NAHB Indices for Multifamily Production and Vacancies

This document explains the methodology underlying NAHB's two summary multifamily indices—the **Multifamily Production Index (MPI)** and the **Multifamily Vacancy Index (MVI)—**presented the formulas on which they are based, and discussed in general terms how they track stand measures of multifamily activity produced by the federal government. The appendix explains how the formulas where derived through a process of analyzing Census data and responses to NAHB's quarterly Multifamily Market Index survey and Census data.

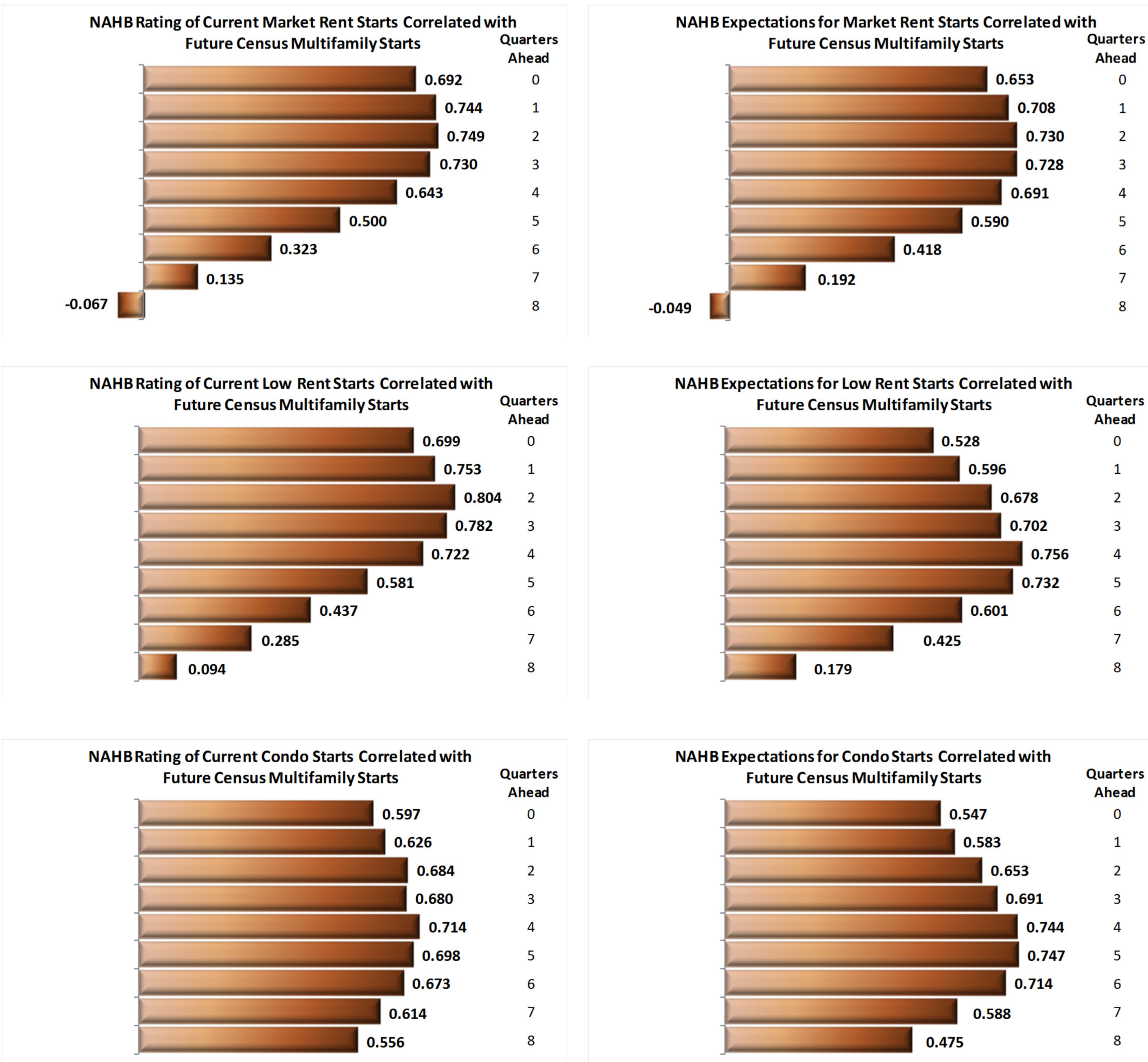
This derivation was accomplished into two stages: 1) choosing the time frame to analyze 2) constructing index formulas that performed well (in a sense described below) within that frame.

Choosing the Time Frame

The basic premise is that responses to the MMI survey should have some ability to predict standard measures of multifamily activity. If that is so, the first question to ask is over what time frame.

The history of the MMI survey is now long enough to make it possible to investigate this premise. Although the series are still not long enough to make all time series procedures feasible, it is possible to compute the correlation (a statistic that ranges in absolute value from 0 to 1—where 1 indicates a perfect relationship between two variables and 0 indicates no relationship at all) between various MMI indices and current and future values of Census starts and vacancy measures. Figure A1 shows how component NAHB multifamily production indices correlate with the Census multifamily starts rate—currently and out to eight quarters in the future.

Figure A1
Correlations Between NAHB Multifamily Market Indices and Future Starts



A general pattern to all the charts in Figure A1 is that the contemporary correlations (i.e., between the NAHB index and the census multifamily starts rate 0 quarters into the future) are relatively modest, but increase as starts are considered a few quarters into the future. In other words, the NAHB indices tend to predict future starts better than they explain starts on a contemporary basis.

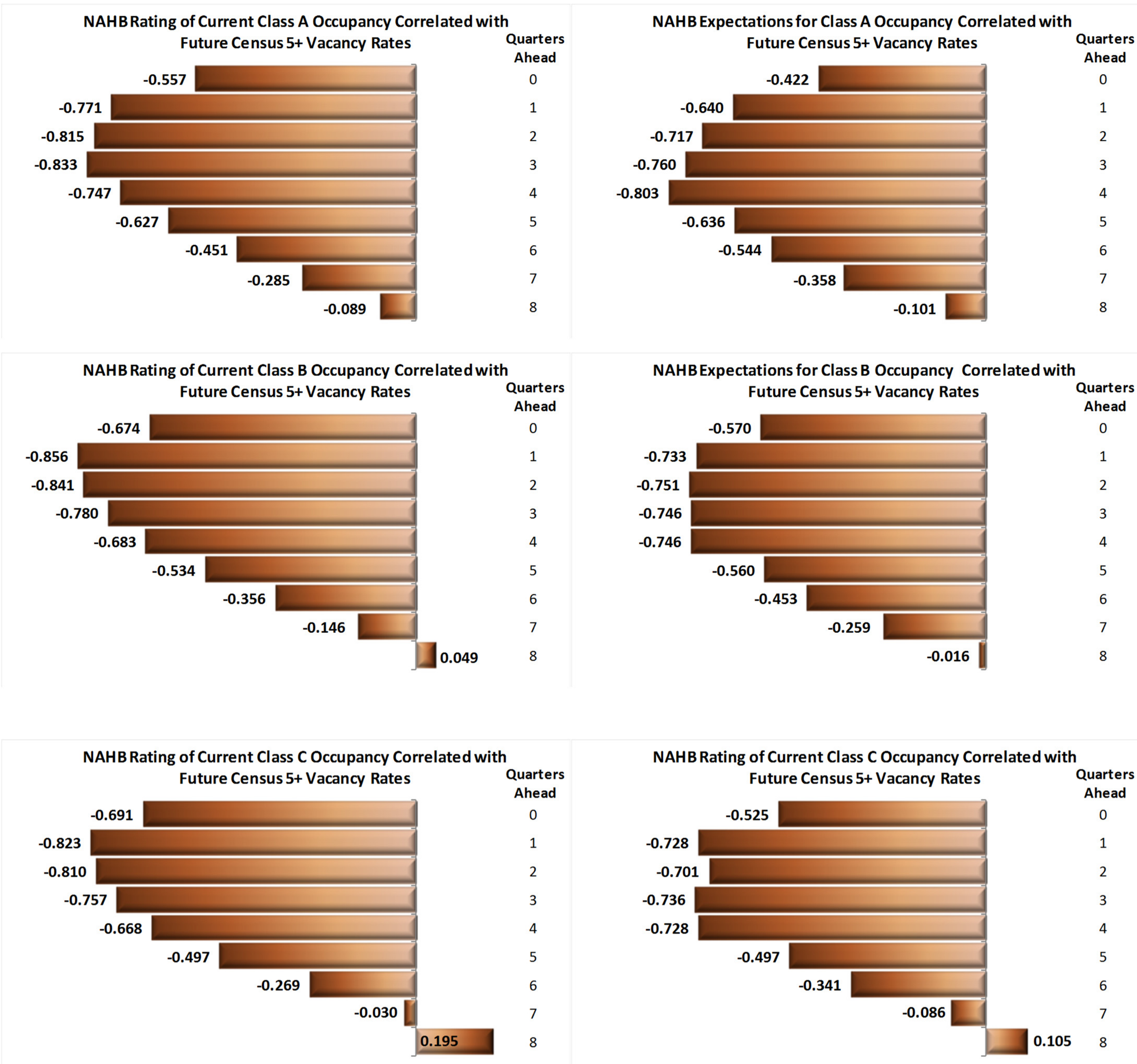
For NAHB member ratings of current market rate or low rent production, the correlations are highest—over .73—with starts one to three quarters in the future, with a peak of .80 for the low rent starts index with census starts two quarters into the future. After the third or fourth quarter, the correlations decay fairly rapidly and are close to zero eight quarters into the future.

The NAHB index based on expectations for market rent production performs roughly the same as the index based on ratings of current market rent production, except that each of the expectations-based correlations is somewhat lower than its current-conditions counterpart. For the NAHB index based on expectations for low rent starts, the correlations are not only generally lower than the current-condition equivalents, but the peak is shifted further into the future, with correlations in excess of .70 occurring with census multifamily starts three to five quarters in the future.

Correlations of NAHB condo indices with census multifamily starts do not show the same pattern and are below .7 up to three quarters into the future. The highest correlations occur between the NAHB condo expectations index and census multifamily starts four and five quarters ahead.

Similarly, Figure A2 shows how the NAHB multifamily rental occupancy indices correlate with the Census vacancy rate in buildings with five or more apartments. Again, the charts show the NAHB index correlated with the Census measure currently and out to eight quarters in the future.

Figure A2
Correlations Between NAHB Multifamily Market Indices and Future Vacancy Rates



These correlations are generally negative, because they are comparing an *occupancy* index to a Census measure of *vacancy*, so the expected relationship is, in fact, negative (i.e., when one measure increases, the other tends to decline).

As was the case for most of the production indices, the NAHB occupancy indices show a relationship with the relevant Census measure that is relatively modest concurrently, but becomes stronger as the Census measure in considered a few quarters into the future, then decays to near zero by the eighth quarter ahead. This general pattern is prevalent across all six graphs shown in Figure A2.

The three current-condition indices show a particularly strong relationship with Census five-plus vacancies one to three quarters ahead. These correlations are (in absolute value) uniformly above .75, and above .80 in two of the three quarters in each graph.

The expectations indices produce correlation patterns that are similar, but not quite as strong, and with peaks shifted slightly into the future. In one case the correlation has not yet reached .70 (in absolute value) with the Census five-plus vacancy rate one quarter ahead, and in all three cases it remains above .70 with the vacancy rate four quarters ahead.

In summary figures A1 and A2 show that contemporaneous correlations between NAHB indices and the relevant Census measure are relatively modest. "Current Condition" MMIs produce relatively strong correlations of at least .7 (in absolute value), sometimes .8, with the only exception being slightly lower correlations for the condo production index. "Expectations" MMIs produce correlations with Census measures that are somewhat weaker and peak somewhat further into the future.

Constructing Index Formulas

Based on the above results, NAHB decided to consider two summary MMIs, based on correlations with multifamily starts and five-plus vacancy rates 1 to 3 quarters into the future. Use of this time frame will tend to give smaller weight to the MMI expectations components, which tend to be correlated with the Census measures at longer lags. However, series that correlate well at shorter lags are preferable, especially for the graphical presentations that NAHB anticipates will constitute the primary use of the summary indices. In addition, given the amount of space available to discuss the MMI in a typical press release, construction of more than two summary indices—for example, two based on current conditions plus two additional based on expectations components—was considered impractical.

In order to derive specific formulas for the summary indices, four-equation latent variable models of the following form were considered:

$$Y_{t+1} = b_1X + e_1$$

$$Y_{t+2} = b_2X + e_2$$

$$Y_{t+3} = b_3X + e_3$$

$$X = \sum w_iM_i$$

where Y_{t+1} is a Census measure of multifamily activity one quarter into the future, X is the summary index entering the model as a latent variable, and M_i are component indices from the NAHB MMI survey. e1, e2, and e3 are error terms. b1, b2, b3, and w_i are coefficients to be estimated. The w_i are weights that define the index, estimated subject to the constraint that $\sum w_i = 1$.

In other words, each model estimates a summary index as a weighted sum of component indices, the weights estimated based on their correlation with the Census measure one to three quarters into the future, all three considered simultaneously. The CALIS procedure in SAS was used to estimate each model.

Results for five models are shown in Figure A3. The models differ in terms of the Census measure used as an independent variable to be predicted by the index, and the particular NAHB MMI components used to define the index. MMI components based on the first or "starts" part of the questionnaire are used when the dependent variable is the Census measure of multifamily starts. MMI components based on the "rental occupancy" are used when the dependent variable is the Census five-plus rental vacancy rate.

Figure A3
Estimates of Latent Variable Models Defining Summary MMIs

	Model I 3 variable production index	Model 2 4 variable production index	Model 2 6 variable production index	Model 4 3 variable vacancy index	Model 5 6 variable vacancy index
Dependent Variable	Multifamily Starts	Multifamily Starts	Multifamily Starts	5+ Vacancy Rate	5+ Vacancy Rate
Manifest Variable Equations: Coefficient on Latent Variable Index with Standard Error					
Dependent Variable 1 Quarter Ahead	0.9285	0.9274	0.9274	-0.8537	-0.8537
	(0.1340)	(0.1356)	(0.1356)	(0.1090)	(0.1090)
Dependent Variable 2 Quarters Ahead	0.9899	0.9936	0.9936	-0.8701	-0.8701
	(0.1160)	(0.1163)	(0.1163)	(0.1039)	(0.1039)
Dependent Variable 3 Quarters Ahead	1.0445	1.0522	1.0522	-0.8662	-0.8662
	(0.0959)	(0.0945)	(0.0945)	(0.1051)	(0.1051)
Latent Variable Equation: Definition of the Index as a Function of NAHB MMI Components					
Current Rating of Starts - Low Rent	0.3533	0.3572	0.3572		
Current Rating of Starts - Market Rent	0.2806	0.2687	0.2687		
Current Rating of Starts - For Sale	0.3661	0.2398	0.2398		
Expectations Next 6 mo - Low Rent			1.72 E-17		
Expectations Next 6 mo - Market Rent			9.92 E-18		
Expectations Next 6 mo - For Sale		0.1342	0.1342		
Current Rating of Occupancy - Class A				0.3632	0.3632
Current Rating of Occupancy - Class B				0.4630	0.4630
Current Rating of Occupancy - Class C				0.1738	0.1738
Expectations Next 6 mo - Class A					-3.47 E-18
Expectations Next 6 mo - Class B					7.45 E-18
Expectations Next 6 mo - Class C					3.70 E-19
Akaike Information Criterion	57.72	56.75	55.27	-4.22	-11.52

Analysis of candidate vacancy indices based on the results in Figure A3 is comparatively straightforward, as models 4 and 5 are effectively identical. In the 6 variable model, the weights on the current condition variables are the same as in the 3 variable model to four decimal places, and the weights on the expectations components are zero to more than 15 decimal places.

Moreover, when estimating the 6 variable model, it was necessary to constrain the weights to be greater or equal to zero; otherwise, they turned out to be negative on some of the expectations components. In the 3 variable model, the only constraint imposed was the basic requirement that the weights to sum to 1.0, and this constraint had only a small rescaling effect on the final weights.

All of this argues for an **MVI** defined as a weighted sum of the 3 "current condition" components, with the weights as shown in Model 4 in Figure A3.

Results for the production index models (1, 2, and 3 in Figure A3) were similar, but not identical. Again, in the 3 variable model, the only constraint that had to be imposed was the basic requirement that weights to sum to 1.0, and imposition of this constraint had only a small effect on the final weights.

However, in the 6 variable model, estimated weights were effectively zero for only two of the three expectations components. This suggested the possibility of a 4 variable model, based on all three current condition and one expectation component (for condo production), which is shown as Model 2 in A3. From a practical standpoint, Models 2 and 3 are identical. In the 6 variable model, the weights on the current condition variables and one expectation variable that appears in both models are the same to four decimal places, and estimated the weights on the two additional expectations components that appear in the 6 variable model are zero to more than 15 decimal places. Again, it was necessary to add additional constraints to the 6 variable model to avoid producing negative weights on some of the expectations components.

On balance, NAHB economists chose Model 1 over Model 2, primarily because consistency between the formulas for the **MVI** and **MPI** seemed desirable for public presentations, and because there was little practical difference between the indices produced by these two models.

Beyond estimating coefficients of the models and their standard errors, the CALIS procedure computes a large number of goodness of fit statistics. The most commonly cited one, the Akaike Information Criterion (AIC), is included in Figure A3. The AIC is intended as a tool for choosing between models, where the model that produces the smallest AIC should be preferred.

It is interesting to note that the AIC (as well as most of the others computed by CALIS, including those designed to penalize models with superfluous parameters more heavily, such as Schwarz's Bayesian Criterion or Mulaik's Parsimonious Goodness of Fit Index), consistently prefer the model with a greater number of variables—even though these models may employ additional constraints that are strongly binding and assign negligible coefficients to some of the variables. This perhaps raises a question about reliance on standard goodness of fit statistics. Final choice of formulas for the **MPI** and **MVI** were based on inspection of the estimated results and rejection of those that assigned effectively zero weights to some index components, rather than an evaluation of particular goodness of fit criteria.