

Supplementary Statistical Analysis

Jamie Bryan Hall

For each indicator, the latest figure and its one-year, five-year, and 10-year changes are easy to understand in terms of raw data, but we need supplementary statistical analysis to determine whether its observed trend may be merely random variation in the data.

To determine the appropriate regression model to use for each indicator, we must identify which indicators are trend stationary and which are possibly nonstationary. A trend-stationary indicator shows random movement around a trend line with a tendency to return to that trend line over time, while a nonstationary indicator follows a random walk (possibly with drift).¹ We first calculate the augmented Dickey-Fuller test statistic under the null hypothesis that the indicator follows a random walk with drift. For p-values less than 0.1 (i.e., when there is less than 10 percent chance that as extreme a value of the test statistic would be observed if the null hypothesis were true), we reject the null hypothesis and deem the indicator to be trend stationary. Of the 31 indicators, 16 are trend stationary and 15 are possibly nonstationary.

While we use the full series of available data for each indicator, the *Index* highlights

recent trends. We allow older data to lose statistical influence gradually over time by calculating geometrically decaying importance weights with a common ratio of 0.8. For example, data from 10 years prior to the latest year will receive a weight of $0.8^{10} \approx 0.134$ times the weight of the data from the latest year. This choice of common ratio means that average age of the data used, weighted by its importance in the regression model, is about five years prior to the latest year, the same weighted average age as if we had used equally weighted data from the latest and 10 previous years but with far less sensitivity to the behavior of the indicator five–10 years prior to the latest year.²

For each trend-stationary indicator, we then regress the data against time, allowing for the possibility that the deviations from the trend line depend on those from the previous period and may not be normally distributed. This is accomplished by estimating an ARIMA (1, 0, 0) model³ with robust standard errors using our importance weights. For each regression, we report the p-value of the test statistic for the trend parameter under the null

-
1. The quintessential example of a nonstationary time series is the number of “heads” minus the number of “tails” in a series of coin tosses. Someone who, following several consecutive heads, states that he or she is “due” for tails on the next toss is implicitly and incorrectly assuming that the series is stationary.
 2. We examined the sensitivity of the regression model results to the choice of common ratio in the range from 0.7 to 0.9 and found that it has little effect on the statistical significance of most of the estimated trend parameters.
 3. An ARIMA (p, d, q) model represents autoregressive integrated moving average with parameters p, d, and q and is the primary class of model used in time series analysis. The model may be extended in a variety of ways, and an explanation of the methods used to select an appropriate model structure is beyond the scope of this book.

hypothesis of a zero trend. Eight of 16 indicators have a p-value less than 0.1, indicating a non-zero trend.

For each nonstationary indicator, we estimate a regression model of the year-to-year change in the available data, allowing for the possibility that the errors depend on those from the previous period and may not be normally distributed. This is accomplished by estimating an ARIMA (0, 1, 1) model with robust standard errors using our importance weights. For these regressions, we report the p-value of the test statistic for the constant parameter under the null hypothesis of a zero constant. With p-values less than 0.1, eight of

15 indicators show a non-zero constant parameter, which is analogous to a non-zero trend parameter for a trend-stationary indicator.

Overall, 16 of 31 indicators in the *Index* show a statistically significant trend, nine of which are on the right track and seven of which are on the wrong track, while 15 currently show no clear statistically significant trend. The comprehensive table on pp. 100 and 101 reports these results, which we calculated using the statistical software package Stata 13.

Jamie Bryan Hall is a Senior Policy Analyst in the Center for Data Analysis, of the Institute for Economic Freedom, at The Heritage Foundation.