FUNDAMENTAL RESEARCH

For. Sci. 64(1):1–14
https://doi.org/10.5849/FS-2017-020

economics

Projecting Housing Starts and Softwood Lumber Consumption in the United States

Jeffrey P. Prestemon, David N. Wear, Karen L. Abt, and Robert C. Abt

New residential construction is a primary user of wood products in the United States; therefore, wood products projections require understanding the determinants of housing starts. We model quarterly US total, single-family, and multifamily housing starts with several model specifications, using data from 1979 to 2008, and evaluate their fit out of sample, 2009–14. Goodness-of-fit statistics show that parsimonious models outperform general models in out-of-sample predictions. Monte Carlo simulations of total housing starts to 2070 project median starts ranging from 0.86 million/year at 0% real gross domestic product (GDP) growth to 1.91 million/year at 5% real growth, with 90% uncertainty bounds ranging from 0.52 to 2.13 million/year. Assuming that future GDP growth equals the average rate observed over 1990–2015, there is less than 9% probability that housing starts will exceed 2.0 million in any given year, 2016–35. Results show no evidence of structural change in the determinants of total or single-family housing starts coincident with the recession of 2007–09. Using these housing projections in a softwood lumber consumption model shows that GDP growth slower than 2% is consistent with stagnant or declining median softwood lumber consumption.

Keywords: construction activity, lumber consumption, econometrics, Monte Carlo

In the United States, new residential construction consumes a large share of domestic wood products output, including one third of all lumber (Howard and Jones 2016) and two fifths of wood-based structural panels such as softwood plywood and oriented strandboard (APA 2010). New residential construction is cyclical and connected to similar variation in the broader economy (Leamer 2007, Glaeser et al. 2008, Agnello and Schuknecht 2011). This cyclical and variable nature of housing starts and gross domestic product (GDP) growth is carried through to changes in softwood lumber consumption (Figure 1). The most recent recession, which followed a run-up in national average housing prices (Figure 2), has also been suspected of inducing a structural shift in the housing market related to demographic changes (Pitkin and Myers 2008, Anundsen 2015, Myers 2016). Increased delinquent mortgage rates (Figure 3) and the resulting tightening in lending requirements and rates of loan application denials may be linked to changing housing demand over the long run (Federal Reserve Board 2016, Vojtech et al. 2016). If the housing market has changed, then the shift would have implications for wood products markets.

The primary objective of this research is to develop a parsimonious reduced-form model of residential construction activity (measured as housing starts) in the United States. A reduced-form model of residential construction could be a useful addition to integrated assessment models of the forest sector (e.g., Buongiorno 2014) that require projections of key demand factors (Ince et al. 2011). To demonstrate the potential utility of the reduced-form modeling approach, we use a housing starts model in two ways. First, we use it to project residential construction activity over the coming decades under varying assumptions about overall economic growth in the United States, 2015–70. This provides an assessment of possible ranges of residential construction that would drive wood products demands under alternative assumptions about economic growth. Second, we incorporate it into a projection of potential ranges of softwood lumber consumption in the United States (e.g., Song et al. 2011) over the same time span. Monte Carlo methods are used to generate median levels and probability bands for housing starts and softwood lumber consumption, given assumptions about economic growth, over these future decades.

The following sections describe our methods, including assessments of the time series properties of housing starts and related variables, specification of the alternative housing starts models, specification of reduced-form softwood lumber consumption (quantity) models for the United States, and the Monte Carlo approaches used to project median levels and variability in starts and wood products consumption. In the *Results* section, we describe the equation estimates and the Monte Carlo simulation outcomes. In the *Conclusion*, we lay out the implications of the study for integrated assessments of the wood products sector and suggest follow-on research.

Manuscript received February 10, 2017; accepted August 10, 2017; published online September 28, 2017.

Affiliations: *Jeffrey P. Prestemon (jprestemon@fs.fed.us), Forestry Sciences Laboratory, USDA Forest Service, Southern Research Station, PO Box 12254, Research Triangle Park, NC 27709. David N. Wear (dwear@fs.fed.us) and Karen L. Abt (kabt@fs.fed.us), USDA Forest Service, Southern Research Station Robert C. Abt (bobabt@ncsu.edu), College of Natural Resources, North Carolina State University.*

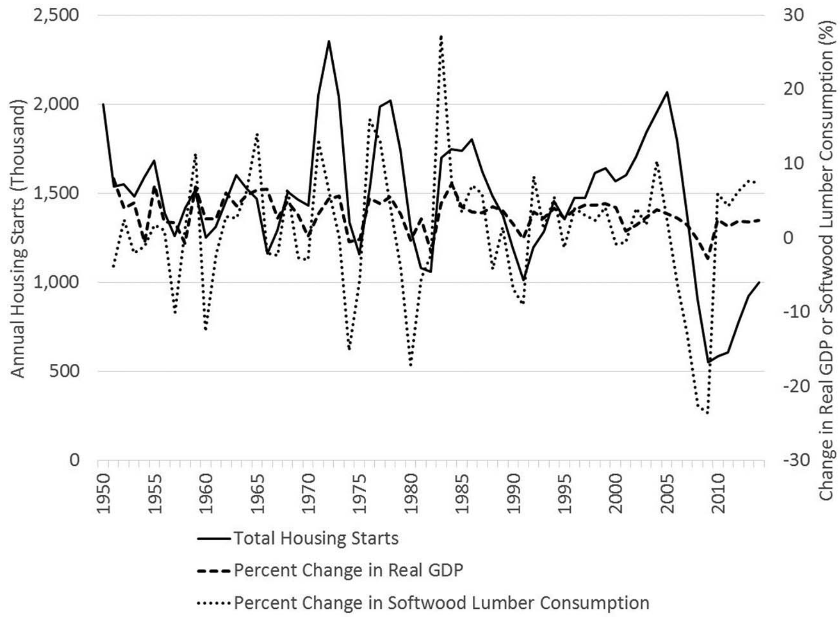

Figure 1. Annual, nationwide total (single-family plus multifamily, two or more units per structure) housing starts in the United States, 1950–2015, percentage change in real US GDP, and percentage change in US softwood lumber consumption. (Sources: starts, 1959–2015: US Bureau of the Census (2016); starts, 1950–58: Siskind [1979]; GDP: US Bureau of Economic Analysis [2016]; softwood lumber consumption: Howard and Jones [2016], Random Lengths [2016]).

Methods

Housing starts and prices are determined at the equilibrium of housing demand (h^D) and supply (h^S). Theory indicates and empirical evidence confirms (e.g., de Leeuw 1971, Mankiw and Weil 1989, Goodman and Thibodeau 2008) that housing demand responds negatively to house prices and positively to income. This literature also indicates that housing demand can be modeled, at fine and large spatial and temporal scales, to include demographic factors, interest rates, tax policies, and credit access (e.g., Glaeser et al. 2008). Therefore, we assume that demand for new housing is a function of house prices (p), household income (y), credit conditions (c), lending interest rates (r), demographic factors (a), and taxes (property and income taxes; g). Quantities of housing starts demanded in period t are

$$h_t^D = f(p_t, y_t, c_t, r_t, a_t, g_t) \quad (1)$$

Theory and empirical evidence (e.g., Blackley 1999, Ball et al. 2010) suggest that the supply of new housing can be specified as a function of house prices, prices of construction inputs (w), land constraints (l), and regulations (z):

$$h_t^S = f(p_t, w_t, l_t, z_t) \quad (2)$$

A reduced-form expression of the quantity of housing starts assumes that $h^D = h^S = h$ (i.e., Equations 1 and 2 are equal) at equilibrium and solves for endogenous price and quantity. Therefore, the reduced form equation for the quantity of housing is (a similarly specified equation for house prices could also be expressed):

$$h_t = f(y_t, c_t, r_t, a_t, g_t, w_t, l_t, z_t) \quad (3)$$

We expect that h is a positive function of y; a negative function of r, g, w, and l; and an undefined function of c, a, and z. Equations 1–3 apply to demand for and supply of single-family and multifamily starts, although the parameters of such functions could differ for the two models.

Although compact, the reduced form specification brings with it several questions, including what are the most appropriate spatial and temporal scales and functional forms and how might data be transformed. Equations 1–3 could be estimated at a local scale or could be aggregated and estimated at progressively larger spatial scales but under strong homogeneity assumptions regarding the effects of the included variables. The choice of scale is typically influenced by the availability of data. Over time spans of a few decades, some variables may not vary significantly, so that a time series model of, say, Equation 3, would place the effects of a, g, w, l, and z into an intercept or, if they are presumed to be trending consistently over time, captured in aggregate with a time trend. Alternatively, heterogeneity in the relationships described by Equations 1–3 could be addressed using a constant elasticity functional form with log transformation of variables.

Management and Policy Implications

This study finds that US residential construction can be modeled simply with a reduced-form equation that relates housing starts to US economic growth and mortgage delinquencies. This parsimonious relationship is amenable to incorporation within forest sector models. Projections of starts to 2070 using model estimates show that housing starts are likely to average less than 1.5 million per year under plausible assumptions regarding long-run economic growth in the United States. This probable future for the US construction sector implies that markets for some categories of wood products in the United States are likely to experience moderate future growth. For example, softwood lumber consumption growth would average 0.6% per year given economic growth continuing at the average rate observed from 1990 to 2014.

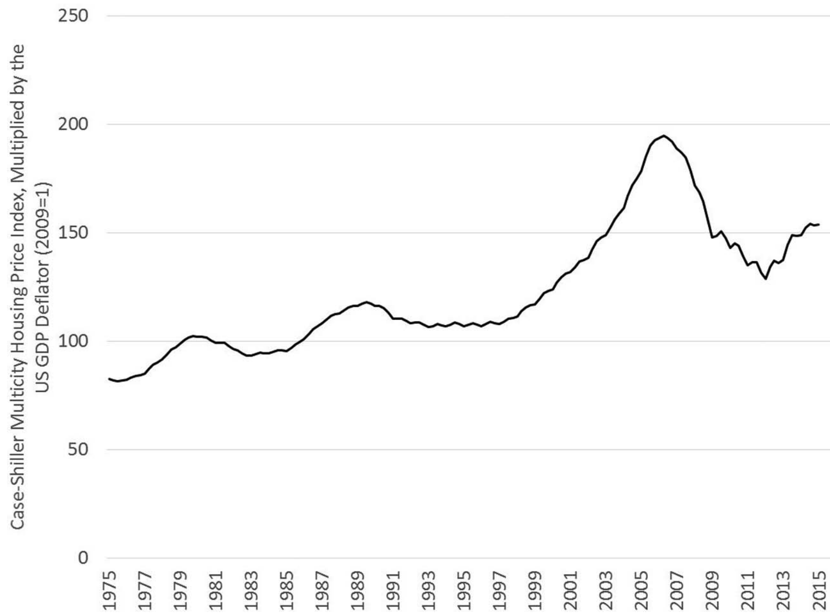

Figure 2. Case-Shiller Multicity House Price Index in the United States, 1975q1 to 2015q3, multiplied by the US GDP deflator (2009 = 1). (Source: US Federal Reserve Bank of St. Louis [2015]).

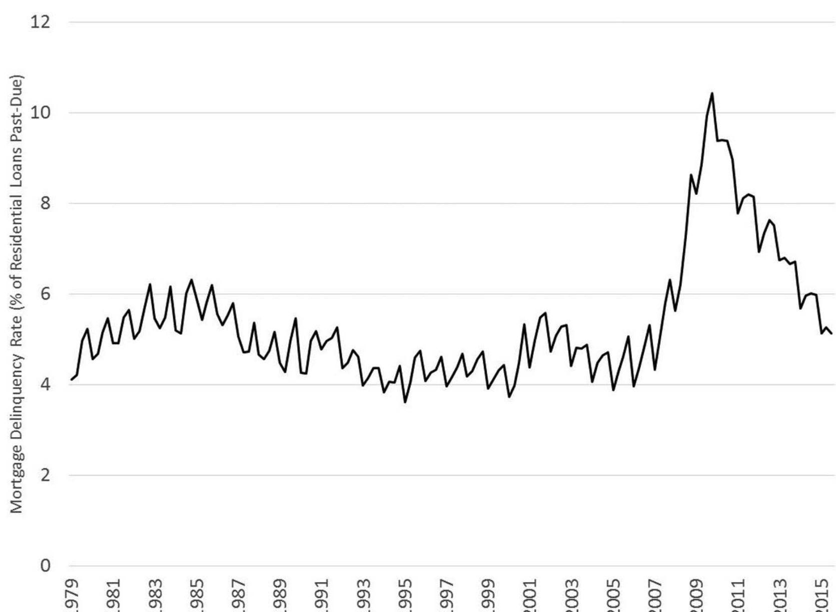

Figure 3. Delinquency rate for all residential mortgages combined, total past due, United States, 1979q1 to 2015q3 (data not seasonally adjusted). (Source: Mortgage Bankers Association [2015]).

Housing Starts Empirical Specifications

In this study, we model Equation 3, with the dependent variable defined as housing starts aggregated to the national level (all 50 states) in the United States (i.e., excluding Puerto Rico and US territories; see US Bureau of the Census, 2017a). We assume that the effects of changes in a, g, w, l, and z on starts can be captured by aggregate expressions of these variables at the national level. We estimate separate and combined equations for single-family and multifamily starts. This is motivated by three factors: multifamily start shares have varied over time (multifamily housing starts ranged from 11% to 48% of total quarterly starts, 1963– 2015, and averaged 35% from 2013 to 2015; US Bureau of the Census 2016a), single-family and multifamily starts might respond differently to causal variables shown in Equation 3, and the quantity of wood used in single-family structures is larger on a per-family basis than that used in multifamily dwellings. For example, in 2013, the average floor area was 103 m^2 in a multifamily unit and 241 m^2 in a single-family home (Howard and Jones 2016).

Equation 3 (summed across all locations, i), with capitalized variable names to indicate national aggregates, becomes

$$H_t = f(Y_t, R_t, C_t, A_t, G_t, W_t, \tau_t) \quad (4)$$

where Y is real-dollar US GDP, R is the mortgage interest rate (percent), C is the mortgage delinquency rate (in percent, approximating credit conditions), A is total US population, G is the average

Table 1. Data sources.

Variable	Units	Citation
Housing starts	Thousands/yr	US Bureau of the Census (2016a)
GDP	\$/yr	US Bureau of Economic Analysis (2016)
GDP deflator	Chained, 2009 = 100	US Bureau of Economic Analysis (2016)
Population, 2010–15		US Bureau of the Census (2016c)
Population, 1970–2009		US Bureau of the Census (2014)
Construction wages	Dollars/week	US Bureau of Labor Statistics (2015b)
Tax rate at \$60,000 annual income	%	Tax Foundation (2015)
Mortgage delinquency rates	%	Mortgage Bankers Association (2015)
Producer price indices, concrete products	1982 = 100	US Bureau of Labor Statistics (2016)
Producer price indices, softwood plywood	1982 = 100	US Bureau of Labor Statistics (2016)
Producer price indices, softwood lumber	1982 = 100	US Bureau of Labor Statistics (2016)
Producer price indices, oriented strandboard	1982 = 100	Spelter (2015)
Producer price indices, all commodities	1982 = 100	US Bureau of Labor Statistics (2015a)
Mortgage interest rates	%	Freddie Mac (2016)
Unemployment rate	%	US Bureau of Labor Statistics (2015c)
Softwood lumber consumption, 1960–2014	Million board feet/yr	Howard and Jones (2016), Random Lengths (2016)
Producer price indices, industrial electric power	1982 = 100	US Bureau of Labor Statistics (2015d)
Federal funds rate	%	Federal Reserve Bank of St. Louis (2015)

Table 2. ADF, PP Tau, and KPSS unit root tests.

Variable	ADF 1979q2–2008q4	ADF 1979q2–2014q4	PP Tau 1979q2–2008q4	PP Tau 1979q2–2014q4	KPSS 1979q2–2008q4	KPSS 1979q2–2014q4
ln(Starts$_t$), total	−3.20**	−3.17**	−4.32***	−3.19**	0.199	0.352*
ln(Starts$_t$), single-family	−3.01**	−2.90**	−3.88***	−3.16**	0.594**	0.29
ln(Starts$_t$), multifamily	−1.97	−2.81*	−2.61*	2.72*	0.439*	0.519**
ln(Real GDP$_t$)	−0.88	−1.15	−0.74	−1.22	1.56***	1.49***
ln(Population$_t$)	0.32	−0.49	0.53	−0.56	1.58***	1.52***
ln(Mortgage Rate$_t$)	−0.95	−0.49	−0.88	−0.44	1.42***	1.39***
ln(Real Wage Rate$_t$)	−1.70	−1.39	−1.82	−2.15	1.11***	0.883***
ln(Tax Rate$_t$)	−0.54	−0.67	−0.56	−0.68	0.813***	1.08***
ln(Delinquency Rate$_t$)	−1.42	−2.37	−3.19**	−2.38	0.353*	0.459*
ln(PPI Concrete$_t$)	0.00	−0.42	0.92	0.10	1.5***	1.46***
ln(PPI SW Plywood$_t$)	−2.87*	−3.17**	−1.51	−1.22	1.39***	1.33***
ln(PPI SW Lumber$_t$)	−1.64	−2.13	−1.43	−1.33	1.23***	1.01***
ln(PPI OSB$_t$)	−3.29***	−4.00**	−2.20	−2.47	0.679**	0.637**
ln(Unemployment Rate$_t$)	−2.10	−3.04**	−1.75	−1.99	0.901***	0.257
ln(Real GDP$_t$/Population$_t$)	−1.03	−1.18	−0.94	−1.26	1.55***	1.46***
ln(PPI$_t$)	−1.84	−1.48	−1.20	−1.01	1.49***	1.45***
ln(SW Lumber C$_t$)	−1.45	−2.18	−1.50	−1.87	0.991***	0.263

Note: * Indicates statistical significant at 10%, ** at 5%, and *** at 1%. PPI, producer price index; OSB, oriented strandboard; SW, softwood.

marginal federal income tax rate (percent) for a household annual income level of \$60,000 at constant (2012) dollars (deflated by the Consumer Price Index for all urban consumers [deemed most appropriate for deflating household expenditures]; US Bureau of the Census, 2017b), W is the wage of construction workers deflated by the chained GDP deflator, and τ is a time trend that indexes gradual changes in unmodeled factors (unrelated to population) affecting housing demand as well as building regulations and land constraints affecting supply (e.g., Saiz 2010). The \$60,000 income level used to calculate tax rates was based the median income of first-time home buyers (National Association of Home Builders 2017).

Because we model a time series of housing starts and use quarterly observations, we explicitly recognize seasonality and the order of integration of all variables (where Table 1 lists the variables in the study and Table 2 reports results of unit root tests). Augmented Dickey-Fuller (ADF) tests (Dickey and Fuller 1979, Said and Dickey 1984), Phillips-Perron (PP) Tau tests (Phillips and Perron 1988; both with the null hypothesis that a series is nonstationary, i.e., contains a unit root), and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) tests (Kwiatkowski et al. 1992; null hypothesis of stationarity, i.e., no unit root), done over a time span from the beginning of our modeled series to the bottom of the last recession and then to the end of our time series, for reasons explained later in the *Methods* section, indicate that quarterly, nationwide total housing starts are stationary. Single-family starts are likely to be stationary, with only data through 2008q4 indicating possible nonstationarity. Multifamily starts series are less likely to be stationary, according to all three tests and all time spans tested, although the evidence is ambiguous, with the ADF and PP-Tau weakly rejecting (at 10%) a unit root and the KPSS tests only weakly (10%) or more strongly (5%) rejecting a null of stationarity. The findings that total and single-family starts are stationary make sense from a perspective that starts are effectively a gross housing stock change (e.g., Ball et al. 2010). However, the tests do suggest that a near-unit-root process exists in housing starts, meriting a single-quarter lag of the starts to be included in the empirical specifications. Because the possible near-unit-root situation in small samples such as ours also presents issues of size distortions in the KPSS test (Caner and Kilian 2001, Müller 2005)—tests being oversized, rejecting the null at a rate greater than the nominal significance levels—the rejection of stationarity is weak. Given that the time series of starts demonstrates seasonality at the quarterly time step (tests not reported but easily observable in plots of quarterly data), we additionally adjust Equation 4 to include quarterly (seasonal) dummies ($\mathbf{D}_t$). Finally, unit root tests also confirm that real GDP, mortgage interest rates, construction wages, mortgage delinquency rates, and the marginal federal income tax rate at median household incomes are nonstationary, necessitating first-differencing to achieve stationarity. Thus, we model nationwide housing starts, with superior dots indicating changes from quarter $t - 1$ to t, as

$$H_t = f(H_{t-1}, \mathbf{D}_t, \dot{Y}_t, \dot{R}_t, \dot{A}_t, \dot{C}_t, \dot{G}_t, \dot{W}_t, \tau_t) \quad (5)$$

An unbiased, consistent estimate of Equation 5 using historical data would generate a set of independently and identically distributed random errors, ϵ_t^H.

Estimates of Equation 5 can also be used to evaluate a question of structural change. We apply a Chow breakpoint test (Chow 1960), examining whether the parameters of the estimated model differ before and after some proposed breakpoint. Our proposed breakpoint was 2008q4 to 2009q1, at the deepest point in the last US recession. The resulting test statistic is distributed $F(\kappa, T - 2\kappa)$, where κ is the number of regressors in the model and T is the number of observations over the whole data set.

Because the model will be used as a basis of projections, we evaluate alternative specifications with respect to out-of-sample performance. In particular, we estimate models over data from 1979q2 through 2008q4 and then forecast to 2014q4. This long span of out-of-sample conditions for model testing can reveal the effects of biases resulting from aggregation and omitted variables.

We also estimate a version of Equation 5 that is parsimonious but not more biased than a model that is (according to theory) fully specified. Parsimony is desirable because it reduces the number of independent variables also needing projection when housing starts are projected into the future. Therefore, expressions of Equation 5 progress from the general to the parsimonious. Parsimony is achieved by applying a model assessment and evaluation process as suggested by Gauch (1988) using data splitting and model assessment with a goal of identifying an unbiased, parsimonious predictive model. To accomplish this, we start by estimating a fully specified model and assess its goodness of fit, including bias. We then drop insignificant variables and reassess bias. Next, we drop significant variables that are not easily projected and assess the effects on bias until we arrive at a final, parsimonious form that is not more biased than the general specification. We note that a full specification includes potentially endogenous predictors, such as construction wages. Models including these potentially endogenous variables are estimated with instrumental variable methods (two-staged least squares [2SLS]).

Monte Carlo Projection Methods

The primary application of the housing start model is to better understand projected derived demands for wood products. Expectations about future levels of starts and wood product demands can be understood by applying Monte Carlo methods to estimated models. In this study, Monte Carlos are accomplished by (1) randomly sampling from the historical data used to estimate equations of housing starts, wood products, and exogenous predictors of the starts and wood products models; (2) estimating the equations using the randomly drawn historical data; (3) projecting the exogenous predictors to some future date; (4) projecting the starts models and the wood products models to the future date and recording those projections; (5) repeating steps 1–4 for many iterations; and (6) summarizing the results of the projected starts and wood products. The projection period in this study runs from 2015 to 2070. Each Monte Carlo projection consists of 1,000 iterations.

Projections of housing starts require projections of all exogenous variables that explain starts and wood products demands. As will be shown in the *Results* section, we selected a model of starts that includes only changes in GDP (i.e., GDP growth) and residential mortgage delinquencies as exogenous predictors. As such, the projections require assumptions, or models, of the future evolution of GDP growth and delinquencies. For real GDP growth, we opted for a simple autoregressive specification of order K (an AR(K) model) in first-differences of GDP, housing starts, and residential mortgage delinquencies. The inclusion of first-differences of housing starts in this model is justified by the idea that new construction has a small but non-negligible effect on overall US economic output (e.g., Moench and Ng 2011) whereas inclusion of delinquencies could be similarly justified. Seasonality is also evident in GDP growth; therefore, we additionally include quarterly dummies in the GDP specification. The change in the natural log of real GDP is expressed as a function of a constant, its own lagged quarterly changes, quarterly dummies, and lagged quarterly changes in housing starts. We use a general-to-specific model selection process to identify which lagged difference terms and quarter dummies appear in a final, parsimonious specification. The initial specification is

$$\dot{Y}_t = \alpha + \beta_k \dot{\mathbf{Y}}_{t-k} + \delta \mathbf{D}_t + \omega_l \dot{\mathbf{H}}_{t-l} + \epsilon_t^{\dot{Y}} \quad (6)$$

where the elements of $\boldsymbol{\beta}_k$ are coefficients on lagged GDP changes; $\boldsymbol{\delta}$ is a vector of quarterly dummy coefficients corresponding to quarters 1, 2, and 3; $\boldsymbol{\omega}_l$ is a vector of parameters measuring how changes in housing construction levels affect GDP growth; and $\epsilon_t^{\dot{Y}}$ is an independent and identically distributed random error. Evident in Equation 6 is that lagged changes in residential mortgage delinquencies were not significant explainers of GDP growth; therefore, they were dropped from the final specification.

The intercept of the GDP growth Equation 6 needs to be adjusted if we seek to model alternative rates of real GDP growth into the future. The intercept in a first-difference model measures the historical rate of change of the dependent variable. Therefore, we adjust the intercept to project GDP into the future under assumed growth rates that differ from the historical rate. The adjustment can be done by converting the intercept to a function of the coefficients on the lagged changes and the quarterly dummies:

$$\alpha^g = [(1 + g)^{0.25} - 1 - (0.25 \textstyle\sum_{Q=1}^{3} \delta_Q)\,(1 - \sum_{k=1}^{K} \beta_k)^{-1}](1 - \sum_{k=1}^{K} \beta_k) \quad (7)$$

where g is the decimal rate of assumed (projected) annual average real GDP growth over the 55-year span of the projection (2015–70), α^g is the intercept used in projecting GDP growth into the future, δ_Q represents the parameters of the $\boldsymbol{\delta}$ vector from Equation 6, and other variables and parameters are as previously defined. In our simulations, we add random quarterly errors to GDP growth, taken from random draws of $\epsilon_t^{\dot{Y}} \sim N(0, \delta^{\dot{Y}})$, on the basis of the regression results for Equation 6.

In this study, projected real GDP growth was varied in seven separate Monte Carlo projections that differed only in their assumptions of future real GDP growth. The first six projected annual growth in 1% increments, from 0% to 5%. The seventh projected real GDP annual growth at 2.4%, which was the historical average rate observed between 1990 and 2014, which we contend is a plausible outlook for the future of US economic growth (e.g., Gordon 2016). A projected set of random draws of $\epsilon_t^{\dot{Y}}$ does not guarantee that the average annual growth rate in a Monte Carlo projection will match the assumed rate. To ensure that annual real GDP growth over the entire projection, 2015–70, matches the assumed rates, we (1) generate, using the GDP growth equation and housing starts equation, a series of random changes in logarithmically transformed real GDP (Equation 6); (2) add the random changes to logarithmically transformed real GDP, $\hat{Y}_t = \hat{Y}_{t-1} + \hat{\dot{Y}}_t$; (3) calculate the average quarterly deviation of the randomized rate of quarterly growth over the random real GDP projection as $\bar{g}_q = \left(\frac{1}{134}\right)(\hat{Y}_{2070q4} - \hat{Y}_{2015q3})$; and (4) generate an adjusted random realization that matches assumed growth as $\hat{\hat{Y}}_t = \hat{Y}_t - \bar{g}_q$. Because lagged changes in housing starts are part of the GDP growth equation (Equation 6), steps 1–4 are repeated 3 times for each Monte Carlo iteration to allow both the starts projection and the adjusted GDP growth projection to converge to a stable random projection.

To project the residential mortgage delinquency rate, we developed a statistical model specified as a function of exogenous variables. As in the case of real GDP, we started from a general specification of the mortgage delinquency rate and dropped insignificant variables to arrive at a parsimonious specification. Given that the unit root tests of this variable indicated possible stationarity (Table 2), the general specification related the level of the mortgage delinquency rate to its lagged level, four lagged changes in the level, four lagged changes in real GDP, four lagged levels of total housing starts, and quarterly dummies:

$$C_t = \mu_0 + \mu_1 C_{t-1} + \sum_{j=1}^{4} \delta_j \dot{C}_{t-j} + \sum_{j=1}^{4} \gamma_j \dot{Y}_{t-j} + \sum_{j=1}^{4} \tau_j H_{t-j} + \delta \mathbf{D}_t + \epsilon_t^C \quad (8)$$

The final specification of Equation 8 did not include housing starts but did include lagged changes in delinquency rates and lagged changes in real GDP.

The United States has historically imported a substantial share (on average, 28% between 1979 and 2013) of softwood lumber domestically consumed. These imports and domestic production have been driven in part by softwood lumber demand from the construction sector (e.g., Song et al. 2011). To quantify softwood lumber consumption, we formulate a reduced-form softwood lumber quantity equation that is derived from equilibrium supply and demand. Softwood lumber demand derives from housing starts as well as other components of the economy that demand softwood lumber as an input. These other components include repairs and renovations of the existing housing stock, commercial construction, manufacturing, and shipping, which we proxy with real GDP. Demand for lumber is also influenced by the price of substitutes in construction. Although including in a softwood lumber demand model a variable that indexes residential improvements and repair activity might be preferred, we note that consistently reported quarterly time series data on such a variable are not available from government data sources. (We tested inclusion of the real value of house maintenance and repairs [US Bureau of the Census 2016b] in Equation 8, obtained from the US Census Bureau, but this variable was statistically insignificant, did not significantly affect the magnitudes of the estimated parameters of either starts or real GDP, and covered a shorter time series than that available for other models. We dropped further consideration of this variable in Equation 8. We contend that improvements and repair spending is likely to be captured by the included real GDP growth variable.) Softwood lumber supply is a function of lumber price as well as the price of inputs to lumber production, such as mill wages, electricity, and timber. Softwood lumber demand and supply are defined as follows:

$$\begin{aligned} Q_{L,t}^D &= f(P_{L,t}^{US},\ P_{L,t}^{M},\ H_t,\ Y_t,\ \mathbf{N}_t) \\ Q_{L,t}^S &= g(P_{L,t}^{US},\ \mathbf{F}_t) \\ M_{L,t}^S &= h(P_{L,t}^{M},\ \mathbf{F}_t) \end{aligned} \quad (9)$$

where $Q_{L,t}^D$ is the quantity of derived softwood lumber demanded in the United States in period t, $Q_{L,t}^S$ is the quantity of softwood lumber supplied by the domestic US market in period t, $M_{L,t}^S$ is the net import quantity of softwood lumber in period t, $\mathbf{F}_t$ is a vector of lumber production input prices in period t, $\mathbf{N}_t$ is a vector of substitutes for lumber in construction, and other variables are as previously defined. At equilibrium, softwood lumber consumption equals the sum of domestic production and net imports. A reduced-form equation, factoring out own price ($P_{L,t}^{US}$), leads to

$$Q_{L,t} = f_2(P_{L,t}^{M},\ H_t,\ Y_t,\ W_t,\ \mathbf{R}_t) \quad (10)$$

Recognizing that the import price and some input prices may also be endogenous in this reduced-form specification, a more parsimonious version of Equation 10 would be

$$Q_{L,t} = f_3(H_t,\ Y_t,\ W_{1,t},\ \mathbf{R}_{1,t}) \quad (11)$$

where $W_{1,t}$ is the subvector of input prices considered exogenous in derived lumber demand and $\mathbf{R}_{1,t}$ is the subvector of input prices considered exogenous in softwood lumber manufacture.

This model was specified using a logarithmic transformation and, given that an ADF test could not reject a null of a unit root (Table 2), we specify the model in first-differences in Equation 12, with superior dots indicating changes from quarter $t-1$ to t:

$$\ln \dot{Q}_t = f(\dot{H}_t,\ \dot{Y}_t,\ \dot{R}_t,\ \dot{W}_t) \quad (12)$$

Equations 11 and 12 are abstracted from what might be considered fully specified consumption quantity equations, which could include $\mathbf{F}_t$ and $P_{L,t}^M$. Other studies have shown (e.g., Song et al. 2011) that structural softwood lumber demand specifications have been successfully estimated with own and substitute prices. Although their inclusion could improve model fit, the additional variables would represent new challenges in projecting lumber demand into the future.

Results

Housing Starts Models

Housing starts equation estimates, from least to most parsimonious, are shown in Table 3 (total housing starts), Table 4 (single-family housing starts), and Table 5 (multifamily housing starts). Models 1–3 apply instrumental variables methods (2SLS) whereas the remaining models are estimated with least squares. All specifications apply a White's (1980) correction for residual heteroscedasticity. Standard errors of regression were similar across all model specifications for total, single-family, and multifamily starts categories. Across all models, in-sample R^2 values ranged from 0.89 to 0.93,

Table 3. Equation estimates for total housing starts, general to parsimonious specifications, by model.

Variable	Model 1	2	3	4	5	6	7
Constant	0.30 (0.37)	0.48 (0.26)*	0.49 (0.26)*	0.44 (0.25)*	0.43 (0.25)*	0.43 (0.25)*	0.42 (0.28)
$\ln(\text{Starts}_{t-1})$	0.93 (0.07)***	0.89 (0.04)***	0.89 (0.04)***	0.90 (0.04)***	0.90 (0.04)***	0.90 (0.04)***	0.90 (0.05)***
Quarter 1 dummy	−0.12 (0.08)	−0.09 (0.04)**	−0.09 (0.04)***	−0.09 (0.03)**	−0.09 (0.03)**	−0.09 (0.03)**	
Quarter 2 dummy	0.41 (0.08)***	0.47 (0.03)***	0.47 (0.03)***	0.46 (0.02)***	0.46 (0.02)***	0.46 (0.02)***	0.46 (0.02)***
Quarter 3 dummy	0.086 (0.068)	0.15 (0.03)***	0.15 (0.03)***	0.14 (0.02)***	0.14 (0.02)***	0.14 (0.02)***	0.14 (0.02)***
$\ln(\text{GDP}_t)-\ln(\text{GDP}_{t-1})$	2.82 (3.45)	4.58 (1.17)***	4.59 (1.16)***	4.51 (1.17)***	4.54 (1.18)***	4.54 (1.18)***	4.14 (1.26)***
$\ln(\text{Population}_t)-\ln(\text{Population}_{t-1})$	6.38 (12.08)	1.79 (5.54)					
$\ln(\text{Mortgage Rate}_t)-\ln(\text{Mortgage Rate}_{t-1})$	−0.67 (0.25)***	−0.54 (0.20)***	−0.53 (0.19)***	−0.50 (0.18)***	−0.53 (0.19)***	−0.53 (0.19)***	
$\ln(\text{Wage Rate}_t)-\ln(\text{Wage Rate}_{t-1})$	0.47 (0.96)	−0.42 (0.67)	−0.41 (0.67)				
$\ln(\text{Tax Rate}_t)-\ln(\text{Tax Rate}_{t-1})$	−0.25 (0.15)	−0.17 (0.17)	−0.17 (0.17)	−0.18 (0.16)			
$\ln(\text{Delinquency Rate}_t)-\ln(\text{Delinquency Rate}_{t-1})$	−0.71 (0.25)***	−0.62 (0.13)***	−0.62 (0.13)***	−0.64 (0.13)***	−0.65 (0.13)***	−0.65 (0.13)***	−0.36 (0.08)***
$\ln(\text{PPI Concrete}_t)-\ln(\text{PPI Concrete}_{t-1})$	−1.04 (1.56)						
$\ln(\text{PPI SW Plywood}_t)-\ln(\text{PPI SW Plywood}_{t-1})$	0.72 (0.58)						
$\ln(\text{PPI SW Lumber}_t)-\ln(\text{PPI SW Lumber}_{t-1})$	0.31 (1.51)						
$\ln(\text{PPI OSB}_t)-\ln(\text{PPI OSB}_{t-1})$	−0.43 (0.47)						
Standard error of regression	0.091	0.080	0.079	0.078	0.079	0.079	0.084
R^2	0.89	0.91	0.91	0.91	0.91	0.91	0.90
Durbin's H-statistic	0.48	0.44	0.45	0.46	0.41	0.41	1.91*

Note: Values in parentheses are standard errors. * Indicates statistical significance at 10%, ** at 5%, *** at 1%. PPI, producer price index; OSB, oriented strandboard; SW, softwood.

Table 4. Equation estimates for single-family housing starts, general to parsimonious specifications, by model.

Variable	Model 1	2	3	4	5	6	7
Constant	0.06 (0.29)	0.21 (0.23)	0.21 (0.23)	0.20 (0.22)	0.17 (0.22)	0.13 (0.22)	0.26 (0.26)
$\ln(\text{Starts}_{t-1})$	0.96 (0.06)***	0.92 (0.04)***	0.93 (0.04)***	0.93 (0.04)***	0.93 (0.04)***	0.94 (0.04)***	0.92 (0.05)***
Quarter 1 dummy	−0.06 (0.08)	−0.04 (0.04)	−0.04 (0.04)	−0.03 (0.04)	−0.03 (0.04)		
Quarter 2 dummy	0.48 (0.09)***	0.54 (0.04)***	0.54 (0.04)***	0.53 (0.02)***	0.53 (0.02)***	0.54 (0.02)***	0.51 (0.02)***
Quarter 3 dummy	0.095 (0.077)	0.17 (0.04)***	0.17 (0.04)***	0.16 (0.02)***	0.16 (0.02)***	0.16 (0.02)***	0.16 (0.02)***
$\ln(\text{GDP}_t)-\ln(\text{GDP}_{t-1})$	2.02 (3.40)	3.96 (1.23)***	3.95 (1.23)***	3.91 (1.23)***	3.92 (1.23)***	3.99 (1.22)***	3.15 (1.30)**
$\ln(\text{Population}_t)-\ln(\text{Population}_{t-1})$	10.10 (16.27)	4.28 (4.27)					
$\ln(\text{Mortgage Rate}_t)-\ln(\text{Mortgage Rate}_{t-1})$	−0.92 (0.28)***	−0.74 (0.21)***	−0.72 (0.20)***	−0.71 (0.19)***	−0.73 (0.20)***	−0.73 (0.20)***	
$\ln(\text{Wage Rate}_t)-\ln(\text{Wage Rate}_{t-1})$	0.57 (1.12)	−0.23 (0.76)	−0.21 (0.75)				
$\ln(\text{Tax Rate}_t)-\ln(\text{Tax Rate}_{t-1})$	−0.23 (0.16)	−0.16 (0.17)	−0.16 (0.17)	−0.16 (0.16)			
$\ln(\text{Delinquency Rate}_t)-\ln(\text{Delinquency Rate}_{t-1})$	−0.82 (0.27)***	−0.71 (0.14)***	−0.71 (0.14)***	−0.72 (0.13)***	−0.73 (0.13)***	−0.63 (0.08)***	−0.60 (0.09)***
$\ln(\text{PPI Concrete}_t)-\ln(\text{PPI Concrete}_{t-1})$	−1.23 (1.92)						
$\ln(\text{PPI SW Plywood}_t)-\ln(\text{PPI SW Plywood}_{t-1})$	0.97 (0.64)						
$\ln(\text{PPI SW Lumber}_t)-\ln(\text{PPI SW Lumber}_{t-1})$	0.22 (1.57)						
$\ln(\text{PPI OSB}_t)-\ln(\text{PPI OSB}_{t-1})$	−0.47 (0.55)						
Standard error of regression	0.10	0.08	0.08	0.08	0.08	0.08	0.09
R^2	0.89	0.93	0.93	0.93	0.93	0.92	0.91
Durbin's H-statistic	0.08	0.27	0.30	0.34	0.33	0.32	2.50***

Note: Values in parentheses are standard errors. * Indicates statistical significance at 10%, ** at 5%, *** at 1%. PPI, producer price index; OSB, oriented strandboard; SW, softwood.

with best-fitting models found for single-family starts, least for multifamily. Significant residual serial correlation was found in only the most parsimonious specification, as measured by Durbin's H-statistic (Durbin 1970).

Across all starts categories and model specifications, there is serial dependence in housing starts; that is, the previous quarters' (lagged) starts coefficients ranged from 0.89 to 0.96. Seasonality is evident and significant, with spring and summer quarters (2 and 3) having higher starts levels and winter having lower starts levels compared with the fall quarter (quarter 4, the omitted dummy). Real GDP growth is a significant and positive predictor of starts. Changes in population in the United States are not statistically significant contributors to housing starts. Changes in the mortgage rate are negatively and significantly (at 1%) related to housing starts, and this effect is consistent across all specifications tested. Another consistently negative and significant (at 1%) predictor of total starts is the loan delinquency rate. In contrast, we did not find statistically significant relationships with starts for (instrumented) construction wages or changes in the marginal tax rate. Other input prices in construction, including concrete, softwood plywood, softwood lumber, and oriented strandboard, were not significantly related to starts levels. The insignificant finding on construction wages and other input prices could be due to a lack of good instruments for these variables. Our instruments included the change in the unemployment rate (aggregate, US), lagged input prices, and per-capita real GDP.

Model forecasts out of sample (Table 7) showed that, as the models became more parsimonious, goodness-of-fit statistics improved. Statistics shown in Table 7 are based on the logarithmically transformed dependent variable and its forecast. Model specification 7 was the best fitting, as measured by the root mean squared error (RMSE) and bias. Both statistics were lowest with this specification, with the exception of multifamily starts, in which the specification that included the change in the mortgage interest rate generated

Table 5. Equation estimates for multifamily housing starts, general to parsimonious specifications, by model.

Variable	Model 1	2	3	4	5	6	7
Constant	0.20 (0.32)	0.24 (0.17)	0.22 (0.15)	0.21 (0.15)	0.22 (0.15)	0.22 (0.15)	0.15 (0.14)
$\ln(\text{Starts}_{t-1})$	0.94 (0.06)***	0.93 (0.03)***	0.93 (0.03)***	0.93 (0.03)***	0.93 (0.03)***	0.93 (0.03)***	0.93 (0.03)***
Quarter 1 dummy	−0.25 (0.17)	−0.24 (0.06)***	−0.24 (0.06)***	−0.23 (0.06)***	−0.23 (0.06)***	−0.23 (0.06)***	
Quarter 2 dummy	0.36 (0.28)	0.32 (0.04)***	0.33 (0.04)***	0.30 (0.03)***	0.30 (0.03)***	0.29 (0.03)***	0.34 (0.03)***
Quarter 3 dummy	0.149 (0.215)	0.12 (0.03)***	0.12 (0.03)***	0.09 (0.03)***	0.09 (0.03)***	0.09 (0.03)***	0.11 (0.03)***
$\ln(\text{GDP}_t)-\ln(\text{GDP}_{t-1})$	4.32 (2.14)**	4.68 (1.52)***	4.65 (1.51)***	4.66 (1.49)***	4.72 (1.49)***	4.56 (1.50)***	5.59 (1.67)***
$\ln(\text{Population}_t)-\ln(\text{Population}_{t-1})$	−5.51 (18.71)	−4.18 (13.93)					
$\ln(\text{Mortgage Rate}_t)-\ln(\text{Mortgage Rate}_{t-1})$	−0.14 (0.29)	−0.14 (0.26)	−0.16 (0.26)	−0.10 (0.24)	−0.12 (0.24)		
$\ln(\text{Wage Rate}_t)-\ln(\text{Wage Rate}_{t-1})$	−1.76 (6.96)	−0.78 (0.58)	−0.80 (0.58)				
$\ln(\text{Tax Rate}_t)-\ln(\text{Tax Rate}_{t-1})$	−0.15 (0.25)	−0.13 (0.19)	−0.14 (0.19)	−0.15 (0.16)			
$\ln(\text{Delinquency Rate}_t)-\ln(\text{Delinquency Rate}_{t-1})$	−0.10 (0.21)	−0.48 (0.20)**	−0.48 (0.20)**	−0.50 (0.20)**	−0.52 (0.20)**	−0.51 (0.20)**	0.23 (0.11)**
$\ln(\text{PPI Concrete}_t)-\ln(\text{PPI Concrete}_{t-1})$	−0.38 (0.47)						
$\ln(\text{PPI SW Plywood}_t)-\ln(\text{PPI SW Plywood}_{t-1})$	0.52 (4.28)						
$\ln(\text{PPI SW Lumber}_t)-\ln(\text{PPI SW Lumber}_{t-1})$	−0.03 (0.32)						
$\ln(\text{PPI OSB}_t)-\ln(\text{PPI OSB}_{t-1})$	0.07 (0.66)						
Standard error of regression	0.13	0.12	0.12	0.12	0.12	0.12	0.13
R^2	0.90	0.91	0.91	0.91	0.91	0.91	0.90
Durbin's H-statistic	−1.03	−1.16	−1.15	−1.23	−1.30	−1.15	−2.08**

Note: Values in parentheses are standard errors. * Indicates statistical significance at 10%, ** at 5%, *** at 1%. PPI, producer price index; OSB, oriented strandboard; SW, softwood.

Table 6. Estimates of models of real US GDP growth rates, the total rate of residential mortgage delinquency, and softwood lumber consumption.

Variable	GDP model $\ln(\text{GDP}_t)-\ln(\text{GDP}_{t-1})$	Mortgage delinquency model $\ln(\text{Delinquency Rate}_t)$	Softwood lumber model $\ln(Q_t)-\ln(Q_{t-1})$
Constant	0.0063 (0.0015)***	0.14 (0.03)***	−0.024 (0.009)**
$\ln(\text{GDP}_t)-\ln(\text{GDP}_{t-1})$			1.23 (0.30)***
$\ln(\text{GDP}_{t-1})$-$\ln(\text{GDP}_{t-2})$	0.33 (0.08)***	−1.93 (0.60)***	
$\ln(\text{GDP}_{t-8})-\ln(\text{GDP}_{t-9})$	−0.20 (0.05)***		
$\ln(\text{GDP}_{t-10})-\ln(\text{GDP}_{t-9})$	0.26 (0.07)***		
$\ln(\text{GDP}_{t-12})-\ln(\text{GDP}_{t-13})$	−0.14 (0.06)**		
Quarter 1 dummy		−0.12 (0.02)***	
Quarter 2 dummy	0.0042 (0.0013)***	−0.019 (0.008)**	
Quarter 3 dummy	−0.010 (0.003)***		
$\ln(\text{Delinquency Rate}_{t-1})$		0.94 (0.02)***	
$\ln(\text{Delinquency Rate}_{t-4})-\ln(\text{Delinquency Rate}_{t-5})$		0.40 (0.08)***	
$\ln(\text{Housing Starts}_t)-\ln(\text{Housing Starts}_{t-1})$[a]	0.029 (0.007)***		0.44 (0.04)***
$\ln(\text{Housing Starts}_{t-3})-\ln(\text{Housing Starts}_{t-4})$	0.0074 (0.0031)**		
Standard error of regression	0.0056	0.044	0.030
R^2	0.45	0.96	0.91
Durbin's H-statistic	−1.86*	−0.42	−0.11
Observations	146	142	36

Note: Values in parentheses are standard errors. * Indicates statistical significance at 10%, ** at 5%, *** at 1%. The subscript *t* in the dependent and independent variables refers to the quarter for the housing starts and mortgage delinquency rate models and the year in the softwood lumber model.
[a] Single-family starts for the softwood lumber consumption equation.

Table 7. Out-of-sample model assessments, general to parsimonious specifications, for total housing starts, single-family starts, and multifamily starts.

Model number	Model type	Total starts RMSE	Total starts Bias	Single-family starts RMSE	Single-family starts Bias	Multifamily starts RMSE	Multifamily starts Bias
1	2SLS	0.68	−0.38	1.11	−0.62	0.64	−0.21
2	2SLS	0.63	−0.43	0.57	−0.36	0.57	−0.24
3	2SLS	0.62	−0.44	0.57	−0.38	0.54	−0.22
4	OLS	0.56	−0.41	0.53	−0.36	0.50	−0.19
5	OLS	0.54	−0.41	0.50	−0.35	0.49	−0.20
6	OLS	0.54	−0.41	0.44	−0.32	0.47	−0.19
7	OLS	0.44	−0.35	0.43	−0.32	0.54	−0.12

Note: OLS, ordinary least squares.

better fit statistics. Although these dynamic forecasts out of sample, keeping parameter estimates constant, display a tendency to forecast too high, we made no bias adjustments to the Monte-Carlo-based forecasts. The quarter 1 dummy was dropped in model specification 7 for total housing starts and multifamily housing starts and specifications 6 and 7 for single-family housing starts. This was done because models estimated over the longer time span, to 2014q4, showed the quarter 1 dummy to be statistically insignificant.

We also tested the conjecture that the 2008/2009 recession resulted in a structural change in housing markets (Anundsen 2015, Myers 2016) by testing the hypothesis that pre- and postrecession parameter estimates are equivalent. The Chow breakpoint test for a structural break in the parameters of the total starts model (specification 7) showed that there was no statistically significant structural change in total starts ($P = 0.27$) or single-family starts ($P = 0.56$). For the estimate of specification 7 of the multifamily starts, the test found in favor of a breakpoint ($P = 0.02$).

GDP Growth Model Estimates

We estimated the GDP growth model using quarterly data from 1979q1 to 2015q3 (Table 1). The selected model is reported in Table 6. Estimated with a correction for heteroscedasticity, this model has an R^2 of 0.45. Predictors in the final specification include lagged GDP growth rates at lags 1, 2, 8, 9, and 12 and lags 1 and 4 of the first-differences of housing starts. The quarter 2 and quarter 3 dummies were significant and included. The estimate also includes a statistically significant intercept of 0.0063. This intercept captures the average rate of quarterly GDP change whereas the coefficient on the quarter 2 dummy, 0.0042, indicates that the second quarter of each year tends to have a higher GDP growth rate than other quarters and the coefficient on quarter 3 (-0.010) indicates that the third quarter growth tends to be slightly lower than growth the rest of the year. Lagged housing starts changes show that housing is positively related to future GDP growth.

Mortgage Delinquency Model Estimates

The mortgage delinquency rate models used quarterly data from 1979q1 to 2015q3 (Table 1). The parsimonious version of Equation 8, reported in Table 6, includes the lagged level of the delinquency rate, the first lag of GDP growth rate, quarter 1 and quarter 2 dummies, and the fourth lag of the change in the delinquency rate. The estimate shown in Table 6 uses all observations, but model selection occurred using data only through 2008q4. The estimate was also tested in out-of-sample conditions, estimating the model over data through 2008q4 and forecasting through 2015q3. Predictions out of sample (and back-transforming the logarithmic prediction, including a bias correction done by adding half of the variance of the regression estimate) showed that the level of the mortgage delinquency rate tracked observed rates over 2009q1 through 2015q3 with a positive bias in percentage points of 0.31 and a RMSE of 0.53. The model estimated over the whole sample had negligible autocorrelation as measured by the H-test and a high degree of explanatory power ($R^2 = 0.96$).

In our projections of mortgage delinquency rates, we added random changes to the mortgage delinquency predicted by the final equation shown in Table 6 by adding a normally distributed random error with a mean of zero and a standard deviation equal to the standard error of the regression as reported in Table 6. The lagged GDP growth rate in this projection was the adjusted random quarterly growth generated by the GDP projection at the assumed GDP growth rate. The randomly generated mortgage delinquency rate was restricted to fall between the minimum (3.62%) and maximum (10.44%) levels observed between 1979q2 and 2015q3; therefore, these bounds were an assumed range of plausibility in projected future years. Therefore, when the equation projected a higher rate than 10.44% for a given quarter, the value for the quarter was set at 10.44%; when it projected a lower rate than 3.62%, the value for the quarter was set at 3.62%.

Softwood Lumber Consumption Model Estimates

We tested a fully specified reduced-form model of softwood lumber consumption (Equation 12) with annual data, 1960–2014. This model was specified using a logarithmic transformation and, given that tests for a unit root (Table 2) were consistent with a unit root process, estimated in first-differences. The most general specification of this model showed that the price index for concrete products (an alternative building material), the price index of electricity (an input cost), and the federal funds rate were not statistically significant drivers of softwood lumber consumption. The final model is reported in Table 6. The only variables statistically significant at 5% or stronger were total housing starts and the GDP growth rate. In an alternative version that separately included single-family and multifamily housing starts as predictors, only single-family starts were significant at 5% or stronger. Hence, two competing models are evaluated here: total housing starts and GDP growth (Model 1) and single-family housing starts and real GDP growth (Model 2).

In parsimonious models using data from 1960 to 2000, all variables were significant at stronger than 1%. Model 1 had an in-sample R^2 of 0.83, Model 2 had an in-sample R^2 of 0.87, and neither version had statistically significant residual autocorrelation. Both models were then tested for their forecast performance out of sample, 2001–14. In the out-of-sample goodness-of-fit evaluation, predictions were back-transformed and converted to levels by adding half of the variance of the equation errors before exponentiation of the predicted quantity in natural logarithms. Model 1 had a RMSE of 3,679 and a bias (tending to overpredict) of 3,815 whereas Model 2 had analogous statistics of 2,595 and 2,690. On the basis of these results, Model 2 is preferred.

Previous research is mixed regarding explanatory variables in a model of softwood lumber. Song et al. (2011) excluded GDP in their model of softwood lumber demand in the United States whereas Buongiorno (2015) reports specifications of total (coniferous plus nonconiferous) lumber demand internationally as a function of GDP but not housing starts or other indices of construction. On the other hand, Ince et al. (2011) included GDP and single family starts in their softwood lumber demand specification. Because a model that included both variables explained more variation (R^2 of 0.86 versus 0.79 for one that excluded real GDP and 0.44 for one that excluded housing starts but included real GDP for data 1960–2014), in our model estimates we opted for a specification that included both. We also tested a version that included housing stock, in addition to starts and real GDP, but because inclusion of housing stocks would require a separate projection of housing stocks, we opted not to pursue this model.

Given its out-of-sample performance, the reduced-form softwood lumber model (Model 2), which included single-family housing starts and real GDP growth, was used to simulate the effects of alternative rates of GDP growth and housing starts from 2015 to 2070 on softwood lumber consumption using the starts and GDP projections described in the previous sections. The final model is reported in Table 6, with parameters estimated over the available sample of annual data corresponding to the quarterly data used for the other equations (1979–2014, because 2015 was not available at the time of this study). In the Monte Carlo simulations, the projections were made using annual data randomly drawn with replacement (the same years of the randomly drawn quarterly observations for the housing starts Monte Carlo simulations) to introduce parametric uncertainty into the softwood lumber consumption projections.

Monte Carlo Simulation Results

We project housing starts and softwood lumber consumption using the most parsimonious specification of total starts (Table 3, Model 7) and use the specifications for GDP, mortgage delinquency rate, and softwood lumber consumption summarized in Table 6. Results of the Monte Carlo projections of total US housing starts from 2015 to 2070 are shown in two figures. Figure 4 shows the

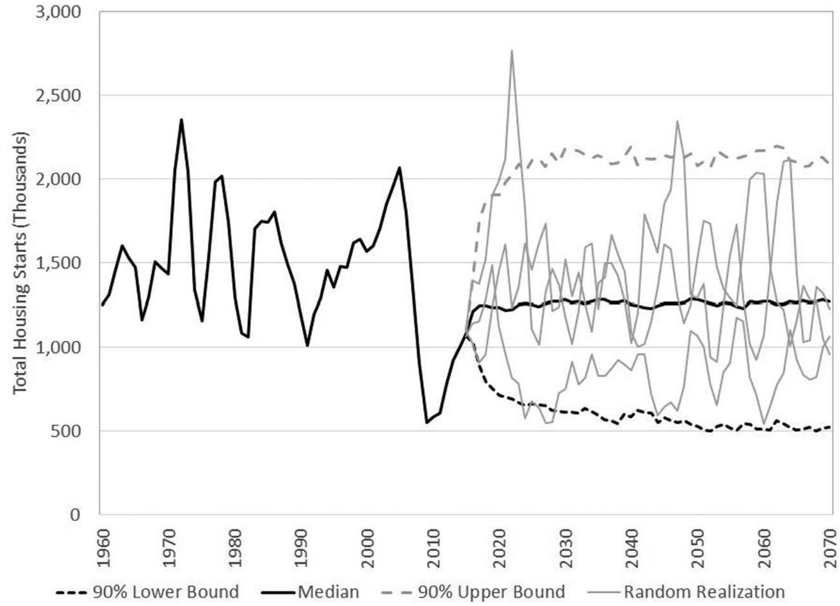

Figure 4. Housing starts: historical through 2014 and projections to 2070 using Monte Carlo methods, most parsimonious model, total (single-family plus multifamily housing), aggregate United States, assuming a real average annual GDP growth rate (2015–70) of 2.4%, including upper and lower bounds of a 90% uncertainty band and four Monte Carlo random realizations.

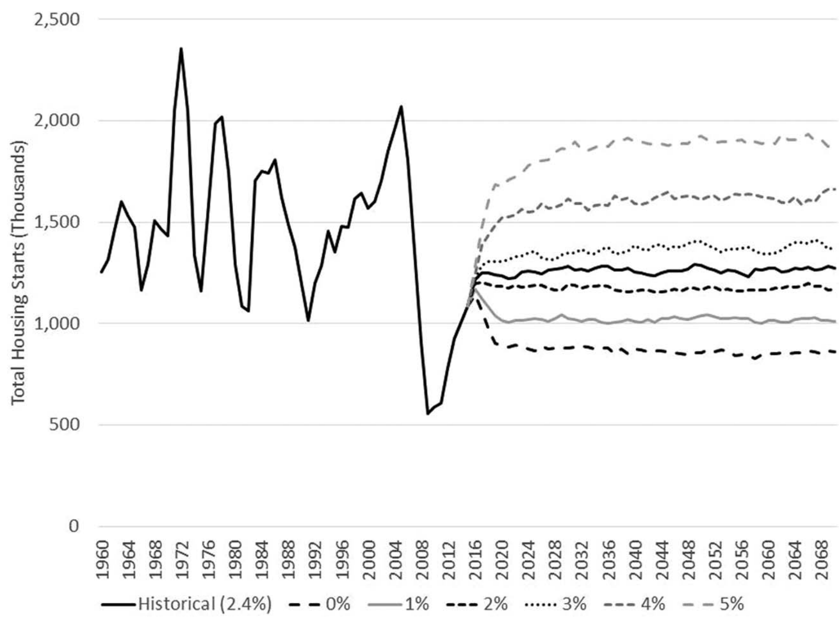

Figure 5. Housing starts: historical through 2014 and projected median levels to 2070 using Monte Carlo methods, most parsimonious model, total (single-family plus multifamily housing), aggregate United States, assuming alternative real average annual GDP growth rates (2015–70) from 0% to 5%.

projected median and 90% uncertainty bounds for total US housing starts using the assumed long-run average annual GDP growth rate of 2.4%. Median starts converge to a long-run level of 1.24 million under this GDP growth rate assumption. This figure also includes four random draws of possible futures for housing starts, which serve to illustrate two phenomena: future starts are likely to follow patterns of historical starts, including large swings in levels from one year to the next, because of the sensitivity of starts to real GDP growth, and future starts levels would be expected to swing between highs and lows that depart substantially from the projected median levels, given an average real GDP growth rate, and could drift to low or high levels that persist for long stretches. This persistence is due to the near-unit-root process evident in housing starts.

Median projections of housing starts across variations in the assumed GDP growth rate, from 0% to 5%, are displayed in Figure 5. Varying assumed rates of GDP growth results in projections of long-run median housing starts ranging from 0.86 million at 0% real GDP growth to 1.91 million at 5% real GDP growth. This range indicates that each additional percentage point of real GDP growth generates an additional median of approximately 200,000 annual housing starts.

We also used the housing starts model to generate projections of

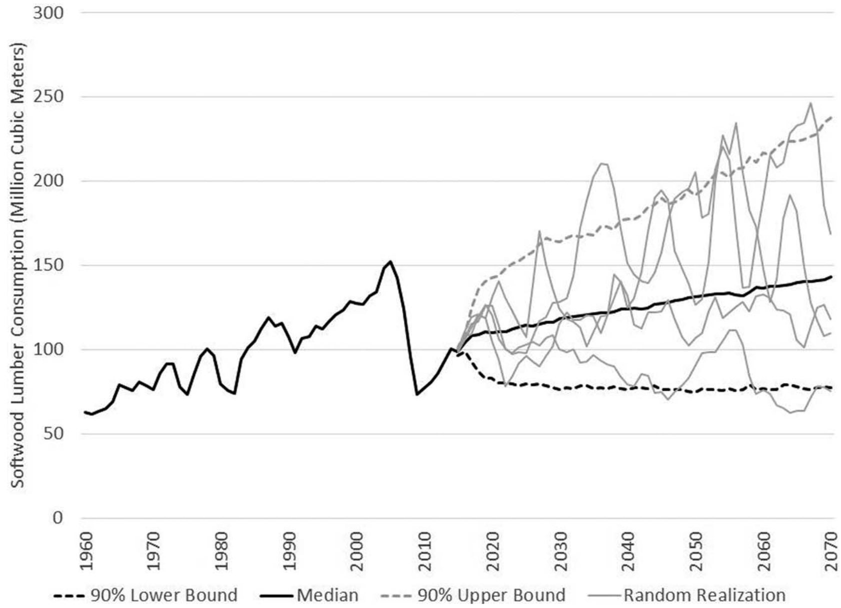

Figure 6. Softwood lumber consumption: historical through 2014 and projected median levels to 2070, Monte Carlo methods, most parsimonious housing starts model coupled with a reduced-form softwood lumber consumption model, aggregate United States, assuming a real average annual GDP growth rate (2015–70) of 2.4%, including upper and lower bounds of a 90% uncertainty band and four Monte Carlo random realizations.

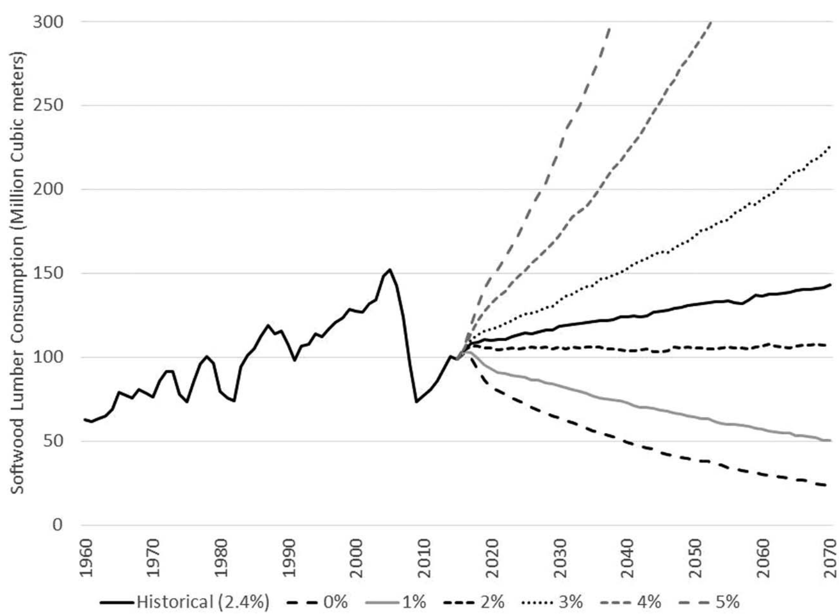

Figure 7. Softwood lumber consumption: historical through 2014 and projected median levels to 2070, Monte Carlo methods, most parsimonious housing starts model coupled with a reduced-form softwood lumber consumption model, aggregate United States, assuming alternative real average annual GDP growth rates (2015–70) from 0% to 5%.

softwood lumber consumption, 2015–70 (Figures 6 and 7). Projections at assumed annual real GDP growth of 2.4% are shown in Figure 6, including upper and lower 90% confidence limits. This figure also includes four random iterations of the Monte Carlo to illustrate how such consumption might evolve into the future. Figure 7 shows how the median levels of softwood lumber consumption would differ across different assumptions about real GDP growth rates.

Figure 6 shows that, at an average real GDP growth rate of 2.4%, median softwood lumber consumption in the United States would rise from the 2014 observed level of approximately 100 million m^3 to approximately 120 million m^3 in the 2030s and then to approximately 140 million m^3 in the 2060s. By 2070, 90% lower and upper uncertainty limits are 78 million m^3 and 238 million m^3, respectively. The simulation shows an increase in median consumption of approximately 0.6% per year between 2015 and 2070. This rise, in contrast to the plateauing of housing starts, derives from the net of a negative trend in consumption, which is 2.4%, as measured by the intercept in the estimated equation in first-differences and a positive and elastic relationship with real GDP growth. The relationship with real GDP implies a 1.23% increase in softwood lumber consumption for each 1% increase in annual GDP. It should be noted that the overall effect of real GDP growth on softwood lumber consumption comes through two channels: the indirect effect of real GDP growth through housing starts and the direct effect of real GDP growth.

The elastic relationship of softwood lumber consumption to real GDP growth explains why high rates of sustained GDP growth (say, 4% and 5% as shown in Figure 7) would generate extreme increases in consumption over time and why median consumption would decline by 70% by 2070 with 0% GDP growth. Assuming annual GDP growth that is 2% or lower would lead to median softwood lumber consumption not significantly higher than the level observed in 2014 (see Figure 7). Median softwood lumber consumption would exceed 300 million m^3 by the early 2050s at 4% average annual real GDP growth and would exceed 300 million m^3 by the late 2030s if annual real GDP growth averaged 5%. Historically, the United States has not experienced stretches of real GDP growth rates of over 3% for more than 5 years since 1979. Furthermore, real GDP growth has trended downward since 1950, with a 5-year moving annual average of approximately 5% at that time to approximately 2% in the last 5 years. Our model has been parameterized over a time span, 1979–2014, in which real GDP growth averaged 2.64%; therefore, these parameters reflect consumption patterns in construction and other applications that reflect technology and consumer tastes and preferences over that time span.

Discussion

Construction activity in the United States is a primary consumer of solidwood products. For sector analysts and policymakers, accurate assessments of future demands for such products require accurate projections of housing starts. Models of the sector could be more accurate if simple specifications of starts are used. Our study has shown that simpler specifications of housing starts generate more accurate out-of-sample forecasts of starts. The superior performance of parsimonious models simplifies long-run projections by limiting the need for additional projections of driving variables.

We found no evidence of a recession-induced structural break in housing markets (measured as total housing starts and measured as single-family housing starts) based on an explicit test for differences in pre- and postrecession parameters in the housing starts model. However, multifamily housing starts might have undergone a significant structural change. The lesson for sector modelers is that, although the 2007–09 recession was deep and induced large changes in construction activity and wood products demand in the United States, statistical models of total or single-family housing starts that are based on historical data can be an accurate representation of the essential features of the housing sector, aiding in an understanding of likely futures. Our simulation of that future, using Monte Carlo methods, indicates that future residential construction levels would settle at a long-run median of just under 1.3 million total annual starts, given continued economic growth at recent historical rates, although levels in specific years could vary widely from the median, 90% of the time ranging between 0.5 and 2.1 million.

Our additional statistical modeling also revealed how softwood lumber consumption in the United States depends elastically on the growth of GDP and on residential construction. Simulations showed that, at recent historical rates of GDP growth, median softwood lumber consumption would rise by an additional 0.6% each year through 2070, achieving a level by 2070 that is 45% higher than in 2014. However, zero economic growth in the United States would be consistent with a –2.4% annual change in softwood lumber consumption. Although a projection of 55 years imposes a constraint that the parameters will not change, and although economic growth of 0% is unlikely, we contend that long-run average annual GDP growth of 3% or less is more consistent with the economy of the future (Gordon 2016). Therefore, a future of modest ($<$1% per year) growth in softwood lumber consumption would be expected.

Previous studies (e.g., Ince et al. 2011, National Association of Home Builders 2017, Fannie Mae 2016, Freddie Mac 2016) have projected US housing starts using models of housing "needs," which are driven by assumptions regarding household formation rates by demographic age group, residential vacancy rates, and housing unit destruction rates. Some projections by those analysts suggested rapid recovery of housing starts to prerecession levels or higher. Ince et al. (2011) projected that housing starts would increase from 2010 levels to approximately 1.3 million in 2020 and to approximately 2.3 million by 2060 in one scenario. Using the parsimonious single-family model developed here, this projected change in starts would require real annual GDP growth of 5.5% in the 2020s, rising to 11.5% by 2060, rates that are far beyond recent historical rates of GDP growth. None of the other forecasts exceed the long-run average of 1.45 million starts per year. Our empirical models of housing starts provide a mechanism for projecting future housing starts that would be consistent with their historically stable relationship with GDP growth and mortgage delinquency rates, and they provide a counterpoint to needs-based projections of housing in the United States.

Our statistical model estimates indicated that, after accounting for real GDP, changes in population in the United States do not provide additional explanation of variability in housing starts. We tested an alternative statistical model for total housing starts, specified identically as Model 7 shown in Table 3. This model substituted real GDP per capita instead of real GDP and it included population. As in Models 1 and 2 shown in Table 3, real population changes were not statistically significant.

Conclusions

Making long-term projections of possible futures entails conditioning those futures on plausible evolution of the forest sector and the US and world economies. Gordon (2016) outlines a future of slower US growth than previously observed in the United States. Such a vision is consistent with projections to 2100 for the United States (and other wealthy countries) provided by the International Institute for Applied Systems Analysis (IIASA) and the National Center for Atmospheric Research (IIASA 2016). For the United States, under the five Shared Socioeconomic Pathways (SSP)'s projections of real US GDP done by IIASA, growth projections for 2010–70 range from 1.04% to 2.62% per year, with a mean of 1.76%. Therefore, analysts using the SSP-based projections of the US forest sector might expect starts levels, based on our research (Figure 5), at median levels of less than 1.3 million per year. Likewise, assuming our lumber consumption projections are representative of the future at this GDP growth rate, softwood lumber consumption growth is likely to remain low (i.e., $<$1% per year), on average, into the foreseeable future.

The conclusions we have reached regarding the future evolution of starts and, in our example, softwood lumber consumption critically depend on assumptions of unbiased parameter estimates (and hence unbiased projections). Our specifications of softwood lumber consumption quantity omitted information about softwood lumber import prices and the prices of some lumber production inputs.

They also omitted information about the state of lumber demand in Canada, the primary import source. (An alternative specification that included Canada's housing starts found that changes in Canadian starts did not, at 10% significance or stronger, affect consumption of lumber in the United States.) These omissions in the reduced-form statistical models could have led to upward biases in the effects of GDP growth on consumption. We omitted this information to simplify the projection modeling and to demonstrate approximately how alternative housing and economic growth futures would translate into lumber sector outcomes. To the extent that the statistical models that we used contained omitted variables biases, our projections of the softwood lumber consumption futures should be understood as preliminary and demonstrative of what could be done.

Our study uses changes in GDP to represent changes in economic growth, but using GDP represents two limitations of this study. First, housing investment, of which new housing construction is a part, is a component of GDP. Thus, there is some potential endogeneity because a change in GDP is the sum of changes in housing starts and other changes in the components of GDP. Our estimated GDP growth model included lagged changes in housing starts, and these changes were positively related to GDP growth, revealing one way that growth is connected to the sector. Although housing investment is only 6% of GDP, a reduction of this investment by 50% could comprise a notable component of the change in GDP. Therefore, a suggested line of future inquiry might be to estimate housing starts models in which contemporaneous GDP changes are considered endogenous, not exogenous, as we have done. Second, our model for projecting GDP explains only 45% of the variation in the change in GDP, implying that there might be additional factors influencing GDP that we did not model but could have; exploration of those additional factors is another line of research that could aid in producing more accurate projections.

Although not directly comparable because of differences in scenario assumptions, we note that our housing starts projections are somewhat lower than projections derived from the housing needs formulation based on demographic changes. The needs formulation is not independent of economics, but if the projected needs were actually higher than the equilibrium quantity of new housing, then other changes would occur to equilibrate housing markets, including perhaps a reduction in the destruction rate of housing, a decline in the vacancy rates, or subdivision of existing structures into multiple smaller units. Therefore, one topic of future research could be to explore how needs-based formulations could include these additional variables. Another approach would be to investigate how to include explicit representations of demographic variables in aggregate econometric representations of housing starts at the national level or at finer spatial scales than modeled in our study.

Literature Cited

AGNELLO, L., AND L. SCHUKNECHT. 2011. Booms and busts in housing markets: Determinants and implications. *J. Hous. Econ.* 20(3):171–190. doi:10.1016/j.jhe.2011.04.001.

ANUNDSEN, A.K. 2015. Econometric regime shifts and the US subprime bubble. *J. Appl. Econ.* 30(1):145–169. doi:10.1002/jae.2367.

APA. 2010. *Structural panel & engineered wood yearbook*, APA Economics Report E176. APA, Tacoma, WA.

BALL, M., G. MEEN, AND C. NYGAARD. 2010. Housing supply price elasticities revisited: Evidence from international, national, local and company data. *J. Hous. Econ.* 19(4):255–268. doi:10.1016/j.jhe.2010.09.004.

BLACKLEY, D.M. 1999. The long-run elasticity of new housing supply in the United States: Empirical evidence for 1950–1994. *J. Real Estate Finan. Econ.* 18(1):25–42. doi:10.1023/A:1007781228328.

BUONGIORNO, J. 2014. Global modeling to predict timber production and prices: The GFPM approach. *Forestry* 88:291–303. doi:10.1093/forestry/cpu047.

BUONGIORNO, J. 2015. Modeling some long-term implications of CO_2 fertilization for global forests and forest industries. *For. Ecosys.* 2:29. doi:10.1186/s40663-015-0054-3.

CANER, M., AND L. KILIAN. 2001. Size distortions of tests of the null hypothesis of stationarity: Evidence and implications for the PPP debate. *J. Int. Money Financ.* 20:639–657. doi:10.1016/S0261-5606(01)00011-0.

CHOW, G.C. 1960. Tests of equality between sets of coefficients in two linear regressions. *Econometrica* 28(3):591–605. Available online at www.jstor.org/stable/1910133.

DE LEEUW, F. 1971. The demand for housing: A review of cross-section evidence. *Rev. Econ. Stat.* 53(1):1–10. doi:10.2307/1925374.

DICKEY, D.A., AND W.A. FULLER. 1979. Distribution of the estimators for autoregressive time series with a unit root. *J. Am. Stat. Assoc.* 74(366): 427–431. doi:10.1080/01621459.1979.10482531.

DURBIN, J. 1970. Testing for serial correlation in least-squares regression when some of the regressors are lagged dependent variables. *Econometrica* 38(3):410–421. doi:0012-9682(197005)38:3<410:TFSCIL>2.0.CO;2-0.

FANNIE MAE. 2016. *Housing forecast, October 2016.* Available online at www.fanniemae.com/resources/file/research/emma/pdf/Housing_Forecast_101816.pdf; last accessed Dec. 14, 2016.

FEDERAL RESERVE BOARD. 2016. *The July 2016 senior loan officer opinion survey on bank lending practices.* Available online at www.federalreserve.gov/boarddocs/snloansurvey/; last accessed Sept. 21, 2016.

FREDDIE MAC. 2016. *Economic and housing research outlook: October 2016.* Available online at www.freddiemac.com/finance/pdf/201610-Outlook-10%2020%2016.pdf; last accessed Dec. 9, 2016.

GAUCH, H.G., JR. 1988. Model selection and validation for yield trials with interaction. *Biometrics* 44(3):705–715. doi:10.2307/2531585.

GLAESER, E.L., J. GYOURKO, AND A. SAIZ. 2008. Housing supply and housing bubbles. *J. Urban Econ.* 64(2):198–217. doi:10.1016/j.jue.2008.07.007.

GOODMAN, A.C., AND T.G. THIBODEAU. 2008. Where are the speculative bubbles in US housing markets? *J. Hous. Econ.* 17(2):117–137. doi:10.1016/j.jhe.2007.12.001.

GORDON, R.J. 2016. *The rise and fall of American growth: The U.S. standard of living since the Civil War.* Princeton Univ. Press, Princeton, NJ. 784 p.

HOWARD, J.L., AND K.C. JONES. 2016. *U.S. timber production, trade, consumption, and price statistics, 1965–2013.* USDA Forest Service, Res. Pap. FPL-RP-679, Forest Products Laboratory, Madison, WI. 91 p.

INCE, P.J., A.D. KRAMP, K.E. SKOG, H.N. SPELTER, AND D.N. WEAR. 2011. *U.S. Forest Products Module: A technical document supporting the Forest Service 2010 RPA Assessment.* USDA Forest Service, Res. Pap. FPL-RP-662, Forest Products Laboratory, Madison, WI. 61 p.

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS. 2016. *SSP public database, Version 1.1.* Available online at tntcat.iiasa.ac.at/SspDb; last accessed Jan. 23, 2017.

KWIATKOWSKI, D., P.C.B. PHILLIPS, P. SCHMIDT, AND Y. SHIN. 1992. Testing the null hypothesis of stationarity against the alternative of a unit root. *J. Econometrics* 54(1–3):159–178. doi:10.1016/0304-4076(92)90104-Y.

LEAMER, E.E. 2007. *Housing is the business cycle.* NBER Working Paper No. 13428. Available online at www.nber.org/papers/w13428; last accessed Dec. 9, 2016.

MANKIW, N.G., AND D.N. WEIL. 1989. The baby boom, the baby bust, and the housing market. *Reg. Sci. Urban Econ.* 19(2):235–258. doi:10.1016/0166-0462(89)90005-7.

Moench, E., and S. Ng. 2011. A hierarchical factor analysis of U.S. housing market dynamics. *Economet. J.* 14(1):C1–C24. doi:10.1111/j.1368-423X.2010.00319.x.

Mortgage Bankers Association. 2015. *National delinquency survey.* Data obtained by special request on Dec. 1, 2015.

Müller, U.K. 2005. Size and power of tests of stationarity in highly autocorrelated time series. *J. Econometrics* 128(2):195–213. doi:10.1016/j.jeconom.2004.08.012.

Myers, D. 2016. Peak Millennials: Three reinforcing cycles that amplify the rise and fall of urban concentration by millennials. *J. Hous. Pol. Debate* 26(6):928–947. doi:10.1080/10511482.2016.1165722.

National Association of Home Builders. 2017. *Characteristics of home buyers.* Special Studies, July 1, 2015 (Updated November 3, 2015). Available online at www.nahbclassic.org/generic.aspx?sectionID=734&genericContentID=246591&channelID=311; last accessed Jan. 23, 2017.

Phillips, P.C.B., and P. Perron. 1988. Testing for a unit root in time series regression. *Biometrika* 75(2):335–346. doi:10.1093/biomet/75.2.335.

Pitkin, J., and D. Meyers. 2008. *U.S. housing trends: Generational changes and the outlook to 2050. Special Report 298: Driving and the built environment: The effects of compact development on motorized travel, energy use, and CO_2 emissions.* Paper prepared for the Committee on the Relationships Among Development Patterns, Vehicle Miles Traveled, and Energy Consumption. Transportation Research Board and the Division on Engineering and Physical Sciences, Washington, DC. 30 p.

Random Lengths. 2016. The 2016 Yearbook. Random Lengths, Portland, OR.

Said, S.E., and D.A. Dickey. 1984. Testing for unit roots in autoregressive moving average models of unknown order. *Biometrika* 71(3):599–607. doi:10.1093/biomet/71.3.599.

Saiz, A. 2010. The geographic determinants of housing supply. *Q. J. Econ.* 125(3):1253–1296. doi:10.1162/qjec.2010.125.3.1253.

Siskind, D. 1979. *Revised housing starts, 1910–1958.* Draft report. Available online at www.michaelcarliner.com/files/Data/CB82Hous_Start_Revised_History1945-82.pdf; last accessed Oct. 2, 2014.

Song, N., S.J. Chang, and F.X. Aguilar. 2011. U.S. softwood lumber demand and supply estimation using cointegration in dynamic equations. *J. For. Econ.* 17(1):19–33. doi:10.1016/j.jfe.2010.07.002.

Tax Foundation. 2015. *U.S. Federal individual income tax rates history, 1862–2013 (nominal and inflation-adjusted brackets*). Available online at taxfoundation.org/article/us-federal-individual-income-tax-rates-history-1913-2013-nominal-and-inflation-adjusted-brackets; last accessed Dec. 15, 2015.

US Bureau of the Census. 2014. Population and housing unit estimates datasets. Available online at www.census.gov/programs-surveys/popest/data/data-sets.html; last accessed Sep. 6, 2017.

US Bureau of the Census. 2016a. *Housing units started: United States.* Available online at www.census.gov/econ/currentdata/dbsearch?program=RESCONST&startYear=1959&endYear=2014&categories=STARTS&dataType=TOTAL&geoLevel=US&notAdjusted=1&submit=GET+DATA; last accessed Jan. 30, 2016.

US Bureau of the Census. 2016b. *Expenditures for residential improvements and repairs, seasonally adjusted annual rate in millions of dollars.* Available online at www.census.gov/construction/c50/histtab1.pdf; last accessed June 19, 2017.

US Bureau of the Census. 2016c. *National population totals datasets: 2010–2016.* Available online at www2.census.gov/programs-surveys/popest/datasets/2010-2015/national/totals/; last accessed Sep. 5, 2017.

US Bureau of the Census. 2017a. *How the data are collected (building permits survey).* Available online at www.census.gov/construction/nrc/how_the_data_are_collected/bps.html; last accessed Jan. 17, 2017.

US Bureau of the Census. 2017b. *Why does BLS provide both the CPI-W and CPI-U?* Available online at www.bls.gov/opub/btn/volume-3/why-does-bls-provide-both-the-cpi-w-and-cpi-u.htm; last accessed Jan. 17, 2017.

US Bureau of Economic Analysis. 2016. *Current-dollar and "Real" gross domestic product.* Available online at www.bea.gov/national/xls/gdplev.xls; last accessed Jan. 29, 2016.

US Bureau of Labor Statistics. 2015a. *Prices-producer, commodity data including "headline" FD-ID indexes.* Available online at www.bls.gov/data/; last accessed Nov. 6, 2015.

US Bureau of Labor Statistics. 2015b. *Average weekly earnings of production and nonsupervisory employees, not seasonally adjusted. Employment, hours, and earnings from the Current Employment Statistics Survey (National).* Available online at www.bls.gov/data/; last accessed Nov. 6, 2015.

US Bureau of Labor Statistics. 2015c. *Labor force statistics from the Current Population Survey: Unemployment rate (16 years and over).* Available online at data.bls.gov/timeseries/LNS14000000; last accessed June 24, 2015.

US Bureau of Labor Statistics. 2016. *Prices-producer, commodity data including "headline" FD-ID indexes.* Available online at www.bls.gov/data/; last accessed Sept. 6, 2016.

US Federal Reserve Bank of St. Louis. 2015. *Series CSUSHPINSA.* Available online at research.stlouisfed.org/fred2; last accessed June 19, 2015.

Vojtech, C.M., B.S. Kay, and J.C. Driscoll. 2016. *The real consequences of bank mortgage lending standards*. Federal Reserve Board Office of Financial Research Working Paper 16–05. Available online at www.financialresearch.gov/working-papers/files/OFRwp-2016-05_Real-Consequences-of-Bank-Mortgage-Lending-Standards.pdf; last accessed Sept. 21, 2016.

White, H. 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroscedasticity. *Econometrica* 48(4):817–838. doi:10.2307/1912934.