

DYNAMIC SCORING: UPDATE AND RECOMMENDATIONS

Norton Francis, Chief Economist

- The pilot program for dynamic scoring ended in 2005 with inconclusive results. The REMI model was not a practical model for dynamic scoring of legislation during the legislative session.
- Following presentations to the committee, there was no clear indication that the project should continue and there was no additional funding so at this point the state no longer has a contract with REMI to provide dynamic scoring modeling.
- Many other states have had similar experience with dynamic scoring. While the federal government uses dynamic scoring and analysis, the open borders of state economies make the analysis much more difficult at the state level.
- REMI and other models can be used for economic impact analysis of large projects or significant changes in legislation. REMI is very expensive and requires dedicated trained analysts but there are other options available.
- DFA has renewed interest in REMI and DFA and LFC have discussed ways to make the REMI output useful for scoring legislation. One idea is to use the consensus revenue group framework to develop “rules-of-thumb” coefficients that could easily be applied to changes in the tax system. This could work with REMI output as well as other input-output models.
- At the present time, dynamic scoring of legislation is not being considered for the 2008 session.

MEMORANDUM

TO: Senator John Arthur Smith, Chair LFC
CC: LFC members
David Abbey, Director LFC
FROM: Norton Francis, Chief Economist LFC
DATE: 12/7/07
SUBJECT: DYNAMIC SCORING

Several times during the interim the subject of dynamic scoring of legislation and proposals has come up and this memo serves to address the issue as comprehensively as possible. Laird Graeser at the Department of Finance and Administration and Tom Clifford at New Mexico Tax Research Institute provided input for this report.

Executive Summary

Dynamic scoring is a way of analyzing government policies to take into account economic impacts from changes in behaviors the policies create. In contrast to static scoring, which assumes no economy wide changes, dynamic scoring lets a policy work within the economy, follows the “ripples” and reports a cumulative impact. The value of such scoring is hotly contested. The federal government requires dynamic scoring and reports from the Joint Committee on Taxation have shown only modest feedback effects. Other studies have shown more significant feedbacks particularly on tax changes regarding capital investments. States have proceeded cautiously with very few of them doing any meaningful dynamic analysis. The main problem is that the economy simply has too many variables to be able to isolate the changes from a particular policy.

Dynamic scoring has been studied by state agencies and LFC since 1995 and the results have been inconclusive. In the 90’s, the Department of Finance and Administration (DFA) and Taxation and Revenue Department (TRD) developed a state model that was not a practical model. In recent years, DFA, TRD, the Department of Transportation (DOT), and LFC contracted with an econometric consulting firm to develop dynamic scoring capabilities. Following presentations to the committee, there was no clear indication that the project should continue and there was no additional funding so at this point the state no longer has a contract with REMI to provide dynamic scoring modeling.

DFA has since begun to revisit the model, which DOT continues to use, to do economic impact analysis without trying to incorporate the dynamic fiscal analysis. The model, provided by Regional Economic Modeling, Inc (REMI), is a very sophisticated tool that can forecast the impacts of significant projects (i.e. a uranium enrichment facility) or policy changes (i.e. a major tax cut) on the economy. DFA, TRD and LFC have discussed ways to use those results to estimate the fiscal impacts and that conversation is ongoing.

Options:

- Resume contact with REMI for dynamic economic impact modeling and use consensus revenue group expertise to determine fiscal impacts. REMI costs about \$17,000 per year by extending the DOT contract or \$50,000 to \$75,000 to resume the old contract.
- Purchase a static input-output model such as IMPLAN or US Bureau of Economic Analysis Regional Input-Output Multipliers (RIMS) to do simple economic impact analysis and use consensus group expertise to determine fiscal impacts. IMPLAN would cost \$3,000 and a RIMS table would cost \$225 per region.
- Contract with outside consultants to analyze selected projects or legislation and report to committee. Cost would vary by project.
- Acknowledge in fiscal impact reports that there may be offsetting dynamic effects of proposed legislation but that there is too much uncertainty to adequately address and continue to provide static score only.

BACKGROUND REPORT ON DYNAMIC SCORING

History. Beginning in 1995, TRD and DFA contracted with the University of New Mexico's Economics Department for their experts to develop a computable general equilibrium model of the state's economy. This was a very complicated model that attempted to simulate the action of the state's economy through the use of mathematical equations and calculations. This project was interesting, but largely inconclusive. The agencies involved in its development were unable to maintain the model or use it in any practical way.

In 2003, the Legislature authorized the LFC, the Department of Finance and Administration (DFA), the Taxation and Revenue Department (TRD) and the Department of Transportation (DOT)—all members of the consensus revenue group—to conduct a two year pilot program of dynamic scoring. DFA requested proposals for an econometric model and the agencies chose Regional Economic Modeling Inc (REMI).

The 2004 session was the first session where the model was used. REMI consultants assisted with calibrating their Policy Insight model to NM's economy. Initially, there were problems with the calibration and the model did not seem to be able to produce credible results in a timely fashion. However, the model was used to demonstrate the dynamic effects of the 2003 personal income tax rate cuts and was included in the LFC Post-session Review in April of 2004.

The results of dynamic scoring of the fiscal impacts of the personal income tax rate cuts of 2003 were not significantly different from the static analysis performed by TRD and LFC (see attachment A). The dynamic forecast could show, however, employment, income and population effects of the legislation, variables which are not included in the static analysis.

In the 2005 session, there were no requests for analysis from the legislature on specific bills. One of the issues that came up and was cited by the Department of Transportation in the original FIR for HB 28 in 2003 was that it was unclear how to deal with substitute bills. The conference substitute of House Bill 410 in the 2005 Session was a combination of several different bills, many of which individually did not meet the \$10 million threshold but in combination exceeded the threshold. The timing of the substitute, towards the end of the session, also made any analysis impossible using REMI for use during the session.

In follow up meetings and trainings with REMI, it has become clear that REMI is not designed for *fiscal* impact analysis but more generally *economic* impact analysis. This means that it works well, with the caveat that significant training is required, for estimating the changes in economic output, employment, income, population and a host of other economic variables that may result from a particular policy choice. However, it does not do a good job analyzing the changes in specific tax revenue that may result from a particular policy choice. The fiscal impact model within the REMI model has not

proven to be useful to revenue estimators in New Mexico even with a high degree of calibration from REMI consultants.

Part of the enabling legislation required the LFC to evaluate the success of the pilot. LFC staff provided a brief to the committee for a hearing on October 25, 2007, which included many of the points made here. The hearing included presentations by James Jimenez, then-Secretary of DFA, and Jan Goodwin, Secretary of TRD. At that time, Secretary Jimenez reported that the “project was useful but less than hoped” and that if the committee wanted to continue with the project additional funding would have to be made available to DFA to support it. **There was no clear indication that the project should continue and there was no additional funding so at this point the state no longer has a contract with REMI to provide dynamic scoring modeling.**

Even though the pilot was not successful in providing the state with a usable model to add dynamic scoring to the toolbox of revenue analysts, there are options to consider. DFA is currently revisiting REMI and other tools to see if dynamic scoring can be resurrected. The initial assessment is that the REMI model has been substantially improved from the model used in the 2004 and 2005 sessions. The model is now accessible to state economists and could be used during the interim between legislative sessions to develop a library of plausible project types, such as an expansion or contraction of the national laboratories, a major production facility similar to the Uranium Reprocessing plant in Eunice, or a major expansion of the film industry in the state.

A modification to dynamic scoring can also be achieved with modest resources. An input-output model like IMPLAN (which is significantly less expensive than REMI and requires no subscription) or the Bureau of Economic Analysis’ RIMS data can be combined with the consensus revenue group’s knowledge of tax structure to create a consensus rule-of-thumb. Currently, revenue analysts use rules-of-thumb calculations to determine the static impacts of revenue legislation. These calculations may include the amount of income the average individual is likely to spend on taxable purchases. This calculation could be applied to the resulting income generated by an input-output model to estimate the amount of gross receipts tax the impact would generate. Similar models have been used by UNM’s Bureau of Business and Economic Research to evaluate industrial revenue bonds and by ImpactDataSource to estimate the impact of projects for the New Mexico Finance Authority. Since the DOT currently owns the REMI PolicyInsight model and will make it available to LFC and DFA analysts, this “rule-of-thumb” approach can move forward at no additional cost to the state, outside of the time analysts will take to specify and run scenarios.

Calculating dynamic impacts. There are several questions revenue analysts must ask when trying to estimate a particular policy change on the economy.

- Who is the target of the tax change? What industry or income group is likely to be impacted by the change?
- What is the purpose of the tax change? Is it to induce a specific company to locate or is it to keep existing companies?

- Will the policy influence behavior to either take advantage of the policy or to avoid the policy? If it is a rate increase, do taxpayers switch to activities not subject to the tax?
- Is it to provide incentives to invest or save?
- If a tax cut, how will it be paid for? If a tax increase, what will the proceeds be used for?
- What is the term of the change? One time or recurring? Phased-in?

The answer to each question is an assumption that will affect the results of the analysis. Many of these questions have to be answered in a static analysis as well but only to select the model to be used to provide an estimate. These assumptions in a dynamic scoring context become more complex as varying elasticities need to be estimated.¹ A personal income tax rate cut may have an identical static impact as a gross receipts tax rate cut but the impact on the economy will be very different. Who pays each tax is different and how a taxpayer responds is different.

Example of economic feedback using rule-of-thumb calculations. Assume a \$100 million rebate to taxpayers costs the government \$100 million under static scoring. However, if the \$100 million is spent in the economy by the recipients, there are multipliers that will lower the cost of the rebate because additional revenue will be generated by the economic activity. As the details show, the fiscal analysis relies on a set of key assumptions that are derived from various sources and usually are fairly blunt instruments.

Example: \$100 Million Tax Rebate

Static fiscal impact: \$100 million

Dynamic fiscal impact: \$ 94.5 million

Key assumptions for dynamic fiscal impact:

- 92.3% spent; 7.7% saved/invested
- 65% spent on taxable GRT
- Output multiplier = 1.5
- Income multiplier = .56 million income per \$1 million output
- Jobs multiplier = 15 jobs per million
- **Tax rebate is not offset by reduced government expenditures**

The above example shows that the cost of a tax rebate is less than the amount of the rebate when economic impacts are taken into account. The \$5.5 million in savings comes from the additional economic activity arising from the increased consumption of goods and services. Implicit in this calculation are statistics about jobs (1,384), income to New Mexicans (\$51.7 million) outside of the tax cut, and indirect output (\$46.15 million). Since this is a one time rebate, the analysis would be complete at this stage. If this were a

¹ Elasticity measures the responsiveness of one variable to a change in another. . For example, demand for cigarettes is largely thought to be elastic with respect to the price because an increase in the cigarette tax, which raises the price, will cause a decrease in the demand for cigarettes.

tax cut that amounted to \$100 million, it would “pay” for itself over time with increased economic activity. The payoff period would be about 18 years assuming a completely static economy, a relatively unrealistic assumption.

If the example were a reduction in taxes without expiration, the analysis would require some econometric model that can incorporate the change into a long run forecast. A reduction of \$100 million in personal income tax, for example, would pay for itself over time with additional economic activity, all other things remaining equal. The assumption that the rebate is not offset by reduced government expenditures is an important distinction: government expenditures have a multiplier and if those expenditures are reduced by \$100 million, the net cost may be greater than the \$100 million static cost.

An analysis of the 2003 tax cuts using the regional economic analysis model REMI under a pilot project showed the static fiscal impact was not materially different from the dynamic fiscal impact. The dynamic analysis showed a net loss of employment since it assumed that the revenues would be offset by lower expenditures. The pilot was considered by the consensus revenue group unsuccessful due to the sensitivity of assumptions like this.

The assumptions are the cause of most of the complexity and controversy. Assumptions are made that may not suit the situation. In the example above, a 1.5x multiplier was used which made the example simple to calculate. This is a rough average of the multiplier for all NM economic activity. If the rebate only went to low-income households, the multiplier might be slightly higher since more of it would be spent locally but also lower since more would be spent on low multiplier industries like retail. Also, the assumption about saving changes with the level of income: higher incomes save and invest more which generally is a leakage from the economy. Even the tax rates are subject to controversy since there are assumptions implicit about the level of deductions and credits, exemptions, etc, that are arguable.

Academic Debate on Dynamic Scoring. Dynamic scoring is a way of analyzing government policies to take into account economic impacts from changes in behaviors the policies create. In contrast to static scoring, which assumes no economy wide changes, dynamic scoring lets a policy work within the economy, follows the “ripples” and reports a cumulative impact. The value of such scoring is hotly contested. The federal government requires dynamic scoring and reports from the Joint Committee on Taxation have shown only modest feedback effects. Other studies have shown more significant feedbacks particularly on tax changes regarding capital investments. States have proceeded cautiously with very few of them doing any meaningful dynamic analysis. The main problem is that the economy simply has too many variables to be able to isolate the changes from a particular policy. Even a \$100 million personal income tax cut, assuming that it all stays in the state and is spent on taxable items, would only add 0.2 percent to taxable gross receipts in New Mexico. This amount is essentially lost in the normal fluctuations in taxable gross receipts.

Another problem is illustrated in the preceding sentences: assumptions about what a policy change will do directly influence the result. Even in static scoring, assumptions are constantly challenged: with dynamic scoring, the range of assumptions increases significantly. These two challenges hinder analysis on how a particular policy change will impact any part of the economy.

“There is no convincing evidence that either aggregate labor supply or saving responds in a significant way to taxes, and the evidence regarding business investment is mixed.”² – Joel Slemrod, PhD.

Slemrod in the same lecture presents a hierarchy of responses to a change in tax policy. The most significant response is in the timing of a taxable activity. For example, some of the federal tax stimulus package of 2003 was designed to encourage businesses to invest in equipment earlier rather than later. If the investment was already planned and simply moved forward to take advantage of the accelerated depreciation, there was no long term impact but rather a distortion that helped the short term economy. It is important to note that Slemrod is discussing impacts at the national level and not the state level. At the state level, there is significantly more mobility of labor and capital so responses are fundamentally different.

The second response in the hierarchy is avoidance, evasion or shifting. This happens when the policy causes businesses to shift from a taxable activity to a non- or lower-taxable activity. For example, a company may decide to hire a lawyer on staff as an employee rather than an outside legal firm to avoid paying gross receipts tax on the lawyer’s work. Another example may show up as a result of the recent personal income tax rate cuts: a firm may decide to organize as an S-corporation or limited partnership to avoid the higher corporate income tax rate (if the company has over \$500,000 in taxable income).

The final response and the source of the quotation above is the economic response through the labor supply and/or savings rate. Slemrod argues that even those this response is the least significant there is still *some* response and so by ignoring it through static analysis guarantees that the response is zero.

“There are two main arguments for dynamic scoring. Argument one is that it is better to have an imprecisely right answer that recognizes behavioral responses than a precisely wrong answer, the precisely wrong answer being that tax cuts or changes have absolutely *no* effect on the macro economy.”³

Four Examples of Behavioral Responses

1. Tax Rebate. A tax rebate will have a very short term impact on the economy, the size of which is determined by the size of the rebate. Recently, a rebate of \$107 million was sent to NM taxpayers with the average size of rebate was \$125 per

² Slemrod, Joel. “The Dynamic Tax Economist,” 11th Annual Erwin N. Griswold Lecture to the American College of Tax Counsel January 25, 2003.

³ Ibid.

- taxpayer. Some of the taxpayers likely spent the rebate on taxable items but it is also likely that a large number of taxpayers either saved the rebate and did not adjust their spending at all or spent the rebate on non-taxable transactions like rent. In the former case, the rebate would have a multiplier effect and create additional tax revenue that would offset the cost of the rebate. In the latter cases, there would be little to no revenue feedback.
2. PIT rate cuts. Responding to a personal income tax rebate changes by taxpayer as well. The latest rate cuts were primarily targeted at upper income brackets and so likely only had a modest impact on their consumption, which makes up the largest share of taxable transactions. One of the selling points of the rate cuts was that it would entice large companies to locate their headquarters here since their highest paid managers would have a lower tax rate. The data suggests that this has not happened.
 3. Working Families Credit. This credit is targeted at low income working families and the likely economic impact is greater since the savings rate is generally very low amongst lower income earners.
 4. Locomotive Fuel GRT deduction. This deduction is largely thought to have been passed in order to entice a single company to establish operations in New Mexico. The behavioral response in this case is the company moving here and quantifying the impact on the economy (and subsequent revenues) is probably the simplest of the examples since there is data on how much employees earn and how many will be hired as a result of the relocation.

Other Experience with Dynamic Scoring. Both other states and the federal government have experimented with dynamic scoring or analysis and the experiences have been mixed. California's experience with dynamic scoring is a familiar story. They set up a pilot in the 90's with one of the universities. After investing a significant sum of money into creating and compiling a model (they had their own proprietary model), they realized two things: it takes at least an FTE at a high salary to maintain and operate the model and the results were not meaningful to the policy makers. They discontinued use of the model in 2000 though they may start it up again in the near future.

There have been several studies on dynamic scoring at both the federal and state levels. At the federal level, the US Treasury, the Congressional Budget Office and the Joint Committee on Taxation have all researched, published and testified on how dynamic scoring works for the US budget process.

Recently, the US Treasury did an analysis of the impact of extending tax cuts enacted in 2001 and 2003.⁴ In the introduction, two potentially controversial assumptions are stated: that capital has limited international mobility and that resources are fully employed. Many economists think that there is a disconnect in the economy defined by markets and corporate profits and that economy which most people encounter and it has to do with the free flow of global capital. In other words, by making an assumption that capital has limited mobility, any assumption about capital investment will likely be overstated. The results show a net increase in GDP of about 0.7 percent over the long term and a 0.5

⁴ _____. "A Dynamic Analysis of Permanent Extension of the President's Tax Relief" United States Department of the Treasury Office of Tax Analysis, July 25, 2006.

percent increase in the 10 year budget window but those are based on the federal government lowering spending to pay for the tax cuts. Jason Furman, writing for Slate, says that the long term of 20 years means that the tax cuts added 0.04 percent annually to the economy, a relatively insignificant impact.⁵ The range of results using some sensitivity parameters suggests virtually no impact to a 1.2 percent impact on GDP. Furman also notes that tax revenues (i.e. the feedback effects that pay for the tax cuts) are not estimated.

The Congressional Budget Office makes a distinction between “dynamic scoring” that has a particular meaning in federal budgeting and “dynamic analysis,” which is what the CBO does. In a recent presentation to the American Economics Association, Benjamin Page, a CBO economist, reported that different models used to analyze federal tax policy yielded contradictory results.

The range of estimates produced under CBO’s different assumptions for forward-looking behavior highlights the difficulties in producing a single point estimate for dynamic scoring. Doing so would involve sorting through results that are in many cases very different in magnitude or even sign, with little firm basis to choose among them. (For example, estimates produced using CBO’s textbook growth model, which incorporates no forward-looking behavior, suggested that adopting the President’s proposals for fiscal year 2004 would have lowered GDP by an average of 0.2 percent over the following five years; estimates produced using CBO’s infinite-horizon model indicated those same proposals would have raised GDP by between 0.2 and 0.9 percent.) Any such process would by its nature be intensely controversial.⁶

Options. All is not lost with the dynamic scoring pilot. Below are a few options to provide decision makers with useful information about the economy wide impacts of major projects or policies.

- LFC and DFA could contract with REMI to resume use of the model and collaborate on fiscal impacts implied by the model output. The last estimate for cost was \$20,000 per year for sufficient copies for LFC and DFA.
- LFC and DFA could use IMPLAN to provide static economic impacts (output, jobs and compensation) and collaborate on fiscal impacts as described above. The estimated cost for two copies of the IMPLAN software and the data is approximately \$3,000 with an additional cost of \$1,700 for updating the data every two to three years.
- LFC and DFA could select major projects or legislation to analyze in detail during the interim culminating in a joint report. Some input-output model would still be required though inexpensive multipliers from the US Bureau of Economic Analysis could be used.

⁵ Furman, Jason. “The Last Laffer,” *Slate* (<http://www.slate.com/id/2146868>), July 31, 2006.

⁶ Page, Benjamin. “CBO’s Analysis of the Macroeconomic Effects of the President’s Budget,” Presentation to American Economic Association, 2005.

- LFC could select major projects or legislation and contract with an outside consultant such as UNM, NMSU or NMHU to prepare a report and presentation.
- An option worthy of consideration given the uncertainty and controversy over assumptions is to acknowledge that there may be offsetting dynamic effects but continue to only provide static impacts.