No. 10-75
December 2010

WORKING PAPER

ASSESSING THE QUALITY OF REGULATORY ANALYSIS: A New Evaluation and Data Set for Policy Research

By Jerry Ellig and John Morrall

MERCATUS CENTER
George Mason University

Assessing the Quality of Regulatory Analysis: A New Evaluation and Data Set for Policy Research

December 15, 2010

Principal Investigators

Jerry Ellig
Senior Research Fellow
Mercatus Center at George Mason University
3301 N. Fairfax Dr., #450
Arlington, VA 22201
jellig@gmu.edu

John Morrall
Affiliated Senior Scholar
Mercatus Center at George Mason University
3301 N. Fairfax Dr., #450
Arlington, VA 22201
jmorrall3@aol.com

Research Team*

Mark Adams
Jarrod Anderson
David Bieler
Katelyn Christ
Christina Forsberg
Stefanie Haeffele-Balch
Allison Kasic
Ben Klutsey
Gabriel Okolski
Tate Watkins

* In addition to the research team that helped us evaluate agency Regulatory Impact Analyses and *Federal Register* preambles, we gratefully acknowledge research assistance from Dan Claybaugh, Gwen D'Luzansky, and Krista Speicher. We thank Randy Lutter and participants at the Society for Government Economists and Society for Risk Analysis meetings in 2010 for helpful comments and discussions on earlier drafts of this paper.

Assessing the Quality of Regulatory Analysis: A New Evaluation and Data Set for Policy Research

Abstract

Congress and the executive branch have attempted to improve the quality of regulatory decisions by adopting laws and executive orders that require agencies to analyze benefits and costs of their decision options. This paper assesses the quality and use of regulatory analysis accompanying every economically significant regulation proposed by executive-branch regulatory agencies in 2008 and 2009. It considers all analysis relevant to the topics covered by Executive Order 12866 that appears in the Regulatory Impact Analysis document or elsewhere in the *Federal Register* notice that proposes the rule.

Our research team used a six-point qualitative scale to evaluate each regulation on 12 criteria grouped into three categories: (1) Openness: How easily can a reasonably intelligent, interested citizen find the analysis, understand it, and verify the underlying assumptions and data? (2) Analysis: How well does the analysis define and measure the outcomes the regulation seeks to accomplish, define the systemic problem the regulation seeks to solve, identify and assess alternatives, and evaluate costs and benefits?; and (3) Use: How much did the analysis affect decisions in the proposed rule, and what provisions did the agency make for tracking the rule's effectiveness in the future?

We find that the quality of regulatory analysis is generally low, varies widely, and did not change much with the change of administrations between 2008 and 2009. The principal improvements across all regulations occurred on the Openness criteria. Budget or "transfer" regulations, which define how the federal government will spend money or collect revenues, have much lower-quality analysis than other regulations. Use of analysis is correlated with its quality, and use of analysis fell in 2009 after controlling for the quality of the analysis. Regulations implementing Recovery Act spending programs have better provisions for retrospective analysis than other transfer regulations.

Keywords: regulatory impact analysis, benefit-cost analysis, regulatory review, regulation

JEL categories: D61, D73, D78, H11, H83, K23, L51, P16

Introduction

For nearly four decades, presidential administrations have required executive-branch agencies to conduct some type of economic impact analysis when they issue major regulations. Since 1993, President Clinton's Executive Order 12866 has laid out the fundamental analytical steps agencies must take. The very first section of the executive order states that agencies must identify the problem they are trying to address and assess its significance, examine a wide range of alternatives to solve the problem, assess the costs and benefits of the alternatives, and choose to regulate only when the benefits justify the costs. Analytical requirements are especially rigorous for "economically significant" regulations, defined as regulations that "have an annual effect on the economy of $100 million or more or adversely affect in a material way the economy, a sector of the economy, productivity, competition, jobs, the environment, public health or safety, or state, local or tribal government or communities" (EO 12866, Sec. 2(f)(1)). Office of Management and Budget (OMB) Circular A-4, issued in September 2003, offered more detailed guidance on "best practices" in regulatory analysis (OMB 2003).

Despite executive orders and detailed guidance, the quality of agencies' regulatory analysis has been inconsistent at best:

- Several studies compared agencies' ex-ante predictions of regulatory benefits and costs with ex-post estimate of actual benefits and costs (Harrington et. al. 2000, OMB 2005, Harrington 2006). These studies found that, in the past, ex-ante estimates tended to overestimate both benefits and costs.

- In a series of papers, Robert Hahn developed and applied a yes/no checklist to evaluate whether agencies' Regulatory Impact Analyses have included a series of major elements that OMB expects them to include. The evaluations focused on final regulations issued by health, safety, and environmental agencies (Hahn and Dudley 2007, Hahn et. al. 1990, Hahn and Litan 2005, Hahn, Lutter, and Viscusi 2000). Surveying the evidence, Hahn and Tetlock (2008, 82–83) conclude that economic analysis has not had much impact, and the general quality of regulatory analysis is low. "Nonetheless," they note, "in a world where regulatory impacts are frequently measured in the billions of dollars, margins matter. Thus, economists should pay more attention to how economic analysis can contribute to improving benefits and costs on the margin."

- Belcore and Ellig (2008) employed a qualitative scoring approach to assess the quality of regulatory analysis at the Department of Homeland Security during its first five years; they conclude these analyses have been seriously incomplete but improved over time.

Most recently, Ellig and McLaughlin (2010) developed a 12-point qualitative framework to assess both the quality and use of regulatory analysis in federal agencies. They evaluated the quality and use of regulatory analyses of "economically significant" rules that were reviewed by OMB's Office of Information and Regulatory Affairs (OIRA) in 2008 and proposed in the

Federal Register.[1] The evaluation criteria are drawn from Executive Order 12866, OMB Circular A-4, and pre-existing scholarship on regulatory scorecards.[2] Ellig and McLaughlin found that the average quality of the 2008 regulatory analyses is low, both the quality and use of regulatory analysis vary widely, and there are significant opportunities for improvement through the diffusion of best practices. They also found that better analyses are more likely to be used in agency decisions, but only one-fifth of the regulatory analyses in 2008 appeared to have any effect on regulatory decisions (based on information agencies supplied in the preamble).

This study utilizes the Ellig and McLaughlin method to evaluate the quality and use of regulatory analysis for economically significant regulations proposed by executive-branch agencies in 2009. This is of interest for several reasons. First, a comparison of 2008 and 2009 would help identify whether the change of presidential administrations had any effect on the quality or use of regulatory analysis. Second, the Obama administration proposed in February 2009 to revise Executive Order 12866 (OMB 2009a); evaluating the quality and use of regulatory analysis in the Obama administration prior to the revision establishes a baseline to gauge the effects of any changes. Third, extending the evaluation to 2009 and subsequent years builds a larger data set, which may allow us to draw more reliable general inferences about the relative quality of analysis at different agencies or for different types of regulations.

Our principal findings include:

Quality is mostly unchanged in 2009. The average score for regulations proposed in 2008 and 2009 was virtually the same—27 points out of a possible 60. The most significant improvements occurred on Openness criteria, such as online accessibility of regulatory analyses and clarity. On average, explanations of how regulatory costs affect prices of goods and services also improved. Very modest improvements occurred in evidence of regulatory benefits and analysis of the distribution of benefits.

Analysis is less-widely used in 2009. Higher-quality analysis is more likely to be used in regulatory decisions. But for any given level of quality, regulatory agencies were less likely to use the analysis in 2009 than in 2008. This change is disturbing, because one of the most important reasons for doing regulatory analysis is so that decision makers can somehow use it to make better decisions. Of course, good regulatory analysis is also important for reviewers (like OMB) and stakeholders.

Quality is generally low. In both years, the average score is less than half of the possible 60 points. The highest-scoring regulation in 2008 earned 43 out of 60 possible points, equivalent to a grade of C. The highest-scoring regulation in 2009 earned 48 out of 60 possible points, equivalent to a B–.

---

[1] Economically significant regulations require an extensive Regulatory Impact Analysis (RIA) that assesses the need, effectiveness, benefits, costs, and alternatives for the proposed regulation. (EO 12866 Sec. 6(a)(3)(C))

[2] The qualitative evaluation method is based on the Mercatus Center's *Performance Report Scorecard*, a 10-year project that assessed the quality of federal agencies' annual performance reports required under the Government Performance and Results Act of 1996. For the most recent results, see McTigue et. al. (2009).

Diffusion of best practices could generate substantial improvement. In 2009, scores ranged from a high of 48 points to a low of just 3 points. In 2008, scores ranged from a high of 43 points to a low of 7 points. For each of our 12 criteria, at least one regulation earned the highest possible score of 5. But for 11 of our 12 criteria, less than a handful of regulations receive a 5. The fact that the highest-scoring regulation in 2009 resulted from collaboration between two agencies also suggests wider sharing of best practices can improve regulatory analysis.

Transfer regulations have worse analysis. Budget or "transfer" regulations, which determine how the federal government will spend or collect money, receive much lower scores. On average, transfer regulations received only 17 points in 2008 and 20 points in 2009, compared to an average of 32–34 points for non-transfer regulations.

Greatest strength: Accessibility on the Internet. Scores on this criterion averaged 4.06 out of 5 possible points in 2009 and 3.53 out of 5 possible points in 2008. These far exceeded average scores on any other evaluation criterion.

Greatest weaknesses: Retrospective analysis and identification of systemic problem. Few regulations or analyses set goals, establish measures, or provide for data gathering to assess the effects of the regulation after it is implemented. Few analyses provide a coherent theory and empirical evidence of a market failure, government failure, or other systemic problem the regulation is supposed to solve.

1. Evaluation Protocol

We evaluated the quality and use of regulatory analysis using 12 criteria grouped into three categories—Openness, Analysis, and Use:

1. Openness: How easily can a reasonably intelligent, interested citizen find the analysis, understand it, and verify the underlying assumptions and data?

2. Analysis: How well does the analysis define and measure the outcomes or benefits the regulation seeks to accomplish, define the systemic problem the regulation seeks to solve, identify and assess alternatives, and evaluate costs and benefits?

3. Use: How much did the analysis affect decisions in the proposed rule, and what provisions did the agency make for tracking the rule's effectiveness in the future?

Figure 1 lists the 12 criteria. Appendix 1 provides additional detail on the kinds of questions considered under each criterion. For a more extensive explanation and justification of this evaluation method, see Ellig and McLaughlin (2010). Individual "Report Cards" showing all scores and scoring notes for each regulation are available at www.mercatus.org/reportcard.

Ten of the 12 evaluation criteria closely parallel the Regulatory Impact Analysis checklist released by the Obama administration on November 3, 2010 (OMB 2010). This is not surprising, since both the administration's checklist and the Mercatus evaluation criteria are based on Executive Order 12866 and OMB Circular A-4. Appendix 2 presents a crosswalk chart comparing the OMB checklist with the 12 criteria used in this paper.

The principal Mercatus evaluation criteria not mentioned in the Obama administration's checklist are two criteria that assess whether the agency provided for retrospective analysis of the regulations' actual effects after it is adopted: criterion 11 (Measures and Goals) and criterion 12 (Retrospective Data). Although ex post, retrospective analysis has not received as much attention as ex ante analysis of proposed regulations; section 5 of Executive Order 12866 states that agencies should conduct retrospective analysis. OMB (2005) has recommended it repeatedly; most recently, OMB (2009b, 45) stated, "[W]e recommend that serious consideration be given to finding ways to employ retrospective analysis more regularly, in order to ensure that rules are appropriate, and to expand, reduce, or repeal them in accordance with what has been learned." The Government Performance and Results Act arguably requires retrospective analysis of regulations (Brito and Ellig 2009). It is a major area of regulatory analysis where the United States lags other industrialized nations (OECD 2009, 92).

Figure 1: Regulatory Analysis Assessment Criteria

Openness

1. **Accessibility:** How easily were the Regulatory Impact Analysis, the proposed rule, and any supplementary materials found online?
2. **Data Documentation:** How verifiable are the data used in the analysis?
3. **Model Documentation:** How verifiable are the models and assumptions used in the analysis?
4. **Clarity:** Was the analysis comprehensible to an informed layperson?

Analysis

5. **Outcomes:** How well does the analysis identify the desired benefits or other outcomes and demonstrate that the regulation will achieve them?
6. **Systemic Problem:** How well does the analysis identify and demonstrate the existence of a market failure or other systemic problem the regulation is supposed to solve?
7. **Alternatives:** How well does the analysis assess the effectiveness of alternative approaches?
8. **Benefit-Cost Analysis:** How well does the analysis assess costs and compare them with benefits?

Use

9. **Some Use of Analysis:** Does the preamble to the proposed rule or the Regulatory Impact Analysis present evidence that the agency used the analysis?
10. **Cognizance of Net Benefits:** Did the agency maximize net benefits or explain why it chose another option?
11. **Measures and Goals:** Does the proposed rule establish measures and goals that can be used to track the regulation's results in the future?
12. **Retrospective Data:** Did the agency indicate what data it will use to assess the regulation's performance in the future and establish provisions for doing so?

Scoring Standards

For each criterion, the evaluators assigned a score ranging from 0 (no useful content) to 5 (comprehensive analysis with potential best practices). Thus, each analysis has the opportunity to earn between 0 and 60 points. In general, the research team used the guidelines in table 1 for scoring. Because the Analysis criteria involve so many discrete aspects of regulatory analysis, we developed a series of sub-questions for each of the four Analysis criteria and awarded a 0–5 score for each sub-question. These scores were then averaged to calculate the score for the individual criterion.

Table 1: What Do the Scores Mean?

5	**Complete analysis of all or almost all aspects, with one or more "best practices"**
4	**Reasonably thorough analysis of most aspects and/or shows at least one "best practice"**
3	**Reasonably thorough analysis of some aspects**
2	**Some relevant discussion with some documentation of analysis**
1	**Perfunctory statement with little explanation or documentation**
0	**Little or no relevant content**

Caveats and Qualifications

At the outset of this project, we had to address a seemingly simple question: What counts as a "regulatory analysis"? Most previous research focuses on the document required by OMB that is explicitly named the "Regulatory Impact Analysis" (Hahn and Dudley 2007, Hahn et. al. 1990, Hahn and Litan 2005, Hahn, Lutter, and Viscusi 2000). We adopted a broader definition that includes the entire preamble to the proposed rule, the freestanding document or section of the preamble labeled Regulatory Impact Analysis, and additional "technical support documents" that sometimes accompany a Regulatory Impact Analysis. Since different agencies organize their material in different ways, this approach helped ensure that we were fair to all agencies and included all material relevant to the topics a good regulatory analysis is supposed to address. We also needed to read the entire preamble to assess whether the agency used the results of the regulatory analysis or made provisions to conduct retrospective analysis in the future.

Given resource constraints, any evaluation project like this faces a fundamental choice between breadth and depth of the assessment. We assess whether the Regulatory Impact Analysis and preamble to the proposed rule make a reasonable effort at covering the major elements of regulatory analysis. Commenters on earlier versions of this paper who have detailed knowledge of particular regulations have usually told us that our evaluations seem too lenient. Others with more specialized knowledge will likely have additional important critiques of individual regulations, especially related to the quality, completeness or use of the underlying science. We have opted for less depth in favor of greater breadth. To the best of our knowledge, this is the

most-detailed assessment of the quality of regulatory analysis for all economically significant regulations proposed in a two-year period.

Finally, we caution the reader about drawing direct policy conclusions about particular regulations based on our analysis. Criteria 1–8 only evaluate the quality of regulatory analysis. We do not evaluate whether the proposed rule is economically efficient, fair, or otherwise good public policy.

The same caveat applies to the Use criteria. Criteria 9 and 10 assess the extent to which analysis of the regulation's outcomes or benefits, the systemic problem, the alternatives, and costs informed the agency's decisions about the regulation. On these criteria, we took great pains to avoid imposing the value judgment economists often make: that the agency should choose the most economically efficient alternative, as determined by a comparison of quantified benefits and costs. If an agency used some analysis of a regulation's benefits to make decisions, even if it did not consider costs or efficiency, it could receive some points on criterion 9. Similarly, if an agency demonstrated that it was fully cognizant of the net benefits of alternatives, but explicitly rejected the alternative with the greatest net benefits in favor of some other alternative for clearly articulated reasons, it could receive points on criterion 10. As a result, an agency can earn points on these two criteria even in cases where it is prohibited by law from considering costs, such as the EPA's national ambient air quality standards. We believe this approach is consistent with the spirit of Executive Order 12866 (sec. 1), which identifies multiple factors in addition to efficiency that are supposed to guide agency decisions: "[I]n choosing among regulatory approaches, agencies should select those approaches that maximize net benefits (including potential economic, environmental, public health and safety, and other advantages; distributive impacts; and equity), unless a statute requires another regulatory approach."

Criteria 11 and 12 assess the extent to which the agency demonstrated its willingness to evaluate the regulation's actual effects in the future. Ideally, agencies would articulate goals, measures, and data that they could use to assess both realized benefits and costs, thus assessing the regulation's economic efficiency. In practice, so few regulations include any provisions for retrospective analysis that the handful of high scores occur in cases where agencies have at least identified goals, measures, and data that could be used to assess the regulation's effectiveness.

Improving the transparency of regulatory documents and the quality of regulatory analysis are necessary but not sufficient to improve public policy. Nevertheless, stakeholders or the agencies themselves may find these analyses useful as a starting point for identifying weaknesses in agency analyses. For example, if an agency has identified only one or two closely related regulatory alternatives, stakeholders may be able to identify additional alternatives that may accomplish the goal at a lower cost.

2. Results for 2009

2.1 Best and Worst Analyses

Table 2 lists all 42 economically significant proposed regulations for 2009. The best analysis was for the combined Environmental Protection Agency–Department of Transportation regulation on greenhouse gases from light-duty vehicles and Corporate Average Fuel Economy (CAFE) standards. This regulation received the highest total score (48 points) as well as the highest Analysis score (18 points). The two agencies collaborated on developing the regulation and the analysis. The regulatory analysis discusses the "conundrum" associated with the identified market failure. The agencies recognize that their estimates of the private benefits of increased fuel efficiency outweigh private costs, yet consumers do not voluntarily purchase as many fuel-efficient cars as economic rationality would suggest. This sort of disclosure should prove invaluable to stakeholders who wish to comment more extensively on the merits of the rule that requires increases in fuel efficiency. The result suggests that more extensive sharing of best practices could improve the quality of regulatory analysis. This regulation received a score six points higher than the next-best regulation in 2009 and five points higher than DOT's CAFE regulation in 2008.

Capturing second place in 2009 are three energy-efficiency regulations from the Department of Energy and the Department of Homeland Security's regulation limiting concentrations of live organisms permitted in discharged ballast water from ships.

The three worst analyses came from the Department of Education (General and Non-Loan Programmatic Issues, 14 points) and the Department of Energy (Weatherization Assistance, 10 points; Loan Guarantees for Projects that Employ Innovative Technologies, 5 points). Like most of the low-ranking regulations, all three of these are budget or "transfer" regulations. Transfer regulations, italicized in table 2, outline how the federal government will spend money, set fees, or administer spending programs. Most of these regulations score poorly, continuing a trend observed in 2008 (Ellig and McLaughlin 2010, 14–15).

The best analysis in 2009 received 48 points, or 80 percent of the maximum possible score. The worst received just five points (8 percent). The range of scores widened compared to 2008. In 2008, scores ranged from seven points to 43 points. If these were student papers, the best one in 2009 would have received a B-, and the best one in 2008 would have received a C.

2.2 Summary Statistics

Table 3 summarizes average total scores and scores on the three categories of criteria for 2008 and 2009. The average score in 2009 was 27.02 points out of a possible 60, or 45 percent. The average for 2008 was 27.31, virtually the same. The very low t-statistic indicates that the difference is not statistically significant; for all practical purposes, the averages are the same.[3]

[3] In plain English, that means the total scores for 2008 and 2009 are like two sets of ping pong balls pulled at random out of the same bucket; any difference in the averages is random chance. There is likely no difference at all between the total scores for the two years.

Table 2: Scores for 2009 Proposed Regulations

Proposed Rule	RIN	Department	Total	Openness	Analysis	Use
Greenhouse Gases from Light-Duty Vehicles	2060-AP58	DOT/EPA	48	15	18	15
Energy Conservation: Small Electric Motors	1904-AB70	DOE	42	16	14	12
Energy Efficiency Standards for Commercial Clothes Washers	1904-AB93	DOE	40	14	14	12
Energy Efficiency Standards for Pool Heaters etc.	1904-AA90	DOE	40	14	14	12
Living Organisms in Ships' Ballast Water Discharged in U.S. Waters	1625-AA32	DHS	40	15	15	10
Nutrition Labeling of Single-Ingredient Products	0583-AC60	USDA	38	14	16	8
Title V Greenhouse Gas Tailoring Rule	2060-AP86	EPA	38	15	11	12
Emissions From New Marine Compression-Ignition Engines	2060-AO38	EPA	37	15	16	6
Portland Cement NESHAP	2060-AO15	EPA	35	17	11	7
Greenhouse Gas Mandatory Reporting Rule	2060-AO79	EPA	34	12	10	12
Migratory Bird Hunting	1018-AW31	Interior	34	13	12	9
Emission Standards, Reciprocating Internal Combustion Engines	2060-AP36	EPA	33	14	11	8
End Stage Renal Disease Prospective Payment System	*0938-AP57*	*HHS*	*32*	*13*	*9*	*10*
Lead; Opt-out and Recordkeeping Provisions	2070-AJ55	EPA	32	16	13	3
Primary National Ambient Air Quality Standard for Nitrogen Dioxide	2060-AO19	EPA	32	11	14	7
Motor Vehicle Safety Standards, Ejection Mitigation	2127-AK23	DOT	31	12	11	8
School Improvement Grants	*1810-AB06*	*ED*	*31*	*11*	*7*	*13*
Primary National Ambient Air Quality Standard for Sulfur Dioxide	2060-AO48	EPA	30	12	12	6
Medical Examination of Aliens	0920-AA26	HHS	28	14	12	2
Positive Train Control	2130-AC03	DOT	26	10	7	9
Prospective Payment Skilled Nursing Facilities	*0938-AP46*	*HHS*	*26*	*11*	*8*	*7*
Electronic Health Record Incentive Program	*0938-AP78*	*HHS*	*25*	*13*	*7*	*5*
Home Health Prospective Payment System	*0938-AP55*	*HHS*	*25*	*11*	*8*	*6*
Prospective Payment System for Inpatient Rehabilitation Facilities	*0938-AP56*	*HHS*	*25*	*15*	*5*	*5*
Hospital Inpatient and Long-Term Care Prospective Payment System	*0938-AP39*	*HHS*	*24*	*14*	*5*	*5*
Hazard Communications Standard	1218-AC20	DOL	24	13	7	4
Outpatient Prospective Payment	*0938-AP41*	*HHS*	*24*	*13*	*6*	*5*
Race to the Top Fund	*1810-AB07*	*ED*	*23*	*9*	*5*	*9*
Revisions to Payment Policies Under the Physician Fee Schedule	*0938-AP40*	*HHS*	*23*	*11*	*6*	*6*
State Fiscal Stabilization Fund Program	*1810-AB04*	*ED*	*23*	*13*	*5*	*5*
Renewable Fuels Program	2060-A081	EPA	21	11	6	4
Special Community Disaster Loans Program	*1660-AA44*	*DHS*	*20*	*11*	*6*	*3*
Investing in Innovation	*1855-AA06*	*ED*	*19*	*11*	*4*	*4*
Hospice Wage Index for FY 2010	*0938-AP45*	*HHS*	*18*	*9*	*4*	*5*
Housing Trust Fund Program	*2506-AC23*	*HUD*	*18*	*10*	*3*	*5*
Revisions to the Medicare Advantage Program	*0938-AP77*	*HHS*	*18*	*9*	*4*	*5*
Credit Assistance for Surface Transportation Projects	*2105-AD70*	*DOT*	*17*	*11*	*5*	*1*
Expansion of Enrollment in the VA Health Care System	*2900-AN23*	*VA*	*17*	*11*	*3*	*3*
Children's Health Insurance Program (CHIP)	*0938-AP53*	*HHS*	*15*	*8*	*1*	*6*
General and Non-Loan Programmatic Issues	*1840-AC99*	*ED*	*14*	*8*	*2*	*4*
Weatherization Assistance Program	*1904-AB97*	*DOE*	*10*	*6*	*3*	*1*
Loan Guarantees for Projects that Employ Innovative Technologies	*1901-AB27*	*DOE*	*5*	*3*	*2*	*0*
Averages			27.02	12.00	8.38	6.64

Note: Regulations in red italics are budget or “transfer” regulations.

Some slight shifts in scores may have occurred in two of the categories between 2008 and 2009. The average Analysis score was largely unchanged. The average Openness score increased by about one point–from 11.04 in 2008 to 12 in 2009. The average Use score fell by about a point, from 7.73 in 2008 to 6.64 in 2009. These differences are statistically significant at the 85 percent confidence level. This is suggestive, but not nearly as strong an indicator as the 95 percent confidence level economists normally use as the standard to infer a likely relationship. Based on this comparison of averages for all kinds of regulations, perhaps the transparency of regulatory analysis increased in 2009, and actual use to make decisions may have decreased, but the difference is not clear enough to tell for sure.

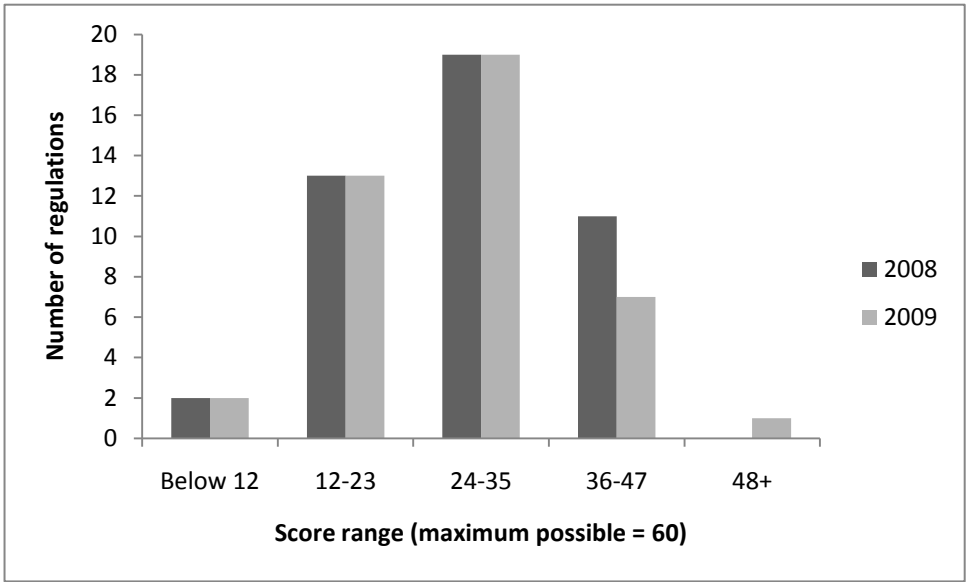
Figure 2 shows that the distribution of scores was roughly the same in both years. The only differences are that the joint DOT/EPA regulation received a score of 48 in 2009, and several more regulations in 2008 received scores in the 36–47 range.

Table 3: Average Scores, 2008 vs. 2009

	2008 (n=45)	2009 (n=42)	Change	T-stat.
Total Score	27.31	27.02	–0.29	0.14
Openness	11.04	12.00	0.96	1.46
Analysis	8.53	8.38	–0.15	0.16
Use	7.73	6.64	–1.09	1.48

Maximum possible total score = 60. Maximum possible score on each category = 20.

Figure 2: Distribution of Scores



2.3 Average Scores by Criterion

Table 4 shows the average score for each criterion in 2008 and 2009. For each criterion, at least one regulation earned the highest possible score of 5 in most cases. Best practices, however, are not widely shared. The "# Earning Highest Score" column demonstrates that, except for Availability, very few regulations earn a score of 5 on any individual criterion. The "Theoretical Highest Score" is the score a hypothetical regulation could have earned if it had incorporated all of the best practices identified that year. For 2009, the highest-scoring regulation is much closer to the theoretical highest score than in 2008.

Table 4: Scores by Criterion

Criterion	2008 Average Score	2008 Highest Score	2008 # Earning Highest Score	2009 Average Score	2009 Highest Score	2009 # Earning Highest Score
1. Accessibility	3.53	5	12	4.06	5	14
2. Data Documentation	2.24	5	1	2.50	5	5
3. Model Documentation	2.33	5	3	2.62	5	1
4. Clarity	2.93	5	3	2.83	4	10
5. Outcome Definition	2.36	5	2	2.38	5	1
6. Systemic Problem	1.80	5	1	1.60	4	4
7. Alternatives	2.29	5	1	2.21	5	1
8. Benefit-Cost Analysis	2.09	4	3	2.19	5	1
9. Some Use of Analysis	2.44	5	2	2.24	5	1
10. Considered Net Benefits	2.20	5	2	1.62	5	4
11. Measures and Goals	1.36	5	1	1.29	4	1
12. Retrospective Data	1.73	5	1	1.50	4	2
Total	27.31	43		27.02	48	
Theoretical Highest Score*		59			56	

Very few of the score changes between 2008 and 2009 are statistically significant.[4] Moreover, changes in averages for some criteria appear to be driven by the changing mix of regulations rather than an actual change in the quality of agencies' analysis. An accurate assessment of changes, therefore, requires separate consideration of transfer and non-transfer regulations.[5]

[4] Summary statistics for all criteria, and the sub-questions for criteria 5–8, are in appendix 3.
[5] Statistically significant changes in averages for the entire set of regulations, without distinguishing between transfer and non-transfer regulations, are in appendix 4.

2.4 Transfer vs. Non-Transfer Regulations

Several previous studies using 2008 data, as well as table 2, demonstrate that the quality and use of analysis for transfer regulations is well below the quality and use of analysis for non-transfer regulations (Ellig and McLaughlin 2010, McLaughlin and Ellig 2010). Indeed, OMB (2008, 12–17) observes that although transfer regulations generate social costs via mandates, prohibitions, and price distortions, agencies do not usually estimate the social benefits and costs of transfer regulations.

Table 5 confirms that the quality and use of analysis for transfer regulations is much lower in both 2008 and 2009. In 2008, for example, the average total score for transfer regulations (17 points) is 47 percent below the average score for non-transfer regulations (32 points). Similarly, in 2009 the average total score for transfer regulations (21 points) is 40 percent below the average total score for non-transfer regulations (34 points). These differences occur for Openness, Analysis, and Use. Openness has the smallest gap, but even there, transfer regulations score 20–30 percent below non-transfer regulations.

Table 5: Transfer vs. Non-Transfer Regulations, Average Scores

	Transfer 2008 (n=15)	**Non-Transfer 2008** (n=30)	**Difference**	**T-stat.**
Total Score	17.07	32.43	15.37	8.03
Openness	8.6	12.27	3.67	4.16
Analysis	3.53	11.03	8.53	8.71
Use	4.93	9.13	4.20	4.99
	Transfer 2009 (n=22)	**Non-Transfer 2009** (n=20)	**Difference**	**T-stat.**
Total Score	20.54	34.15	13.65	6.84
Openness	10.5	13.65	3.15	4.32
Analysis	4.91	12.20	7.29	8.9
Use	5.14	8.3	3.16	3.18

All differences are statistically significant at greater than the 99 percent level of confidence. Maximum possible total score = 60. Maximum possible score on each category = 20.

Because transfer regulations generally receive lower scores, a shift in the mix of transfer vs. non-transfer regulations could affect changes in average scores from one year to the next. In 2008, there were 15 proposed economically significant transfer regulations, accounting for 33 percent of proposed economically significant regulations. In 2009, there were 22 proposed economically significant transfer regulations, accounting for 52 percent of proposed economically significant regulations. The increase mostly reflects five regulations proposed in 2009 that implemented provisions of the American Recovery and Reinvestment Act. Thus, one might expect that the average quality and use of regulatory analysis would be lower in 2009 than in 2008 simply because more transfer regulations were proposed in 2009.

Table 6: Score Changes on Individual Criteria and Questions, Transfer vs. Non-Transfer Regulations

	2008 (n=30)	**2009** (n=20)	**Change**	**T-stat.**
Non-Transfer Regulations				
Total Score	32.43	34.15	1.72	0.94
Openness	12.27	13.65	1.38	1.91*
Criterion 1 – Availability	3.30	3.95	0.65	1.69*
Criterion 2 – Data Documentation	2.63	3.15	0.52	1.66*
Criterion 3 – Theory and Model Documentation	2.83	3.30	0.47	1.49
Analysis	11.03	12.20	1.17	0.20
Criterion 5 – Outcomes	3.10	3.55	0.45	1.63
Question 5D – Evidence Regulation Will Affect Outcome	2.40	3.15	0.75	1.88*
Criterion 8 – Cost-Benefit Analysis	2.60	3.10	0.5	2.15**
Question 8C – Effects on Prices of Goods and Services	1.70	3.30	1.60	3.91***
Question 8G – Calculates Cost-Effectiveness	1.43	2.35	0.92	2.35**
Question 8I – Incidence of Benefits	2.07	2.95	0.88	2.33**
Use	9.13	8.3	–0.83	0.35
Transfer Regulations				
Total Score	17.07	20.55	3.48	1.70*
Openness	8.60	10.50	1.90	2.11**
Criterion 3 – Theory and Model Documentation	1.33	2.00	.67	1.88*
Criterion 4 – Clarity	1.80	2.45	.65	2.37**
Analysis				
Criterion 5 – Outcomes	0.87	1.31	0.45	1.61
Question 5A – Articulate Desired Outcome	1.80	2.45	0.65	1.52
Question 5D – Evidence Regulation Will Affect Outcome	0.20	1.00	0.80	2.86***
Criterion 6 – Systemic Problem	0.60	1.00	0.40	1.79*
Question 6B – Coherent Theory of Systemic Problem	0.47	0.86	0.40	1.64
Question 7A – List Alternatives	1.07	1.91	0.84	2.18**
Criterion 8 – Cost-Benefit Analysis	1.07	1.36	0.30	1.51
Use	4.93	5.14	0.20	0.83

Statistical significance: *90 percent ** 95 percent
Maximum possible score on individual criteria or questions = 5.

Table 6 shows changes in mean scores calculated separately for transfer and non-transfer regulations. We report statistics for individual criteria or questions only when the differences approach statistical significance.

For non-transfer regulations, there are very few improvements. Average Openness scores improved from 12.27 points to 13.65 points. The difference is almost statistically significant at the 95 percent level. Within the Analysis category, there is weak evidence of improvement on criterion 5 (Outcomes), largely because agencies provided more evidence that the regulation will accomplish the intended outcomes. Criterion 8 (Cost-Benefit Analysis) also saw improvement due to better scores on three questions: question 8C (Effects on Prices of Goods and Services), question 8G (Evaluation of Cost-Effectiveness) and question 8I (Incidence of Benefits). These changes are consistent with the administration's goals of improving the transparency of the regulatory process, identifying benefits of regulation, and expanding the focus on distributional issues. We caution, however, that the changes are quite small, and the improvements under the Analysis category mostly just move the average scores closer to 3.

Transfer regulations show slightly more improvement than non-transfer regulations. The average Openness score improved, largely due to increases in scores on criterion 3 (Theory and Model Documentation) and criterion 4 (Clarity). The improvement on criterion 4 is actually significant at the 98 percent level. All four Analysis criteria saw higher average scores in 2009 than in 2008. However, all of these scores remained well below 2 in 2009. This indicates only that more analyses presented a small amount of discussion or evidence relevant to these criteria instead of saying nothing. While these improvements are certainly welcome, the low levels of the scores indicate that analysis of transfer regulations has a long way to go before it is as good as the analysis of non-transfer regulations.

We draw the following conclusions from this breakdown between transfer and non-transfer regulations:

- The only category of criteria that appears to have improved for both transfer and non-transfer regulations is Openness.

- The few improvements in the Analysis criteria for non-transfer regulations seem consistent with the Obama administration's regulatory priorities.

- Improvements in some of the Analysis criteria for transfer regulations largely reflect the presence of some content or assertions where previously there were none.

- Regulators made little commitment to retrospective analysis of regulations proposed in either year.

2.5 Total Scores by Agency

Another way to control for factors that might affect the average quality or use of regulatory analysis is to break scores down by agency. Some agencies may do a better job of

analysis than others. Some may tackle analytical problems that are inherently more difficult. Yet others may have different mixes of transfer regulations and non-transfer regulations. Table 7 presents average scores by agency for 2008 and 2009, with and without transfer regulations.

When all regulations are included, five agencies increased their average total scores in 2009, and five agencies reduced their average total scores. When transfer regulations are excluded, four agencies increased their average total scores in 2009, and four agencies reduced their average total scores. Given that most agencies proposed small numbers of economically significant regulations, few agencies proposed comparable numbers of economically significant regulations in both years, and six agencies proposed economically significant regulations only in 2008, it is difficult to infer any general pattern of improvement or deterioration from these results.

However, it is clear that the presence or absence of transfer regulations in a given year has a big effect on some agencies' scores. Scores for the Departments of Energy, Homeland Security, Transportation, and Health and Human Services climb noticeably in one or both years when transfer regulations are excluded. Omitting transfer regulations, Energy and Homeland Security leapfrog Agriculture, EPA, and Interior in the 2009 rankings, and HHS edges past Labor.

Table 7: Average Total Scores by Agency

All Regulations	2009 Average Score	# of Regulations	2008 Average Score	# of Regulations	2008-09 Change
Joint DOT/EPA	48.0	1	NA	0	NA
USDA	38.0	1	28.0	1	+10.0
Interior	34.0	1	27.3	4	+6.7
EPA	32.5	9	39.5	2	–7.0
DHS	30.0	2	38.0	2	–8.0
Energy	27.4	5	27.0	1	+0.4
DOT	24.7	3	32.3	6	–7.6
Labor	24.0	1	34.1	6	–10.1
HHS	23.6	12	20.7	11	+2.9
Education	22.0	5	22.0	2	0
HUD	18.0	1	41.0	1	–23.0
Veterans	17.0	1	10.0	1	+7.0
Justice		0	35.0	3	NA
Treasury		0	27.0	1	NA
Fed Acquisition		0	24.0	1	NA
State		0	13.0	1	NA
Defense		0	12.0	1	NA
SSA		0	7.0	1	NA
Non-Transfer Regulations	**2009 Score**	**# of Regulations**	**2008 Score**	**# of Regulations**	**2008-09 Change**
Joint DOT/EPA	48.0	1	NA	0	NA
Energy	40.7	3	27.0	1	+13.7
DHS	40.0	1	38.0	1	+2.0
USDA	38.0	1	28.0	1	+10.0
EPA	32.5	9	39.5	2	–7.0
Interior	34.0	1	27.3	4	+6.7
DOT	29.0	2	32.3	6	–3.3
HHS	28.0	1	29.0	2	–1.0
Labor	24.0	1	34.1	6	–10.1
HUD		0	41.0	1	NA
Justice		0	35.0	3	NA
Treasury		0	27.0	1	NA
Federal Acquisition		0	24.0	1	NA

Maximum possible average total score = 60.

5. Use of Analysis

Previous research found that use of the analysis was positively correlated with the quality of the analysis in 2008. Scores on criteria 9–12, which evaluate use of analysis, are positively correlated with the Analysis score and overall quality, defined as the sum of the Openness and Analysis scores, criteria 1–8 (Ellig and McLaughlin 2010). An additional year gives us a larger data set to test whether this relationship still held and whether it changed in 2009.

5.1 Total Use Score

Table 8 shows the results from regressing the Use score on the Quality score, along with several control variables. A one point increase in the Quality score is associated with a 0.25–0.31 point increase in the Use score, and this correlation is highly statistically significant. The result also seems quantitatively significant. The standard deviation of Quality is 6.86; a one-standard-deviation change in Quality implies about a two-point change in Use. Given that the mean Use score is 7.21, variation in Quality seems to explain a great deal of the variation in Use.[6]

The Year 2008 dummy tests whether Use scores tend to be different in 2008 and 2009. It shows that Use is about 1.3 points higher in 2008, after controlling for Quality. This result indicates a 1.3-point shift in the intercept of the regression equation. One might also speculate that the slope of the line might be different in the two years. When we run the same regressions using Quality × Year as an explanatory variable instead of the year dummy, we get roughly the same results with a bit worse statistical fit.[7]

The year appears to make a big difference, considering that the mean Use score is only 7.21 and its standard deviation is 3.45. However, it would be a mistake to portray the first year of the Obama administration as a retreat from stellar use of analysis in the Bush administration. Figure 3 shows the distribution of Use scores in 2008 and 2009. Neither year shows more than middling use of analysis. The principal difference is that the middle class shrinks in 2009, with more regulations that either fail to use the analysis or make only a passing reference to it.

Models 3 and 4 in table 8 include control variables for transfer regulations, to see if tendencies to use analysis differ for this type of regulation. In general, the relationship between Use and Quality seems no different for transfer regulations that for non-transfer regulations. However, the transfer regulations that implement provisions of the American Recovery and Reinvestment Act appear to be marginally more likely to use the analysis. The Use score for these five regulations averages 7 points, compared to an average of 5 points for other transfer regulations in 2009. The difference in averages stems from relatively high Use scores for two Education Department regulations that provide grants to states for education reform: the School Improvement Grants (13 points) and the Race to the Top Fund (9 points). School Improvement Grants earned a relatively high Use score because the regulations focus the grants on education reforms that have research demonstrating their effectiveness, and because the regulation includes

[6] Using only the four Analysis criteria 5–8 as the independent variable produces roughly the same results with a bit worse statistical fit.
[7] Results are in appendix 5.

provisions to gather data and evaluate the effectiveness of the reforms funded by the spending. The Race to the Top fund did not make much use of analysis to create the regulation, but it did establish goals and require states to submit data to evaluate the effectiveness of the reforms funded by the regulation.

5.2 Ex-Ante Use vs. Retrospective Analysis

The total Use score consists of scores for two types of criteria that might be affected differently by the quality of analysis. Criteria 9 and 10 assess the extent to which the agency used the analysis to make decisions in the proposed regulation. Criteria 11 and 12 assess the extent to which the agency provided for retrospective analysis in either the preamble to the regulation or the Regulatory Impact Analysis. To see whether Quality has different effects on these variables, table 9 replicates the regressions in table 8 using criteria 9–10 as a dependent variable and using criteria 11–12 as a dependent variable.

The quality of analysis clearly has a positive, statistically significant correlation with both the use of analysis to craft the regulation and on provisions for retrospective analysis. The effect is about twice as large for the former as for the latter.

The Year dummy variable, however, shows that Quality has a differential effect in 2008 only for use of analysis to craft the regulation. Agencies were no more likely to make provisions for retrospective analysis in 2008 than in 2009. This is perhaps unsurprising, given that Executive Order 12866 and Circular A-4 place little emphasis on retrospective analysis.

Finally, the Transfer dummy variable indicates that agencies were neither more nor less likely to use analysis in crafting transfer regulations or provide for retrospective analysis. The Recovery Act dummy shows that these regulations tend to have better retrospective analysis provisions—again largely because of the higher scores of the two education reform regulations.

These regressions identify some significant correlations, but we are not sure if they imply causation. Perhaps decision makers choose to use analysis when they are confident it is higher quality. Or perhaps analysts prepare better analysis when they are confident the decision makers will use it. Similarly, the higher Use scores in 2008 might reflect a stronger commitment to using regulatory analysis in the Bush administration, but other hypotheses might also explain the difference. To the extent that regulations proposed in 2009 were already in process in 2008, perhaps the Bush administration simply pushed out the regulations that were better-supported by analysis in 2008 and left the rest for the Obama administration to deal with. Alternatively, the difference could just reflect the fact that 2009 was a transition year (perhaps because new members of an administration have to "learn" how to use economic analysis). Forthcoming data on the quality and use of regulatory analysis in 2010 may allow us to test these and other hypotheses. Systematic interviews of federal regulatory personnel, such as those conducted by Williams (2008), could provide additional (and perhaps even better) insights.

Table 8: Quality of Analysis vs. Use of Analysis

Explanatory Variables	Dependent Variable: Use of Analysis Score (Criteria 9–12) (1)	(2)	(3)	(4)
Quality (Criteria 1–8)	0.30 [6.98***]	0.31 [7.28***]	0.27 [3.99***]	0.25 [3.83***]
Year 2008 Dummy		1.34 [2.31***]	1.15 [1.85*]	1.33 [2.14**]
Transfer Regulation			–0.80 [–0.85]	–1.19 [–1.25]
Recovery Act Regulation				2.25 [1.70*]
Constant	1.14 [1.24]	.33 [0.34]	1.64 [0.91]	1.82 [1.02]
N	87	87	87	87
Adjusted R^2	0.36	0.39	0.39	0.40

Ordinary least squares regressions; t-statistics in parentheses.
Statistical significance: ***1 percent **5 percent *10 percent

Figure 3: Use of Analysis Scores by Quintile

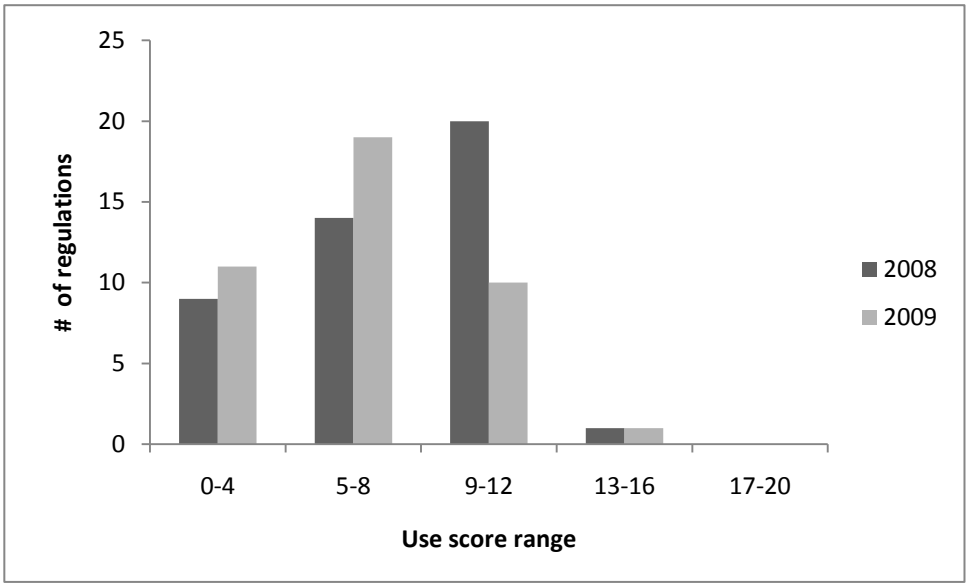


Table 9: Quality of Analysis vs. Separate Scores for Ex-Ante and Retrospective Analysis

Explanatory Variables	Dependent Variable: Ex Ante Use of Analysis (Criteria 9–10) (1)	(2)	(3)	(4)
Quality (Criteria 1–8)	0.20 [6.05***]	0.20 [6.30***]	0.17 [3.46***]	0.17 [3.37***]
Year 2008 Dummy		0.94 [2.18**]	0.83 [1.78*]	0.87 [1.82*]
Transfer Regulation			–0.51 [–0.72]	–0.58 [–0.80]
Recovery Act Regulation				0.45 [0.45]
Constant	0.34 [0.50]	–0.22 [–0.32]	0.60 [0.44]	0.64 [0.47]
N	87	87	87	87
Adjusted R^2	0.29	0.32	0.32	0.31

Explanatory Variables	Dependent Variable: Provisions for Retrospective Analysis (Criteria 11–12) (1)	(2)	(3)	(4)
Quality (Criteria 1–8)	0.11 [3.98***]	0.11 [4.04***]	0.09 [2.19**]	0.08 [2.00**]
Year 2008 Dummy		0.39 [1.06]	0.32 [0.81]	0.47 [1.29]
Transfer Regulation			–0.29 [–0.49]	–0.61 [–1.01]
Recovery Act Regulation				1.80 [2.15**]
Constant	0.79 [1.39]	0.56 [0.91]	1.04 [0.90]	1.18 [1.04]
N	87	87	87	87
Adjusted R^2	0.15	0.15	0.14	0.18

Ordinary least squares regressions; t-statistics in parentheses.
Statistical significance: ***1 percent **5 percent *10 percent

5.3 Use by Individual Agencies

Is the reduction in Use scores widespread, or concentrated in a few agencies? Table 10 sheds light on this question by calculating changes in average Use scores for individual agencies, including and excluding transfer regulations.

Including all regulations, four agencies improved their average Use scores between 2008 and 2009: Interior, Agriculture, Health and Human Services, and Veterans Affairs. Except for Agriculture, all of these improvements were less than one point. Seven agencies saw their average Use scores fall, and all of these reductions exceeded two points. Thus, improvements are small, and reductions are widespread.

Some of these changes were driven by the increased proportion of transfer regulations in 2009. Excluding transfer regulations, four agencies increased their Use scores: Interior, Agriculture, Health and Human Services, and Energy. Interior's score increased by just 0.7 point; all the others increased by at least two points. Four agencies saw their Use scores fall when transfer regulations are excluded: Homeland Security, Transportation, EPA, and Labor. Each of these four reductions was two points or greater. Excluding transfer regulations thus suggests that some agencies had noticeable improvements in their Use scores, while about the same number saw noticeable reductions.

The changing mix of transfer vs. non-transfer agencies had a big effect on results for four agencies: Energy, Homeland Security, Transportation, and Health and Human Services. Excluding transfer regulations actually increases Energy's Use score; with transfer regulations, Energy's Use score falls. Excluding transfer regulations leads to a much bigger increase in Health and Human Services' Use score: a 5.5 point increase instead of a 0.7 point increase. Finally, excluding transfer regulations cuts the reduction in Homeland Security's and Transportation's Use scores by more than half.

The regression equations in tables 8 and 9 show that use of analysis to make decisions about regulations is lower in 2009, even after controlling for transfer regulations. Tabulations in table 10 suggest that the primary reason for the statistically significant decline in Use scores in 2009 appears to be the reductions in Use scores at Transportation and EPA. Of all the agencies whose average Use scores fell, Transportation proposed two regulations in 2009 and EPA proposed nine. No other agency whose Use score for non-transfer regulations fell in 2009 proposed more than one non-transfer regulation in 2009.

In fairness, we should also note that the combined DOT/EPA CAFÉ/greenhouse gas emissions regulation earned the highest Use score in 2009: 15 points. In addition, the caveat we applied to table 7 applies to table 10 as well. Because the number of regulations is so small, it is hard to make reliable generalizations about particular agencies. For that, more years of data are needed.

Table 10: Use by Individual Agencies

All Regulations	2009 Average Score	# of Regulations	2008 Average Score	# of Regulations	2008-09 Change
Joint DOT/EPA	15.0	1	NA	0	NA
Interior	9.0	1	8.3	4	+0.7
USDA	8.0	1	5.0	1	+3.0
Energy	7.4	5	10.0	1	–2.6
EPA	7.2	9	10.5	2	–3.3
Education	7.0	5	9.0	2	–2.0
DHS	6.5	2	12.0	2	–5.5
HHS	5.6	12	5.5	11	+0.1
HUD	5.0	1	10.0	1	–5.0
DOT	4.5	3	10.0	6	–5.5
Labor	4.0	1	8.7	6	–4.7
Veterans	3.0	1	2.0	1	+1.0
Justice		0	11.7	3	NA
Treasury		0	9.0	1	NA
Fed Acquisition		0	4.0	1	NA
SSA		0	3.0	1	NA
State		0	2.0	1	NA
Defense		0	1.0	1	NA
Non-Transfer Regulations	**2009 Score**	**# of Regulations**	**2008 Score**	**# of Regulations**	**2008-09 Change**
Joint DOT/EPA	15.0	1	NA	0	NA
Energy	12.0	3	10.0	1	+2.0
DHS	10.0	1	12.0	1	–2.0
Interior	9.0	1	8.3	4	+0.7
DOT	8.5	2	10.0	6	–2.5
USDA	8.0	1	5.0	1	+3.0
EPA	7.2	9	10.5	2	–3.3
HHS	7.0	1	2.0	2	+5.0
Labor	4.0	1	8.7	6	–4.7
HUD		0	10.0	1	NA
Justice		0	11.7	3	NA
Treasury		0	9.0	1	NA
Federal Acquisition		0	4.0	1	NA

Maximum possible Use score = 20.

6. Conclusions

This study expands on existing research by applying a consistent set of standards to assess the quality and use of regulatory analysis for all economically significant regulations proposed in two different years. We find that the average quality of analysis is not high. The quality and use of regulatory analysis is especially poor for transfer regulations that define how the federal government will spend or collect money. But Regulatory Impact Analyses and *Federal Register* preambles present many examples of best practices that could improve the quality and use of analysis significantly if they were diffused more widely.

Our comparison of regulations in 2008 and 2009 generates several insights relevant to contemporary regulatory policy discussions. We find very little evidence that the quality of regulatory analysis changed between 2008 and 2009. The most significant improvement occurred in accessibility of regulatory analyses on the Internet. While this is a welcome improvement that is consistent with the Obama administration's focus on government transparency, improvements on a few other criteria were generally small and, at best, usually improved average scores from poor in 2008 to middling in 2009. In addition, we find substantial evidence that agencies were less likely to use the analysis to make decisions about proposed regulations in 2009 than in 2008.

This research also raises numerous questions that deserve further inquiry. We have not, by and large, identified why the quality and use of regulatory analysis exhibits the patterns revealed in this paper. For example, it is not obvious why some non-transfer regulations receive better analysis than others. Subject matter, deadlines, differing statutory mandates, explicit policy preferences, or department-specific factors may be part of the explanation.

It is also not clear why the quality of regulatory analysis changed very little between 2008 and 2009. Does this mean career staffers at agencies and/or OIRA consciously promote continuity between administrations? Another factor that may have played a role is that it is likely that the Bush administration focused greater effort on improving the quality of its "midnight" final regulations in 2008 relative to its proposed regulations, while the Obama administration is likely to have placed a greater focus on its own newly proposed regulations. This would suggest that the quality of analysis for proposed rules should have improved in 2009—unless most of the regulations proposed in 2009 were already in the pipeline in 2008. Research on what happened to the quality and use of analysis for final rules might shed further light on this issue.

Our data also indicate a statistically significant reduction in OIRA review time for non-transfer regulations in 2009 (from 66 to 40 days), but not for transfer regulations, which averaged about 35 days in both years. McLaughlin (2010) finds that midnight regulations receive shorter review times at OIRA. Whether OIRA review time impacts quality and use is an area ripe for further research.

Finally, we do not know why the use of regulatory analysis to make regulatory decisions declined in 2009. Indeed, we are not even sure if good analysis leads to use in decisions, or if decision makers' openness to analysis promotes good analysis, or if some third set of factors

causes both of these. Creating consistent data on the quality and use of regulatory analysis is the first step toward answering these questions.

References

Belcore, Jamie, and Jerry Ellig. 2009. "Homeland Security and Regulatory Analysis: Are We Safe Yet?," *Rutgers Law Journal* (Fall), 1-96.

Brito, Jerry, and Jerry Ellig, 2009. "Toward a More Perfect Union: Regulatory Analysis and Performance Management," *Florida State University Business Review* 8:1 (Spring/Summer), 1-55.

Ellig, Jerry, and Patrick McLaughlin. 2010. "The Quality and Use of Regulatory Analysis in 2008," Working Paper, Mercatus Center at George Mason University (June 22), http://mercatus.org/publication/quality-and-use-regulatory-analysis-2008.

Executive Order 12866, *Federal Register* 58:190 (Oct. 4, 1993), 51735-44.

Hahn, Robert W., Jason Burnett, Yee-Ho I. Chan, Elizabeth Mader, and Petrea Moyle. 2000. "Assessing Regulatory Impact Analyses: The Failure of Agencies to Comply with Executive Order 12,866." *Harvard Journal of Law and Public Policy,* 23:3, 859–71.

Hahn, Robert W., and Patrick Dudley. 2007. "How Well Does the Government Do Cost–Benefit Analysis?" *Review of Environmental Economics and Policy,* 1:2, 192–211.

Hahn, Robert W., and Robert Litan. 2005. "Counting Regulatory Benefits and Costs: Lessons for the U.S. and Europe," *Journal of International Economic Law* 8:2, 473-508.

Hann, Robert W. Randall W. Lutter, and W. Kip Viscusi. 2000. *Do Federal Regulations Reduce Mortality?* Washington, DC: AEI-Brookings Joint Center for Regulatory Studies.

Hahn, Robert W., and Paul C. Tetlock. 2008. "Has Economic Analysis Improved Regulatory Decisions?" *Journal of Economic Perspectives* 22:1 (Winter), 67-84.

Harrington, Winston. 2006. "Grading Estimates of the Benefits and Costs of Federal Regulation: A Review of Reviews." Discussion Paper 06-39. Washington, DC: Resources for the Future.

Harrington, Winston, Richard Morgenstern, and Peter Nelson. 2000. "On the Accuracy of Regulatory Cost Estimates." *Journal of Policy Analysis and Management,* 19:2, 297–332.

McLaughlin, P. A. "The Consequences of Midnight Regulations and Other Surges in Regulatory Activity." *Public Choice*, forthcoming 2010.

McLaughlin, Patrick, and Jerry Ellig. 2010. "Does Haste Make Waste in Regulatory Analysis?," Working Paper, Mercatus Center at George Mason University.

McTigue, Maurice, Henry Wray, and Jerry Ellig. 2009. *10th Annual Performance Report Scorecard: Which Federal Agencies Best Inform the Public?* Arlington, VA: Mercatus Center at George Mason University.

Office of Management and Budget. 2010. "Agency Checklist: Regulatory Impact Analysis" (Nov. 3), http://www.whitehouse.gov/sites/default/files/omb/inforeg/regpol/RIA_Checklist.pdf.

_____. 2009a. "Federal Regulatory Review, Request for Comments," *Federal Register* 74:37 (Feb. 26), 8819.

_____. 2009b. *2009 Report to Congress on the Benefits and Costs of Federal Regulations and Unfunded Mandates on States, Local, and Tribal Entities.* Washington, DC: Office of Information and Regulatory Affairs.

_____. 2008. *2008 Report to Congress on the Benefits and Costs of Federal Regulations and Unfunded Mandates on States, Local, and Tribal Entities.* Washington, DC: Office of Information and Regulatory Affairs.

_____. 2005. *Validating Regulatory Analysis: Report to Congress on the Costs and Benefits of Federal Regulations and Unfunded Mandates on State, Local, and Tribal Entities.* Washington, DC: Office of Information and Regulatory Affairs.

_____. 2003. *Circular A-4* (Sept. 17).

OECD. 2009. *Indicators of Regulatory Management Systems*. http://www.oecd.org/dataoecd/44/37/44294427.pdf.

Williams, Richard. 2008. "The Influence of Regulatory Economists in Federal Health and Safety Agencies," Working Paper, Mercatus Center at George Mason University (July), http://mercatus.org/publication/influence-regulatory-economists-federal-health-and-safety-agencies.

Appendix 1

Major Factors Considered When Evaluating Each Criterion

Note: Regardless of how they are worded, all questions involve qualitative analysis of how well the RIA and the *Federal Register* notice address the issue, rather than "yes/no" answers.

Openness

1. How easily were the RIA, the proposed rule, and any supplementary materials found online?

How easily can the proposed rule and RIA be found on the agency's website?
How easily can the proposed rule and RIA be found on Regulations.gov?
Can the proposed rule and RIA be found without contacting the agency for assistance?

2. How verifiable are the data used in the analysis?

Is there evidence that the analysis used data?
Does the analysis provide sufficient information for the reader to verify the data?
How much of the data are sourced?
Does the analysis provide direct access to the data via links, URLs, or provision of data in appendices?
If data are confidential, how well does the analysis assure the reader that the data are valid?

3. How verifiable are the models and assumptions used in the analysis?

Are models and assumptions stated clearly?
How well does the analysis justify any models or assumptions used?
How easily can the reader verify the accuracy of models and assumptions?
Does the analysis provide citations to sources that justify the models or assumptions?
Does the analysis demonstrate that its models and assumptions are widely accepted by relevant experts?
How reliable are the sources? Are the sources peer-reviewed?

4. Was the agency's analysis comprehensible to an informed layperson?

How well can a non-specialist reader understand the results or conclusions?
How well can a non-specialist reader understand how the analysis reached the results?
How well can a specialist reader understand how the analysis reached the results?
Are the RIA and relevant portions of the *Federal Register* notice written in "plain English"? (Light on technical jargon and acronyms, well-organized, grammatically correct, direct language used.)

Analysis

For each Analysis criterion, the lettered sub-questions each receive a score of 0–5, and these are averaged and rounded to produce the score on the criterion.

5. How well does the analysis identify the desired outcomes and demonstrate that the regulation will achieve them?

    A. How well does the analysis clearly identify ultimate outcomes that affect citizens' quality of life?
    B. How well does the analysis identify how these outcomes are to be measured?
    C. Does the analysis provide a coherent and testable theory showing how the regulation will produce the desired outcomes?
    D. Does the analysis present credible empirical support for the theory?
    E. Does the analysis adequately assess uncertainty about the outcomes?

6. How well does the analysis identify and demonstrate the existence of a market failure or other systemic problem the regulation is supposed to solve?

    A. Does the analysis identify a market failure or other systemic problem?
    B. Does the analysis outline a coherent and testable theory that explains why the problem (associated with the outcome above) is systemic rather than anecdotal?
    C. Does the analysis present credible empirical support for the theory?
    D. Does the analysis adequately assess uncertainty about the existence and size of the problem?

7. How well does the analysis assess the effectiveness of alternative approaches?

    A. Does the analysis enumerate other alternatives to address the problem?
    B. Is the range of alternatives considered narrow or broad?
    C. Does the analysis evaluate how alternative approaches would affect the amount of the outcome achieved?
    D. Does the analysis adequately address the baseline—what the state of the world is likely to be in the absence of further federal action?

8. How well does the analysis assess costs and benefits?

    A. Does the analysis identify and quantify incremental costs of all alternatives considered?
    B. Does the analysis identify all expenditures likely to arise as a result of the regulation?
    C. Does the analysis identify how the regulation would likely affect the prices of goods and services?
    D. Does the analysis examine costs that stem from changes in human behavior as consumers and producers respond to the regulation?
    E. Does the analysis adequately address uncertainty about costs?
    F. Does the analysis identify the approach that maximizes net benefits?

G. Does the analysis identify the cost-effectiveness of each alternative considered?
H. Does the analysis identify all parties who would bear costs and assess the incidence of costs?
I. Does the analysis identify all parties who would receive benefits and assess the incidence of benefits?

Use

9. Does the proposed rule or the RIA present evidence that the agency used the Regulatory Impact Analysis?

Does the proposed rule or the RIA assert that the analysis of outcomes, benefits, the systemic problem, alternatives, or costs affected any decisions?
How many aspects of the proposed rule did the analysis affect?
How significant are the decisions the analysis affected?

10. Did the agency maximize net benefits or explain why it chose another option?

Did the analysis calculate net benefits of one or more options so that they could be compared?
Did the analysis calculate net benefits of all options considered?
Did the agency either choose the option that maximized net benefits or explain why it chose another option?
How broad a range of alternatives did the agency consider?

11. Does the proposed rule establish measures and goals that can be used to track the regulation's results in the future?

Does the RIA or *Federal Register* notice contain analysis or results that could be used to establish goals and measures to assess the results of the regulation in the future?
In the RIA or the *Federal Register* notice, does the agency commit to performing some type of retrospective analysis of the regulation's effects?
Does the agency explicitly articulate goals for at major outcomes the rule is supposed to affect?
Does the agency establish measures for major outcomes the rule is supposed to affect?
Does the agency set targets for measures of major outcomes the rule is supposed to affect?

12. Did the agency indicate what data it will use to assess the regulation's performance in the future and establish provisions for doing so?

Does the RIA or *Federal Register* notice demonstrate that the agency has access to data that could be used to assess some aspects of the regulation's performance in the future?
Would comparing actual outcomes to outcomes predicted in the analysis generate a reasonably complete understanding of the regulation's effects?
Does the agency suggest it will evaluate future effects of the regulation using data it has access to or commits to gathering?

Does the agency explicitly enumerate data it will use to evaluate major outcomes the regulation is supposed to accomplish in the future?
Does the analysis demonstrate that the agency understands how to control for other factors that may affect outcomes in the future?

Appendix 2: Crosswalk of 2010 OMB Regulatory Impact Analysis Checklist with Mercatus Regulatory Report Card evaluation criteria

OMB Checklist	Mercatus Evaluation Criteria
Does the RIA include a reasonably detailed description of the need for the regulatory action?	Criterion 6: How well does the analysis demonstrate the existence of a market failure or other systemic problem the regulation is supposed to solve?
Does the RIA include an explanation of how the regulatory action will meet that need?	Criterion 5: How well does the analysis identify the desired outcomes and demonstrate that the regulation will achieve them?
Does the RIA use an appropriate baseline (i.e., best assessment of how the world would look in the absence of the proposed action)?	Criterion 7, question D: Does the analysis adequately assess the baseline—what the state of the world is likely to be in the absence of further federal action?
Is the information in the RIA based on the best reasonably obtainable scientific, technical, and economic information and is it presented in an accurate, clear, complete, and unbiased manner?	Criterion 2: How verifiable are the data used in the analysis? Criterion 3: How verifiable are the models or assumptions used in the analysis? Criterion 4: Was the analysis comprehensible to an informed layperson? *Criterion 3 includes an assessment of whether the models and assumptions are based on peer-reviewed or otherwise reliable publications. However, the Mercatus evaluation does not assess the quality of the underlying science.*
Are the data, sources, and methods used in the RIA provided to the public on the Internet so that a qualified person can reproduce the analysis?	Criterion 1 takes the first step by assessing how easily the RIA itself can be found on the Internet. Criteria 3 and 4 include an assessment of how easily the reader could find the underlying data, sources, and methods from information or links provided in the RIA or the *Federal Register* notice.
To the extent feasible, does the RIA quantify and monetize the anticipated benefits from the regulatory action?	Criterion 5, question 2: How well does the analysis identify how the outcomes are to be measured?

To the extent feasible, does the RIA quantify and monetize the anticipated costs?	Multiple questions under criterion 8 (Benefits and Costs) assess how well the analysis identifies, quantifies, and monetizes costs.
Does the RIA explain and support a reasoned determination that the benefits of the intended regulation justify its costs (recognizing that some benefits and costs are difficult to quantify)?	Criterion 8, question F: Does the analysis identify the approach that maximizes net benefits? Criterion 8, question G: Does the analysis identify the cost-effectiveness of each alternative considered?
Does the RIA assess the potentially effective and reasonably feasible alternatives?	Criterion 7: How well does the analysis assess the effectiveness of alternative approaches?
Does the preferred option have the highest net benefits (including potential economic, public health and safety, and other advantages; distributive impacts; and equity), unless a statute requires a different approach?	Criterion 10: Did the agency maximize net benefits or explain why it chose another option?
Does the RIA include an explanation of why the planned regulatory action is preferable to the identified potential alternatives?	Criterion 9: Does the proposed rule or RIA present evidence that the agency used the Regulatory Impact Analysis? Criterion 10: Did the agency maximize net benefits or explain why it chose another option?
Does the RIA use appropriate discount rates for the benefits and costs that are expected to occur in the future?	Considered under criterion 5, question 2: How well does the analysis identify how the outcomes are to be measured?, as well as several questions about measurement and comparison of benefits and costs under criterion 8 (Benefits and Costs).
Does the RIA include, if and where relevant, an appropriate uncertainty analysis?	Criterion 5, question E: Does the analysis adequately assess uncertainty about the outcomes? Criterion 6, question D: Does the analysis adequately assess uncertainty about the existence and size of the problem? Criterion 8, question E: Does the analysis adequately address uncertainty about costs?

Does the RIA include, if and where relevant, a separate description of the distributive impacts and equity (including transfer payments and effects on disadvantages or vulnerable populations)?	Criterion 8, question H: Does the analysis identify all parties who would bear costs and assess the incidence of costs? Criterion 8, question I: Does the analysis identify all parties who would receive benefits and assess the incidence of benefits?
Does the analysis include a clear, plain-language executive summary, including an accounting statement that summarizes the benefit and cost estimates for the regulatory action under consideration, including the qualitative and non-monetized benefits and costs?	Criterion 4: Was the analysis comprehensible to an informed layperson?
Does the analysis include a clear and transparent table presenting (to the extent feasible) anticipated benefits and costs (qualitative and quantitative)?	Criterion 4: Was the analysis comprehensible to an informed layperson?
Goals and measures to assess results of the regulation in the future – No content.	Criterion 11: Does the proposed rule establish measures and goals that can be used to track the regulation's results in the future?
Provisions for gathering data to assess results of the regulation in the future – No content.	Criterion 12: Did the agency indicate what data it will use to assess the regulation's performance in the future and establish provisions for doing so?

Appendix 3: Summary Statistics on All Criteria and Sub-Questions

2008

Variable	N	Mean	Std. Dev.	Min	Max
Total	45	27.30	9.46	7	43
Openness	45	11.04	3.26	4	18
Analysis	45	8.53	4.48	0	16
Use	45	7.73	3.31	1	14
Criterion 1	45	3.53	1.36	0	5
Criterion 2	45	2.24	1.19	0	5
Criterion 3	45	2.33	1.30	0	5
Criterion 4	45	2.93	1.21	0	5
Criterion 5	45	2.36	1.40	0	5
5A	45	3.31	1.52	0	5
5B	45	2.71	1.74	0	5
5C	45	2.22	1.59	0	5
5D	45	1.67	1.60	0	5
5E	45	2.00	1.86	0	5
Criterion 6	45	1.80	1.47	0	5
6A	45	2.31	1.68	0	5
6B	45	2.00	1.75	0	5
6C	45	1.71	1.59	0	5
6D	45	0.82	1.28	0	5
Criterion 7	45	2.29	1.36	0	4
7A	45	2.78	1.86	0	5
7B	45	1.96	1.45	0	5
7C	45	1.98	1.64	0	5
7D	45	2.04	1.30	0	5
Criterion 8	45	2.09	0.996	0	4
8A	45	2.93	1.16	0	5
8B	45	3.18	1.01	1	5
8C	45	1.38	1.34	0	5
8D	45	1.56	1.47	0	5
8E	45	1.78	1.80	0	5
8F	45	1.91	1.66	0	5
8G	45	1.04	1.17	0	5
8H	45	2.82	1.13	1	5
8I	45	1.60	1.34	0	5
Criterion 9	45	2.44	1.32	0	5
Criterion 10	45	2.20	1.46	0	5
Criterion 11	45	1.36	1.03	0	5
Criterion 12	45	1.73	1.10	0	5

2009

Variable	N	Mean	Std. Dev.	Min	Max
Total	42	27.03	9.37	5	48
Openness	42	12.00	2.82	3	17
Analysis	42	8.38	4.52	1	18
Use	42	6.64	3.56	0	15
Criterion 1	42	4.05	0.85	2	5
Criterion 2	42	2.50	1.50	0	5
Criterion 3	42	2.62	1.23	0	5
Criterion 4	42	2.83	0.88	1	4
Criterion 5	42	2.38	1.43	0	5
5A	42	3.36	1.61	0	5
5B	42	2.52	1.63	0	5
5C	42	2.21	1.60	0	5
5D	42	2.02	1.56	0	5
5E	42	1.76	1.69	0	5
Criterion 6	42	1.60	1.15	0	4
6A	42	2.21	1.70	0	5
6B	42	1.50	1.29	0	4
6C	42	1.21	1.24	0	4
6D	42	0.88	1.31	0	4
Criterion 7	42	2.21	1.42	0	5
7A	42	2.83	1.58	0	5
7B	42	1.86	1.32	0	5
7C	42	1.90	1.76	0	5
7D	42	1.93	1.44	0	5
Criterion 8	42	2.19	1.15	0	5
8A	42	2.83	1.34	0	5
8B	42	3.24	1.32	0	5
8C	42	2.07	1.69	0	5
8D	42	1.60	1.48	0	5
8E	42	1.76	1.59	0	5
8F	42	1.33	1.66	0	5
8G	42	1.24	1.54	0	5
8H	42	3.00	1.17	0	5
8I	42	1.86	1.47	0	5
Criterion 9	42	2.24	1.36	0	5
Criterion 10	42	1.62	1.56	0	5
Criterion 11	42	1.29	0.97	0	4
Criterion 12	42	1.50	1.04	0	4

Appendix 4: Average changes without separating transfer and non-transfer regulations

The table below shows the change in average scores on individual criteria and on sub-questions for the Analysis criteria. We only report average scores whose differences are statistically significant at the 85 percent level or higher. Even for individual criteria or questions, there is very little evidence that average scores changed much between 2008 and 2009. As noted in the text, some of the changes identified below are driven by the increased proportion of transfer regulations in 2009.

Score Changes on Individual Criteria and Questions

	2008 (n=45)	**2009** (n=42)	**Change**	**T-stat.**
Openness				
Criterion 1 – Accessibility	3.53	4.05	0.51	2.10**
Analysis				
Question 6B – Coherent Theory of Systemic Problem	2.00	1.50	–0.50	1.60
Question 6C – Empirical Evidence of Systemic Problem	1.71	1.21	–0.50	1.62
Question 8C – Effects on Prices of Goods and Services	1.38	2.07	0.69	2.13**
Question 8F – Identifies approach that maximizes net benefits	1.91	1.33	–0.58	1.62
Use				
Criterion 10 – Decision Cognizant of Net Benefits	2.20	1.62	–0.58	1.80*

Statistical significance: *90 percent **95 percent
Maximum possible score on any criterion or question = 5 points.

The increase on criterion 1 (Accessibility) indicates that agency regulatory analyses were somewhat easier to find online in 2009 than in 2008. This reflects the fact that regulatory analyses were easier to find on agency websites and *Federal Register* preambles provided clearer information about how to obtain a copy of the Regulatory Impact Analysis. Some of the improvement may also stem from the redesign of the regulations.gov web site, which may have made regulations and accompanying analysis easier to find.

The lower average scores on questions 6B (Coherent Theory of Systemic Problem) and 6C (Empirical Evidence of Systemic Problem) suggest that agencies may be somewhat less likely to demonstrate that proposed regulations actually address a market failure, government failure, or other systemic problem in 2009. Average scores were already quite low in 2008; this weakness may have gotten even weaker in 2009.

The higher average score on criterion 8C (Effects on Prices of Goods and Services) indicates that agencies were more likely in 2009 to discuss the effects of regulatory costs on the prices of goods and services. This is something that agencies usually do either reasonably well or pretty poorly; there are few mid-range scores. The increase from 1.38 to 2.07 implies that this improvement occurred only for a few regulations, or that agencies provided just a bit more discussion or evidence in place of unsupported assertions.

The lower scores on question 8F (Identifies Alternative that Maximizes Net Benefits) and criterion 10 (Decision Cognizant of Net Benefits) suggest that regulatory analyses in 2009 were less likely to assess the net benefits of alternatives, and decision makers were less likely to consider net benefits when choosing among alternatives. Agencies usually do these things either reasonably well or not at all, so this shift suggests that fewer regulations in 2009 identified or considered net benefits of alternatives.

Appendix 5: Use vs. Quality Employing Quality x Year Interaction Variable

Explanatory Variables	Dependent Variable: Use of Analysis Score (Criteria 9-12) (1)	(2)	(3)	(4)
Quality (Criteria 1-8)	0.30 [6.98***]	0.28 [6.26***]	0.23 [3.67***]	0.22 [3.41***]
Year 2008 Dummy X Quality		0.06 [2.21***]	0.05 [1.79*]	0.06 [1.98**]
Transfer Regulation			–0.88 [–0.95]	–1.28 [–1.34]
Recovery Act Regulation				2.07 [1.57]
Constant	1.14 [1.24]	1.06 [1.18]	1.64 [0.91]	2.70 [1.63]
N	87	87	87	87
Adjusted R^2	0.36	0.39	0.38	0.40

Ordinary least squares regressions; t-statistics in parentheses.
Statistical significance: ***1 percent **5 percent *10 percent