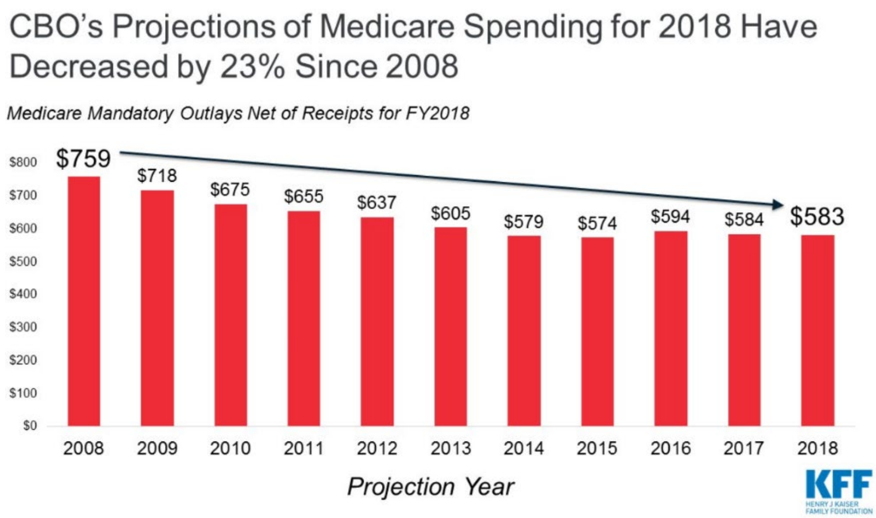
CHAPTER 7: SENSITIVITY ANALYSIS

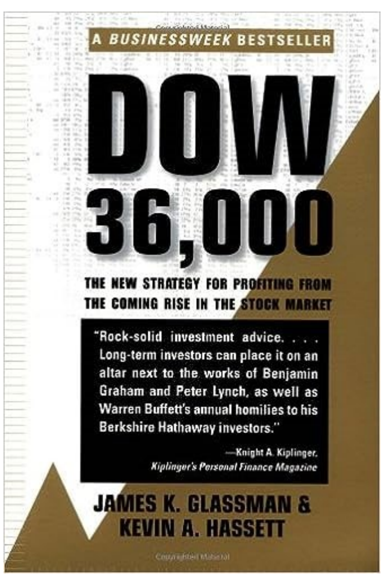
Overview

It is hard to make accurate predictions, especially in rapidly-changing fields such as health care. The graph below shows the Congressional Budget Office's predictions of health care spending in 2018 by the year when the prediction was made. As you can see, predictions before 2013 were off by a wide mark. Given the difficulty of making accurate predictions, it is important to convey uncertainty to policymakers who consider predictions when making decisions.



Confidence

Forecasters should describe how confident they are in their predictions. Economist Charles Manski writes[1]:



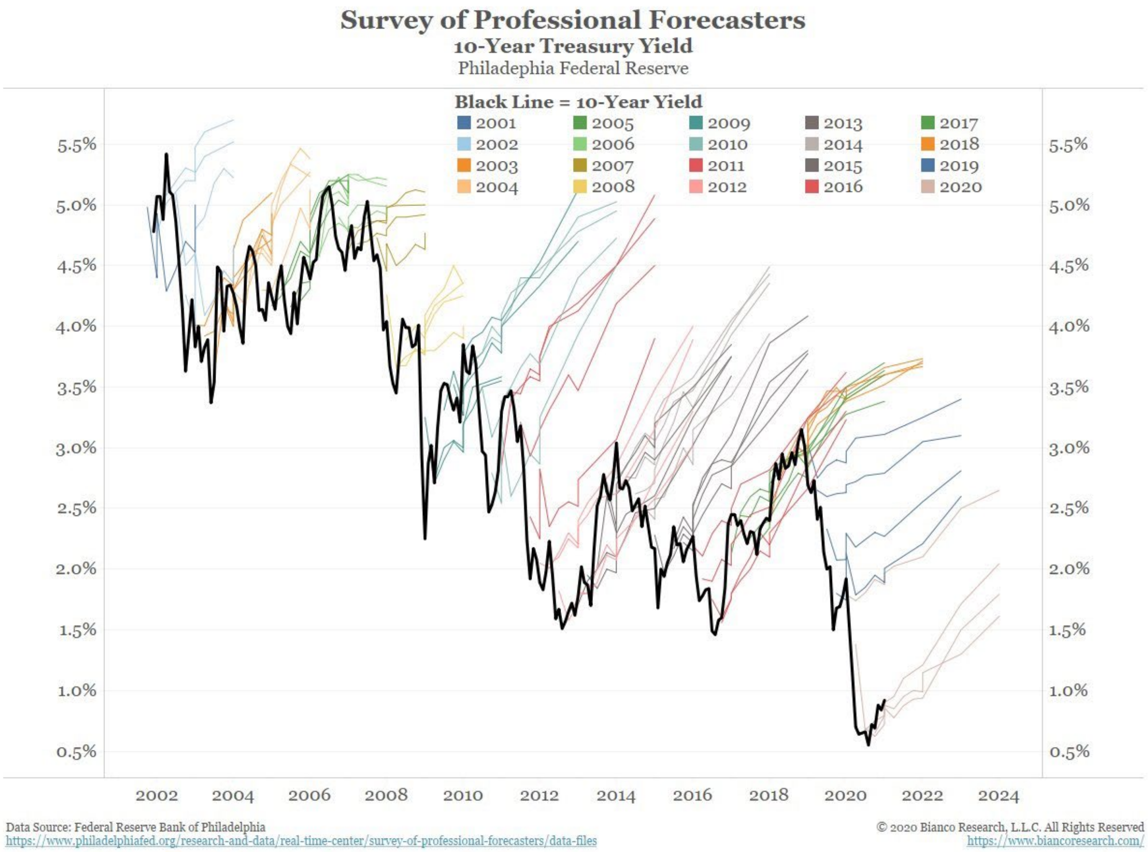
> Analysts produce estimates without qualification based on the belief that policymakers demand certainty and are unable to cope with ambiguity. But forecasts that do not express uncertainty ultimately will risk their credibility and policymakers see time and time again that the forecast is not accurate.

Unfortunately, analysts who make bold, confident predictions tend to receive more attention than those who hedge their bets. Confidence sells. Journalist James Glassman and economist Kevin Hassert, who recently stepped down as the head of Trump's Council of Economic Advisors, wrote a book in 1999 called Dow 36,000, predicting that the Dow Jones Industrial Average would soon reach 36,000 (at the time it was around 11,000). The stock market crashed soon after the book was published.

[1] Manski CF, Policy analysis with incredible certitude. *The Economic Journal* 2011.

The index's value as of this writing is 22,773. The authors received a lot of publicity for their prediction. They received some criticism, but their careers have flourished.

The moral of the story: if you are interested in fame or fortune, it is better to write a book called *Dow 36,000* than one titled *The Dow May Reach 36,000, But There Is Also A Good Chance It Will Hover Between 10,000 and 20,000 For Awhile.* But, if you are policy analyst, you should be more circumspect. Predictions and recommendations should be commensurate with the quality of the evidence.

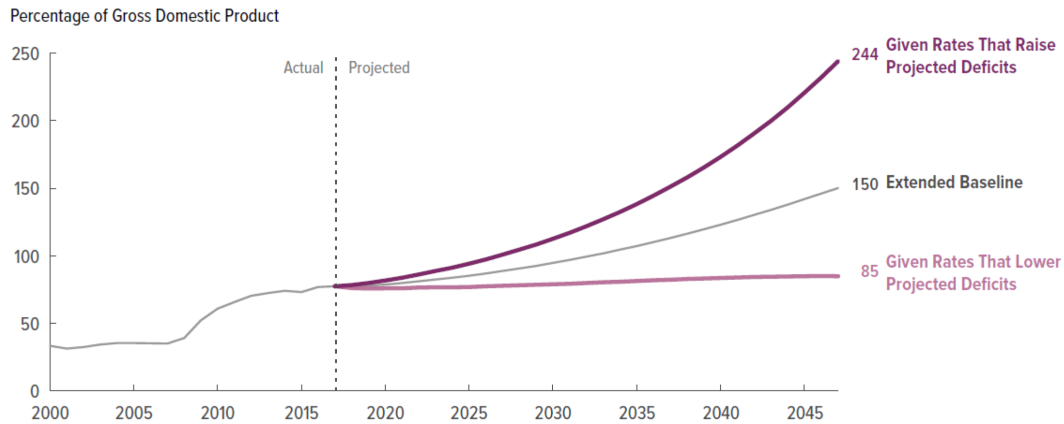


Agencies and analysts differ in how they convey uncertainty. It its Long Term Budget Outlook[2], the Congressional Budget Office estimates federal debt under difference scenarios.

[2] Congressional Budget Office. The 2017 Long-Term Budget Outlook. March 2017.

Figure 10.

Federal Debt Given Different Labor Force Participation Rates, Productivity Growth Rates, Federal Borrowing Rates, and Rates of Excess Cost Growth for Federal Spending on Medicare and Medicaid



The Bank of England uses Monte Carlo simulation to develop ranges of predictions for the inflation rate.[3]

Chart 5.1 CPI inflation projection based on market interest rate expectations, other policy measures as announced

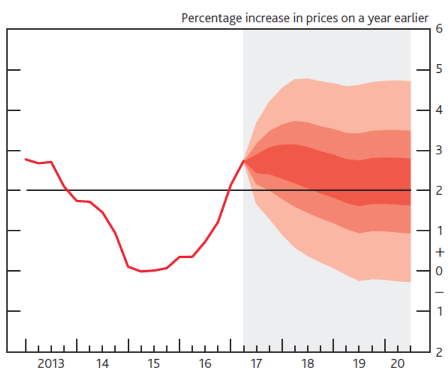
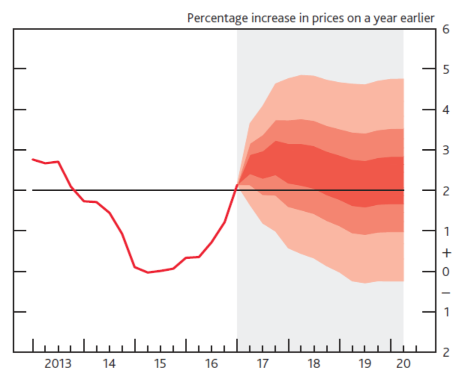


Chart 5.2 CPI inflation projection in May based on market interest rate expectations, other policy measures as announced



Charts 5.1 and **5.2** depict the probability of various outcomes for CPI inflation in the future. They have been conditioned on the assumptions in **Table 5.B** footnote (b). If economic circumstances identical to today's were to prevail on 100 occasions, the MPC's best collective judgement is that inflation in any particular quarter would lie within the darkest central band on only 30 of those occasions. The fan charts are constructed so that outturns of inflation are also expected to lie within each pair of the lighter red areas on 30 occasions. In any particular quarter of the forecast period, inflation is therefore expected to lie somewhere within the fans on 90 out of 100 occasions. And on the remaining 10 out of 100 occasions inflation can fall anywhere outside the red area of the fan chart. Over the forecast period, this has been depicted by the light grey background. See the box on pages 48–49 of the May 2002 *Inflation Report* for a fuller description of the fan chart and what it represents.

[3] Bank of England. *Inflation Report*. August 2017.

Two Types of Uncertainty

The policy analysis Menu Labeling as a Potential Strategy for Combatting the Obesity Epidemic[4] predicts that menu labeling would prevent 39% of the weight gain that would otherwise occur in the population, assuming that labeling causes 10% of patrons to switch to a low calorie meal. The estimate comes from a 2006 paper in the *American Journal of Public Health*.[5] The paper describes the effect thusly:

> When calorie-plus-nutrient information was presented, the percentage of consumers choosing the turkey sandwich (which generally met or exceeded nutrition expectations) increased from 11% to 21%, and it decreased selection of items with higher levels of calories and fat than expected.

(As we discussed previously, this effect should be interpreted as a percentage point change, not a percent change.) There are several sources of uncertainty

- First, it is unclear whether the effect, which reflects respondents' answers to a mail survey, reflects the behavior change we might observe in the real world. Restaurant patrons make decisions on a repeated basis and must eat the food they choose. We might expect the real-world effect to be smaller than 10%. There is uncertainty due to the study design and the generalizability of the effect.

- A second source of uncertainty is statistical uncertainty. Even if the estimate is unbiased, it may differ from the true value because of sampling variation. The sample on which the estimate was based is finite. Standard errors and confidence intervals capture statistical uncertainty.

A sensitivity analysis can address one or both types of uncertainty.

The authors of the analysis present a two-way sensitivity analysis.

---

[4] Simon P, CJ Jarosz, T Kuo, JE Fielding. *Menu Labeling as a Potential Strategy for Combatting the Obesity Epidemic.* County of Los Angeles Public Health. May 2008. http://publichealth.lacounty.gov/docs/menu_labeling_report_2008.pdf
[5] Burton S, Creyer EH, Kees J, Huggins. Attacking the obesity epidemic: the potential health benefits of providing nutrition information in restaurants. *American Journal of Public Health* 2006;96:1669-1675.

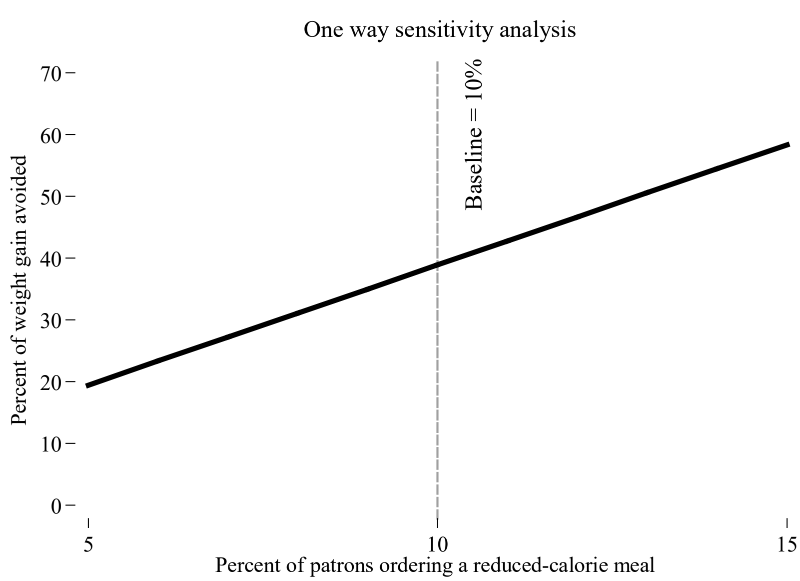
Table 2
Impact of consumer response to menu labeling on the percentage of population weight gain averted: simulation of multiple scenarios of calorie reduction.§

Average Amount of Calorie Reduction per Meal	Percentage of Large Chain Restaurant Patrons Who Purchase a Lower-Calorie Meal as a Result of Menu Labeling				
	10%	20%	30%	40%	50%
25	9.7%	19.4%	29.1%	38.9%	48.6%
50	19.4%	38.9%	58.3%	77.7%	97.2%
75	29.1%	58.3%	87.4%	116.6%	145.7%
100	**38.9%**	77.7%	116.6%	155.4%	194.3%
125	48.6%	97.2%	145.7%	194.3%	242.9%
150	58.3%	116.6%	174.9%	233.2%	291.5%
175	68.0%	136.0%	204.0%	272.0%	340.0%
200	77.7%	155.4%	233.2%	310.9%	388.6%

§ Percentages presented in the table refer to the percentage of population weight gain averted.

They chose to examine only values of the share of patrons choosing a low calorie meal greater than 10%. But, given the likelihood that the 10% estimate is too high, it would have been better to at least include some values below 10%.

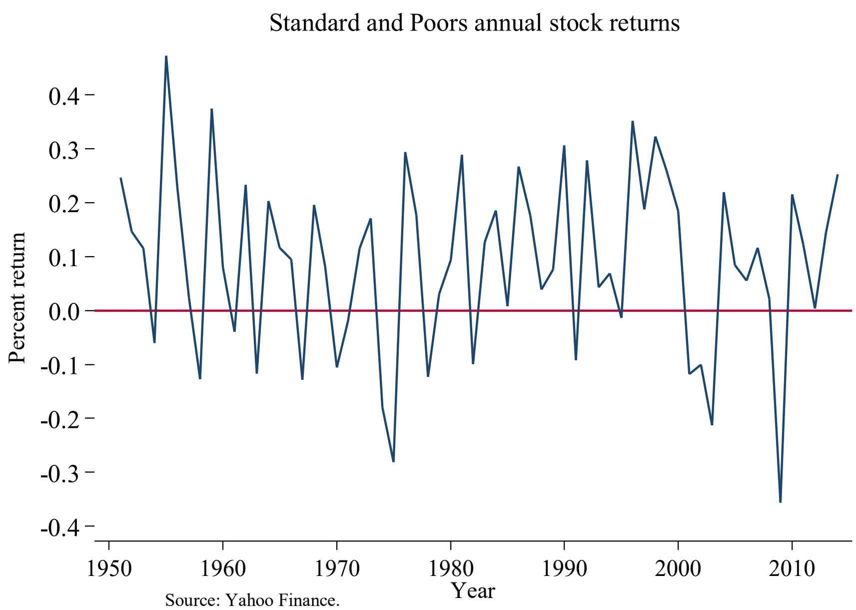
The graph below displays a deterministic, one way sensitivity analysis. I selected the upper and lower bounds, 5 and 15, arbitrarily.



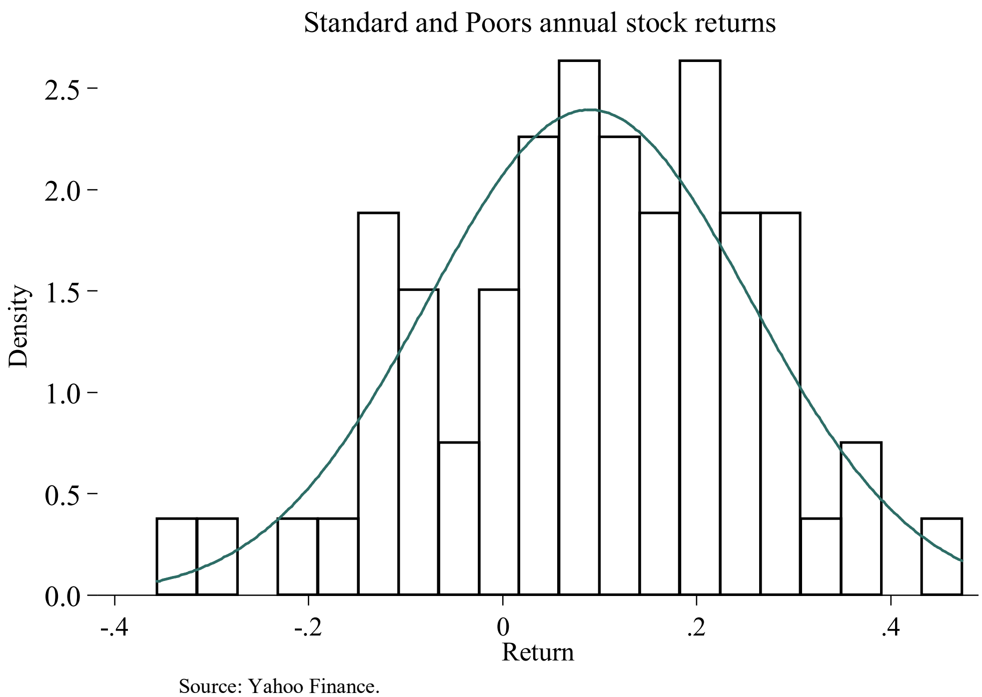
Monte Carlo Simulation

Monte Carlo simulation is a tool for conducting sensitivity analysis that takes advantage of prior information about the distribution of an uncertain outcome or variable. We may not know the eventual value of the outcome, but we do know which outcomes are more or less likely. What will happen to $10,000 if you invest it in the stock market for 10 years? What is the probability that you lose money? You can use Monte Carlo simulation to answer the question.

Here is the annual return on a portfolio invested in the Standard and Poor's 500 index.

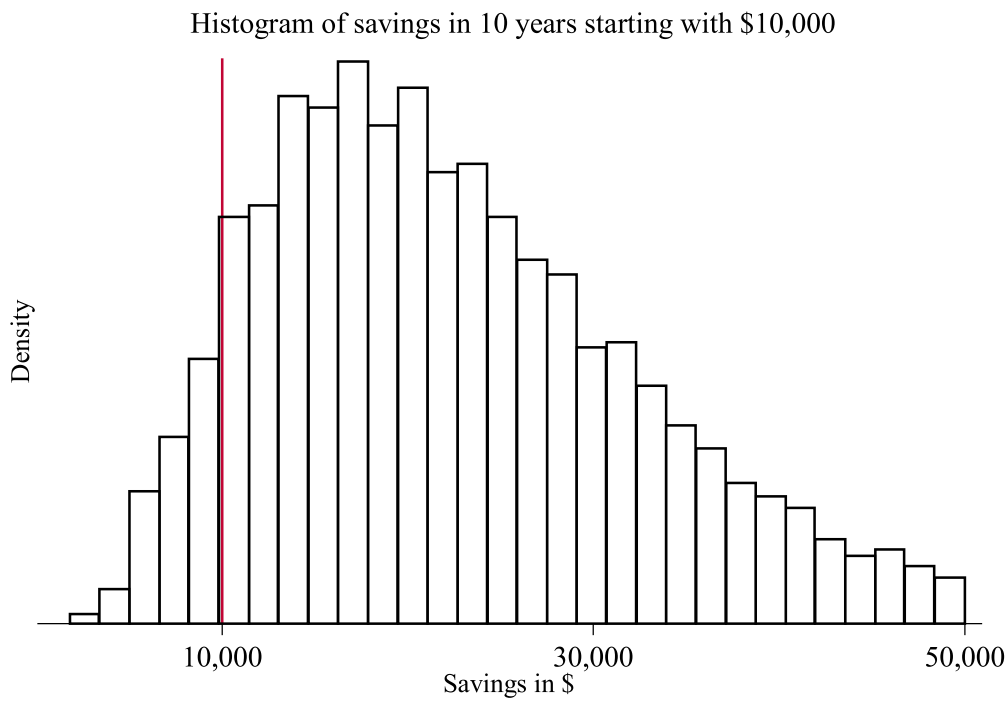


Here is a histogram of the returns.



The mean return was 9%. The return was negative in 17 of 65 years. The market yields a positive return in most years, but on occasion the value of a portfolio will decline by 20% or more.

To calculate the distribution of portfolio values after 10 years, I assume that the distribution of returns in the future will mirror the distribution of returns in the past. I randomly drew a return from the historical distribution, multiply it by the portfolio value, and repeat this process 9 more times (for a total of 10 times, representing 10 years). At the end, I have the predicted value of a simulated portfolio. Then, I repeated the entire process 10,000 times. Each time, I record the value of the portfolio at the end of 10 years. Here is a histogram of the results.



The average portfolio value was $23,000. The portfolio was less than $10,000 in 8% of the simulations.

You could have put the money into a safer investment. If the safer investment returned a steady 2% per year, you would have had about $12,000 at the end of 10 years. The stock portfolio was less than $12,000 in 14% of the simulations.

Deterministic versus probabilistic sensitivity analysis

In a "deterministic" sensitivity analysis, the analyst just examines the sensitivity of results to a range of possible values for an assumption (or assumptions) without worrying about whether some values in the range are more or less likely.

In a "probabilistic" sensitivity analysis, the analyst assumes a distribution for the assumption and randomly samples, via Monte Carlo simulation, from the distribution. The analyst could pick a distribution based on prior knowledge or, if the goal is to capture uncertainty due to sampling variation, a confidence interval.

If you have a good basis for selecting a distribution, then probabilistic sensitivity analysis is preferable, though it is difficult to explain to non-technical audiences. Otherwise, deterministic sensitivity analysis is the default option.

Probabilistic sensitivity analysis for the menu labeling paper

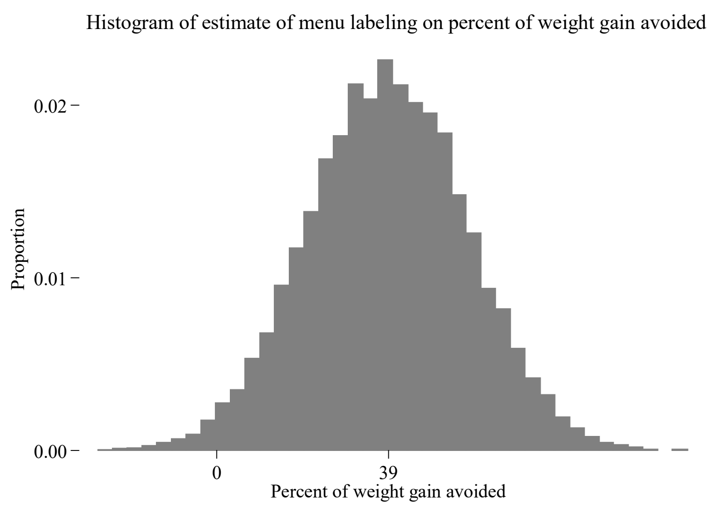
The *American Journal of Public Health* menu labeling paper does not report the confidence interval for the 10% estimate. I calculated it assuming the sample of 240 respondents was evenly split between the treatment and control groups and applying the formula for calculating the standard error for a difference in proportions: = sqrt [p1(1-p1)/n1 + p2(1-p2)/n2]. The upper and lower bounds of the 95% confidence interval are 0.8% and 19%. The standard deviation is 5% (with some rounding).

Suppose we want to conduct a sensitivity analysis to reflect uncertainty in the assumption about the share of consumers who will switch to a low calorie meal due to sampling variability. We could perform a deterministic sensitivity analysis where we vary the assumption between its 95% confidence interval bounds. But not all values between 0.8% and 19% are not equally likely. Instead, the distribution of the estimate is normal, centered on 10% with a standard deviation of 5%. A probabilistic sensitivity analysis, using Monte Carlo simulation, can account for the distribution of the estimate. To conduct the analysis, I

1) took a random draw from a normal distribution with mean 10 and standard deviation 5,
2) recalculated the outcome of the menu labeling model, and
3) repeated this process 1,000 times.

I performed the analysis in Stata, but it is easy to do in Excel. If you were doing this analysis in Excel, you would type "=NORMINV(RAND(),10,5)" in a cell to take a random draw.

Here is a histogram of the percent of weight gain prevented. The distribution of the outcome (percent of weight gain prevented) reflects the distribution of the estimate of the impact of calorie labels on the share of consumers who will select a low calorie meal.



Since the normal distribution is unbounded (values between –∞ and +∞ are possible), you do end up taking some random draws where the value is negative (which would correspond to the situation where menu labeling *decrease*s the share of consumers ordering reduced calorie meals). It would be reasonable to discard these values as part of a sensitivity analysis.

Performing a sensitivity analysis

To perform a one way sensitivity analysis, identify a key assumption that you think has a big influence on the outcome and where you are unsure of the true value. The policy effect assumption is usually a good choice.

Consider the rationale. Is uncertainty due to random, sampling error? In that case, consider a Monte Carlo sensitivity analysis if you know the standard error or confidence interval associated with the assumption or a deterministic sensitivity analysis where you examine the sensitivity of results to a symmetric range centered on the baseline assumption.

Is uncertainty due to a flawed study design? Perhaps you have an assumption about the impact of a policy on an outcome but you believe the study on which it is based may have systematically overstated the effect. In that case perform a deterministic sensitivity analysis. The interval over which you examine the sensitivity of results does not necessarily need to be symmetric or centered on the baseline assumption. If the baseline assumption is Z, then you might perform a sensitivity analysis over the interval 0 to Z (if you think Z overstates the true effect).

You will sometimes see policy analyses present estimates under different scenarios (for example, a “best case” scenario and a “worse case” scenario). This is another approach to sensitivity analysis. It should not be the default approach. In general it is better to first present a single baseline projection.

Put the sensitivity analysis in a separate section. Describe the baseline analysis first. If you try to describe the baseline and sensitivity analyses at the same time, you may confuse your reader.

A sensitivity analysis does not cure or address problems with internal validity or generalizability. It simply shows how results changes as you vary assumptions. That’s it.