Tail Model Risk in a Stress Scenario for House Prices

Camilo Sarmiento*

Guidehouse Inc, Washington DC

*Corresponding Authors: *Camilo Sarmiento, Guidehouse Inc, Washington DC, USA*

Abstract: *This paper incorporates uncertainties of model risk in a stress scenario for house prices. Our approach consists of mapping the Gaussian (or other alternative) distribution quantiles to the quantiles of the empirical distribution using a statistical criterion. The mapping corrects for the presence of fatter tails (relative to the Gaussian distribution) into the re-parameterized distribution. The focus of the analysis is at the tail of the distribution, which is associated with uncertain adverse events. The application of the model to house prices indicates the relevance of our mapping in terms of capturing the home price declines observed in the Great Recession. The derivation of confidence intervals, furthermore, provides buffers for model risk. For example, while regulatory capital may use the point estimate of the quantile of the distribution, risk management organizations may use this confidence intervals to bolster resilience of financial institutions to home price declines.*

Keywords: *House Prices, Model Risk, Fat Tails, Empirical Distribution, Stress Testing*

1. INTRODUCTION

Stress testing provides information on the stability of a system or entity for risk management purposes. Regulators and senior management of major US banks rely on stress testing to guarantee that the bank holding company has sufficient capital to continue operations throughout times of economic and financial stress. For example, the Comprehensive Capital Analysis and Review is an annual exercise by the Federal Reserve to assess whether the largest bank holding companies have sufficient capital to continue operations under stressful economic conditions (see Federal Reserve, 2020).

Scenario forecasting of adverse events is the main analytical foundation of stress testing. Statistically, the modeling of uncertain adverse events relates to capturing a tail of the distribution event. In contrast to the prediction of expected events, the prediction of adverse events could be very sensitive to distributional assumptions, e.g., the Gaussian (Normal) assumption.

Indeed, it is to be expected that the distributional error uncertainty is a main source of estimation error of capital estimates in stress testing. The aim of this paper is to incorporate uncertainties of model risk in stress scenarios.

Our approach consists of mapping the Gaussian (or other alternative) distribution quantiles to the quantiles of the empirical distribution using a statistical criterion: the mapping is implemented if the transformation factors are statistically significant and increase a penalized goodness of fit measure. The mapping corrects for the presence of fatter tails (relative to the Gaussian distribution) into the re-parameterized distribution.

Our mapping also provides a statistical measure or confidence intervals to stress testing. The confidence intervals that pertain to the tail of the distribution extrapolation can then be conciliated with the quantile estimate of the distribution to generate a more comprehensive measure of risk. Importantly, we show the generation of a stress scenario based on the joint distribution of the model risk of tail risk extrapolation and the adverse condition assumptions.

The focus of our proposed method consists of a stress scenario for house prices.[1] Specifically, we start with the data generating process of house prices and then derive a stress scenario. The testing results illustrate the relevance of fatter tails in house price modeling and the effect of model risk on tail risk extrapolation in an adverse scenario.

2. MODELING THE HOME PRICE DISTRIBUTION

Stress testing of home prices is a main factor in regulators evaluation of banks' balance sheets. While home price declines at a national level were considered a low probability event prior to the great recession, it is now considered a plausible adverse event in most stress testing across the financial service industry. The underlying empirical basis for a stress scenario for house prices stems from the data generating process (stochastic differential equation) underlying home prices.

To model home price dynamics, we start with an Ornsten-Uhlenbeck process that embeds the family of stochastic processes with a mean reversion component. The Ornsten-Uhlenbeck process have been successfully adapted from physics to the modeling of financial series (e.g., Vasicek, 1977). More so, the Ornsten-Uhlenbeck process is the most often referenced stochastic differential equation in finance.

The Ornsten-Uhlenbeck process, specifically, models the velocity of a massive Brownian particle under the influence of friction:

$$(1) \quad HP_t = b(HP_{t-1} - c) + \sigma h_t$$

where the long-run mean level of home prices, HP_t, is c; the speed of reversion to the mean is b; and σ is the instantaneous volatility. The underlying structure, in particular, underscores the tendency of the dynamic process to drift toward its long-term mean such that the velocity of the particle tends to mean revert.

A Weiner process with drift is an Ornsten-Uhlenbeck process with the added restriction that b =1:

$$(2) \quad HP_t = HP_{t-1} - c + \sigma h_t.$$

For the modeling of house price dynamics, an extended Ornsten-Uhlenbeck process that incorporates the presence of fat tails (e.g., a lognormal distribution) is:

$$(3) \quad \ln(HP_t/HP_{t-1}) = -bc + (b-1)ln(HP_{t-1}) + \sum_{s=1}^{t} a_s \ln(HP_{t-s}/HP_{t-s-1}) + \sigma h_t$$

where $h_t \sim N(0,1)$, and $\sum_{s=1}^{t} a_s \ln(HP_{t-s}/HP_{t-s-1})$ are short term deviations from the long-term trend.

As in (1), the dynamic process in (3) exhibits a stochastic trend if b = 1(a Weiner process with drift).

3. SCENARIO FORECASTS FOR HOME PRICES: ADVERSE SCENARIO

Estimation of equation 3 uses annual data from the post-War period (1947 to 2018). The use of post-War data is consistent with Federal Reserve Economic Data (FRED) reporting periods for most economic series. The data source is Robert J. Shiller, Irrational Exuberance, 3rd. Edition, Princeton University Press, 2015 (http://www.econ.yale.edu/~shiller/data.htm).

Selection of lag length in (3) uses penalized goodness of fit criteria (AIC, BIC). Specifically, the application of such criteria to (3) yields the following more parsimonious structure:

$$(4) \quad (HP_t/HP_{t-1}) = d + (b-1) \ln(HP_{t-1}) + \sigma h_t.$$

Maximum likelihood estimation of equation 4 yields the coefficient estimates in Table 1.

Table1. *Maximum Likelihood Estimation*

Coefficient	Estimate	Standard Error	t Stat	P-value
D	0.042	0.028	1.544	0.128
B	0.999	0.007	144.77	<0.001

To determine the final model structure, we implement the Augmented Dickey Fuller.[2]The results of the test cannot reject the null hypothesis that d> 0 and b = 1. Therefore, the selected model

---

[1]More than any other sector, housing is the most visible and relevant asset value. Housing is the asset that has the largest impact on household wealth (Case, Quigley, and Shiller, 2001) and housing is a leading indicator of economic activity (Stock and Watson, 1989). As housing goes down, so goes the economy. Lastly, a cyclical period of home price declines is associated with macro instability and a spike on loan defaults (Holt, 2009; Sarmiento, 2012).

[2]The Augmented Dickey-Fuller test is a widely used to test for the presence of a stochastic trend.

representation for year-over-year change in home prices is a stochastic process (unit root with drift), and the final model structure is:

(5) $ln(HP_t/HP_{t-1}) = d + \sigma h_t = .042 + 0.048 h_t.$

Scenario forecasting uses the parameter estimates in (5) to project a probabilistic path of home prices over a specific (one year) time horizon. Different from the projection of the conditional mean, however, scenario forecasting focuses on a quantile of the distribution.

For a one-year ahead forecast, Figure 1 shows a scenario forecasting for house prices that stems from (5). The horizontal axis represents the quantile of the distribution, and the vertical axis reports the expected change in home prices for the respective quantile. For example, for the 50th quantile, the change in home price expected for the next 12months is 4 percent.

From the Figure, Table 2extracts estimates of house prices under adverse scenarios. For example, under the 96th quantile (1-in-25 event), home prices are expected to decline 4.6 percent in the next 12months.National home price declines are, therefore, a plausible occurrence, and a stress event may involve a significant decline in home prices at the national level.
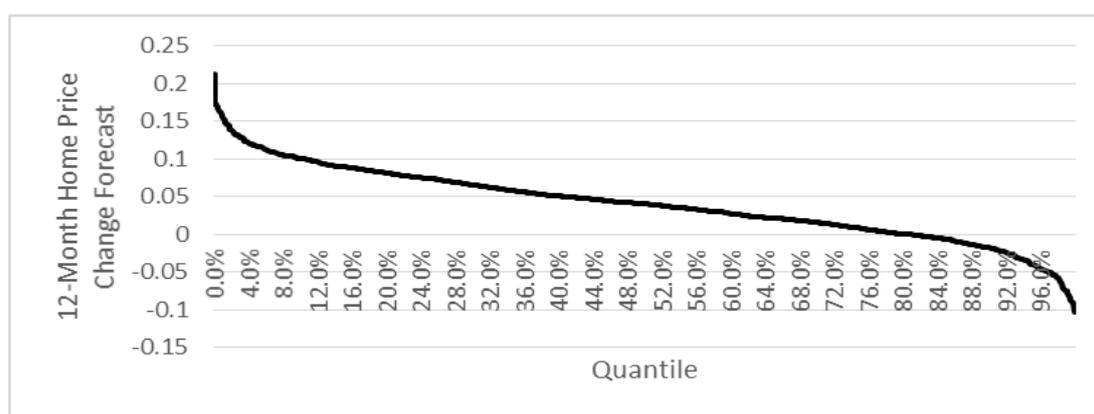

Figure1. *12-Month Scenario Forecast for Home Prices*

Table2. *Adverse Scenario Forecasting*

Adverse Event Odds	1-in-7 Event	1-in-25 Event	1-in-50 Event
HP Growth over 12 Months	-1.0%	-4.6%	-5.8%

4. MODEL RISK ADD-ON TO STRESS TESTING

The characterization of tail risk in Table 2implicitly embeds high uncertainty in regard to whether the Gaussian parameterization of the data captures the actual tail risk, and whether the sampleavailable from historical data is sufficient to infer tail events. For example, there are large differences in the projection of adverse scenarios in Table 2 relative to adverse scenarios that stem from historical simulation (seePérignon and Smith, 2010; Sarmiento, 2020) in Table 3. Historical simulation extracts the empirical distribution of home prices, and the projected adverse scenario in the empirical distribution captures the 9.5 percent declines in home prices observed in 2008 and 2009, while the Gaussian parameterization does not.

Table3. *Adverse Scenario Forecasting Based on the Empirical Distribution*

Adverse Event Odds	1-in-7 Event	1-in-25 Event	1-in-50 Event
HP Growth over 12 Months	0.1%	-3.7%	-9.2%

Still there are no guarantees that the characterization of tail risk in Table 3 is more accurate than the Gaussian assumption in Table 2. We do know, however, that the quantification of home price declines in Table 2 under stress has two main shortcomings. First, it appears to underestimate the likelihood of large home price declines embedded in the data generating process of home prices. Second, Table 2 does not provide information on the reliability or uncertainty embedded in the estimate.

Our proposed method combines the informational value from Tables 2 and 3. In effect, our method corrects for the presence of fatter tails (relative to the Gaussian distribution) into a re-parameterized distribution.

In particular, to create a mapping of the Gaussian (or other alternative) distribution quantiles to the quantiles of the empirical distribution using a statistical criterion, we take the residual h_t of the data generating process in (5) and index the series in terms of error size (rather than time). That is,

$$(6) h_k = h_{f(t)}$$

where k$= f(t)$ if h_t is the kth largest residual error in the sample.

Once we re-indexed the series in terms of size of the residual term, we map the factor in (6) to the standarized normal distribution:

$$(7) h_k = z_k + u_k$$

where $z_k = f^{-1}(N(0,1))$ is the inverse function of the distribution function of the Gaussian distribution.

The mapping in (7) holds only if the underlying distribution in (5) follows a Gaussian distribution; and a more general mapping of the standardized Normal distribution to the empirical distribution is:

$$(8) h_k = b_k + a * z_k + u_k$$

were the introduction of a fatter tail stems from b_k. Specifically, the parameter b_k adds a factor that varies the kurtosis of the distribution and allows for a distribution with fatter tails. The outcome is thus a generalization of the normal distribution fitted to the empirical distribution. The testing of the normal distribution assumption would encompass the condition that $b_k = 0$ and $a = 1$.

Lastly, confidence intervals that pertain to the level of uncertainty can be constructed from the variance associated with the estimate of h_k. That is,

$$(9) \quad Var(h_k) = Var(u_k).$$

The next section shows the estimation methodology of the structure in (8), and the derivation of model risk bounds that uses (9).

5. PARAMETRIZATION OF HOME PRICE DYNAMICS WITH FAT TAILS

The empirical application of (8) combines the informational value results from Tables 2 and 3. In effect, the method corrects for the presence of fatter tails (relative to the Gaussian distribution) into a re-parameterized distribution. Nonetheless, estimation of (8) requires a choice of a functional structure for b_k in terms of k. For maximum flexibility, the rank varying intercept, b_k, is assumed to follow a step function, and the final specification is pinned down using stepwise regression.

The initial model specification uses 20 levels in the step function, and each step is a quantile of the rank order of the observation, i.e., the intercept parameter jumpssequentially across 20 quantiles. For example, observations that determine the intercept in the first quantile correspond to those at the bottom of the distribution, the lower 5th percentile. The second quantile corresponds to those observations that fall between the 5th and 10th percentile of the distribution, and so on for the other quantiles.

Therefore, the econometric specification of (8) in terms of a rank varying intercept parameter is:

$$(10) h_k = b_S(k) + a * z_k + u_k, \; if \; k \in S \; quantile,$$

where $b_S(k) = b_n, \; if \; 0.05\,(n-1)K < k \le 0.05nK$, for K =number of observationsand n $\in\{1,\ldots, 20\}$. For example, $b_S(k) = b_1, if \; k \in \le .05\,K$.

After application of stepwise regression, the model specification in (10) reduces to:

$$(11) h_k = b_{20}(k) + a * z_k + u_k, \text{for k} > 0.95K; \text{and}$$

$$h_k = a * z_k + u_k, \text{otherwise.}$$

Overall, equation (11) encompasses an estimable representation of the rank varying intercept parameter, b_k, in (8). Table 4 presents the parameter estimates of the final specification in (11).

Table4. *Mapping of Normal Distribution to the Empirical Distribution*

Model	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%
Standard Normal	0.925	0.026	35.486	<0.001	0.872	0.977
95th Per-centile Dummy	-0.535	0.104	-5.145	<0.001	-0.743	-0.327
Adjusted R Square = .97						

The results in Table 4 indicate that the Normal distribution does not statistically capture the 95th percentile of the distribution, and the introduction of the mapping in (8) increases model fit, and it is statistically significant at the <.001 percent significant level. Effectively, the mapping reweights the distribution toward the negative tail of the distribution increasing the kurtosis of the distribution.

Figure 2 shows the effect of the reweighting on the model fitat the tail (which in the figure we refer as the modified Normal). The horizontal axis represents the quantile of the distribution, and the vertical axis reports the expected change in home prices for the respective quantile.
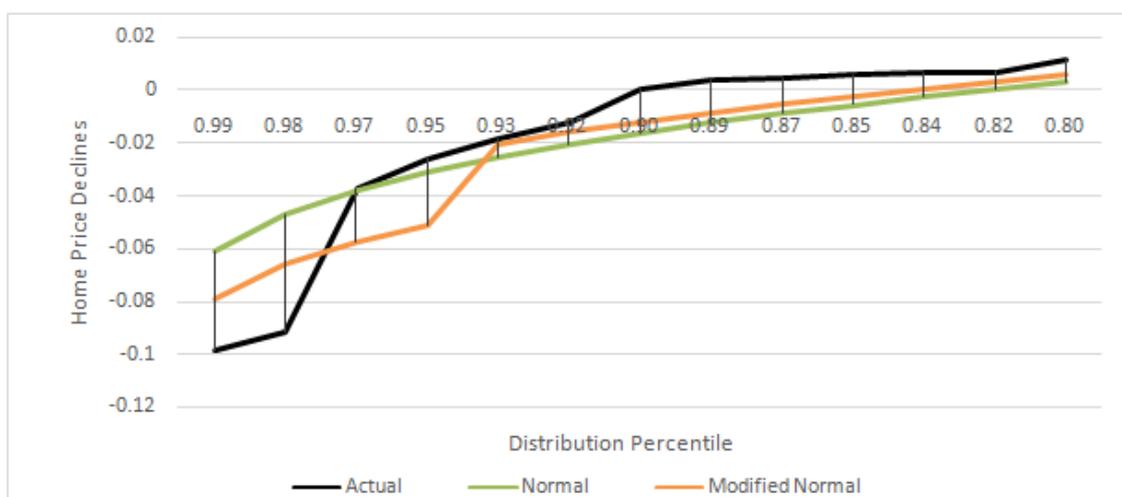

Figure2. *12-Month Scenario Forecast for Home Prices under the Modified Normalversus the Normal Distribution*

Despite the refinement provided in the derivation of stress testing in the figure, it is unclear whether the stress estimate provides a reliable forecast relative to an expected probability event. To provide a sense of the reliability of the forecast, Figure 3 shows the 95thpercent confidence intervals of the forecast of the modified Normal based on the model risk variance in (9).
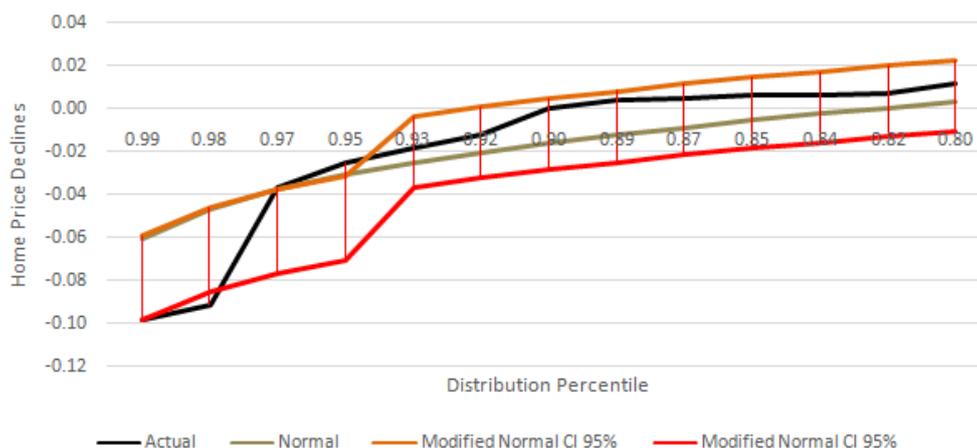

Figure3. *12-Month Scenario Forecast Confidence Intervals for Home Prices under the Modified Normal Distribution*

From Figure 3, the 95th percent confidence intervals of the modified normal projections (derived from Equation 11) embed both the Normal distribution and the empirical distribution.

The confidence intervals in Figure 3also provide buffers for model risk. For example, while regulatory capital may use the point estimate of the quantile of the distribution, risk management organizations may use this confidence intervals to bolster resilience of financial institutions to home price declines. Table 5shows projections of home price declines under adverse scenarios after factoring in a buffer for model risk.

Table5. *Adverse Scenario Forecasting Based on the Modified Normal*

Adverse Event Odds	1-in-7 Event	1-in-25 Event	1-in-50 Event	1-in-75 Event
HP Growth over 12 Months	-0.5%	-5.3%	-6.6%	-7.5%
HP Growth over 12 Months Model Risk Buffers Based on 95% CI	-2.1%	-7.3%	-8.8%	-9.5%

Besides the use of confidence intervals to construct model risk buffers, the model risk can be mapped in terms of confidence intervals that map the inferred quantile from the sample to lower quantiles but with a lower tolerance for model risk. That is,

$$(12) h_k = h_{k-s} + ci_s \, sqrt(Var(u_k))$$

where ci_s is the level of model risk percentile tolerance that makes the adverse scenarios of h_k and h_{k-s} equivalent.

For example, Table 6 maps the stress scenario assumption under the Normal distribution relative to the equivalent risk gradient (stress testing percentile) for home price declines under the modified Normal distributions with and without accounting for the model risk.

Table6. *Mapping the Stress Test Assumption under the Normal to a Distribution with Model Risk Tolerance*

Stress Testing Quantile Assumption under the Normal Distribution	Equivalency Percentile to the Normal Distribution Stress Testing Assumption	
	Modified Normal	Modified Normal Upper 95 CI
0.99	0.97	0.94
0.95	0.94	0.90
0.92	0.93	0.8

Therefore, the definitions of stress scenarios can directly incorporate levels of model risk tolerance.

Overall, the testing results show the while the Normal distribution provides a reasonable representation of the data generating process of home prices on the body of the distribution, however, the Normal distribution differs fromthe empirical distribution at the tail. Our method generates a distribution that conciliates differences between the Normal and the empirical distribution. The method also generates confidence intervals for stress scenarios, and stress testing under different tolerance levels for model risk.

6. GENERALIZATION OF THE METHOD

So far, we have confined the application of the re-parameterization method in (8)to the Gaussian distribution, but it applies to other distributions. In this section, we consider the case of the t-distribution.

The finance literature has explored alternative distributions to the Gaussian distribution (e.g., Akgiray and Booth, 1988; Jansen and De Vries,1991;Pérignon and Smith, 2010), and the most commonly used distribution to capture fat tails in finance is the Student t distribution (see Blattberg and Gonedes, 1974). The use of leptokurtic distribution (such as the t-distribution), however, has had limited application for low frequency applications such as home price modelling. The reason is that the t-distribution fails to capture the shape of the distribution.

Under the home price model in (5), Table 7 shows the results from the calibration of the degrees of freedom of the t-distribution for the residual. From the table, a t-distribution with 20 degrees of freedom minimizes the prediction error of the underlying distribution.

Table7. *Calibration of Degrees of Freedom under the Assumption of t-Distribution*

Degrees of Freedom (DF)	10	15	17	20	22	30	50
Sum of square of the errors	0.044823	0.042673	0.042477	0.042383	0.042384	0.042542	0.042961

The effect to stress testing from the use of a t-distribution rather than a Gaussian distribution is relatively small, and application of the mapping in equation 8 to the t-distribution generates the results in Table 8.

Table8. *Mapping of t-Distribution Distribution to the Empirical Distribution*

Model	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%
Standard Normal	0.90	0.023	35.433	<0.001	0.846	0.945
95th Percentile Dummy	-0.501	0.102	-4.921	<0.001	-0.704	-0.297
Adjusted R Square = .97						

Inspection of Table 8 and a comparison with Table 4 reveals that the mapping in (8) from a parametric distribution to the empirical distribution is quite similar under both the Gaussian and the t-distribution distribution assumption for the residual term in Equation 5.

Overall, the resultant distribution from the application of the mapping in (8) is largely robust to the use of either the Gaussian or a t-distribution and, therefore, the use of a Gaussian distribution as the basis from our mapping for a distribution with fatter tails (and its associated confidence interval) appears to be appropriate for the modeling of stress scenarios for home prices.

7. CONCLUDING REMARKS

This paper incorporated uncertainties of model risk in stress scenarios for house prices. Our approach consisted of mapping the Gaussian (or other alternative) distribution quantiles to the quantiles of the empirical distribution using a statistical criterion. The mapping corrected for the presence of fatter tails (relative to the Gaussian distribution) into the re-parameterized distribution.

Our method also provided measures of model risk uncertainties that stem from both the presence of limited samples for projecting home price dynamics at the tail of the distribution and the uncertainties related to fitting the tail of the distribution to the data.

REFERENCES

[1] Akgiray, V. and G. G. Booth (1988). The Stable-law Model of Stock Returns. *Journal of Business & Economic Statistics*, 6, 51–57.

[2] Blattberg, R.C. and N.J. Gonedes (1974) A Comparison of the Stable and Student Distributions as Statistical Models for Stock Prices. *The Journal of Business*, 47, 244–280.

[3] Case, K. E., Quigley, J.M. AND Shiller, R.J. (2001). Comparing Wealth Effects: The Stock Market versus the Housing Market. *NBER Working Paper* 8606.

[4] Federal Reserve (2020). 2020 Supervisory Scenarios for Annual Stress Tests Required under the Dodd-Frank Act Stress Testing Rules and the Capital Plan Rule.

[5] Holt, J. (2009). A Summary of the Primary Causes of the Housing Bubble and the Resulting Credit Crisis: A Non-Technical Paper. *The Journal of Business Inquiry*, 120-129.

[6] Jansen, D.W. and De Vries C.G. (1991). On the Frequency of Large Stock Returns: Putting Booms and Busts into Perspective. *The Review of Economics and Statistics*, 18–24.

[7] Pérignon, C., & Smith, D. (2010). The level and quality of value-at-Risk disclosure by commercial banks, *Journal of Banking and Finance* 34, 362-377.

[8] Sarmiento, C. (2012). The Role of the Economic Environment on Mortgage Defaults during the Great Recession. *Applied Financial Economics*, 22, 250-258.

[9] Shiller, R.J. (2015). *Irrational Exuberance*, 3rd. Edition, Princeton University Press.

[10] Stock, J. H., and Watson, M.W. (1989). New Indexes of Coincident and Leading Economic Indicators. *NBER Macroeconomics Annual*, 4, 351-409

[11] Vasicek, O. (1977). An Equilibrium Characterization of the Term Structure. *Journal of Financial Economics*, 177-188.

AUTHOR'S BIOGRAPHY

Camilo Sarmiento, Ph.D., has expertise in the areas of econometrics, finance, risk modeling, and macroeconomic forecasting. Camilo has published more than 35 articles in peer-referred journals as well as produced over 60 internal technical reports on model testing and development for top financial institutions. He has also worked as an Assistant Professor, Associate Research Scientist, and Consultant for universities and non-profit organizations. In his spare time, Camilo created the blog, *Insights on Financial News*.