Running the Numbers: Demystifying Regional Economic and Social Data and Analysis

**Presentation to North Carolina
Association of County Commissioners**

Renaissance Asheville Hotel
Asheville, North Carolina
August 16, 2014

John Quinterno
South by North Strategies, Ltd.
Chapel Hill, North Carolina

Warm-Up Exercise: What Is Unemployment?

- Who is included in the count of the **unemployed**?
    - The jobless? Those receiving unemployment insurance? Those on a temporary layoff? Those working less than they want?

- The **unemployed** are all of the members (persons) of the civilian labor force (place of residence) who meet **all three** of the following criteria:
    - Lacked a job during a reference week (12th of month)
    - Were available for work
    - Made 1+ active effort to find a job during the prior four weeks

- Similarly, what is the **unemployment rate**?
    - A percentage that shows the number of unemployed persons as a share of the civilian labor force: (#unemployed ÷ #labor force)*100

2

Learning Objectives

- Provide an overview of the public statistical data system.

- Introduce key concepts needed to interpret regional data.

- Flag key questions to ask when consuming regional data.

- Identify key sources of regional statistical information.

- Illustrate regional applications of economic & social data.

- Describe habits useful to non-expert users of regional data.
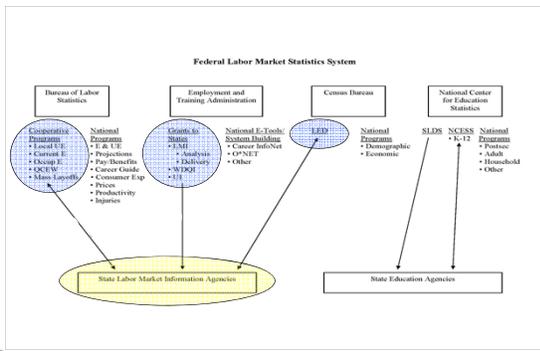
3

Guiding Philosophy

- Regional analysis is a powerful tool for understanding & improving the lives of the people who call a place home.
  - Too many people are scared away by data, or they grant the data too much credit.

- Data are by nature imperfect; there is not one right answer.
  - Everything discussed today simply is a framework for organizing complex phenomena in ways that limited human minds can grasp.

- The power of data & data analysis rests in an observer's ability to find meaning in them.
  - This task is achievable for any curious, educated individual willing to learn some basic data sources, concepts, & techniques.

South by North Strategies, Ltd.

4

The Public Statistical Data System

Federal Labor Market Statistics System

Source: Adapted from Brookings Institution Roundtable on "Putting America to Work," 9/27/10.

South by North Strategies, Ltd.

5

Data Resource Cheat Sheet

- If you want regional data about **economic output & production**, consult the US Bureau of Economic Analysis.
  - Regional Accounts System at www.bea.gov/regional/index.htm

- If you want regional data about **demographic traits of persons & businesses**, consult the US Census Bureau
  - American FactFinder at factfinder2.census.gov

- If you want regional data about **jobs & employment**, consult LMI data from US Bureau of Labor Statistics (BLS) or Labor & Economic Analysis Division (LEAD).
  - Web sites include www.bls.gov (BLS), www.nccommerce.com/lead (LEAD), & www.ncworks.gov (NCWorks Online)

South by North Strategies, Ltd.

6

Data Concept: Absolute & Relative Quantities

- Absolute quantities:

  - A total count of the number of times an item or event occurs (e.g., unemployed).
  - Absolute quantities, while useful, are insufficient for regional analysis.
  - Economic & social issues often are interconnected & context dependent.
  - Absolute values are ill-suited for comparisons.

- Relative quantities:

  - A measure of one item in relation to another item (e.g., the unemployment rate).
  - Examples are ratios, rates, proportions, percentages, percentiles, & index values.
  - Relative quantities permit the standardization of values (e.g., percentages as fraction of 100).
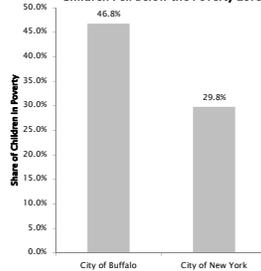 - Relative values enable the drawing of comparisons (size differences).

7

Absolute & Relative Quantities: An Illustration

7
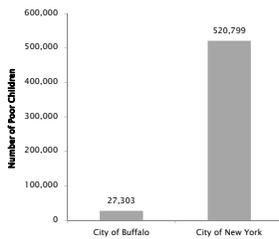
Data Concept: Growth & Change

- A common concern when using public data is seeing how a phenomenon **grows** or **changes** over time.
  - Growth & change, however, can be measured in different ways.

- Change can be gauged in **cross section** or **continuously**.
  - The difference is similar to that between looking at snapshots at & watching a movie.

- Change can be gauged in **absolute** or **relative** terms.
  - Relative values typically are more useful, with perhaps the most common one being **percentage change**.
    - Common analytical mistakes include mixing units, confusing types of changes, overlooking starting levels, & forgetting about "up & back."

9

3

Growth & Change: An Illustration

- Consider changes in the number of **employed persons** between Feb. 10 & Feb. 14 in two labor markets:
    - **Hickory MSA:** + 5,035, rising to 153,228 from 148,193
    - **Winston-Salem MSA:** + 5,431, rising to 223,726 from 218,295

- Winston-Salem had a larger **absolute** change in employment, but Hickory had greater **relative** growth.
    - **Hickory MSA:** 3.4% increase in employment
    - **Winston-Salem MSA:** 2.5% increase in employment

- On an annualized basis, Hickory had a faster growth rate.
    - **Hickory MSA:** employment growth of 0.8% per year
    - **Winston-Salem MSA:** employment growth of 0.6% per year

10

Data Concept: Time Effects

- Economic & social issues are influenced by **time**.

- Data often are **cross sectional** in nature.
    - Public data generally provide snapshots of a phenomenon at points of time, but they are not tracking the same people over time.

- Many economic issues are subject to **seasonality**, or regular patterns that repeat over time.
    - Labor variables have **unadjusted** & **seasonally-adjusted** series.
        - Payroll employment in NC's retail sector fell by 23,800 jobs (-5%) from Dec. 13 to Jan. 14 before adjusting for seasonality & 10,600 jobs (-2.3%) afterward.

- Labor markets also are subject to the **business cycle**.
    - Unemployment falls during expansions & rises in contractions.
    - Nationally, the last contraction ran Dec. 07 to Jun. 09; expansion is underway.

11

Data Concept: Financial Values

- As with other quantities, financial values may be expressed in absolute & relative terms.
    - The average weekly manufacturing wage in 2012 in the **Asheville MSA** was $932; that equaled 91.5% of the statewide average.

- Financial values (e.g., wages) also are influenced by time.
    - In a market economy, the prices of goods & services (e.g., labor) constantly change, both in the aggregate & relative to each other.
        - Changes in the aggregate price level are referred to as **inflation** or **deflation**.

- When viewing values, an observer must differentiate **nominal (current)** values from **real (constant)** ones.
    - An annual salary of $30,000 in 1990 would equal $53,500 in 2013.
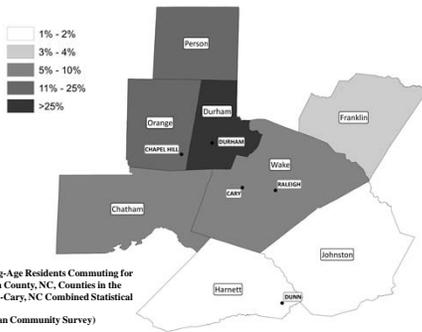
12

Understanding Regions

- A **region** is "an area within the national [state] economy that is sufficiently comprehensive in structure that it can function independently, although … in most practical circumstances it has strong links with the rest of the economy."

- Three common ways of understanding regions:
  - **Territorial perspective:** A region is not simply a discrete physical space but the complex product of economic & social interactions (e.g., the Research Triangle Region).
  - **Functional perspective:** A region is a discrete spatial area containing interconnected places of differing sizes & types (e.g., commuter flows).
  - **Administrative perspective:** A region is a spatial area over which a common political or governmental structure applies (e.g., counties, workforce development boards, regional partnerships).

13

Understanding Regions: An Illustration

- 1% - 2%
- 3% - 4%
- 5% - 10%
- 11% - 25%
- >25%

Person, Durham, Orange, CHAPEL HILL, DURHAM, Franklin, Wake, RALEIGH, CARY, Chatham, Johnston, Harnett, DUNN

Share of Working-Age Residents Commuting for Work to Durham County, NC, Counties in the Raleigh-Durham-Cary, NC Combined Statistical Area 2006-08 (Source: American Community Survey)
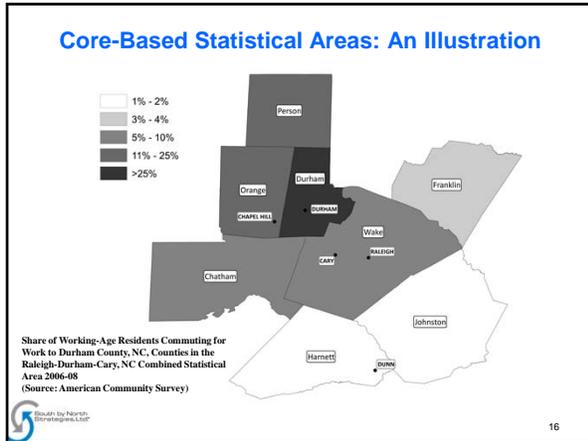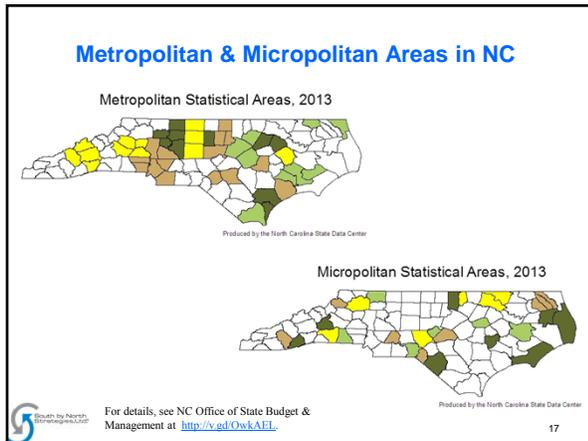
14

Regions as Core-Based Statistical Areas

- A **metropolitan statistical area** (MSA) is a county or set of counties with at least one urbanized area of 50,000 or more residents & a high degree of economic & social connectedness as measured by commuting ties.

- A **micropolitan statistical area** (μSA) is a county or set of counties with at least one urbanized cluster of 10,000 to 49,999 residents & a high degree of economic & social connectedness as measured by commuting ties.

- A **combined statistical area** (CSA) consists of adjacent MSAs or μSAs that have substantial economic ties as measured by employment patterns.

15

Core-Based Statistical Areas: An Illustration

Legend:
- 1% - 2%
- 3% - 4%
- 5% - 10%
- 11% - 25%
- >25%

Share of Working-Age Residents Commuting for Work to Durham County, NC, Counties in the Raleigh-Durham-Cary, NC Combined Statistical Area 2006-08 (Source: American Community Survey)

16

Metropolitan & Micropolitan Areas in NC

Metropolitan Statistical Areas, 2013

Produced by the North Carolina State Data Center

Micropolitan Statistical Areas, 2013

Produced by the North Carolina State Data Center

For details, see NC Office of State Budget & Management at http://v.gd/OwkAEL.

17

Non-Metropolitan Regions: Considerations

- The terms **urban** & **rural** have different meanings in the public statistical system than popular understanding.
    - Distinction is based on population densities of census blocks; this means that "rural" areas may be found in "urban" places.

- For the purposes of regional analysis, **non-metropolitan** better refers to what many people think of as rural.
    - Of 100 NC counties, 74 are in an MSA or μSA; 26 are non-metro.
    - Other useful rural typologies are at http://v.gd/JnsT4N.

- Data exist in NC for counties, cities, WDBs, MSAs, μSAs, CSAs, & selected types of regions.
    - The type of data available varies by type of geography.

18

Framing: What Is the Analytical Problem?

- Every analysis attempts to address one of three problems:
  - Problem of **description** (e.g., size & traits of the unemployed)
  - Problem of **evaluation** (e.g., effects of an unemployment program)
  - Problem of **estimation** (e.g., future changes in unemployment)

- Next step is to define key **concepts** & **measures**.
  - Concepts like unemployment need to be defined in analytical terms.
  - Four types of data measures: **ratio** (e.g., wages); **interval** (e.g., price index); **ordinal** (e.g., county rankings); & **nominal** (e.g., gender).

- Data measures ideally are both **valid** & **reliable**.
  - A **valid** measure fairly measures the concept of interest, & a **reliable** measure yields consistent values over time.

South by North
Strategies,Ltd®

19

Framing: Which Kinds of Data Are Used?

- Public data tend to come from one of three sources:
  - **Enumeration:** Count of every member of a population of interest, a.k.a., a census (e.g., Decennial Census, Economic Census)
  - **Statistical sample:** Data derived from a representative subset of a population of interest (e.g., American Community Survey)
  - **Administrative records:** Data from program files (e.g., wage file)

- All public data are subject to various kinds of errors.
  - Enumerations are prone to **measurement error** (due to design flaws/problems) & **random error** (due to chance).
  - Sample data also are prone to **sampling error**, or the difference between the sample value & the actual population value.

- The use of a **confidence interval** reflects sampling error.

South by North
Strategies,Ltd®

20

Framing: Which Statistical Tools Are Used?

- Regional studies frequently employ **descriptive statistics**.
  - A measure of **central tendency** is one number or score that represents the average value in a group.
    - Most common examples are the **mean** (arithmetic average), the **median** (middle value in an ordered set), & the **mode** (most frequent value).
      - When using financial values, the median often is preferable to the mean.
  - A measure of **dispersion** assesses how much the data do or do not cluster around the mean value.
    - Examples include the **range**, **interquartile range**, **variance**, & **std. deviation**.
  - A measure of **association** shows patterns among variables.
    - **Cross-tabulations** are one way of presenting association (e.g., wages by race).

- Some studies (surveys) also employ **inferential statistics**.
  - Tools to generalize from a sample to a population & test hypotheses.

South by North
Strategies,Ltd®

21

Framing: What Is the Purpose of the Analysis?

- Most often, public officials look to data to help **describe** conditions in a particular region.
    - For instance, how much has poverty grown in past year?

- Sometimes, the goal is to **correlate** different variables.
    - For instance, is there a relationship between the educational attainment of workers & their wages (positive correlation)?
        - **Remember that correlation is not causation!**

- Other times, the purpose is to show if a change in one variable **causes** another one to change.
    - For instance, does a decrease in the duration of unemployment insurance compensation cause workers to leave the labor force?
    - This leads down the path of **regression analysis**.

22

A Brief Review

- The public statistical system is a powerful but confusing source of regional information.

- The challenge with public data is finding meaning.

- There are four core data concepts to keep in mind when considering public data.

- It is most practical to think about public data by regions.

- Asking four key framing questions positions an observer to profit from analyses of quantitative data.

23

Public Data Resource Cheat Sheet

	BEA	Census Bureau	BLS / LEAD
Focus	Economic Output	People & Economy	Labor & Jobs
Key Regional Products	Regional Accounts	Decennial Census, Economic Census, & ACS	LAUS, CES, QCEW, & OES
Data Sources	Admin. Records	Census & Surveys	Census, Surveys, & Admin. Records
Time Lag	Longest Lag	Medium Lag	Short Lag
Revisions	Yes	No	Yes
Limitations	Sampling & Non-sampling Errors	Sampling & Non-sampling Errors	Sampling & Non-sampling Errors
Geography (selected)	State, MSA, μSA, & County	All Geographic Levels	State, County, MSA, μSA, & WDB
Website	www.bea.gov	www.census.gov	www.bls.gov www.ncworks.gov

24

8

A Data Example: Economic Output (1 of 2)

- **Topic:** How did **economic output** in the **Winston-Salem MSA** change during the last **business cycle**?

- **Step 1:** What is the analytical problem?
    - This is a **problem of description**.
    - The key concepts are **output**, **geography**, & **business cycle**.
    - Use standard output measures, the Winston-Salem MSA, & the cycle as having a recession (2007-09) & a recovery (2009-12)

- **Step 2:** Which kinds of data are being used?
    - Most relevant data would be the BEA's **Regional Economic Accounts**, particularly the annual values for 2007-12 (most recent).
    - The values come from a **model** that uses certain assumptions, admin. records, & survey data, all of which are prone to error.

25

A Data Example: Economic Output (2 of 2)

- **Step 3:** Which statistical tools are being used?
    - This example involves simple calculations of absolute & relative change: **absolute numerical change** & **percentage change**.

- **Step 4:** What is the purpose of the analysis?
    - To understand the regional economy's recent performance.

- **Results:**
    - Economic output in Winston-Salem fell by 4.6% during the recession, dropping to $23.2 billion from $24.4 billion (real dollars).
    - Economic output fell another 0.4% during the recovery and is now 5% below the pre-recession level ($23.1 billion vs. $24.4 billion).
    - On a per-capita basis, residents were worse off in 2012.
        - Per capita income down -8.6% or -$3,360 from 2007 ($39,090 to $35,730).

26

A Data Example: Demographic Traits (1 of 2)

- **Topic:** How did **unemployment rates** for **prime-age workers** in the **Asheville MSA** vary by educational attainment in 2012?

- **Step 1:** What is the analytical problem?
    - This is a **problem of description** and/or **problem of correlation**.
    - The key concepts are **unemployment**, **geography**, **prime-age workers** (ages 25-64) & **time**.
    - Use standard labor market concepts (unemployment), standard Census Bureau concepts, Asheville MSA definition, and year 2012.

- **Step 2:** Which kinds of data are being used?
    - Most relevant data would be the Census Bureau's **American Community Survey** estimates for 2007-2012 (most recent).
    - Those values are **survey data**, meaning they are subject to error.

27

9

A Data Example: Demographic Traits (2 of 2)

- **Step 3:** Which statistical tools are being used?
  - This example involves simple calculations of absolute & relative change: **absolute numerical change** & **percentage change**.

- **Step 4:** What is the purpose of the analysis?
  - To understand trends for different segments of the local labor force.

- **Results:**
  - Unemployment rates broke out as follows: Less than HS = 15.8%; High school graduate = 10.4%; Some college or associate's degree = 7%; and Bachelor's degree or higher = 4.1%.
    - Note that additional statistical testing would be needed to determine whether the differences are statistically significant.
  - Rates for every group are higher than they were in 2007.
    - For example, 2007 rate among those with a Bachelor's degree was 2.7%.

A Data Example: Labor Force Trends (1 of 2)

- **Topic:** How has each **labor force component** in the **Raleigh MSA** changed since the onset of the **recovery**?

- **Step 1:** What is the analytical problem?
  - This is a **problem of description**.
  - The key concepts are **labor force components**, **geography**, & **recovery**.
  - For this exercise, use standard labor force definitions, the Raleigh-Cary MSA geography, & the recovery as 2010 through 2013.

- **Step 2:** Which kinds of data are being used?
  - Most relevant data source would be the **LAUS**, particularly the annual values for 2010 & 2013 (seasonally unadjusted only).
  - Those values are **survey data**, meaning they are estimates subject to error, & the data are **cross-sectional** in nature.

A Data Example: Labor Force Trends (2 of 2)

- **Step 3:** Which statistical tools are being used?
  - This example involves simple calculations of absolute & relative change: **absolute numerical change** & **percentage change**.

- **Step 4:** What is the purpose of the analysis?
  - To gauge if local labor market conditions are improving.
  - To identify the size of the population in need of workforce services.

- **Results:**
  - Labor force up 33,000 persons (+5.7%); employed population up 45,000 persons (+8.6%); unemployed population down 12,000 persons (-24%); & unemployment rate down to 6.4% from 8.8%.
  - **Bottom line:** Conditions improved due largely to new members of the labor force finding work, not unemployed ones finding work.

Habits of Good Data Users

- The ability to locate, analyze, & interpret public data is not limited to experts; in fact, non-experts often are the ones tasked with making decisions based on public data.

- Analysis simply is a systematic framework for putting data in forms that limited, fallible human beings can grasp.

- The basic purpose of any analysis is to document regional realities, understand underlying dynamics, & make rational decisions.

- There are seven "habits" of the mind that good data users should strive to cultivate.

31

Habit 1: Know What Needs to Be Known & Why

- Often the desire to use public data is a reactive one.
  - The trigger frequently is a crisis, complaint, or negative report that results in a call to "get some numbers."

- Regional officials typically are interested in four subjects: **persons**, **jobs**, **businesses**, & **economic output**.

- The first step in any analysis is to identify at the outset what needs to be known & why.
  - A study of low-wage workers in an area intended to improve workforce services likely would proceed differently from one aiming to understand shifts in the local economic structure.
  - Being clear upfront limits chances of wasting time & money.

32

Habit 2: Learn Regional Geography

- Economic & social topics typically are regional in nature.
  - Define a region too broadly, & you will include extraneous information, but define it too narrowly, & you will exclude key data.
  - Regional leaders tend to think in terms of political geography, but conflating regions with political units is problematic.

- Core-based statistical areas like MSAs, μSAs, & CSAs are useful geographies for thinking along regional lines.
  - Such geographies are coherent, functional units that are large enough to capture diversity but small enough to be distinct places.
  - As a rule of thumb, use these areas as an analytical starting point.

- Places outside of statistical areas need other frameworks.
  - Coherence should be the minimum standard for defining a region.

33

Habit 3: Befriend Essential Data Sources

- The United States has a fragmented public data system.
  - Absent a central agency, answering questions is a patchwork process.

- The best data sources are those produced by the Census Bureau, BEA, and BLS / LEAD.
  - Useful websites include American FactFinder for Census data (factfinder2.census.gov), Regional Economic Accounts for BEA data (www.bea.gov/regional), & NCWorks Online for labor market data (www.ncworks.gov).

- Best way to improve grasp of resources is to explore them.
  - When in doubt, ask resources in agencies like LEAD, State Data Centers, institutions of higher education, & other experts.

34

Habit 4: Brush Up on Basic Math & Statistics

- "There are three kinds of lies: lies, damned lies, & statistics." ~ Mark Twain
  - Statistics can both improve & cloud understanding.
  - Sometimes confusion stems from deliberate misleading, other times from the imperfect nature of even the best statistics.

- Grasping a few basic quantitative concepts can reduce the odds of being misled, either on purpose or by accident.
  - Core concepts include the difference between absolute & relative quantities; the dynamics of change & growth; the importance of time effects; & the nature of financial values.
  - Many studies involve no more than basic math & statistics.

- Keep the four framing questions discussed earlier in mind.

35

Habit 5: Appreciate the Art of "Satisficing"

- Data analysis is riddled with problems, & answering even the seemingly simplest questions can be quite hard.
  - One data problem is the trade-off between timeliness & accuracy.
  - Another problem is that virtually all data are retrospective in nature.

- Regional leaders attempting to use data to steer policy resemble motorists driving with only a rearview mirror.
  - Public officials therefore need to be comfortable with uncertainty.

- "Satisficing," or making the best out of the limited imperfect data that are available, often is the only real strategy.
  - Satisficing often is the best way to avoid "analysis paralysis," limit personal frustration, extract maximum valuable from what data are available, & otherwise avoid "getting lost in the weeds."

36

Habit 6: Tap the Inner Liberal Arts Student

- A data user should more closely resemble a liberal arts student than a technician.
  - Technical expertise matters, but the process really involves engaging an issue, asking questions, reflecting on the answers, considering multiple perspectives, synthesizing information, communicating insights, & revisiting conclusions.

- This mindset is especially important when setting policy.
  - Technocratic analysis seldom resolves policy arguments as those arguments hinge on rival ideas of shared values & the public good.

- The overarching purpose of regional analysis is less finding a "right answer" & more a framework for understanding & addressing issues of regional importance.

South by North
Strategies, Ltd®

37

Habit 7: Remember the Public Purpose

- Public debates grow heated due to differences of values & interests -- differences personal concerns only exacerbate.
  - The actions of public-sector & quasi-governmental organizations influence the distribution of resources & create winners & losers.

- At the same time, only the public sector is responsible for thinking about the broad public good, thinking for the long term, & minding issues of social equity.

- When using data to set policy & programmatic priorities, civic leaders need to think broadly about all the concerns involved & attempt to balance the diversity of values, opinions, experiences, & needs present in a community.

South by North
Strategies, Ltd®

38

A Review of Learning Objectives

- Provide an overview of the public statistical data system.

- Introduce key concepts needed to interpret regional data.

- Flag key questions to ask when consuming regional data.

- Identify key sources of regional statistical information.

- Illustrate regional applications of economic & social data.

- Describe habits useful to non-expert users of regional data.

South by North
Strategies, Ltd®

39

13

Contact Information

John Quinterno

South by North Strategies, Ltd.

179 East Franklin Street, #294 *Chapel Hill, NC 27514

(919) 622-2392 * **johnq@sbnstrategies.com**

Running the Numbers

More Information Available Online:

www.sbnstrategies.com

www.runningthenumbersbook.com

40