus 50 innovation variances. But for certain empirical applications, enough time-series cross-section data are available to make it feasible to estimate a restricted model of this type.

4.3 Identification and Estimation

We use a two-step triangular factorization method of identification.[24] There are two reasons for using this approach. First, there are theoretical reasons for pre-estimating some of the parameters of the regional impact matrices. In particular, the (assumed) exogeneity of mortgage rates to regions allows identification of the contemporaneous effects of mortgage rates on regional activity. Second, maximum likelihood techniques of estimating structural VAR parameters can quickly become infeasible for HAVAR models with large numbers of micro agents. In our application with

[23] With more time series data, one could explore less restrictive assumptions that would allow for nonzero $\gamma_{l,ij}$ $\forall$ $i \neq j$ (off-diagonal elements). One is to include the *complementary aggregate*, $\widetilde{X}_{it}^a = X_t^a - x_{it}^a$, of regional variables in each regional model. The off-diagonals would become nonzero but there would be no regional heterogeneity. Another possibility is to weight the off-diagonals by measures of the economic relationships among regions, such as geographic distance or input-output structure. For examples of the latter, see Horvath [31] and Horvath and Verbrugge [32]. These additional $9G$ weights could be calibrated from data rather than estimated.

[24] It is important to reemphasize here that: 1) our central ideas on the HAVAR model hold for any identification scheme and any linear economic model; and 2) our main results are not particularly sensitive to the identification scheme. We focus primarily on the dynamic response of income to monetary tightening, and this response is quite robust to alternative specifications, as shown in Christiano, Eichenbaum, and Evans [12].

27 regions, it would be necessary to estimate 327 structural parameters – 3 aggregate and 27 regional variances, plus 18 aggregate and 324 regional impact matrix parameters.[25] Reassuringly, this two-step orthogonolization produces HAVAR dynamics that are qualitatively similar to those of a similarly orthogonalized conventional VAR (as it should).

The goal is to obtain an estimate of the simplified HAVAR contemporaneous impact matrix,

$$\Gamma_0 = \left[\begin{array}{c|c} \Gamma_0^{mm} & \Gamma_0^{ma} \\ \hline \Gamma_0^{am} & \gamma_0^{aa} \end{array}\right] ,$$

using the estimated reduced-form covariance matrix, $\widehat{\Omega} = E(\widehat{V}_t \widehat{V}_t')$. The reduced-form covariance matrix can be decomposed as $\Omega = A\Sigma A'$. Triangular factorization yields the structural innovation covariance matrix, $\Sigma = E(U_t U_t')$, and the lower triangular matrix A is an estimate of Γ_0^{-1}. Because some elements of Γ_0 are identified from the simplifying restrictions and econometric assumptions, we use A along with the restrictions and assumptions to solve for the remaining components of Γ_0.

Before turning to the actual estimation and identification steps, note that the variable ordering in the HAVAR model is now $[P_t \quad Y_t \quad H_t \quad Q_t \quad M_t \quad R_t]$. This ordering differs slightly from the ordering in section 2.1 in that the real housing price now follows income and housing investment. The reason is that Q_t is not a price like P_t, the aggregate price of goods and services, which is likely to move sluggishly. Rather, Q_t is an asset price that is likely to move much more, and more frequently, than goods prices. This reordering alters the preceding model notation somewhat but is not a substantive change.

The two-step identification procedure involves first estimating each of the regional VAR models separately to obtain the 12 parameters in $\Gamma_{0,i}^{am}$ and $\gamma_{0,ii}$ for each of the G regions. Then, taking the regional parameters as given, the remaining 18 aggregate parameters in Γ_0^{mm} and Γ_0^{ma} are estimated in the second step.

Two factors motivate separate estimation of the regional models in step one. First, conventional triangular factorization applied to the HAVAR model will not work properly because it would force an ordering among regions and there is no defensible theoretical justification for such an ordering. Instead, we only want to order the x_{it}^a variables within regions, not across regions. Second, the mortgage rate is assumed to be exogenous to each region. Consequently, the $\Gamma_{0,i}^{am}$ can be obtained from OLS estimates of regional "near" VARs based on the vector $[R_t | x_{it}^a]'$, where only the current

[25] Ultimately, the best solution to the identification problem is to move to a structural model of the regional agents, which would eliminate many of the necessary parameters. We plan to pursue this strategy in future research.

value of R_t enters the model. Assumptions 2 and 3 justify the independent estimation of regional models.

Estimates of the $\gamma_{0,ii}$ can be obtained from the reduced-form covariance matrices of the regional models. However, because the lags of all X_t^m variables influence the reduced-form equations for each region through the national VAR, estimates of V_t^m must be incorporated in the regional covariance matrices. Assumption 1 justifies independent OLS estimation of the three reduced-form VAR of macro variable equations, X_t^m, using only aggregate data, which yields the residuals $\widehat{V}_t^P$, $\widehat{V}_t^M$, and $\widehat{V}_t^R$. Then the regional reduced-form covariance matrices are $\widehat{\Omega}_i = E(\widehat{V}_{it}\widehat{V}_{it}')$, where $\widehat{V}_{it} = [\widehat{V}_t^P \quad \widehat{v}_{it}^y \quad \widehat{v}_{it}^h \quad \widehat{v}_{it}^q \quad \widehat{V}_t^m \quad \widehat{V}_t^R]'$. Triangular factorization of the Ω_i yields model-consistent estimates of $\gamma_{0,ii}$.

In the second step, we construct the HAVAR macro reduced-form covariance matrix, $\widehat{\Omega}^* = E(\widehat{V}_t^* \widehat{V}_t^{*\prime})$, using the weighted sums of the regional $\widehat{v}_{it}^a$ residuals. Triangular factorization of this covariance matrix gives an estimate of Γ_t^*, but this estimate is not fully consistent with the theoretical restriction imposed earlier. However, using this estimate, along with the weighted sums of the $\widehat{\Gamma}_{0,i}^{am}$ and $\widehat{\gamma}_{0,ii}$ estimates, one can solve for the unidentified macro elements of Γ_t^*.

5 Econometric Specification

Data availability constrains our application of the HAVAR model to housing activity. Although housing is conceptually ideal, housing data are only available for a relatively small subset of U.S. regions. Thus, we view our application as illustrative of the potential scope and magnitude of the effects of heterogeneity on aggregate behavior. The aggregate data from this subsample of regions properly reflect these effects *for that subsample* so our conclusions about heterogeneity and aggregation are valid for the subsample. However, the results may not represent actual U.S. behavior accurately because the data are non-representative. For this reason, we cannot (and do not) draw firm conclusions about the actual or optimal conduct of U.S. monetary policy. Nevertheless, our results firmly suggest that the effects of heterogeneity may be important in actual monetary policy models.[26]

We contructed a balanced longitudinal panel of quarterly data for a subsample of 27 U.S. metropolitan statistical areas (MSAs) during the period 1986 to 1996; see the Data Appendix for

[26] In particular, with more complete and representative data it would be appropriate to use HAVAR models for policy analysis. Toward that end, we are investigating the availability of more complete housing data as well as alternative applications of the model using more readily available employment data.

details. The panel includes MSA-level data on housing investment (measured as the quantity of housing starts) and housing prices, plus state-level data on personal income scaled by the MSA population share.[27] The MSAs were selected based on the sample period of their data, which is primarily determined by housing starts; only MSAs with long time series were included in order to obtain a sufficient sample for VAR estimation. Aside from recent papers by Mayer and Somerville [43] and Follain and Harter [20], the MSA-level housing data have not been explored much. Unlike these panel data studies, however, we estimate individual time-series models for each MSA at the quarterly frequency.

The data are transformed as follows. Income y_{it} and housing investment h_{it}, are in levels because they likely share a long-run common trend.[28] Both interest rates, M_t and R_t, are also in levels. Both P_t and $q_{it} = (p_{it}^h/P_t)$, are in growth rates: i.e., aggregate price inflation and real house price appreciation.[29] Standard lag length tests recommend two quarters ($L = 2$) for the HAVAR model variables. Although two lags may seem short compared with other macro models, disaggregated data tend to be much less persistent than aggregate data, and the aggregate data in our panel are a relatively small fraction of the U.S. total.

All VARs are unrestricted and estimated with OLS. The standard policy simulation is monetary tightening, a transitory 100-basis-point shock to the federal funds rate. It is important to note that the dynamic properties of monetary VARs estimated over shorter, more recent sample periods are quite different from those estimated over longer sample periods *regardless of the level of data aggregation* (e.g., total U.S. versus state-level). The income response to a funds rate shock is considerably smaller, slower, and less significant in models estimated with short-sample data.

[27] In addition to starts, national income account measures of housing investment include the quantity associated with quality improvements. However, the quality component does not change much at the high frequencies for which monetary policy is concerned. The business cycle correlation coefficient between detrended starts and detrended residential investment from the national income accounts is 0.9.

[28] The long-run relationship between regional income and housing investment is somewhat tenuous because housing is defined as the quantity of housing starts rather than the quantity of real investment, as in the national accounts. Starts do not incorporate trend increases in house size and quality. In fact, housing starts actually declined on average during the period 1986 to 1996, whereas income and housing investment increased. Nevertheless, we found that the general qualitative and quantitative results on the heterogeneity of the HAVAR aggregate dynamic responses are robust to alternative treatments of nonstationarity, such as first differencing the entire system.

[29] There is also a practical reason for specifying appreciation rather than the real price level. The underlying housing price data are indexes, so the ratio of price indexes does not provide a meaningful measure of the actual relative price level. In fact, real prices are the same for all regions in the base year. However, the difference between the growth rates of the two prices is an accurate measure of the rate of change in the relative price level.

6 Estimation Results and Aggregation Bias

6.1 Regional Models

Figure 5 plots the impulse responses of regional income, investment, and appreciation (y, h, and q) to a monetary tightening. In each panel, the plotted responses are those from the region with the 10th, 50th, or 90th percentile peak, or absolute maximum, response for that variable; thus, the region associated with each percentile may differ across panels. For the reader's reference, Appendix Table 2 lists the peak response and mean lag of each variable for all 27 regions. The gray shaded region in the figure is the band of two standard errors around the median response.[30] Because the shock is a monetary tightening, the peak response is the minimum value of the impulse response after the shock.

The magnitude and duration of the regional responses vary widely. In the median region, income declines about 1/4 percent in about two years. This median response is essentially the same as the one produced by aggregate U.S. data over this sample period; thus it is the sample period, not the subsample of regional data, that is responsible for the shallow income response.[31] However, income in the 10th percentile region income declines nearly 1 percent in about three years, a response about four times larger in absolute value and about 50 percent slower. Variation across regional responses of housing investment and appreciation responses is even larger. The 10th and 90th percentile responses are significantly different from the median responses for up to 10 quarters.

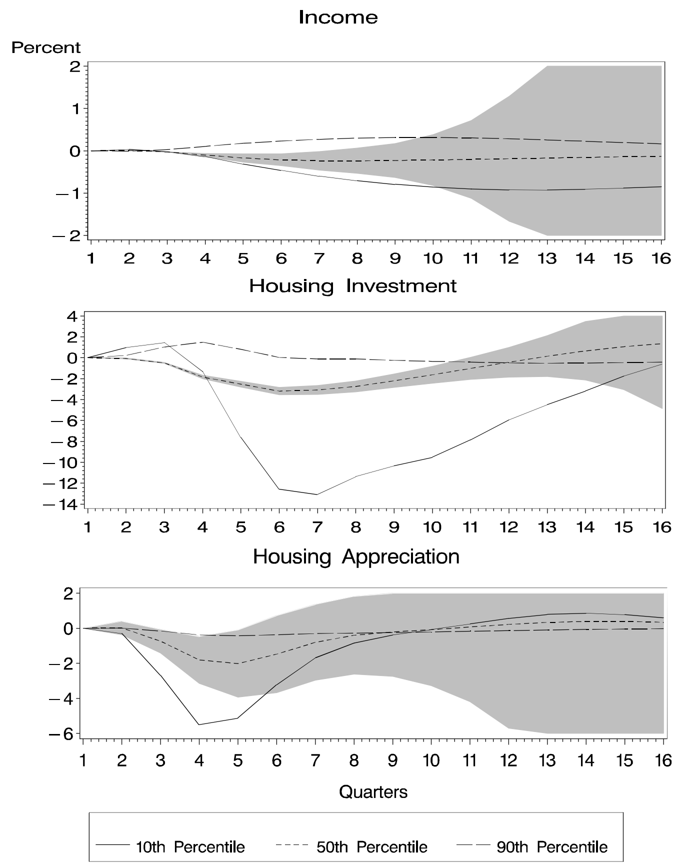
Although income and investment actually decline in most regions, they increase modestly in a few. On the surface, this seemingly counterintuitive result contrasts with the standard income decline reported throughout the monetary VAR literature. However, the regional VARs also include housing appreciation, which the mortgage rate influences the most. Mortgage rate increases lead to immediate, sharp declines in appreciation — much faster responses than in quantities. Depending on the underlying supply and demand elasticities in a region, then, it is plausible that such appreciation movements could yield positive quantity movements, at least in the short run.

The extent and nature of regional heterogeneity generally is consistent with Carlino and De-

[30] The error bands show the extent to which regional heterogeneity differs from sampling error. The fact that substantial fractions of the response distributions lie outside the error bands for many quarters suggests that regional heterogeneity is not merely sampling error around homogeneous parameters.

[31] The weak response of income (or GDP) to a funds rate shock in small monetary-macro models estimated with data beginning after the early 1980s is widely known among applied practictioners but has not been widely publicized.

Figure 5: Regional Impulse Responses



Fina [7], which reports the responses of income in nine Census regions (groups of U.S. states) to monetary tightening. However, because our regions (MSAs) are more disaggregated and because we model income and housing jointly, we find considerably greater heterogeneity — including modest sign differences — in the impulse responses.

Regional heterogeneity could arise for many different reasons. Footnote 13 offers numerous economic explanations from the literature, but there are two main potential sources: 1) heterogeneous sensitivity of regional housing and income to interest rates; and 2) heterogeneity in housing markets, such as supply and demand elasticities. Investigating these sources clearly is a fruitful area for future reseach but will require building a structural model and considerably more data. Appendix Table 2 reports some informative correlations among the peaks and mean lags of the individual regions' impulse responses.[32]

6.2 Aggregate Models

Comparing the impulse responses of the conventional VAR (equation 2), estimated with aggregate data, and the HAVAR (equation 5), estimated with disaggregated data, reveals the effects of aggregation on the dynamic properties of the model. An important issue that arises in HAVAR simulations, but not in conventional VAR simulations, is the setting of initial starting values for each region. The aggregate economy could be in equilibrium (steady state) because all regions are in equilibrium, or because the regional initial conditions (shocks) all "average out." In this section, we set all regions' initial conditions to their equilibrium values, and thus the aggregate economy is also in equilibrium, when the funds rate shock hits. Empirically, this assumption is highly restrictive and counterfactual, so we relax it in section 7.

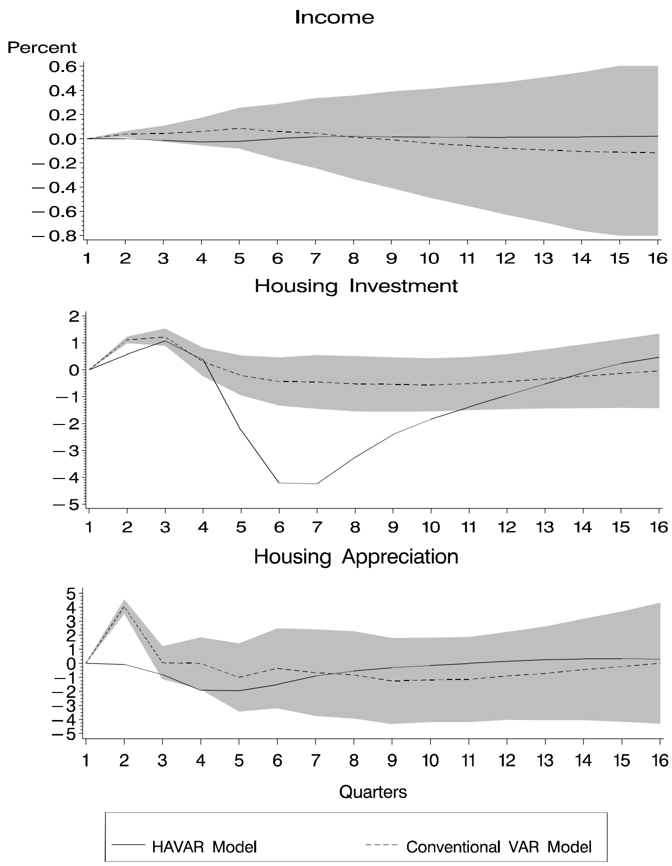
Regarding the comparison of conventional VAR and HAVAR responses, recall that the primary difference between the models is that the conventional VAR has fixed parameters whereas the HAVAR has time-varying parameters due to micro heterogeneity in parameters and size. Thus, the conventional VAR is a special case of the HAVAR where this heterogeneity is assumed away (by setting all micro parameters the same). The relevant econometric issue is whether one can reject this restriction imposed implicitly by the conventional VAR.

Figure 6 plots the impulse responses of aggregate income, investment, and appreciation (Y, H, and Q) to a monetary tightening. Each panel includes the conventional VAR response and the HAVAR response. The gray shaded area is the band of two standard errors around the conventional response. The key result is that the conventional VAR responses differ from the HAVAR responses, and in some cases the differences are quite economically and statistically significant. Two main differences arise between the responses: magnitude and duration.

Magnitude differences are more important for housing variables than income. For income, both responses are small and, for the conventional VAR, mostly insignificantly different from

[32] These correlations reveal several regularities among regional impulse responses. First, regions with relatively large negative reponses tend to have relatively low mean lags; in other words, regions hit hard by monetary shocks tend to recover quickly. Second, regions tend to exhibit relatively large negative responses in all variables or in none; in other words, the interest sensitivity of region tends to be the same in housing and nonhousing. Third, the mean lags of income and housing investment are very highly correlated, but the mean lags of housing prices essentially are unrelated to the mean lags of income and housing investment; in other words, regions are quite heterogeneous in their price adjustment and market clearing dynamics.

Figure 6: Macro Impulse Reponses



zero.[33] The HAVAR response lies near the boundary of the error band for the first year or so, but the economic difference between the responses is negligible over the full horizon. For housing investment and appreciation, however, the two model responses differ greatly. The HAVAR investment response is an order of magnitude larger in absolute value at its peak than the conventional response. For the first two quarters, the HAVAR appreciation response is quite different from the conventional response, which increases sharply in the short run. The appreciation responses differ less thereafter, but the difference is often about 1 percentage point — an economically significant gap. To summarize, aggregation substantially alters the magnitudes of some conventional impulse

[33]It is critical to keep in mind here that a conventional monetary VAR estimated with aggregate U.S. data over our 1986-96 sample produces essentially the same weak and sluggish income response. Apparently, the well-known responses of VARs to monetary shocks are largely shaped by data prior to the early 1980s.

responses relative to the HAVAR responses.

The duration of the conventional and HAVAR aggregate responses differ markedly for the two housing variables. The HAVAR responses of investment and appreciation peak (i.e., reach a minimum) a year or more before the conventional response. Differences in duration are not evident for income because both income response are quite flat.

The longer duration of the conventional responses is a manifestation of the effects of aggregation on adjustment speeds. Aggregate data tend to be much more persistent than disaggregated data, so aggregation tends to bias downward estimates of adjustment speeds, as noted earlier. Thus, the results in Figure 6 indicate that the effects of monetary policy on the economy are actually much swifter than previously thought based on conventional macro VAR simulations.

Although this result contrasts with the conventional wisdom, it is more consistent with the data. Conventional VARs indicate that it typically requires 10 to 12 quarters or more for income to reach its peak (minimum) response after the monetary shock. But in a typical post-war U.S. recession, the peak income decline typically occurs less than eight quarters after the funds rate *begins* to rise and only about a year after the fund rate reaches its peak (maximum).

Differences between the conventional VAR and HAVAR responses may seem implausible to macroeconomists who view the conventional income response as a fundamental empirical fact. However, recall that the HAVAR is the correct benchmark because it nests the conventional VAR. Thus, the differences between the HAVAR and conventional VAR model responses indicate that the restrictions of the conventional model yield dynamic behavior that is inconsistent with the HAVAR model. Moreover, these differences are amplified greatly once we admit heterogeneity in initial conditions, as demonstrated in the next section.

7 HAVAR Model Simulations

7.1 Motivation

The preceding analysis quantifies one important way in which regional heterogeneity and aggregation can influence the dynamic properties of a monetary VAR model. A second important way regional heterogeneity can influence aggregate dynamic behavior is through dependence on initial regional conditions (state dependence). By incorporating regional VARs and imposing aggregation conditions, the HAVAR model provides a natural framework for evaluating this issue.

Impulse responses of conventional VARs are independent of initial regional conditions because the model equations are linear and, implicitly, regions are homogeneous. In the HAVAR model, the regional models are linear but the aggregate model is nonlinear because it exhibits time-varying parameters that depend on regional characteristics. Consequently, the aggregate response depends on the initial value of income, investment, and appreciation in each region.[34] In other words, heterogeneity in initial regional conditions interacts with heterogeneity in regional size and regional parameters to produce an additional source of variation in the dynamic responses of the HAVAR model to monetary policy. The impulse response analysis of section 6.2 implicitly assumed that all regions were in equilibrium at the time of the monetary tightening, in sharp contrast to the results presented in section 2.2.

Sensitivity of dynamic properties to initial regional conditions seems obvious but has not been examined much in the literature. Clearly, monetary policy makers could benefit from knowing whether funds rate changes affect the economy differently in recessions than expansions, for example. Conventional monetary VARs yield the same dynamic responses to monetary policy for every empirical episode.[35] In contrast, the HAVAR model accounts for underlying heterogeneity and provides estimates of the dynamic responses to monetary policy that vary over time.

7.2 Methodology

We simulate monetary tightening with the HAVAR model in situations where the aggregate economy begins in equilibrium (analogous to the conventional VAR) but the regional economies do not. Because aggregate equilibrium is consistent with many configurations of initial regional disequilibrium, it is necessary to perform numerous simulations by randomly drawing multiple sets of initial conditions. By conducting enough simulations we can trace out the ranges of dynamic responses to monetary tightening under alternative initial conditions for regional disequilibrium in income and appreciation.

Further description may help clarify this point. Suppose there are two regions with different sizes and different income responses to a monetary shock. One region is running a positive income

[34] State dependence also arises in nonlinear VARs, such as in Weise [62] where money supply shocks have stronger output effects and weaker price effects when output growth is initially low. Such nonlinear models may describe economic behavior better than linear models, but our goal is to demonstrate how nonlinearity arises inherently from aggregation of heterogeneous agents that behave linearly. Incorporation of nonlinear microeconomic models within an HAVAR type of model would add to the nonlinearity and complexity of the system.

[35] Actually, this point is true for *every* linear conventional macro model used for policy analysis, including the Federal Reserve Board's new FRB/US model.

gap and the other a negative gap, but the weighted average of income gaps is such that there is no aggregate income gap. Then the aggregate response of the economy to a monetary shock will depend on which region received which initial income shock. For example, the aggregate responses when the large region receives a large income shock will differ from the response when the large region receives a small income shock. Moreover, this difference can be amplified or muted by variation in the regions' sensitivities to monetary shocks. The aggregate response when the large region is very sensitive to monetary shocks will differ from the response when the large region is very insensitive to these shocks. Many combinations of region size, sensitivity to monetary shocks, and initial income gaps exist so extensive simulations are needed to identify the range of possible aggregate outcomes.

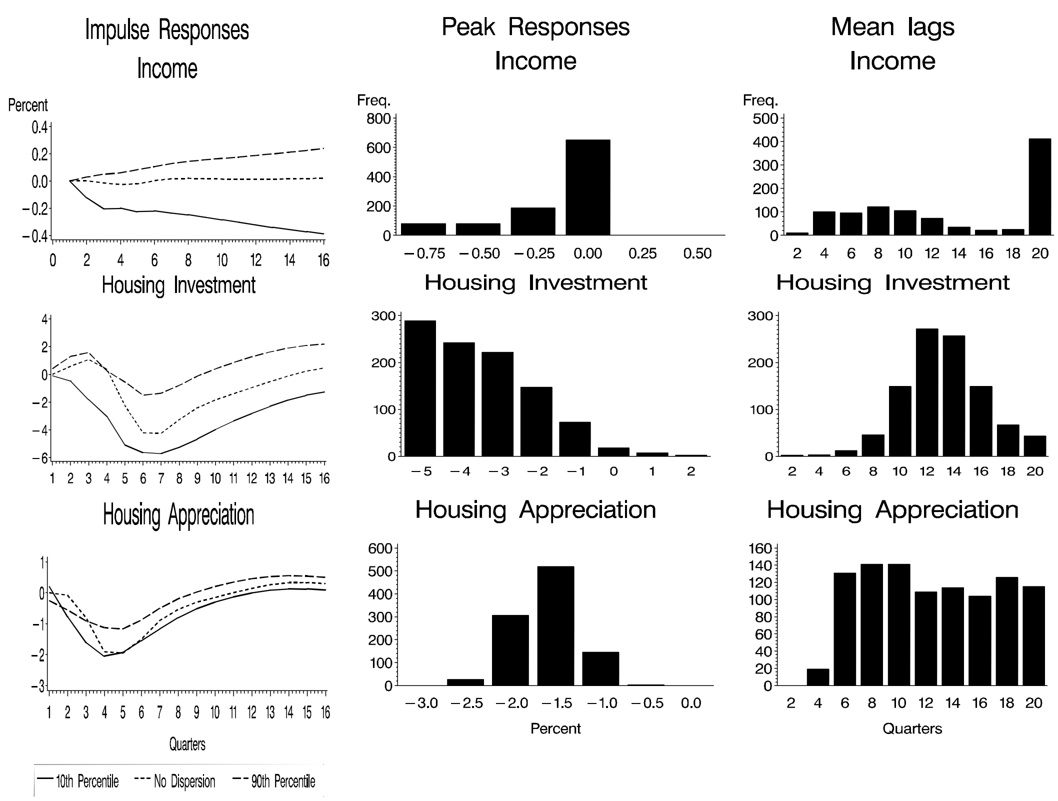
The random draws are obtained from data-consistent estimates of the distributions of regional structural shocks identified by the HAVAR model. We cannot reject the hypothesis that these innovations are distributed *iid* normal with mean zero across time and regions. Random shocks are chosen so that regional variables have heterogeneous initial conditions but aggregate variables begin in equilibrium. To ensure that the aggregate innovation to the HAVAR model remains zero, we weight the draws by region size. Even with 27 regions (draws), the aggregate innovation is not always exactly zero. Thus, when necessary, we adjust the draws such that there is no aggregate effect before simulating.

For each simulation, we draw a set of random initial conditions for income or appreciation in each region and shock the HAVAR model in the first period. One period later, we shock the funds rate equation by 100 basis points for one quarter. We do 1,000 simulations for each experiment. As in Figure 6, we plot the responses of income, investment, and appreciation associated with the simulations that produced the 10th percentile and 90th percentile peak responses. We then compare these extreme responses with those of the HAVAR model with no dispersion in initial regional conditions, that is, the HAVAR responses in section 6.2. We also report the distributions of the mean lags of the impulse responses as a measure of variation in duration, and distributions of the peak responses as a measure of variation in magnitude.

7.3 Regional Income Dispersion

Figure 7 shows the results when initial income conditions vary across regions but initial housing investment and appreciation conditions do not. The magnitude and duration of the HAVAR responses of differ from their no-dispersion baselines. For income, variation in the magnitude of the responses is modest. Variation in duration is substantial, although this may not be particulary meaningful given their flatness and the small differences between the responses. Perhaps more interesting is that the two extreme responses are asymmetric around the no dispersion response. The 10th percentile response declines about twice as much as the 90th percentile response increases, and this difference becomes more meaningful at longer lags.

Figure 7: HAVAR Simulations for Regional Income Dispersion



The income responses demonstrate that the effects of monetary policy are "long and variable." About three-fifths of the responses have a mean lag of up to five years, most of which are between 1-1/2 years to 2-1/2 years. One year is considerable variation in response duration, but a substantial fraction of responses have mean lags even shorter or longer.[36] Most of the peak income responses are modest, but income declines by 1/2 percent or more in one out of every four cases.

[36]The other two-fifths of responses are extremely persistent, like the 10th and 90th percentiles, with mean lags of more than five years (in some cases, much more). These responses may be associated with initial conditions in which the monetary authority would be unlikely to tighten policy, or they may reflect some nonstationarity problems.

The peak responses of housing investment in the 10th and 90th percentiles differ by more than 4 percentage points in a range between -1 and -6 percent. Interestingly, most peak responses occur in the lower end of this range. More than half of the investment mean lags are between 3 and 3-1/2 years, but many mean lags are as short as 2-1/2 years and as long as 4 years.

For appreciation, the peaks of the extreme responses differ by almost 1 percentage point. Interestingly, the no-dispersion response is quite similar to the 10th percentile response for much of the horizon rather than being in between the extreme responses. This asymmetry is a good example of the kind of nonlinear effects that can arise in the aggregate dynamics of models that account for heterogeneity and aggregation. The peak responses of appreciation vary considerably, though not as much as those of investment. However, the duration of the appreciation responses varies more, with the mean lags about uniformly distributed between 1-1/2 years and 5 years.

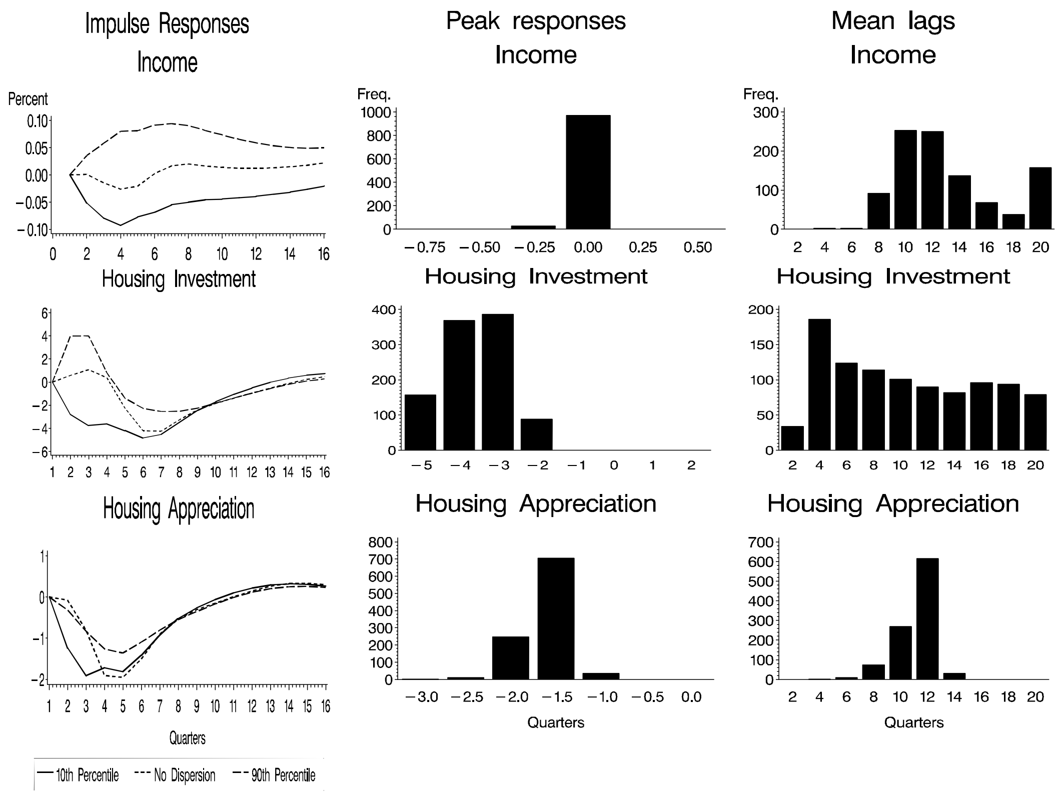
7.4 Regional Appreciation Dispersion

Figure 8 shows the results when initial appreciation conditions vary across regions but initial income conditions do not. The HAVAR simulation results are qualitatively similar to those with dispersion in regional income, so we do not discuss them in detail. The simulations reveal substantial heterogeneity in the magnitude and duration of the dynamic responses. However, the overall message of Figure 8 is that the effects of dispersion in regional appreciation are somewhat less than the effects of dispersion in income, particularly for the income responses.

We also simulated the HAVAR model under a scenario where the coastal regions are experiencing housing booms characterized by unusually high housing appreciation. Specifically, the coastal regions (east and west) are subject to appreciation shocks equal to the 75th percentile of the innovation distribution and noncoastal regions are subject to shocks equal to the 25th percentile. Conceptually, this experiment is designed to approximate situations such as U.S. housing markets during the late 1980s.

Figure 9 shows that monetary tightening is moderately less effective when the economy is experiencing coastal housing booms. Housing appreciation begins about 1 percentage point higher in this simulation. Given the underlying supply and demand elasticities imbedded in the regional VARs, this higher appreciation leads to significantly higher housing investment in the short run; income is higher too, but only modestly. The general patterns of the dynamic responses are similar to the no-dispersion model, but it takes more than a year for differences between the housing point estimates to disappear. This simulation suggests that the monetary authority would have to raise

Figure 8: HAVAR Simulations for Regional Appreciation Dispersion

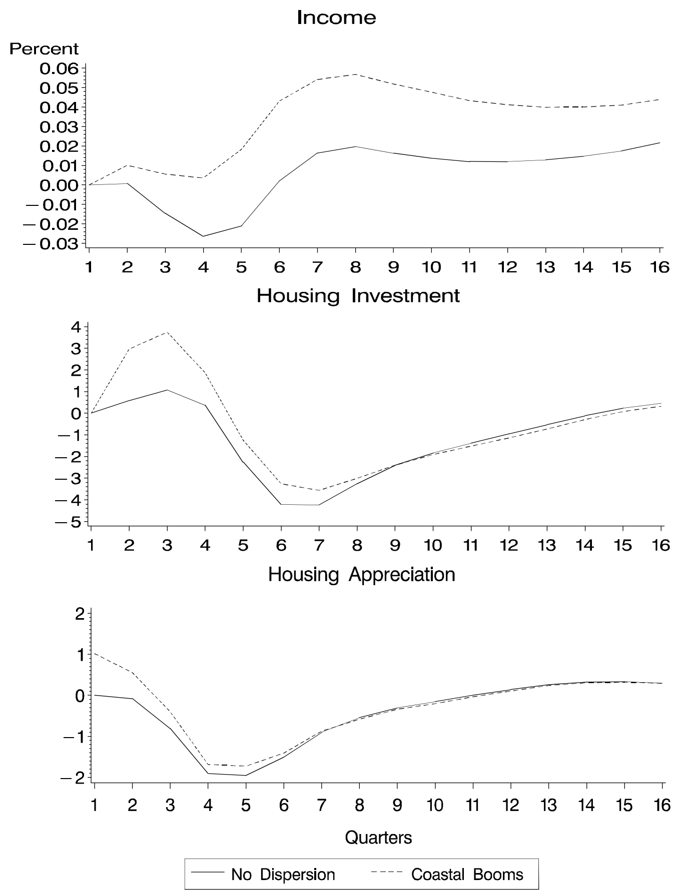


the federal funds rate more with coastal housing booms to achieve the same aggregate responses as when regional appreciation rates are homogeneous.

8 Summary and Conclusions

This paper introduces a new heterogeneous-agent VAR model for monetary policy analysis that incorporates regional heterogeneity in housing markets. Although the underlying regional VAR models are linear, the HAVAR model yields a tractable nonlinear aggregate model whose dynamic properties depend on the distribution of initial economic conditions across regions. For a panel of regional data, the HAVAR model exhibits economically significant variation in the magnitude and duration of dynamic responses to monetary shocks. We interpret the model and observed empirical variation as concrete examples of how the effects of monetary policy on the economy are “long

Figure 9: HAVAR Simulation of Coastal Housing Booms



and variable."

The HAVAR model and results in this paper are a first step toward understanding how monetary policy may affect the economy differentially over time. The methodology could be improved on several dimensions. A model based on dynamic optimizing behavior with deep structural parameters would be preferable, and the HAVAR methodology is consistent with such a structural model, provided it can be represented in restricted-VAR form. Even within the class of VAR models, more structure and alternative identifying restrictions could be entertained. Macroeconomists probably would like a broader focus than housing, and others may prefer a more detailed treatment of the housing market. Empirically, efforts are needed to expand both the time-series and cross-section dimensions of the data base.

Nevertheless, even this relatively simple model offers a step forward for policymakers. Conventional macroeconomic models do not account for heterogeneity, and linear aggregate models predict the same effect of monetary policy actions at each point in time. Our HAVAR model overcomes both of these limitations, providing a relatively easy way to answer the kinds of "what if" questions monetary policymakers ask with regard to changing economic conditions over time.

A Data Appendix

The aggregate data sources are as follows. The aggregate price is the U.S. CPI-U for all items from the Bureau of Labor Statistics. The nominal mortgage rate is the contract rate on commitments for fixed-rate first mortgages from Federal Reserve Statistical Release G.13 The monetary instrument is the nominal effective federal funds rate, also from Release G.13.

The regional data sources are as follows. Housing investment is the MSA-level number of housing starts from the Census Bureau (regional housing stock data are not available).[37] Housing price is the nominal MSA-level index from the Fannie Mae Repeat Transactions Database. Income is state-level personal income from the Census Bureau. Starts and income are converted to per-capita using Census data on MSA-level population, and income is further scaled by the share of MSA population in total state population. Income and housing prices are deflated using the aggregate price, U.S.-level CPI-U, because MSA-level CPI data are not available for all regions, all quarters, and for housing and nonhousing subaggregates.

The database yields a balanced panel for estimation and aggregation analysis that includes 27 regions over 40 quarters from 1986:Q3 to 1996:Q2. Appendix Table 2 lists the regions. A panel with a longer sample period would include many fewer regions, and a panel with more regions would cover a much shorter sample period. We selected this combination of regions and years to obtain sufficient heterogeneity and time series observations.

Aggregate housing market data are the appropriately weighted sums of the regional data, using a balanced panel to ensure exact aggregation. Aggregate data from the 1986-96 panel represent 37 percent of total U.S. personal income and 7.5 percent of total U.S. housing starts.[38] The time series

[37]Theoretically, housing starts is the appropriate data measure but the consensus view in the literature is that starts contain measurement errors that make them less suitable than housing permits. However, Somerville [57] challenges this view, and we find no difference in the empirical results between starts and permits.

[38]The disparity between the shares of income and housing is likely due to the fact that houses tend to be "larger" and of better quality where income is greater. Thus, the share of the housing stock in these regions is probably much

correlation between the panel aggregate and total U.S. detrended data is 0.83 for income and 0.73 for starts. Because the Census housing data are not obtained from a probability sample, it is not possible to quantify sample selection problems.

B Technical Appendix

To see the connection to nonlinear time series models, consider the following simple example motivated by Granger and Newbold [27] (pp. 303-305). Defining

$$\mathbf{X}_t = (X_t, X_{t-1}, \ldots, X_{t-L}, \epsilon_t, \epsilon_{t-1}, \ldots, \epsilon_{t-L}) \ ,$$

then a generalized nonlinear AR(1) is

$$X_t = \beta(\mathbf{X}_{t-1}) X_{t-1} + \epsilon_t \ .$$

where the nonlinear parameter $\beta(\mathbf{X}_{t-1})$ must be "constrained" somehow to be made practical. If $X_t = x_{1t} + x_{2t}$ and $x_{it} = \beta_i x_{i,t-1} + \epsilon_{it}$ for $i = \{1, 2\}$, then we can define

$$\mathbf{X}_t^* = (x_{1t}, x_{1,t-1}, \ldots, x_{1,t-L}, x_{2t}, x_{2,t-1}, \ldots, x_{2,t-L}, \epsilon_{1t}, \epsilon_{1,t-1}, \ldots, \epsilon_{1,t-L}, \epsilon_{2t}, \epsilon_{2,t-1}, \ldots, \epsilon_{2,t-L})$$

so that the macro parameter becomes $\beta(\mathbf{X}_{t-1}^*)$. In this simple two-agent AR(1) case, the macro parameter is

$$\beta(\mathbf{X}_{t-1}^{**}; \beta_1, \beta_2) = \left[\frac{\beta_1 x_{1,t-1} + \beta_2 x_{2,t-1}}{x_{1,t-1} + x_{2,t-1}} \right] = \left[\beta_1 \left(\frac{x_{1,t-1}}{X_{t-1}} \right) + \beta_2 \left(\frac{x_{1,t-1}}{X_{t-1}} \right) \right]$$

where $\mathbf{X}_{t-1}^{**} = (x_{1,t-2}, x_{2,t-2}, \epsilon_{1,t-1}, \epsilon_{2,t-1},)$ because $x_{i,t-1} = \beta_i x_{i,t-2} + \epsilon_{i,t-1}$. Time variation in $\beta(\mathbf{X}_{t-1}^{**}; \beta_1, \beta_2)$ arises from time variation in the micro data shares $(x_{i,t-1}/X_{t-1})$ provided $\beta_1 \neq \beta_2$ (heterogeneous behavior) or $\epsilon_{1,t-1} \neq \epsilon_{2,t-1}$ (heterogeneous shocks), or both.[39] Thus, X_t depends dynamically on the *lagged* micro shocks, $\epsilon_{i,t-1}$, which represent deviations from the micro steady states, x_{i0}.[40]

An interesting implication of this dependence is that a multiplicity of dynamic responses of X_t emerges from any given aggregate shock, $\epsilon_{t-1} = \epsilon_{1,t-1} + \epsilon_{2,t-1}$; or, conversely, any given aggregate response is consistent with a multiplicity of micro shocks. For example, a mean-preserving spread

higher.

[39] A conventional macro VAR implicitly assumes away the effects of both types of heterogeneity.

[40] We thank Simon Gilchrist for helping us discover this important insight.

of micro shocks will alter the aggregate dynamic response of the HAVAR model, as would a moment-preserving redistribution of the micro shocks (it matters where the shocks occur because the micro parameters vary). More generally, HAVAR aggregate dynamic responses depend on the joint distribution of micro shocks and micro parameters. Note, however, that unlike many other nonlinear macro models, there are not multiple *equilibria*.[41] Because the micro models are linear and assumed stable, both the micro and HAVAR models always converge to the same steady state in the long run.

[41] See Aoki [2] for an example with heterogeneous agents.

References

[1] Karim M. Abadir and Gabriel Talmain. Aggregation, persistence and volatility in a macromodel. University of York Discussion Paper No. 01/03, 2001.

[2] Masanao Aoki. Simple model of asymmetrical business cycles: Interactive dynamics of a large number of agents with discrete choices. *Macroeconomic Dynamics*, 2(4):427–42, 1998.

[3] Nathan S. Balke. Credit and economic activity: Credit regimes and nonlinear propogation of shocks. *The Review of Economics and Statistics*, 82(2):344–349, 2000.

[4] Ben S. Bernanke and Alan S. Blinder. The federal funds rate and the channels of monetary transmission. *American Economic Review*, 82(4):901–921, 1992.

[5] Ben S. Bernanke and Ilian Mihov. The liquidity effect and long-run neutrality. NBER Working Paper 6608, June 1998.

[6] Olivier Blanchard and Lawrence Katz. Regional evolutions. *Brookings Papers on Economic Activity*, (1):1–75, 1992.

[7] Gerald Carlino and Robert DeFina. The differential regional effects of monetary policy. *The Review of Economics and Statistics*, 80(4):572–87, November 1998.

[8] Gerald A. Carlino and Robert H. DeFina. Do states respond differently to changes in monetary policy? *Business Review*, pages 17–27, July/August 1999.

[9] Gerald A. Carlino, Robert H. DeFina, and Keith Sill. Sectoral shocks and metropolitan employment growth. Federal Reserve Bank of Philadelphia Working Paper No. 00-9, 2000.

[10] V.V. Chari, Lawrence J. Christiano, and Patrick J. Kehoe. Policy analysis in business cycle models. In Thomas F. Cooley, editor, *Frontiers of Business Cycle Research*, pages 357–391. Princeton University Press, 1995.

[11] Lawrence J. Christiano, Martin Eichenbaum, and Charles L. Evans. The effects of monetary policy shocks: Evidence from the flow of funds. *The Review of Economics and Statistics*, 78(1):16–34, 1996.

[12] Lawrence J. Christiano, Martin Eichenbaum, and Charles L. Evans. Monetary policy shocks: What have we learned and to what end? Federal Reserve Bank of Chicago Working Paper WP-97-18, 1997.

[13] Richard Clarida, Jordi Galí, and Mark Gertler. The science of monetary policy: A new keynesian perspective. *Journal of Economic Literature*, 37(4):1661–1707, 1999.

[14] Todd E. Clark. Employment fluctuations in u.s. regions and industries: The roles of national, region-specific, and industry-specific shocks. *Journal of Labor Economics*, 16(1):202–229, January 1998.

[15] Todd E. Clark and Kwanho Shin. The sources of fluctuations within and across countries. Federal Reserve Bank of Kansas City Working Paper 98-04, 2000.

[16] John H. Cochrane. Shocks. *Carnegie-Rochester Conference Series on Public Policy*, 41:295–364, 1994.

[17] N. Edward Coulson. Sectoral sources of the Massachusetts Miracle and other turning points. Unpublished paper, 2000.

[18] Steven J. Davis and John Haltiwanger. On the driving forces behind cyclical movements in employment and job reallocation. *American Economic Review*, 89(5):1234–1258, December 1999.

[19] James T. Fergus. The effects of abnormal weather conditions on housing construction. Unpublished paper, Federal Reserve Board, 1996.

[20] James R. Follain and Michelle Harter. A time-series approach to the study of supply elasticity of housing. Unpublished paper, 1998.

[21] Benjamin M. Friedman and Kenneth N. Kuttner. Why does the paper-bill spread predict real economic activity? In James H. Stock and Mark W. Watson, editors, *NBER Studies in Business Cycles*, volume 28, pages 213–49, Chicago and London, 1993. University of Chicago Press.

[22] Jeff Fuhrer and George Moore. Inflation persistence. *Quarterly Journal of Economics*, 110(1):127–59, 1995.

[23] A. Ronald Gallant, Peter E. Rossi, and George Tauchen. Nonlinear dynamic structures. *Econometrica*, 61(4):871–907, July 1993.

[24] Marvin Goodfriend and Robert G. King. The new neoclassical synthesis and the role of monetary policy. In Ben Bernanke and Julio Rotemberg, editors, *NBER Macroeconomics Annual*, pages 231–283, Cambridge, MA, 1997. The MIT Press.

[25] Jr Goodman, John L. Aggregation of local housing markets. *Journal of Real Estate Finance and Economics*, 16(1):43–53, 1998.

[26] Clive W.J. Granger. Long memory relationships an the aggregation of dynamic models. *Journal of Econometrics*, 14(2):227–238, oct 1980.

[27] C.W.J. Granger and Paul Newbold. *Forecasting Economic Time Series*. Academic Press, Inc., San Diego, CA, 2nd edition, 1977.

[28] Richard K. Green. Follow the leader: How changes in residential and non-residential investment predict changes in gdp. *Real Estate Economics*, 25(2):253–270, 1997.

[29] James D. Hamilton. *Time Series Analysis*. Princeton University Press, Princeton, NJ, 1994.

[30] Bernd Hayo and Birgit Uhlenbrock. Industry effects of monetary policy in Germany. Working Paper B 14, Center for European Integration Studies, 1999.

[31] Michael Horvath. Cyclicality and sectoral linkages: Aggregate fluctuations from independent sectoral shocks. *Review of Economic Dynamics*, 1(4):781–808, October 1998.

[32] Michael Horvath and Randal Verbrugge. Shocks and sectoral interactions: An empirical investigation. Unpublished paper, 1997.

[33] Geraint Johnes and Thomas Hyclak. House prices and regional labor markets. *The Annals of Regional Science*, 33:33–49, 1999.

[34] Robert G. King and Alexander L. Wolman. Inflation targeting in a St. Louis model of the 21st century. NBER Working Paper No. 5507, 1996.

[35] Gary Koop, M. Hashem Pesaran, and Simon M. Potter. Impulse response analysis in nonlinear multivariate models. *Journal of Econometrics*, 74(1):119–147, 1996.

[36] Owen Lamont and Jeremy C. Stein. Leverage and house-price dynamics in U.S. cities. *RAND Journal of Economics*, 30(3):498–514, 1999.

[37] Eric M. Leeper, Christopher A. Sims, and Tao Zha. What does monetary policy do? *Brookings Papers on Economic Activity*, (2):1–63, 1996.

[38] Arthur Lewbel. Aggregation and simple dynamics. *American Economic Review*, 84(4):905–918, aug 1994.

[39] Marco Lippi. On the dynamic shape of aggregated error correction models. *Journal of Economic Dynamics and Control*, 12(2/3):561–585, June/September 1988.

[40] Stephen Malpezzi. Housing prices, externalities, and regulation in U.S. metropolitan areas. *Journal of Housing Research*, 7(2):209–241, 1996.

[41] Joyce Manchester. The baby boom, housing, and loanable funds. *Canadian Journal of Economics*, 22(1):128–149, 1989.

[42] N. Gregory Mankiw and David N. Weil. The baby boom, the baby bust, and the housing market. *Regional Science and Urban Economics*, 19(2):235–258, 1989.

[43] Christopher J. Mayer and C. Tsuriel Somerville. Land use regulation and the supply of housing: A metropolitan area analysis of the effects of regulation on the construction of new residential structures. Unpublished paper, 1997.

[44] Bennett T. McCallum and Edward Nelson. An optimizing IS-LM specification for monetary policy and business cycle analysis. *Journal of Money Credit and Banking*, 31(3):296–316, 1999.

[45] John Muellbauer and Anthony Murphy. Booms and busts in the uk housing market. *The Economic Journal*, 107(445):1701–1727, 1997.

[46] François Ortalo-Magné and Sven Rady. Housing market fluctuations in a life-cycle economy with credit constraints. Research Paper No. 1501, Graduate School of Business, Stanford University, October 1998.

[47] James M. Poterba. Tax subsidies to owner-occupied housing: An asset-market approach. *Quarterly Journal of Economics*, 99(4):729–752, 1984.

[48] James M. Poterba. Housing price dynamics: The role of tax policy and demography. *Brookings Papers on Economic Activity*, (2):143–203, 1991.

[49] Simon M. Potter. A nonlinear approach to us gnp. *Journal of Applied Econometrics*, 10(2):109–125, 1995.

[50] Simon M. Potter. Time series modelling: An introduction. *Journal of Economic Surveys*, 13(5):505–528, 1999.

[51] Simon M. Potter. Nonlinear impulse response functions. *Journal of Economic Dynamics and Control*, 24(10):1425–1446, 2000.

[52] Danny T. Quah. Aggregate and regional disaggregate fluctuations. CEPR Discussion Paper 1236, 1995.

[53] Julio J. Rotemberg and Michael Woodford. An optimization-based econometric framework for the evaluation of monetary policy: Expanded version. NBER Technical Working Paper No. 233, 1998.

[54] Scott Schuh. Evidence on the link between firm-level and aggregate inventory behavior. Federal Reserve Board FEDS Working Paper 1996-46, November 1996.

[55] Keith Sill. Regional employment dynamics. Federal Reserve Bank of Philadelphia Working Paper 97-28, 1997.

[56] L.B. Smith, K. Rosen, and G. Fallis. Recent developments in economic models of housing markets. *Journal of Economic Literature*, 26(1):29–64, 1988.

[57] Tsuriel Somerville. Builder responses to market changes: Permits vs. starts vs. completions. Unpublished paper, University of British Columbia, May 1999.

[58] James Stock and Mark W. Watson. New indexes of coincident and leading economic indicators. In Olivier Jean Blanchard and Stanley Fischer, editors, *NBER Macroeconomics Annual*, pages 351–394, Cambridge, MA, 1989. The MIT Press.

[59] Henri Theil. *Linear Aggregation of Economic Relations*. North-Holland Publishing Company, Amsterdam, 1954.

[60] P.K. Trivedi. Distributed lags, aggregation and compounding: Some econometric implications. *The Review of Economic Studies*, 52(1):19–35, 1985.

[61] Carl E. Walsh. *Monetary Theory and Policy*. The MIT Press, Cambridge, MA, 1998.

[62] Charles L. Weise. The asymmetric effects of monetary policy: A nonlinear vector autoregression approach. *Journal of Money Credit and Banking*, 31(1):85–108, 1999.

Table 2: Regional Impulse Responses

	Income		Housing			
			Investment		Appreciation	
Region (MSA)	Peak	Lag	Peak	Lag	Peak	Lag
Atlanta, GA	−.78	26.8	−13.10	10.9	−2.14	10.2
Baltimore, MD	−.22	15.1	−2.61	10.3	−.71	10.6
Charlotte, NC	−.36	6.9	−.53	5.9	−.79	8.9
Chicago, IL	−.46	7.4	−4.12	7.2	−1.9	11.8
Colorado Springs, CO	−.49	9.0	−3.31	9.2	−4.65	5.2
Washington, DC	.00	29.1	−10.0	10.8	−.39	10.0
Dallas, TX	−.02	76.5	−13.7	14.3	−4.02	7.4
Denver, CO	−.60	14.2	−8.96	12.5	−4.48	10.1
Fort Lauderdale, FL	−1.22	15.1	−5.83	26.1	−6.47	10.3
Houston, TX	−.02	239.4	−.21	74.7	−1.78	9.2
Kansas City, MO	−.60	115.4	−3.15	14.1	−1.94	15.9
Los Angeles, CA	−.01	13.6	−3.06	19.2	−.82	16.2
Las Vegas, NV	−.03	10.8	−.27	19.1	.00	2.8
Miami, FL	−.09	7.5	−.70	7.4	−2.42	6.8
Minneapolis, MN	−.56	9.7	−8.53	7.4	−2.02	7.9
New Orleans, LA	−.24	5.7	−1.21	10.1	−3.85	6.5
New York, NY	−.32	6.7	−11.60	7.4	−3.52	9.3
Phoenix, AZ	−1.45	14.7	−18.4	9.7	−3.14	9.6
Riverdale, CA	−.02	7.1	−3.23	14.8	−.71	11.2
Sacramento, CA	−.03	11.6	−2.26	13.8	−.80	16.7
Salt Lake City, UT	−.60	12.7	−5.21	7.1	−5.66	11.1
San Diego, CA	.00	17.2	−.60	9.6	−.44	9.9
San Francisco, CA	−.03	118.0	−1.08	58.9	−5.49	6.6
Seattle, WA	−.03	10.9	−.75	14.6	−.49	14.4
St. Louis, MO	−.49	15.6	−1.70	10.0	−2.55	12.0
Tampa, FL	−.18	22.0	−3.70	16.0	−1.97	8.9
West Palm Beach, FL	−.92	9.2	−3.19	32.6	−4.21	10.1
			Correlation Matrix			
Income Peak	1					
Income Lag	.20	1				
Investment Peak	.51	.12	1			
Investment Lag	.15	.82	.28	1		
Appreciation Peak	.52	−.06	.26	−.19	1	
Appreciation Lag	−.03	−.01	.04	−.11	.27	1