

POLICY ANALYSIS

MODERATOR

David Rejeski, Environmental Protection Agency

PRESENTATION 1

H.O. Stekler, Industrial College of the Armed Forces

PRESENTATION 2

Karen Fisher-Vanden, Jae Edmonds, and Hugh Pitcher,
Pacific Northwest Laboratory

PRESENTATION 3

Jerome C. Glenn, United Nations University

Policy Analysis

Forecasting for Large Policy Changes

H.O. Stekler, Industrial College of the Armed Forces

Most forecasters working for the Federal government engage in the typical prediction exercises. We try to determine the effect either of small changes or of events about which we have some knowledge. In other words, we are able to make forecasts using models which replicate the past and which we assume will explain the future. For those incremental changes and for those events about which we have knowledge, this is the appropriate procedure. However, what should we do when there are fundamental structural changes occurring? In those situations the models of the past are probably not reliable for predicting the future. Moreover, we have no experience with the events that are taking place and have great difficulty in explaining the phenomena. We frequently then rely on theoretical analyses which may not yield accurate results in the real world because the assumptions of that theory do not hold in practice.

This paper will do two things. First, I look at the outcomes that were predicted to occur after a major policy change was enacted-airline deregulation. This involves a comparison of what was expected to happen when deregulation occurred with the outcomes we have observed and an explanation for the successes and failures. My thesis is that we must use one of two entirely different approaches to predict the effects of large policy changes. The approach that I will suggest is similar to the methods that the armed forces have used in determining the effects of alternative military strategies-simulated situations which replicate the conditions being investigated or experimentation.

I. Airline Deregulation

Airline deregulation occurred in 1978, but the expected effects of that policy change had been analyzed for a number of years prior to the event. These studies are too numerous to list here, but Levine (1987) has provided a comprehensive review of the major articles and books. Moreover there have been a substantial number of post-event explanations of the reasons that the actual outcomes differed from the predictions. Levine (1987) undertook the most comprehensive investigation of this question, while Kahn (1988) also provided his insights about the surprises that occurred as a result of deregulation. This paper will not replicate the analyses that were undertaken in these papers. Rather we shall only focus on the differences between the predictions and the outcomes, determine why they occurred and whether those discrepancies should have been surprises.

A. Predictions

The basic prediction was that a deregulated airline industry would display all the characteristics of a purely competitive industry and that prices would be substantially lower than they were with regulation. There is general agreement about those forecasts: the first is patently wrong (although there are individuals who argue that the industry is still in transition); the second is correct. Economists were correct in predicting what would happen to prices when the artificial constraints that regulation imposed on the industry were removed. They were unable to predict what the structure of the industry would be once deregulation occurred.

The predictions about the post-regulation structure of the industry included statements that there would be as many as 200 airlines, each operating as few as six aircraft. Prices would be lower and uniform for all customers, with perhaps some discounts for off-peak flights and the airlines which were in existence at the time of deregulation would have a difficult time competing with the new entrants. Entry and exit from the industry would be easy since aircraft were mobile assets. (Kennedy Hearings, 1975.)

Finally, virtually all analysts focused on linear or point to point markets, for those were the markets which prevailed under regulation and were the only situations observed in the unregulated intrastate environments of Texas and California. In other words, the analysts based their forecasts on the experiences that they had and extrapolated into the future on the basis of this model and/or knowledge. They did not attempt to determine whether this would

be a good approach for predicting what would occur after the policy change took place and a new environment was created.

B. Outcomes

The outcomes differed markedly from these forecasts. The U.S. airline industry clearly does not resemble a purely competitive industry. There are only a small number of large firms, and, with perhaps one exception, all were in existence in 1978. In fact, the industry is now at least as concentrated as it was prior to deregulation, with three dominant firms. Although a number of the old line firms--such as Braniff, Eastern and Pan Am no longer exist, most of the new entrants also have failed. Prices are lower as predicted but they are not uniform with all types of restrictive fares in existence. The hub and spoke systems of the major carriers were clearly not foreseen. Why was there such a disconnect between the predictions and the outcomes? Moreover, why were these outcomes such surprises to those who made these predictions?

C. The Predictive Theory

Levine provided a partial answer to this question. He indicated (1987, p.403) that the proponents of deregulation had not considered the relationship between airline markets and the number of firms that would serve them. He, himself, had noted that there was an inconsistency between the assumptions of pure competition and the number of firms that were operating in the various markets, but left it at that. Prior to the actual deregulation of the industry the formal theory (based on the value of information and principal-agent analysis) that Levine used to explain the outcomes that deregulation produced was available in rudimentary form. However, it was not used to predict what might happen once the regulatory constraints were removed from the industry. The economists who analyzed the industry primarily focused on the costs that regulation imposed on consumers and the economy. They also were concerned about the absence or presence of economies of scale, i.e. the costs of producing a given number of seat miles and the ease of entry. Several years after regulation was abandoned, a new theory about contestable markets was proposed. This theory basically states that, if specific conditions prevail, merely the threat of competitive entry would be sufficient to produce the competitive market result. However, the structural changes that have occurred are inconsistent with this theory. These include hub domination, fare structures that are complex, computerized reservation systems that have been used for strategic advantage, frequent flier programs, etc.

Thus, we can conclude that prior to deregulation the theory which could have predicted what would happen once regulation was abandoned was still not fully developed and definitely not used for analysis while the formal theory which was proposed after the event was inappropriate. "It is difficult to be reassuring about the effects of phenomena that you did not predict and can not explain." (Levine, 1987 p.396.) Is this an appropriate way to undertake public policy?

Economic theory usually makes a number of simplifying assumptions in order to generate broad general principles or generalities. It can not generally predict exactly how the principals in a market will react. However, the economists who were analyzing the industry should have realized that the market participants who were present prior to deregulation would endeavor to protect their interests once the regulatory constraints were lifted. In other words, the institutional arrangements that existed or might develop and the possible strategic reactions of existing firms were not sufficiently well analyzed.

In fact, market participants and financial analysts were better able to predict what would happen in the years immediately after deregulation than were those economists who engaged in theoretical speculations. Moreover, many of the structural changes which actually occurred were foreshadowed in, an even cursory reading of, the trade literature. While it is dangerous to look for corroborating evidence after the fact, some of the evidence is quite clear and should have caught the attention of the analysts. A few examples will suffice to illustrate this point.

D. Some Industry Views

The evidence suggests that the advocates of deregulation were surprised by the creation of hub and spoke airline systems. Yet the testimony of United Airlines at the Kennedy Committee Hearings indicated that the firm considered the connecting flights as an important means of filling its longer distant flights. In other words, they were not merely concerned with point to point transportation. That airline indicated that deregulators were

conducting their analyses based on city pairs, for the CAB regulated on that basis but in reality the country's system was definitely not linear. (Aviation Week, April 7, 1975, p.21.) This lack of concern with actual institutional situations may have been the factor that prompted that airline to complain that there was too little management experience among the deregulators. (Aviation Week October 13, 1975, p.90.) United was not the only firm that was aware of the potential of hubs, for Aviation Week (February 17, 1975, pp.20, 22) indicated that both United and American were focusing attention on their hubs in Chicago and Dallas, respectively. Moreover, Boeing indicated that hubbing had been the norm with respect to airline operations and that point to point service was a relatively new phenomenon. (AW November 29, 1976, p.7.)

Similarly, the potential value of using computer systems for setting differential fares was general knowledge prior to deregulation. TWA noted that discounts did not have to be the same on every flight (AW February 17, 1975, p.22) while United indicated that it was possible to develop capacity controlled fares. (AW August 15, 1977, p.24.) Another surprise was the increased role that travel agents have played. Yet, there had been general discussions about permitting the agents to have direct access to the airlines' computerized reservation systems and of providing cash incentives to the agents for ticketing on specific (international) carriers. (AW September 22, 1975 p.28; December 1, 1975 p.27; February 14, 1977 p.24.) Finally, there were observations that customers were relying to an increasing extent on travel agents because fares had become more complex. (AW February 20, 1978 p.25.)

As for the structure of the industry, United envisioned the number of trunk lines being reduced to three by the year 2026 with an increased emphasis on international travel. There would be a large number of commuter airlines serving the domestic market. Thus that company foresaw some of the trends that have occurred since 1978. (However, it was wrong in one respect, for it predicted that there would be no discount fares.) It was also noted that exit is not costless for airline planning takes a long time and there are key connecting points. (AW April 7, 1975, p.21; February 16, 1976, p.7). According to the perceived theory prevailing prior to deregulation, predatory pricing should not occur, but it did and its potential was recognized by industry observers and financial analysts. (AW August 4, 1975; March 28, 1977, p.9.)

With hindsight it is always possible to find some analyst who correctly foresaw some of the trends that have occurred. That was not the objective of citing the aforementioned examples, for one could have also found a number of industry executives who failed to identify the observed trends. Rather, the idea was to note that there was a wide divergence of views with respect to the likely outcome of deregulation between the academic community and knowledgeable industry leaders. When such a wide discrepancy exists, it might be advisable to adopt another approach to analyze the impact of a proposed policy change.

Obviously it would not have been possible to engage in large scale experimentation either prior to or during deregulation, but there is another technique that can take account of institutional or industry knowledge, for predicting the likely outcomes of large policy changes. This is a technique which would use the technical expertise of the academic and the institutional knowledge of the industry analyst. The approach would be to adapt the methodology used by the military to analyze the possible outcomes of adopting a particular strategy or tactic, war games.

II. War Games

A war game is an analytical device which attempts to recreate the conditions of an actual historical battle or war. Alternatively, it might create the probable or possible situations which might prevail in a hypothetical future war. The players, who are knowledgeable about the institutional conditions, deploy and employ their forces and resources using alternative strategies to determine whether the outcomes would remain unchanged even if the combatants had made different decisions. It also permits an analysis of the conditions under which the outcomes would have been changed.

The individuals who take part in the game play the roles of all of the opponents who actually did (or possibly might) participate in a conflict. The rules of the game require each player or side to make a move, which might involve the movement of forces, the initiation of an attack, or the creation of a new strategic alliance. When all the players have made their moves the outcomes must be determined. Since the interactive moves involve uncertainties and an element of randomness, the outcomes are clearly not deterministic. The rules of the game determine the outcomes and in every case involve some probability generating function. Statistical tables are provided in the non-computer version of these games while the black box rules embodied in the computer determine those results. Upon the adjudication of that move, the process is repeated until the game is completed.

These games are not limited to single battles or theaters of war. In fact, one commercial game involves the recreation of the entire social, economic, political and military structure of the 14th century as the players become involved in the 100 Years War. This is an online interactive computer game involving 300 participants. It is obvious that these games can be very complicated and can embody many of the institutional factors that cannot be captured by abstract models.

The advantages of these games are obvious. By placing players in interactive competitive situations, they permit analysts to uncover and evaluate alternative strategies that actual competitors might employ in real life situations. In other words the games take over where theory ceases. In addition, the sensitivity of possible outcomes to alternative assumptions can be tested. This methodology might reduce the observed tendency of individuals to place unrealistically high confidence in their own theory, judgement or scenario. (Tversky and Kahneman, 1982; Ascher, 1993).

III. Application to Policy Changes

In my opinion, this methodology could have been applied to an analysis of the deregulated airline market. If, prior to deregulation, a game had been developed to analyze the new conditions and if the players had been individuals knowledgeable about the institutional factors, the results of the game might have been more realistic than the actual predictions. Then there would have been far fewer surprises about the strategies of the market participants, for the players would have assumed the roles of executives of new and old airlines, travel agents, various classes of consumers, the government, etc.

I would, therefore, recommend that such games be used whenever large policy changes might be implemented. These games can be developed relatively quickly by individuals versed in the art of devising war games. The participants should be people who are intimately connected with the policy under consideration and who might be affected by it. (One drawback of using executives of companies that might be impacted by such policy changes is that the firm's true strategies might not be revealed. An alternative would be to use former industry leaders, who have no vested interest in the outcomes). Repeated plays of the game would reveal the range of possible outcomes and would assist both the executive and legislative branches of the government in enacting an appropriate policy and to take account of the possible loopholes that were discovered before the policy was actually implemented.

This proposal has some elements in common with the Delphi method of forecasting, but it also deviates from that technique. The Delphi method asks knowledgeable sources what outcomes are likely to occur, and when those responses have been received, they are all revealed to all these experts. The range of outcomes is then discussed and usually a consensus prediction is obtained. The similarity to a war game is that experts are used in both cases. The difference is that in the game the players are placed in an interactive competitive environment and must make strategic decisions in real time. The latter approach is likely to provide a more realistic prediction of what the real outcome will actually be.

The suggestion that policy changes be modelled through a war game-like analysis is not at all that radical. Business schools have frequently used computer games as decision making devices for analyzing the strategies of firms. What is new is the suggestion that such an analytic device be used in the public policy arena for predicting the outcome of a significant change in policy. The new health policy is the obvious candidate for such a mode of analysis. Moreover, in this situation it might be possible to combine this mode of forecasting with some experimentation, since alternative policies might be tried in different states.

IV. Conclusions

In this paper I have tried to suggest that the customary forecasting procedures might not be appropriate when large policy changes are being considered. The airline deregulation policy was used as an example to illustrate this point, for there were a large number of differences between what was predicted to occur and what actually happened after deregulation was implemented. Experimentation and simulations were suggested as alternative methods for predicting the outcomes of large policy changes.

References

- Ascher, William (1993), "The Ambiguous Nature of Forecasts in Project Evaluation: Diagnosing the Over-Optimism of Rate of Return Analysis," International Journal of Forecasting, 9, 109-115.
- Kahn, Alfred E. (1988), "Surprises of Airline Deregulation," American Economic Review, Paper and Proceedings, 78, 316-322.
- Levine, Michael E. (1987), "Airline Competition in Deregulated Markets: Theory, Firm Strategy and Public Policy," Yale Journal on Regulation, 4, 393-494.
- Tversky, A., and Kahneman, D. (1982), "Judgment Under Uncertainty: Heuristics and Biases," in D. Kahneman, P. Slovic and A. Tversky, Eds. Judgment Under Uncertainty: Heuristics and Biases (Cambridge University Press, Cambridge).

The Second Generation Model of Energy Use, The Economy, and Greenhouse Gas Emissions

Karen Fisher-Vanden, Jae Edmonds, Hugh Pitcher, Dave Barns, Richard Baron, Sonny Kim, Chris MacCracken, Elizabeth L. Malone, Ronald D. Sands, and Marshall Wise, Pacific Northwest Laboratory

1.0 Introduction

The global change problem has been a major issue on the public agenda since 1988, when a combination of events including a drought in the United States midwest, Congressional hearings, and an international meeting in Toronto Canada on the subject brought the issue to the forefront of public consciousness. The Toronto meeting set an arbitrary goal of reductions in fossil fuel carbon emissions of 20% for developed nations. The 1992 Framework Convention on Climate Change called on nations to stabilize greenhouse emissions at 1990 levels. And in 1993 the President of the United States pledged the United States to stabilize national greenhouse emissions at 1990 levels by the year 2000. In this paper, we explore policy options for achieving reductions of carbon emissions and their economic implications using the Second Generation Model (SGM) of human activities and global change, a new, computable general equilibrium (CGE) model designed to analyze global change problems. Further, we explore some of the greenhouse emissions consequences of policy proposals such as a BTU tax.

We begin our discussion with a brief description of the model, describe the derivation of reference year parameters, proceed to develop reference case assumptions, describe the resultant reference case scenario, and then explore the costs and benefits of selected policy options for stabilizing greenhouse emissions.

2.0 Model Description

In this section we provide a brief overview of the model and its numerical embodiment of the United States economy. The theoretical structure of SGM Version 0.0 is described in detail in Edmonds et al. (1993). The development of estimates for parameters is documented in Fisher-Vanden et al. (1993).

2.1 Origins

The SGM's intellectual roots can be traced to the modeling work of Edmonds, Reilly, and Barns, who developed and exercised the Edmonds-Reilly-Barns model of long-term global energy and greenhouse gas emissions (Edmonds and Reilly, 1985; Edmonds and Reilly, 1986; Edmonds and Barns, 1991). This model (member of the first generation of models) focused on emissions of energy related greenhouse gases. The model was simple and transparent, but had long time steps, did not consider non-energy related greenhouse gas emissions and sinks, and completely neglected impacts on human systems from global change. The SGM is structured to provide both increased scope (including both energy and non-energy related emissions activities and a framework designed to assess impacts of global change on human activities), and finer resolution of human activities (including a 5-year time step, enhanced technology descriptions, and an integrated demographic module).

The analysis reported here uses SGM Version 0.0. While 20 global regions are being developed for the global implementation of the SGM, only the United States module is used in this exercise.

2.2 Overview of SGM Version 0.0

In summary, the SGM is a member of the CGE class of models. The SGM Version 0.0 has nine producing sectors:

1. Agriculture
2. Crude Oil
3. Natural Gas Production
4. Coal Production
5. Uranium Production and Refinement
6. Electric Power Generation

7. Oil Refining
8. Natural Gas Transformation and Distribution
9. Other Production

We note that seven of the nine sectors of the economy in SGM Version 0.0 are energy sectors. The emphasis on the energy sector in SGM Version 0.0 reflects the central role energy plays in the determination of emissions of greenhouse gases to the atmosphere.

In addition to the producing sectors of the SGM, there are four final demand sectors:

1. Households
2. Government
3. Investment
4. Net Export

Producing sectors use goods and services produced by other producing sectors, itself, and primary factors of production to produce net output. The three primary factors of production are:

1. Land
2. Labor
3. Capital

There is a market, which is cleared by the price mechanism, for each producing sector of the economy. Similarly there is a market for each of these three primary factors of production.

The interactions of the main components of the SGM are shown in Figure 1, where the seven energy sectors are included in the energy box. In general energy, agricultural, and other production sectors supply net quantities of goods and services for each other and for final consumption by households, the government, investment, and net exports. Primary factors of production - land, labor and capital - are also used by the producing sectors. Labor and land are assumed to be owned by members of the household sector who supply them to the market. Capital is associated with producing sectors. Net profits are returned to the household sector, while taxes are collected from producers and households and provide the revenues for local, regional, and national governments.

Producing sectors are assumed to make decisions regarding production and investment with the objective of maximizing expected wealth. Governments are assumed to produce government services, including education, national defense, and general services, according to prescribed production functions. Households behave as if they are maximizing utility with regard to the allocation of resources to current consumption; however, the supply of savings and labor is modeled by exogenously specified rules.

All produced goods and services, and primary factors of production pass through markets which are assumed to clear in every period via a price mechanism. Estimates of gaseous emissions are computed in the model via the application of emissions coefficients to inputs and outputs of processes. Variation in these coefficients reflect variation in the emissions characteristics of alternative human activities.

In the remainder of this section, we will briefly highlight some of the more salient features of the SGM.

2.3 Subsectors and Technologies

Within the nine producing sectors of SGM Version 0.0, several sectors have subsectors and technologies defined. Each subsector, defined for a specific sector, is assumed to produce a homogeneous good which becomes part of the total production of the sector. For example, the Electric power generating sector has six subsectors defined: Oil, Gas, Coal, Biomass, Nuclear, and Solar/Hydro. Subsectors defined for the SGM Version 0.0 are shown in Table 1.

Table 1: Subsector in SGM Version 0.0

Producing Sector	Number of Subsectors	Description of Subsectors
1. Agriculture	1	Agriculture
2. Crude Oil Production	5	Crude oil is treated as a depletable resource with five grades modeled as subsectors.
3. Natural Gas Production	5	Natural gas is treated as a depletable resource with five grades modeled as subsectors.
4. Coal Production	5	Coal is treated as a depletable resource with five grades modeled as subsectors.
5. Uranium Production and Refinement	5	Uranium is treated as a depletable resource with five grades modeled as subsectors.
6. Electric Power Generation	6	Power generation by mode is defined for Oil, Gas, Coal, Biomass, Nuclear, and Solar/Hydro.
7. Oil Refining	3	Conventional Oil Refining, Coal Liquefaction, and Biomass Liquefaction.
8. Natural Gas Transformation	3	Natural gas transformation and distribution is defined for Natural Gas, Coal Gasification, and Biomass Gasification.
9. Other Production	1	Other Production

Just as sectors can have subsectors, subsectors can have technologies. Technologies are alternative modes for producing the product of the subsector. For example, there may be two technologies for producing electricity with natural gas (a subsector of Electric Power Generation): Conventional combustion, and Combined Cycle combustion. The framework of the model is sufficiently flexible to allow an arbitrary number of technologies to compete.

2.4 Natural Resources

Natural resources are treated explicitly in the SGM, which identifies two forms: depletable and renewable. Depletable resources are consumed in use, for example, fossil fuels. Renewable resources are not consumed in use, for example, land. This distinction is important to the treatment of energy production and transformation, as well as agricultural activities.

Fossil fuels and uranium are treated as depletable natural resources. They are divided into two categories, Resources and Reserves. Reserves are those energy sources whose location is known, which are producible using present technologies under present and anticipated economic conditions, and which investment for extraction purposes has been made. (Some crude oil, discovered conventional, and some Alaskan oil are examples of crude oil reserves.) Resources include all energy sources, including those which are known by location and those whose existence is inferred, those which are producible under present technologies and economic conditions, and those which may require greater incentives to exploit. Total resources include reserves. (In the SGM, reserves are created as economically recoverable resources are discovered). In the SGM, energy production from depletable natural resources occurs only from reserves. The rate of production from reserves depends on the amount of productive capacity put in place with discovery, and prices of inputs and outputs.

Biomass, solar power, and hydroelectric power are renewable resources. Biomass feedstocks are treated as a component of the agricultural production. Total domestic production is therefore constrained both by the total

available arable land, and the competition for the use of that land for other purposes. Biomass feedstocks are either transformed to liquids or gases, or consumed as a solid by electric utilities.

Electric utilities also consume solar and hydro power. Hydro power capacity is fixed for the purposes of this analysis. No resource constraint is placed on the solar power component, a source of electricity generation based on photovoltaic array technology.

Land is modeled as a separate factor of production. The identification of land reflects its potential importance as a constraint in emissions reduction strategies which utilize biomass or carbon sequestration by trees.

2.5 Expectations, Capital Formation, and Productivity

In the SGM the demand for capital by producing sectors depends on expected profits. Expectations of profits in turn depends on both the technology which describes the relationship between inputs and outputs, and expected prices over time, for inputs and outputs, including capital. The SGM provides a variety of options for describing the formation of price (and tax/subsidy) expectations, including assumptions that

1. Prices will remain at current levels over the expected life of the equipment (assumed in this exercise),
2. Prices will change at rates which reflect behavior over a prior period of experience, and
3. Price expectations are exogenously specified (which can be used to generate a rational expectations price path).

The last capability can be used to explore behavior when future prices are known. We note also that different combinations of expectation formation for market prices and taxes/subsidies can be constructed. Thus, the formulation of price expectations is not restricted to the assumption of perfect knowledge about future events.

Sectors and subsectors with the highest expected profits experience the greatest realized investments, although within a sector a logit function distributes investment resources across all subsectors which are economically attractive. Investment opportunities which have zero or negative expected contributions to wealth (after allowing for taxes and subsidies) receive no investment funds. The inclusion of subsidies in the expected contribution to wealth calculation is important.

The SGM uses a "putty-clay" specification of capital. That is, once an investment occurs, capital is permanently associated with that particular application. Thereafter, capital does not move from one sector to another. This is a particularly useful assumption for modeling energy applications where capital investments are highly specialized. It is also not particularly restrictive to the rest of the economy, which is modeled as a single aggregate sector.

Using a vintage approach to modeling capital also means that capital investments become a fixed cost of production. Thus, existing vintages continue to operate as long as they can cover their operating expenses, even if the rate of profit differs from that anticipated at the time the original investment was made. We note here that other factors of production are not fixed; they can be varied to either expand or contract output from an existing facility. Other factors of production also move freely among alternative applications.

Because a production function is associated with each capital investment by application, factor productivity can be manifest as both embodied and disembodied changes in input requirements. That is, combustion efficiencies may be embodied in the physical plant and vary depending upon the date of installation. On the other hand, agricultural productivities may be weakly associated with capital stocks, and more closely associated with changing management practices and seedstocks. The SGM can represent both types of productivity change.

2.6 Demographics

Estimates of population and its structure are developed within the SGM by a demographic module. The SGM is therefore capable of creating an array of internally consistent estimates of population and critical details such as the associated size of the work age population disaggregated by gender. The SGM demographic module builds population estimates from assumptions about age specific fertility rates, survival rates, and net immigration rates.

2.7 Emissions

The SGM was explicitly designed to provide estimates of gaseous emissions from all human activities, including those associated with energy, agriculture, and industrial processes. Not only does the model yield estimates for carbon emissions, but it also tracks emissions of CO, CH₄, VOCs, N₂O, NO_x, and SO₂. These emissions are associated with specific human activities, and where appropriate, with specific technologies.

2.8 Data Sources and Calibration

Parameters for the United States module of the SGM are derived so as to be consistent with behavior of the United States economy in 1985, investment behavior over the prior 26 years, and physical flows of energy resources in 1985, as well as the distribution of resources and reserves in that year, the demographic structure of the United States population, and emissions rates by human activity in 1985.

A variety of statistical information was used to develop calibration parameter estimates for the SGM. Table 2 contains a summary of these data, their sources, and an indication of their use.

Table 2: Data Sources for SGM Version 0.0 Calibration to 1985 United States Economy

Data Type	Data Source	Use
1985 Input-Output Table	DOC (1990)	Provide statistics for flows between different economic activities in 1985.
1985 Energy Flows	EIA (1991)	Provide benchmark flows of energy throughout the economy for 1985.
Investment Patterns 1959-1985	DOC (1987)	Provides pattern of capital accumulation by sector.
Resource and Reserve Estimates	EIA (1992) DOE (1991a)	Provide total resource constraints on fossil fuels and uranium, distinguish between reserves and resources, provide a basis for allocation of reserves and resources to economic grades.
Demographic Profile	DOC (1992)	Estimate fertility, survival rates, and net immigration in the demographic module.
Land	DOC (1992) USDA (1992)	Estimate payments to land by sector of the economy.
Emissions	DOE(1991a)	Provide estimates of emissions fluxes by human activity.

The methods used to transform data into model parameters are documented in detail in Fisher-Vanden, et al (1993).

3.0 Reference Case Assumptions: 1985-2030

We begin with the analysis of policy options for controlling United States greenhouse gas emissions by constructing a reference case against which to view policy derived impacts. We firmly believe that no reference case that we could construct today is capable of representing the future. An alternative to a single or multiple reference cases is the development of probabilistic scenarios. But while these map out a range of alternative futures systematically, they suffer from problems of specifying input parameter probability distributions as well as problems with covariance among exogenous variables. Furthermore, such scenarios lead to difficulties in analyzing and communicating results. We therefore chose the use of a reference case which is reproducible, and for which policy induced variations can be explained.

The SGM United States module requires input assumptions in six different areas: Demographics, Energy Resources, Productivity Change, Nuclear Power, International Trade, Fiscal Policy, and Emissions Coefficients. We discuss each in this section.

3.1 Demographic Assumptions

The demographic module in SGM Version 0.0 develops population estimates using exogenous fertility, survival, and migration rates. Population is differentiated by gender and 5-year age group.

Figure 2 shows the demographic profile of the U.S. for our benchmark year of 1985 (Department of Commerce, 1992). The bulge in the chart for the age groups of 20-40 years corresponds to the "baby boom" generation. Even by 1985, it is clear that the cohorts that follow are much smaller. By the first quarter of the next century, the population bulge will have reached retirement age, and the fraction of population that is available for the labor supply may be much smaller than it is currently.

In SGM version 0.0, fertility rates, survival rates, and migration rates are determined exogenously. We have specified initial values for these rates based on data for 1985. We also included a terminal value for each of these rates as a modeling parameter. Both the terminal values and the time to reach the terminal values can be specified by the user. Table 3 shows the initial values of the demographic rates which are assumed to persist throughout the modeled period.

Table 3: Demographic Rates (per thousand people)

Age	Fertility Rates	Male Birth Fraction	Female Death Rates	Male Death Rates	Migration Rates
0-4	0.0	.51	2.35	2.89	2.80
5-9	0.0	.51	0.20	0.3	2.30
10-14	0.0	.51	0.20	0.30	2.30
15-19	53.6	.51	0.50	1.10	2.30
20-24	111.5	.51	0.60	1.70	5.60
25-29	113.4	.51	0.60	1.70	3.70
30-34	73.7	.51	0.80	1.90	2.30
34-39	27.9	.51	1.10	2.30	2.30
40-44	4.8	.51	1.80	3.20	2.30
45-49	0.0	.51	2.90	5.20	1.40
50-54	0.0	.51	4.60	8.50	1.20
55-59	0.0	.51	7.30	13.50	0.50
60-64	0.0	.51	11.20	20.80	0.50
65-69	0.0	.51	16.70	31.00	0.50
70-74	0.0	.51	26.00	47.70	0.50
75+	0.0	.51	75.89	100.97	0.50

Source: Department of Commerce (1992).

3.2 Depletable Energy Resources and Reserves

By definition the resource base is the total quantity of a resource which could be produced over all time using any conceivable technology. It includes both discovered and undiscovered quantities. It is therefore the physical constraint on cumulative production. As indicated earlier, we distinguish energy resources for oil, gas, and uranium by grades reflecting variation in the ease of extraction which in turn drives costs. Five grades are defined for this exercise. The estimates of the energy content available in each of the five grades (excluding previously consumed quantities), as well as the relative cost of extracting each grade compared to Grade 1 are provided in Table 4.

Table 4: U.S. Energy Resources Remaining and Relative Costs by Grade

	Grade 1	Grade 2	Grade 3	Grade 4	Grade 5	Total
Crude Oil (EJ)	80	170	208	253	379	1,090
Relative cost	1.0	1.4	1.8	2.4	3.4	
Natural Gas (EJ)	313	152	271	450	203	1,388
Relative cost	1.0	1.85	2.7	3.84	16	
Coal (EJ)	5,978	7,837	11,755	25,000	27,500	78,070
Relative cost	1.0	1.3	2.7	5.3	9.0	
Uranium (tons)	4,338	8,115	13,211	0	0	25,664
Relative cost	1.0	2.2	3.2	3.2	3.4	

Source: Based on estimates derived in Edmonds and Reilly (1985), and updated with information from DOE (1991a).

In the United States coal is by far the most abundant conventional fossil fuel resource. In fact, the domestic supply of coal in the cheapest grade--Grade 1--is enough to sustain current levels of production for nearly 300 years. Natural gas is the next most abundant domestic source of energy, although much of its supplies are in the more expensive grades. Gas resources in the first three grades are sufficient to sustain current production for approximately the next 35 years; therefore, the supply of relatively inexpensive natural gas becomes a binding constraint over the year 2030 study horizon of the model. Currently about half of the oil consumed in the U.S. is imported, so the implications of consumption on domestic resources are not as straightforward. Under current production levels, the supply of oil from the first three grades is clearly a constraint over the next 40 years. However, the future price of foreign oil plays a major role in determining domestic investment in the more expensive grades. Even so, domestic supplies of the more costly grades of oil are limited.

3.3 Productivity Change

One of the more important assumptions affecting the development of a reference description of the United States economy is the set of assumptions which govern productivity change in the variety of human activities in the SGM. Productivity changes in the SGM in either of two ways, through smooth changes in production function coefficients associated with new investment options, which we refer to as general productivity growth, or by the introduction of a new technology option at a specific point in time.¹

Assumed rates of change for general, Hicks neutral, productivity in new investments are given in Table 5.

¹ Similarly the model can remove technologies from the available set; e.g., when new source performance standards are introduced.

Table 5: Assumptions of Exogenous Change in Total Factor Productivity for New Investments

Sector	Productivity Growth Assumption (%/yr)
1. Agriculture	0.6
2. Crude Oil Production	0.6
3. Natural Gas Production	0.6
4. Coal Production	0.6
5. Uranium Production and Refinement	0.6
6. Electric Power Generation	0.2
7. Oil Refining	0.0
8. Natural Gas Transformation	0.0
9. Other Production	0.6

Since the Other Production Sector produces approximately 85 percent of total output, the productivity growth rate of this sector was adjusted to attain a growth forecast of the U.S. economy similar to what was assumed in the National Energy Strategy. Lack of detailed knowledge of variations in future productivity growth between sectors resulted in the assumption of no variation in productivity growth rates between the first five sectors and the Other Production Sector. Future versions of the SGM will incorporate current research in this area to better define variations in productivity growth between these sectors.

Productivity growth rates in two of the three energy transformation sectors are set to 0.0% per year to maintain energy balances between inputs and outputs. These three sectors--Electric Power Generation, Oil Refining, and Natural Gas Transformation--differ from the other producing sectors because they use large quantities of energy as inputs; most of which is passed through as output instead of being consumed. For example, the amount of energy leaving the gas distribution sector is nearly equal to the energy coming in, with almost no opportunity for improvement in the ratio of energy output to energy input. In these sectors changes in productivity are modeled as discrete changes in available production functions over time.

Although productivity improvements in the Electric Power Sector are handled mainly through the incorporation of new technologies in the SGM, small productivity improvements resulting from changes in management practices warrant the inclusion of a small neutral productivity growth assumption. This value (0.2) was chosen in order to attain results within the Electric Power Sector similar to those found in AEO93.

This analysis allows five new technologies to be introduced after 1990: liquids from biomass, gases from biomass, solar electric power, clean coal, and new gas turbine. Cost assumptions used to estimate production function parameters for these five technologies are given in Table 6.

Table 6: Annual Operating Costs for Hypothetical Plants

	Solar (1982 MM\$)	Liquids from Biomass (1987 MM\$)	Gases from Biomass (1986 MM\$)	Clean Coal Technology (Atmospheric Fluidized Bed Combustion) (1989 MM\$)	New Gas Turbine Technology (Natural Gas Combined Cycle) (1989 MM\$)
Plant Capacity	1,000 MW	1000 dry tons per day	1000 dry tons per day	500 MW	300 MW
Materials	\$16.50	\$27.52	\$15.67	\$38.57	\$15.39
Labor	\$10.10	\$6.92	\$3.67	\$18.15	\$7.24
Capital (Production)	\$187.10	\$14.00	\$10.88	\$650.00	\$180.00
Capital (Distribution)	\$90.60	---	\$4.17	\$210.32	\$126.22
Total Cost (without taxes)	\$304.30	\$48.44	\$34.39	\$917.05	\$328.85
Indirect Business Taxes	\$16.40	\$1.85	\$1.14	\$9.83	\$4.77
Total Cost (with taxes)	\$320.70	\$50.29	\$35.53	\$926.88	\$333.62
Break-Even Price (1982\$)	\$0.11 per kwh	\$15.71 per GJ	\$8.71 per million BTU	\$0.056 per kwh	\$0.046 per kwh

Sources: Derived from Elliott, et al. (1990), Wan and Fraser (1990), DOE (1991a)

The two biomass technologies purchase their feedstock (wood) from the SGM's agricultural sector, thereby competing with other agricultural products for land. Future versions of the SGM will split the Agriculture sector into various crop sectors as well as a biomass feedstock sector.

The SGM allows for electricity sector investments in solar photovoltaics, a clean coal technology (Atmospheric Fluidized Bed Combustion), and a new gas turbine technology (Natural Gas Combined Cycle) beginning in the year 1990. New capital equipment is required for both generation and distribution. Operation and maintenance costs are very small relative to capital costs. Cost assumptions for solar are somewhat optimistic for the near term.

3.4 International Trade

Because world trade occurs, a single region model is not closed. Closing the SGM requires assumptions about the region's interactions with the rest of the world. These are given in Table 7. The crude oil market is assumed to be open, and the model can import or export as much as it wishes at an exogenously specified price. (This price path is taken from DOE (1991b)). The trade account is assumed to be balanced. This is accomplished by setting net exports of the Other Production sector, Sector 9, to the negative of the sum of net exports from other sectors.

Table 7: Closure Assumptions

Sector	Closure Assumptions
1. Agriculture	1985 net exports fixed
2. Crude Oil Production	World price path given exogenously
3. Natural Gas Production	1985 net exports fixed
4. Coal Production	1985 net exports fixed
5. Uranium Production and Refinement	1985 net exports fixed
6. Electric Power Generation	1985 net exports fixed
7. Oil Refining	1985 net exports fixed
8. Natural Gas Transformation	1985 net exports fixed
9. Other Production	Set to balance trade account

3.5 Fiscal Policy

Budgets for all government activities taken together, including federal, state, and local, are assumed to be balanced. That is, federal deficits are assumed to be offset by state and local surpluses.

4.0 The Reference Case

The assumptions discussed in the preceding section generate a reference case trajectory for the United States economy, energy system, and greenhouse gas emissions. We discuss each in turn.

4.1 The Reference Case Economy

While the rate of population growth and level of economic activity are input assumptions for many models, they are results of the SGM. We begin by discussing trends in population and economic activity. Figure 3 shows the real gross national product (GNP) and its three principal components: consumption, investment, and government spending (recall that net exports are zero by assumption). GNP grows throughout the period of analysis, more than doubling between 1985 and 2030. We note, however, that the rate of growth slows distinctly after the year 2010. The rate of population growth slows to 0.29 percent/yr after 2010 compared with the post world war II period average of 1.01 percent/yr. By 2030, population is actually declining, even given relatively high immigration rates, and a decline in mortality rates.

In order to understand the economic forces at work within the SGM, it is useful to examine why GNP growth falls from 1995 to 2030. The decline is steady from 1995 to 2025 when growth rates appear to stabilize at about a half percent per year. This drop occurs primarily because of the fall in the size of working age population (Figure 4), from 182.5 million in 2010 to 167.8 million by 2030, an 8.0 percent decline. This leads to only a 2.9 percent reduction in the number of workers (149.2 million to 144.9 million) during the period 2010 to 2030 as labor force participation increases from 81.8 percent to 86.4 percent, and wage income per worker per year rises from \$33,190 per year to \$44,380 per year. As might be expected given the scarcity of labor, the capital/labor ratio increases over the same time frame from \$136.9 to \$165 million per 1000 workers, continuing the trend seen earlier (Table 8).

Table 8: Factors affecting Employment

Year	Work Age Population (millions)	At Work (millions)	Percent At Work	Wage Annual (\$1000*)	Capital Labor Ratio (\$1000* per worker)
1985	158.8	109.8	69.1	20.28	103.0
1990	163.0	116.3	71.3	21.88	108.7
1995	167.0	125.0	74.8	24.74	114.8
2000	172.5	134.0	77.6	27.56	122.8
2005	179.0	143.0	79.9	30.29	128.0
2010	182.5	149.3	81.8	33.19	136.9
2015	181.4	151.2	83.3	36.06	145.2
2020	177.3	149.8	84.5	38.78	154.0
2025	171.9	147.1	85.6	41.72	160.1
2030	167.8	144.9	86.4	44.38	165.0

* in 1982 baseyear (1985) dollars.

Since labor is scarce, the price of labor rises over time, leading to a substitution away from labor. This is seen in the much larger decline in the ratio of labor to output than in the ratio of capital to output. Figure 5 presents the path of prices for factors of production capital, labor, and land. Not only does the price of labor rise, but the fraction of the working age population at work rises as seen in Table 8.

The decline in the work age population has another route for affecting growth rates. Part of the determinant of the desired investment is the change in the rate of potential growth of the economy, namely the change in worker productivity and the change in working age population.

While productivity, measured as GNP/Worker, continues to grow rapidly (Table 9), working age population peaks in 2010 and then declines, thus depressing growth in investment. Examination of the components of GNP makes this clear, as investment becomes static after 2015.

This leads to an increase in the average age of capital, slowing down the rate at which technical change in realized in the economy. (By assumption, neutral technical change occurs at a constant pace in each industry).

Table 9: Factors Affecting the Capital Stock

Year	Capital Stock (\$trillion)	Capital/Output Ratio	Interest Rate	Output/Worker (\$1000)
1985	11.30	3.096	.0312	33.3
1990	12.63	3.066	.0313	35.4
1995	14.34	3.013	.0314	38.1
2000	16.45	3.024	.0312	40.6
2005	18.31	2.977	.0309	43.0
2010	20.44	3.006	.0306	45.6
2015	21.99	3.020	.0305	48.2
2020	23.00	3.029	.0304	50.7
2025	23.59	3.019	.0304	53.1
2030	23.87	2.968	.0303	55.5

4.2 The Reference Case Energy System

Total primary energy use increases steadily over the modeling period with an overall increase of approximately 55 percent between the years 1990 and 2030 (Figure 6). A period of rapid increase (averaging more than one percent annually) occurs during the period 1990 and 2005 mirroring the steady growth in the economy. A relative modest growth (less than half a percent average annual rate) in primary energy consumption occurs between 2005 and 2020 as a result of a slowdown in the economy, and an increase in oil and gas energy prices caused by the depletion of the less costly grades of domestic resources. This leads to a decrease in the ratio of energy use to GNP during this period. A period of accelerated growth in the last 10 years of the analysis (an average annual growth of slightly less than one percent per year) occurs due to a shift to electricity consumption increasing the consumption of coal.

Oil remains the dominant fuel over this period (despite the fact that its price is assumed to rise by more than 50 percent between the years 1990 and 2010) rising to almost 50 exajoules per year by the year 2015. Consumption is relatively stable throughout the remainder of the period of analysis. The escalation in oil consumption leads to a substantial increase in oil imports beginning in the year 1990.

The only primary energy carrier whose price remains relatively stable is coal. We note that the price of electricity rises by about 50 percent between the years 1990 and 2010, but the introduction of improved technologies and a decline in coal prices lead to a reduction of cost by more than one third between 2010 and 2030.

The relative stability of the price of coal compared to other fuels leads to an increase of domestic coal production and consumption which rises by more than 50 percent between 1990 and 2030. Natural gas consumption increases by 40 percent between 1990 and 2015, but depletion of the least expensive grades of domestic natural gas resources after 2015 causes a sharp increase in the price of natural gas and thus stabilizes consumption.

The assumption that no new nuclear powerplants are built leads to a stable pattern of nuclear electricity production in the near term, but after the year 2010 production declines steadily until all nuclear facilities are retired in the year 2025.

Electric power consumption as a whole expands steadily over the period of analysis. Larger increases in consumption of electric power occur after 2015, when a fall in the relative price of coal causes a shift towards electricity consumption.

4.3 Reference Case Greenhouse Gas Emissions

The focus of this analysis is to assess policy options for stabilizing United States greenhouse gas emissions by the year 2000 at 1990 levels. The reference emissions provide an important point of departure. We begin by examining anthropogenic carbon emissions. In the United States carbon is emitted predominantly by fossil fuel combustion, though total emissions include emissions from other human activities, notably cement manufacture and agriculture.² The time profile of emissions of carbon is shown in Figure 7. In the reference case emissions are anticipated to rise relatively rapidly between the years 1990 and 2000, at an average rate of 1.6 percent per year. The somewhat more rapid ramp-up in emissions than in total energy use is the consequence of an increased fraction of coal in the energy mix. While the rate of growth of carbon emissions declines somewhat over time, the continuing increase in the share of energy use in the form of coal results in a growth in carbon emissions that remains above that of total energy use.

Emissions from other greenhouse gases are displayed in Figure 8. Since the majority of carbon monoxide (CO) emissions are a result of oil combustion activities, we see the growth in these emissions mirroring that of oil consumption. Since we see a steady increase in oil consumption over the modeling period, we also see a steady rise in CO emissions. Other greenhouse gas emissions also experience a steady increase over the modeling period. SO_x emissions stabilize after 2000 but rise again after 2020, mirroring coal consumption during this period.

5.0 Stabilizing United States Greenhouse Gas Emissions

The issue which we consider in this paper is strategies for stabilizing United States emissions of greenhouse gases at 1990 levels by the year 2000. Due to the complexity of comparing the various gases' impact on the earth's radiation balance, we chose to focus our analysis in this paper on the stabilization of fossil fuel carbon emissions; the simplest concept to measure and regulate³.

Two strategies for stabilizing carbon emissions were analyzed: (1) constant and variable taxes on fossil fuels based on each fuel's carbon content; and (2) variable energy taxes on fossil fuels based on each fuel's energy (BTU) content.

5.1 Carbon Taxes

The imposition of carbon taxes is the most direct price instrument to reduce carbon emissions. A policy of carbon taxes was modeled in the SGM as a tax on the production of fossil fuels applied as an additive tax based on the proportion of each fuel's carbon content (i.e., \$ per ton of carbon (TC)). Carbon taxes affect carbon emissions and economic factors by raising the price of fossil fuel by the amount of the additive tax. Based on the known carbon content of various fuels, Table 10 shows how a \$135/TC carbon tax (the lowest constant tax rate which would achieve carbon emissions stabilization) is translated into price increases for each of these fuels. The price increase is largest for coal since coal contains the largest concentration of carbon per 10⁶ Btu.

² We note that the dominant non-energy related emissions of carbon are from ruminant livestock, and are given off in the form of CH₄.

³The concept of a global warming potential (GWP) was developed by the Intergovernmental Panel on Climate Change (IPCC) and elsewhere, to provide a metric by which to compare the various gases. But problems in defining these coefficients have emerged which have been most acute with regard to those gases with the greatest indirect effects on climate change; precisely those gases for which the GWP is most important.

Restricting the concept of greenhouse gas emissions narrowly implies that some potentially cost effective measures to reduce net carbon emissions to the atmosphere might be missed, implying higher than necessary costs for achieving any emissions reduction objective.

Table 10: Taxes Based on Carbon Content of Fuels

Fuel	\$135/TC Carbon Tax
Crude Oil (\$/barrel)	\$24.64
Crude Oil (\$/10 ⁶ Btu)	\$2.84
Crude Natural Gas (\$/1000 ft ³)	\$2.03
Crude Natural Gas (\$/10 ⁶ Btu)	\$1.96
Coal (\$/ton)	\$75.60
Coal (\$/10 ⁶ Btu)	\$3.38
Gasoline (\$/gal)	\$0.35

Carbon Emissions

The SGM was used to determine the carbon tax rate required in each period to offset carbon emissions above the 1990 emissions level. These tax rates are shown in Figure 9. Tax rates begin at \$43/TC in 1995 and rise steadily until 2015 when a peak tax rate of \$140/TC is reached⁴. All revenue received from the carbon tax is recycled back to households as additions to personal income. Higher carbon tax rates after 2015 are not required due to a fall in the ratio of carbon (in TgC) to real GNP, implying a shift to more energy-efficient methods of production and less carbon-intensive fuel use.

Figure 10 compares the level of carbon emissions under a "business-as-usual" scenario with various "flat" carbon tax and "stabilization" carbon tax options. Assuming a carbon emissions reduction goal to maintain carbon emissions below the 1990 level, the SGM shows that a flat carbon tax of no less than \$135/TC is required.

Impacts on GNP and Consumption

Figure 10 also shows that the \$135/TC flat tax case results in an exceedance of the emissions reductions required in the years 1990 to 2030 to stabilize emissions. This will result in greater economic losses than the "stabilization" tax case. Figures 11 and 12 show the percent loss in GNP and consumption, respectively, resulting from a \$135/TC "flat" carbon tax and "stabilization" carbon tax. As expected, the \$135/TC tax case results in much lower growth rates than the "stabilization" case in the beginning years but converges with the "stabilization" case in later years.

It is interesting to note that the peak variable tax rate (\$140/TC in the year 2015) is higher than the constant tax rate (\$135/TC). This is attributable to the fact that the constant tax rate of \$135/TC is built into earlier investment decisions (by way of expected prices) while the tax trajectory is not known in the case of the variable tax, thus not allowing producers to make the most optimal investment decisions. However, the difference between the two tax rates (\$135/TC and \$140/TC) is small, implying that the additional knowledge of future tax rates makes small differences in investment decisions.

The "flat" and "stabilization" carbon tax cases result in similar impacts, differing only in the size of the impacts; therefore, it would suffice to concentrate on the impacts of one. Since the "stabilization" tax case achieves stabilization of carbon emissions with the least economic cost, any future comparisons will concentrate on this carbon tax case.

⁴ Using a 1985 price for crude oil, natural gas, and coal of \$4.15, \$2.26, and \$1.15 per MBTU, respectively (Annual Energy Review, 1991), a \$43/TC carbon tax would amount to an increase in price ranging from 22 percent for oil to 93 percent for coal while a \$140/TC carbon tax would amount to an increase ranging from 71 percent for oil to 304 percent for coal.

Impacts on Primary Energy Consumption

Results from the SGM show a dual effect on primary energy consumption resulting from carbon taxes - a reduction in total primary energy consumption and a shift to lower carbon emitting fuels. Reductions in total primary energy consumption range from 8 percent in 1995 to 19 percent in 2030, while the share of coal consumed (out of all primary fuels) falls from 26 percent to 18 percent in 2030. This would indicate that the carbon taxes had a greater effect on reducing overall energy consumption than shifting energy consumption from higher carbon emitting fuels to lower ones. The Electricity Sector experiences similar impacts with the reduction in total fuel consumption outweighing the shift from more carbon-intensive fuels to less carbon-intensive ones. It is interesting to note, however, that although electricity fuel consumption drops by 28 percent in 2030, total electricity generated only drops by 9 percent. This can be attributed to a shift to less carbon intensive fuels, which are associated with more energy efficient methods of generation.

The variable carbon tax also has effect of making biomass fuels economically attractive in the later years. In the reference case, no biomass fuel production occurs throughout the modeling period 1990 to 2030. With the implementation of carbon taxes, we see both liquid and gaseous fuels from biomass being produced. By 2025, gaseous fuels from biomass make up 17 percent of total gas produced, while liquid fuels from biomass only make up 0.4% of total refined liquids produced.

Impacts on Energy Resources

Further evidence that carbon taxes have a greater impact on energy efficiency than on substitution of less carbon-intensive fuels can be seen in the depletion of energy resources. Instead of seeing a more rapid depletion of less carbon-intensive fuels (e.g., natural gas) than the reference case, production of these fuels falls with the imposition of carbon taxes, leading to less depletion of these resources. In addition, as would be expected, production of the more carbon intensive fuels is lower with the imposition of carbon taxes, leading to less depletion of these resources also.

Ratios

A shift to more energy efficient modes of production in reaction to carbon taxes is also apparent when we look at the ratio of energy use (in EJ) to GNP (in trillion \$). Differences from the reference case range from 8.1 percent in 1995 to 18.3 percent in 2030. To a much lesser extent, the fossil fuel use (in EJ) to energy use (in EJ) ratio shows a shift to less carbon-intensive fossil fuels with a drop in this ratio of 2.1 percent in 2030 from the reference case, proving that carbon emissions reductions were obtained mainly through energy efficiency measures. Lastly, since population projections do not differ between cases, the fall in the ratio of GNP to population resulting from carbon taxes mirrors the loss in GNP. This loss ranges from 0.4 percent in 1995 to 0.7 percent in 2030.

5.2 Energy Tax

Another policy option which has the effect of reducing overall carbon emissions is a tax on energy use. The Clinton Administration, as part of its budget proposal, suggests a tax based on the BTU content of the various energy sources (i.e., all fossil fuels, and nuclear and hydro electric power generation). Points of tax collection would be at the minemouth for coal, refinery for crude oil, pipelines for natural gas, utilities for hydro and nuclear power, and points of importation for imported electricity and petroleum products. Nonconventional energy sources such as solar and wind would be excluded from the tax. Clinton's BTU energy tax calls for a basic rate of 25.7 cents per million BTUs with a supplemental oil tax of 34.2 cents per million BTUs (amounting to a 59.9 cents per million BTUs tax on oil). This energy tax is equivalent to \$28.50, \$17.70, and \$10.10 per tonne of carbon for oil, gas and coal, respectively⁵. In comparison with the carbon taxes discussed above, the Clinton Administration's energy tax is relatively small. SGM results show small impacts by 2030 as a result of the proposed Clinton energy

⁵The conversion of an energy tax to a carbon tax was done by dividing the \$/BTU energy tax by the amount of carbon per BTU in each fuel. The amount of carbon per BTU is highest for coal and lowest for natural gas; therefore, given a basic energy tax rate of \$0.257/BTU across all fuels, coal has the lowest \$/TC carbon tax rate. Oil has the highest carbon tax rate since it includes a \$0.342/BTU supplemental tax.

tax with a 0.18 percent decline in GNP, a 3 percent decline in primary energy consumption, and a 3 percent fall in total carbon emissions.

In this modeling exercise, we impose the general structure of Clinton's energy tax on a policy to stabilize carbon emissions to the 1990 level. That is, the basic rate of 25.7 cents per million BTUs is increased until carbon emissions are stabilized in each period at the 1990 emissions level. (The supplemental oil tax is also increased in the same proportion as the basic rate). In the SGM, the energy tax is imposed on the refined oil, refined gas, coal and uranium sectors, and the hydro electric power subsector beginning in 1995. As in the carbon tax stabilization case, all tax revenue from the energy tax is recycled to the household sector as additions to personal income.

Carbon Emissions

As in the variable carbon tax case, the SGM was used to determine the energy tax rate required in each period to offset carbon emissions above the 1990 emissions level. These tax rates are shown in Figure 13. The energy tax rate trajectory follows a similar path to that of the carbon tax with rates rising steadily from 1995 to 2015 when the peak rate is reached. Energy tax rates begin at \$1.03/MBTU in 1995 and rise steadily until 2015 when a peak tax rate of \$3.53/MBTU is reached⁶.

Figure 14 shows total tax revenue from both stabilization cases (carbon tax versus energy tax). As is expected, total tax required to stabilize carbon emissions is lower in each period with a carbon tax policy option. From a tax standpoint, carbon taxes are more effective at reducing carbon emissions since this tax targets the carbon content of fuels.

Figure 15 compares the stabilization carbon tax rates and energy tax rates (converted to \$/TC) applied to the coal sector. We see a larger impact from a carbon tax on this sector in the years 1995 to 2015 because this tax has the effect of reducing use of the more carbon-intensive fuels first; notably coal. As coal is driven out in the beginning periods, carbon emission reductions in later years is obtained by reductions in other fuels. An energy tax, on the other hand, has the effect of reducing overall energy use and, due to the supplemental tax on oil, reducing refined oil consumption. Carbon emission reductions are achieved in each period by targeting reductions in the use of each energy source with a emphasis on refined oil. In Figure 15, carbon taxes have a greater impact on the coal sector in the early years since reductions in coal use is targeted with this policy option while an energy tax does not specifically target use of this fuel. In the later years, carbon emission reductions are still obtainable from the coal sector with an energy tax since this policy option did not drive out coal use in the beginning years like a carbon tax. Therefore, we see a shift where a energy tax has a larger impact than a carbon tax on the coal sector in the later years.

Impacts on GNP, Consumption, and Investment

Due to the higher total tax imposed on the economy with a stabilization energy tax, the impacts on GNP and consumption are slightly greater with this tax option than with a carbon tax as shown in Figure 16. Reductions in GNP from a stabilization energy tax range from 0.6 percent in 1995 to 1.7 percent in 2015 while reductions from a stabilization carbon tax range from 0.4 percent in 1995 to 1.2 percent in 2015. Reductions in consumption follow a similar track with reductions ranging from 0.1 percent in 1995 to 1.1 percent in 2015 with an energy tax while a carbon tax results in reductions of 0.02 percent in 1995 to 0.8 percent in 2015.

The energy tax has a greater impact on investment as compared to the carbon tax case especially in later years. Reductions in 1995 investment are 1.28 percent with an energy tax and 1.02 percent with a carbon tax while reductions in 2020 are 2.22 percent with an energy tax and 0.97 percent with a carbon tax. Since the model allows for shifts in capital inputs only in the long-term (the "putty-clay" assumption), it makes sense that investment under a carbon tax would shift to less carbon-intensive energy production. With the energy tax option, however, all energy sources are targeted (with emphasis on refined oil); therefore, the effects of this tax would be to decrease overall investment in energy production.

⁶Assuming a 1985 price paid by electric utilities for refined oil, gas, and coal of \$4.35, \$3.43, and \$1.65 per MBTU, respectively (Annual Energy Review, 1991), an energy tax of \$1.03/MBTU amounts to an increase in price ranging from 24 percent for oil to 62 percent for coal while an energy tax of \$3.53/MBTU amounts to an increase in price ranging from 81 percent for oil to 214 percent for coal.

Impacts on Primary Energy Consumption

Figure 17 shows the impacts of both stabilization tax options on primary energy consumption. Both taxes result in similar impacts on energy consumption with the energy tax resulting in slightly larger impacts. Since an energy tax affects all sources of energy, we would expect this policy option to result in lower overall energy use when compared to a carbon tax policy option.

With both tax cases, we see a shift in electricity fuel consumption from oil and coal to natural gas and renewables. This is because the percent increase in price of oil and coal is greater than that of gas in both tax cases causing a shift to gas consumption⁷. The difference between the two cases is out of which fuel most of the shifting is taking place. Due to the supplemental tax on refined oil with the energy tax, we see most of the decline in electricity fuel consumption coming from refined oil; in fact, refined oil use in the electricity sector is completely phased out by 2000 under an energy tax whereas, under a carbon tax, oil is not phased out until after 2025. With a carbon tax, most of the decline in electricity fuel consumption comes from coal. In either case, the amount of electricity generated is essentially the same implying that each tax option has the same effect on electricity generation.

As with carbon taxes, we see the emergence of gaseous fuels from biomass with an energy tax with gas from biomass making up 14 percent (compared to 17 percent with the carbon tax) of gas production in 2025. No production of liquids from biomass occurs since the energy tax does not distinguish between feedstocks of the refined liquids in its supplemental tax on this sector. (We see a larger amount of the biomass fuels being produced with the carbon tax since these fuels are exempt from any carbon tax; unlike in the energy tax case where all sources of energy are affected).

Impacts on Energy Resources

In both tax cases, we see a faster depletion of natural gas than the reference case. This is due to a shift from oil and coal use in both cases to a more energy-efficient and less carbon-intensive use of energy; namely natural gas. As explained above, this shift is due to the larger percent increases in price of oil and coal than gas in both tax cases.

Ratios

As with a carbon tax, an energy tax results in more energy-efficient and less carbon-intensive energy use. As would be expected, since an energy tax targets all energy sources and does not target those which are more carbon-intensive, we see a slightly faster decline in the ratio of energy to GNP with an energy tax than with a carbon tax. In addition, we see a slightly slower decline in the ratio of carbon to energy with an energy tax than with a carbon tax for the same reason. Lastly, due to the larger reductions in GNP with an energy tax, we see a slightly smaller decrease in the ratio of GNP to population with a energy tax than with a carbon tax.

6.0 Conclusions

The purpose of this paper was to demonstrate some of the capabilities of the SGM and to specifically show its effectiveness as a long-term model of energy use, the economy, and greenhouse gas emissions. The model was designed to assess global emissions from all human activities and the major direct and indirect consequences of potential policies to reduce emissions, capabilities that no currently existing model possesses. As obvious from the model design, the SGM is capable of a wide array of modeling exercises which have not been presented in this paper. Examples include (1) various tax revenue recycling options; (2) reforestation as a policy option;

⁷Although the reasons why oil and coal experience a larger increase in price than gas under a carbon tax is straightforward (since the tax is based on carbon content, the more carbon-intensive fuels will face a larger tax), the reasons under an energy tax are more elusive. In the baseyear, coal is the cheaper fuel (in \$/BTU) as compared to gas. (Conversely, more energy (BTU) per dollar is obtainable with coal). Imposing a constant energy tax (\$/BTU) on both coal and gas, we see that the tax increases the price of coal more than gas since the baseyear price of coal is lower. (This can be seen by dividing the \$/BTU energy tax by the \$/BTU baseyear price of coal). The oil price increase is due to the same reason in addition to a supplemental oil tax.

(3) emissions trading as a policy option; (4) new investment in nuclear energy; (5) investment tax credits; (6) new source compliance; and (7) implementation of new energy technologies and retrofits, to name a few.

Future versions of the SGM will further expand on the capabilities of the model. Future expansions of the model include the provision of energy services and traditional biomass fuels, and the addition of a detailed description of the agriculture and health sectors to enable the model to consider the consequences of atmosphere and climate change in detail.

FIGURE 1

The Flow of Goods and Services in the SGM

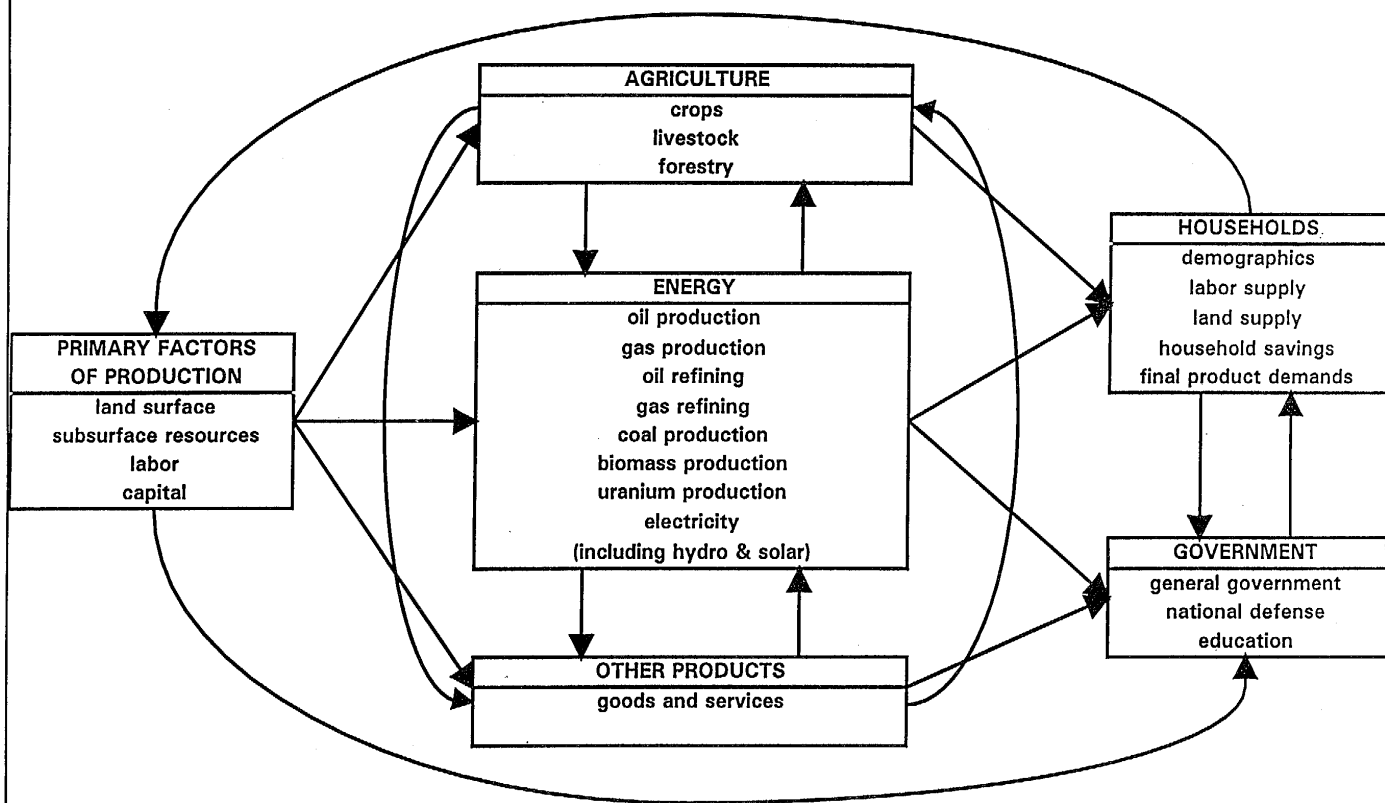


FIGURE 2: 1985 Population

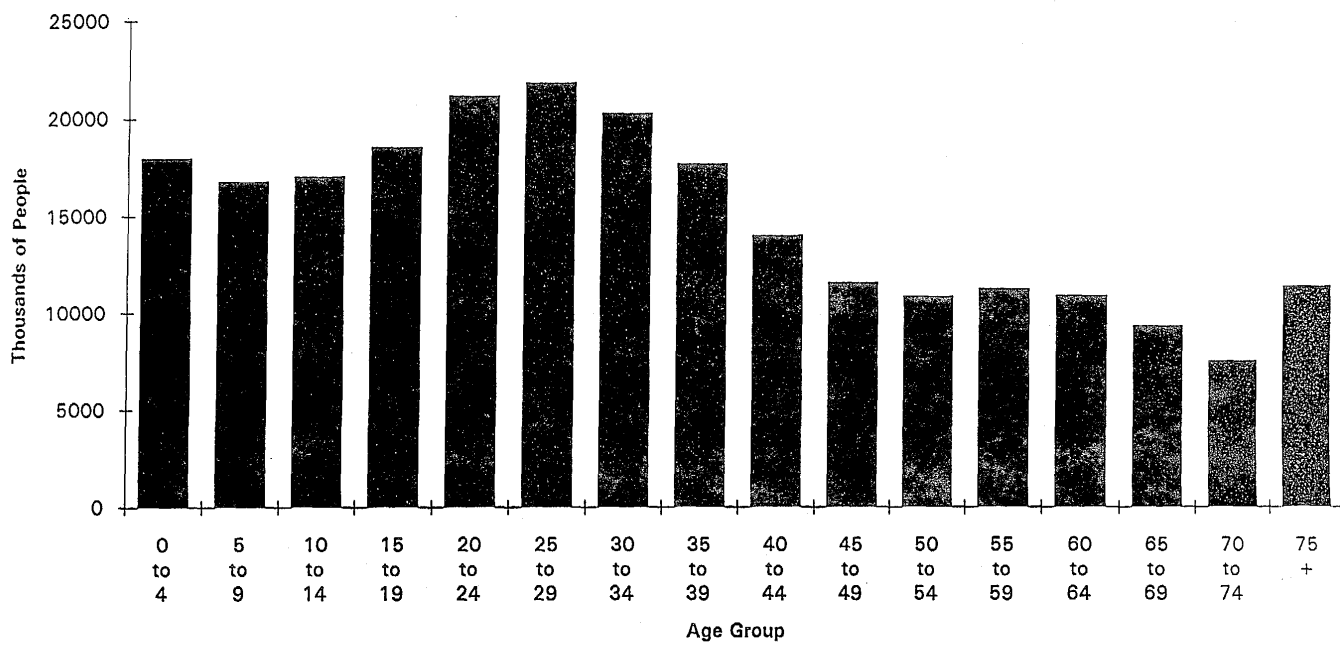


Figure 3: Components of Real GNP (1982\$): Historical and Reference Case

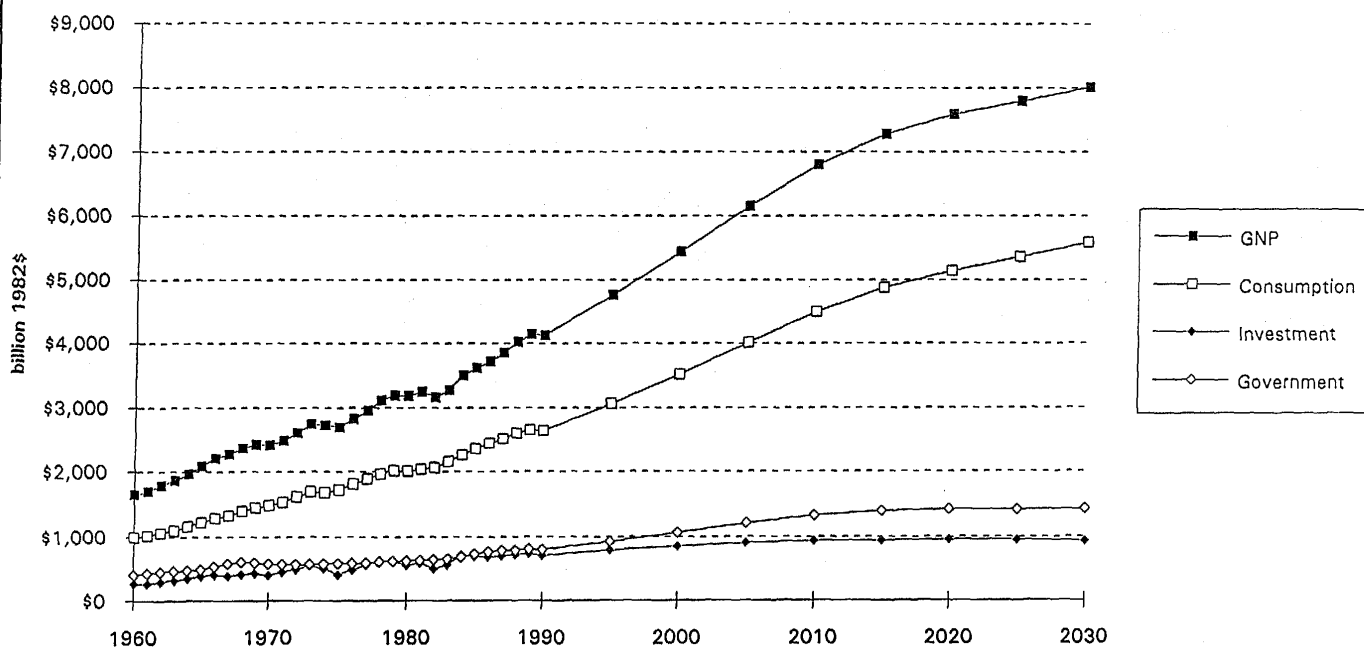


FIGURE 4: Working Age Population

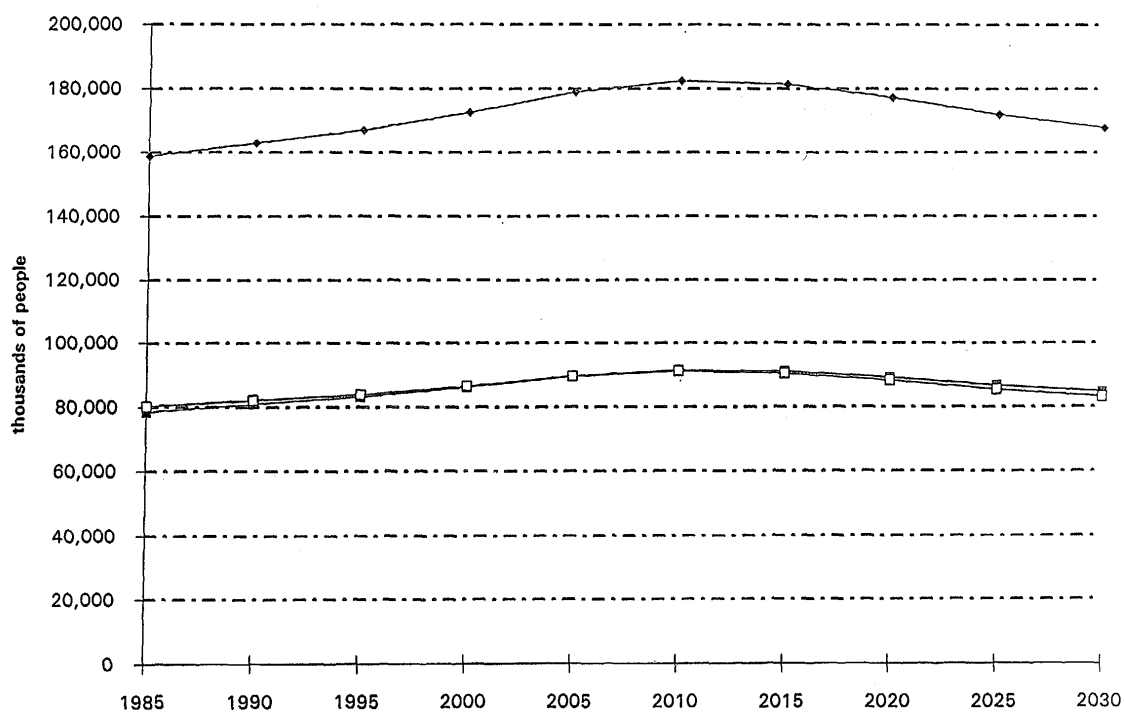


FIGURE 5: Prices of Primary Factors of Production (Base Year Prices)

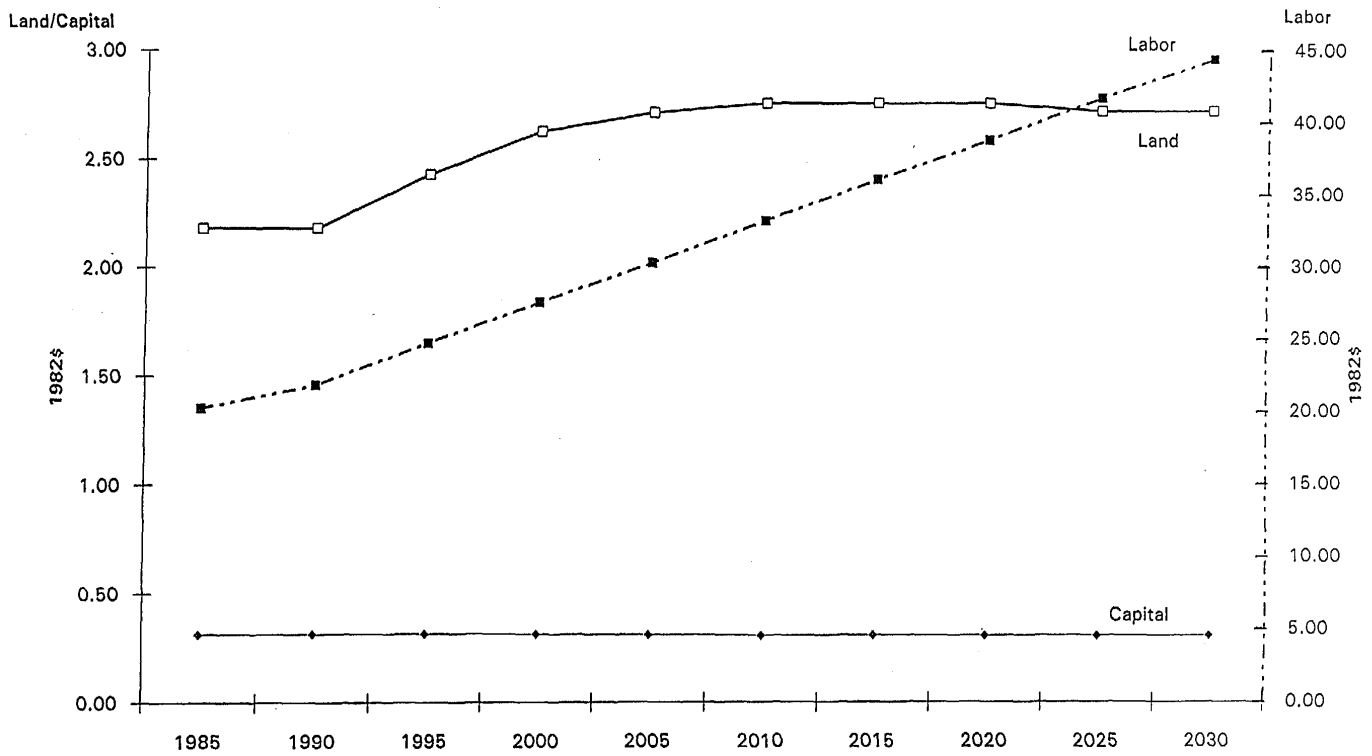


FIGURE 6: Primary Energy Consumption

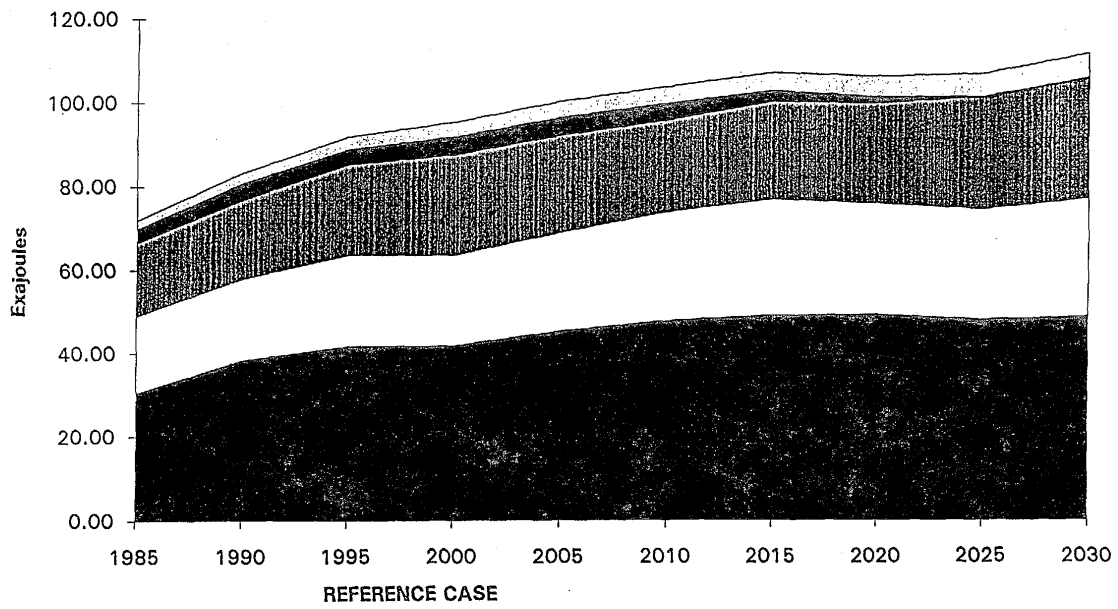


Figure 7: Co2 Emissions (TgC) : Historical and Reference Case

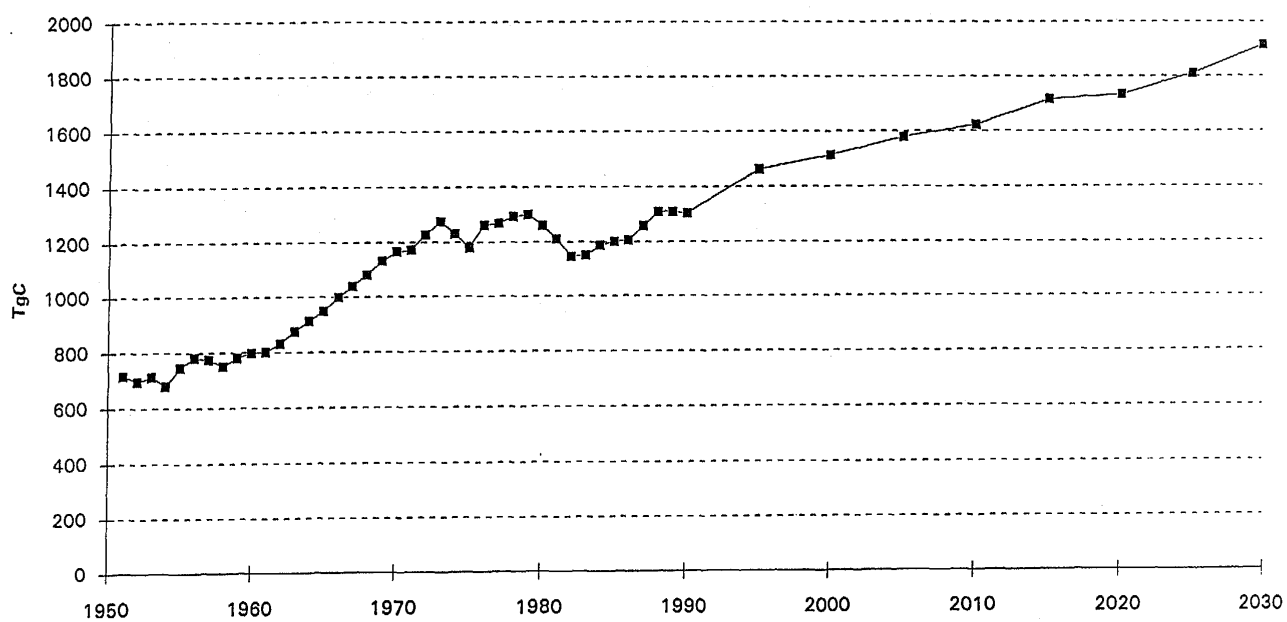


FIGURE 8: Other Greenhouse Gas Emissions

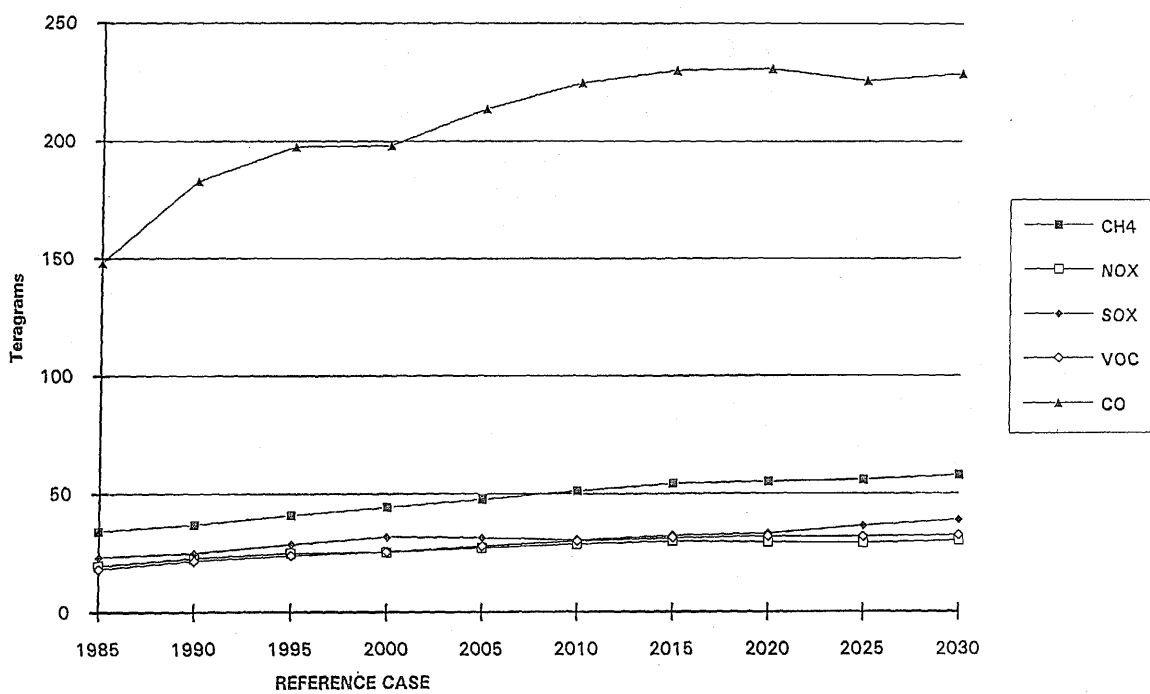


FIGURE 9: Carbon Tax for CO₂ Stabilization to 1990 Level

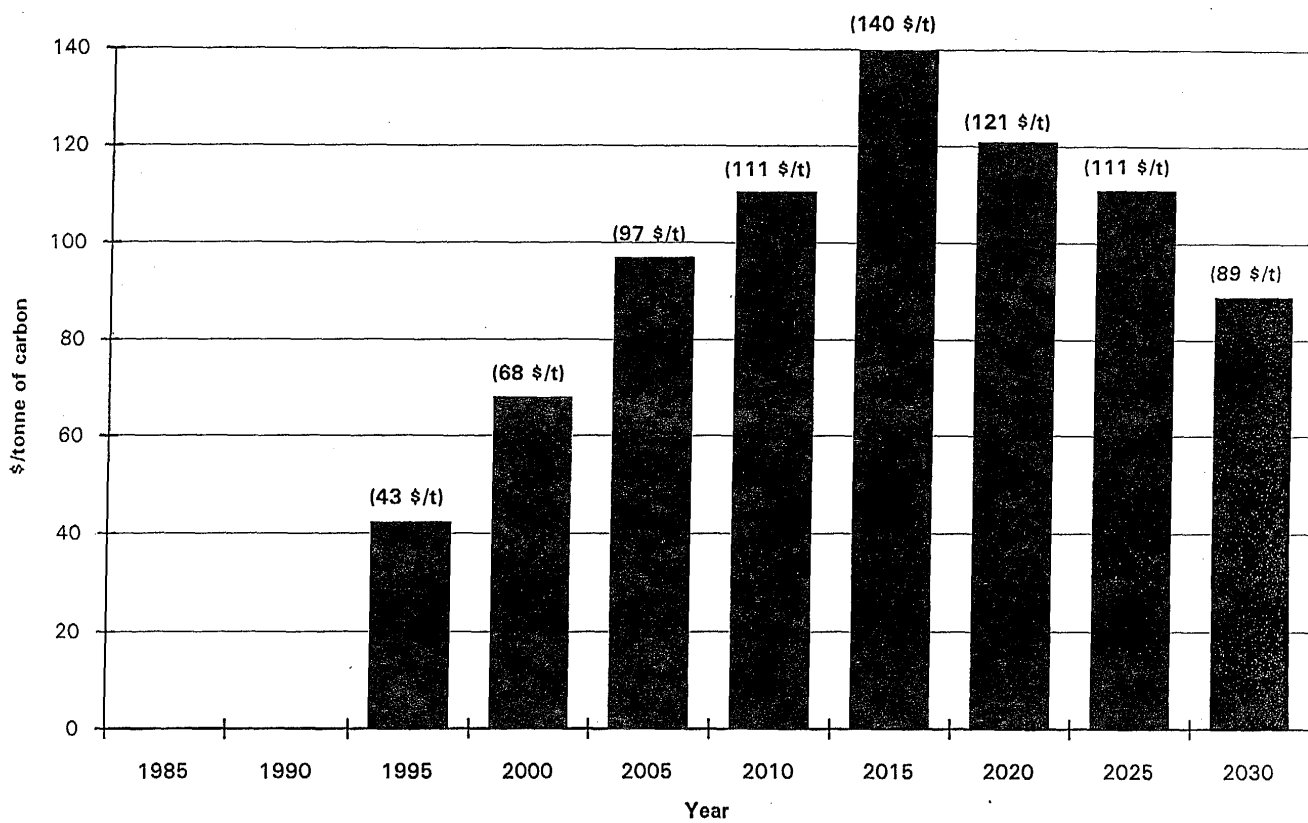


FIGURE 10: Total Carbon Emissions

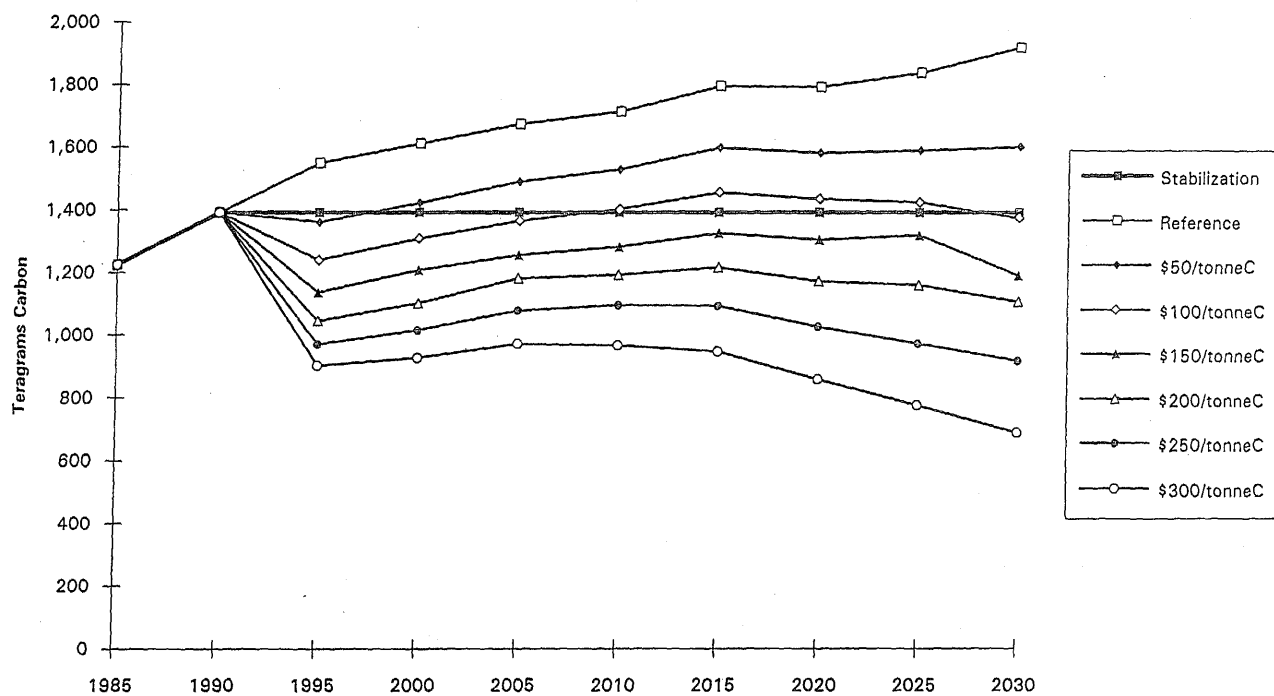


FIGURE 11: % Reduction in Real GNP from Reference Case

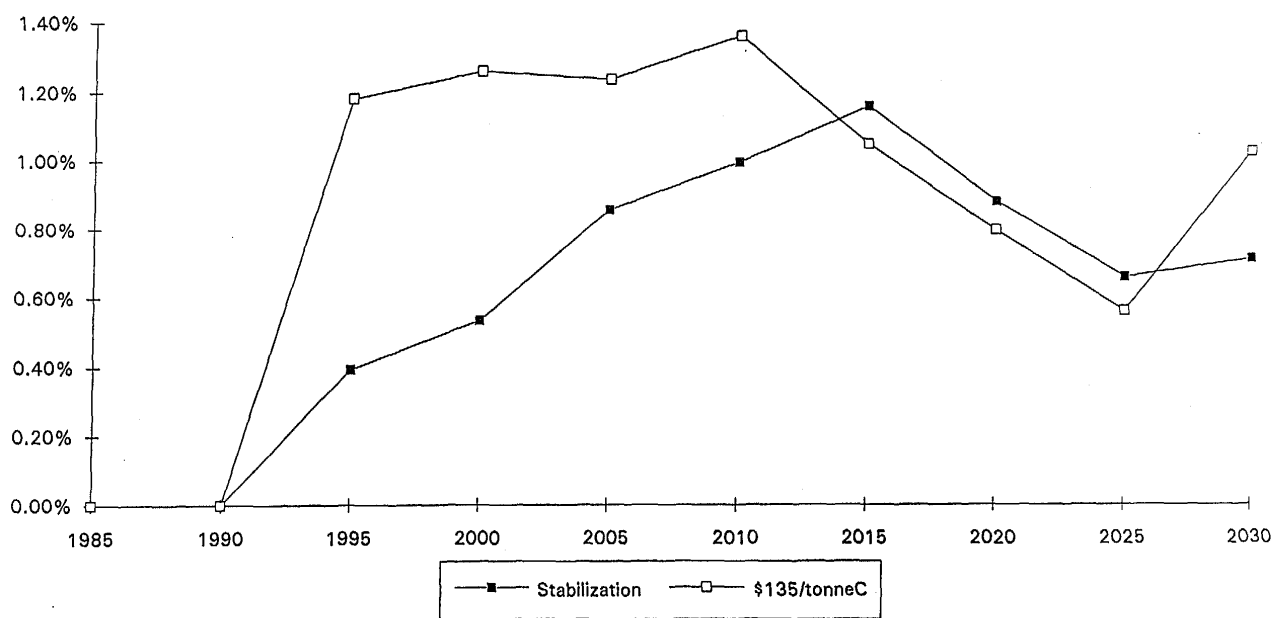


FIGURE 12: % Reduction in Real Consumption from Reference Case

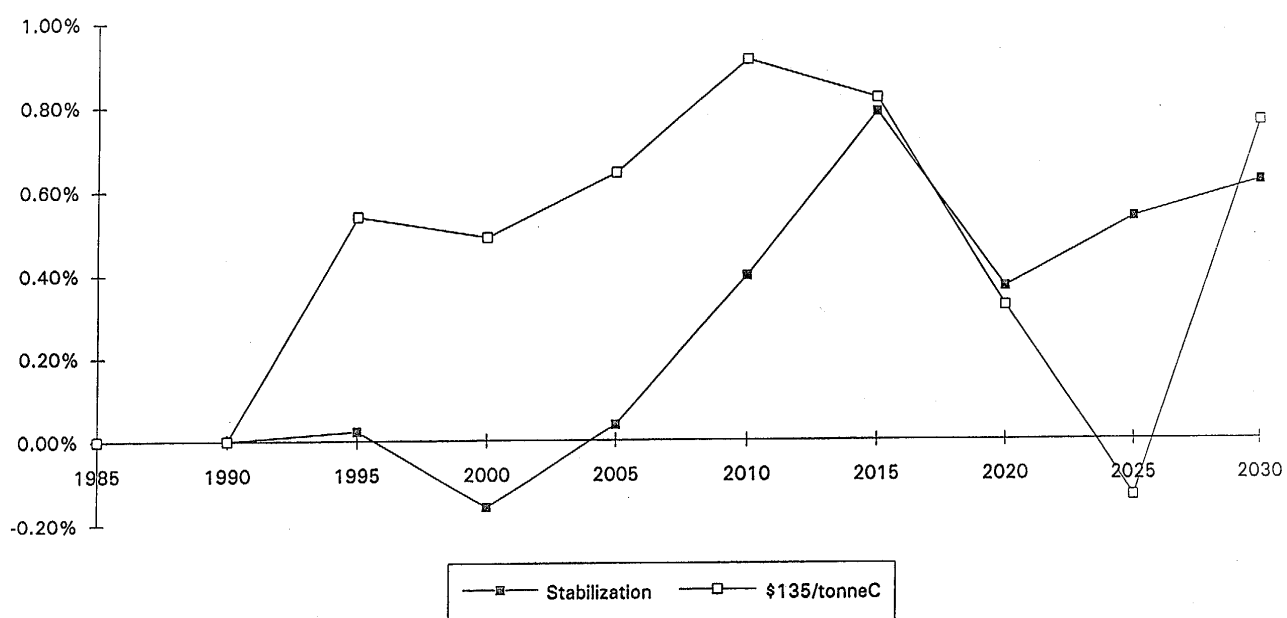


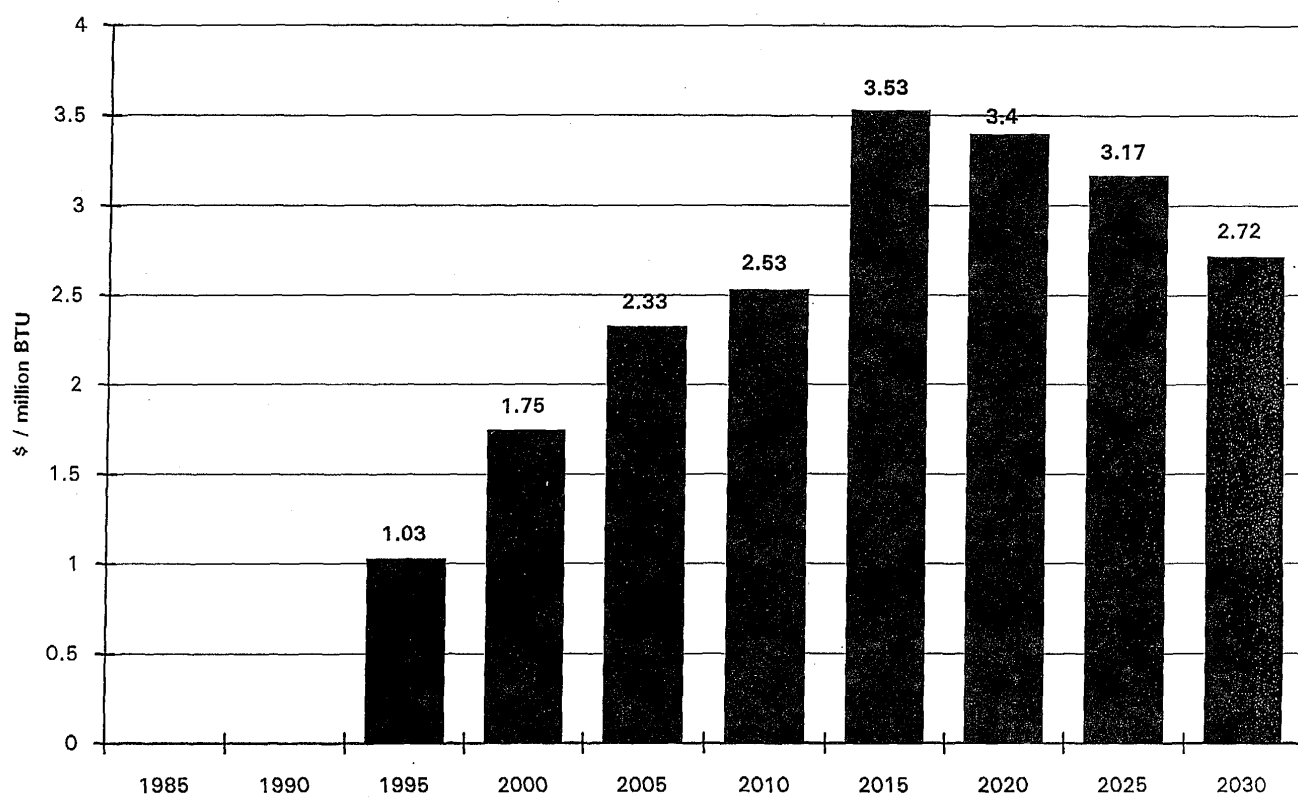
FIGURE 13: Energy Tax for CO₂ Stabilization to 1990 Level

FIGURE 14: Tax Revenues

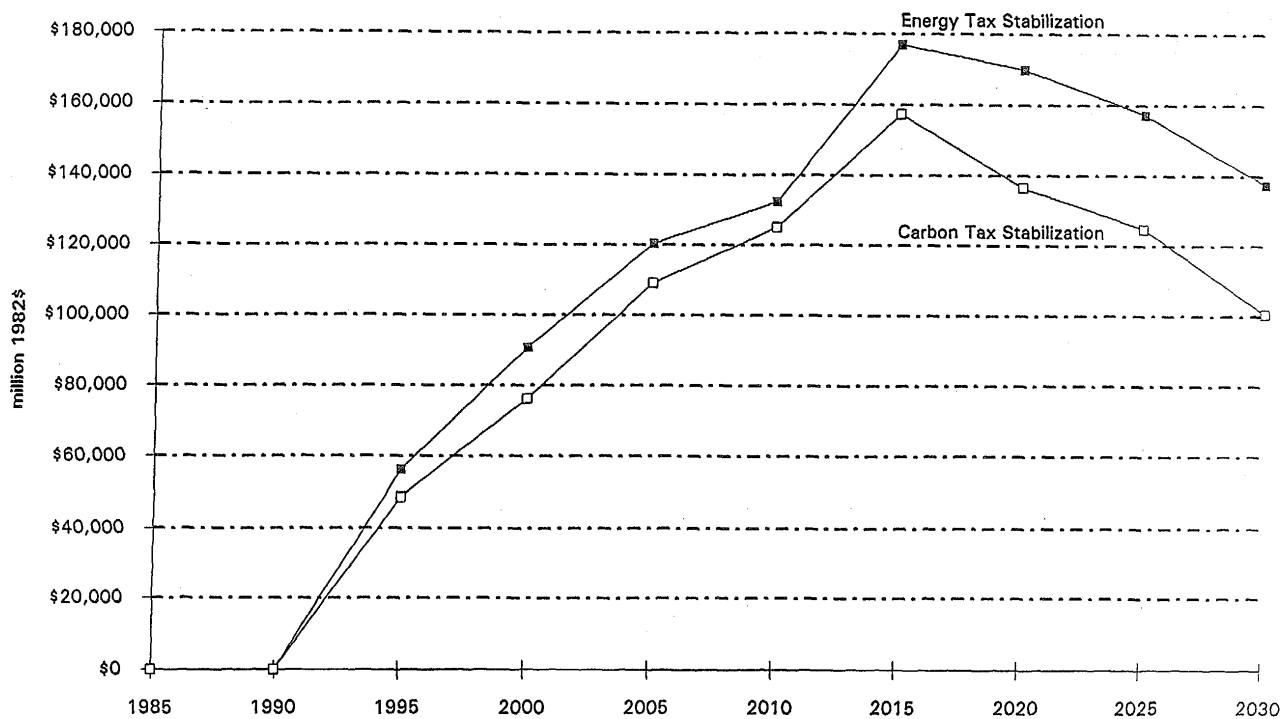


FIGURE 15: Taxes for CO₂ Stabilization to 1990 Level (Energy tax converted to carbon tax based on carbon content of coal.)

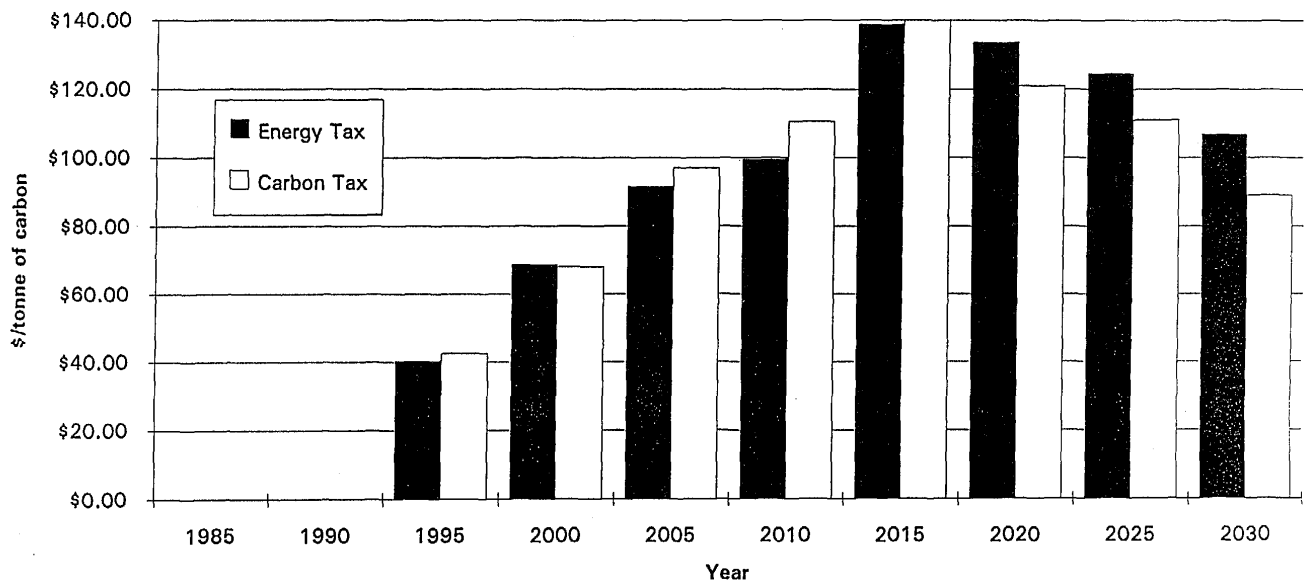


FIGURE 16: % Reduction in GNP from Reference Case (Base Year Prices)

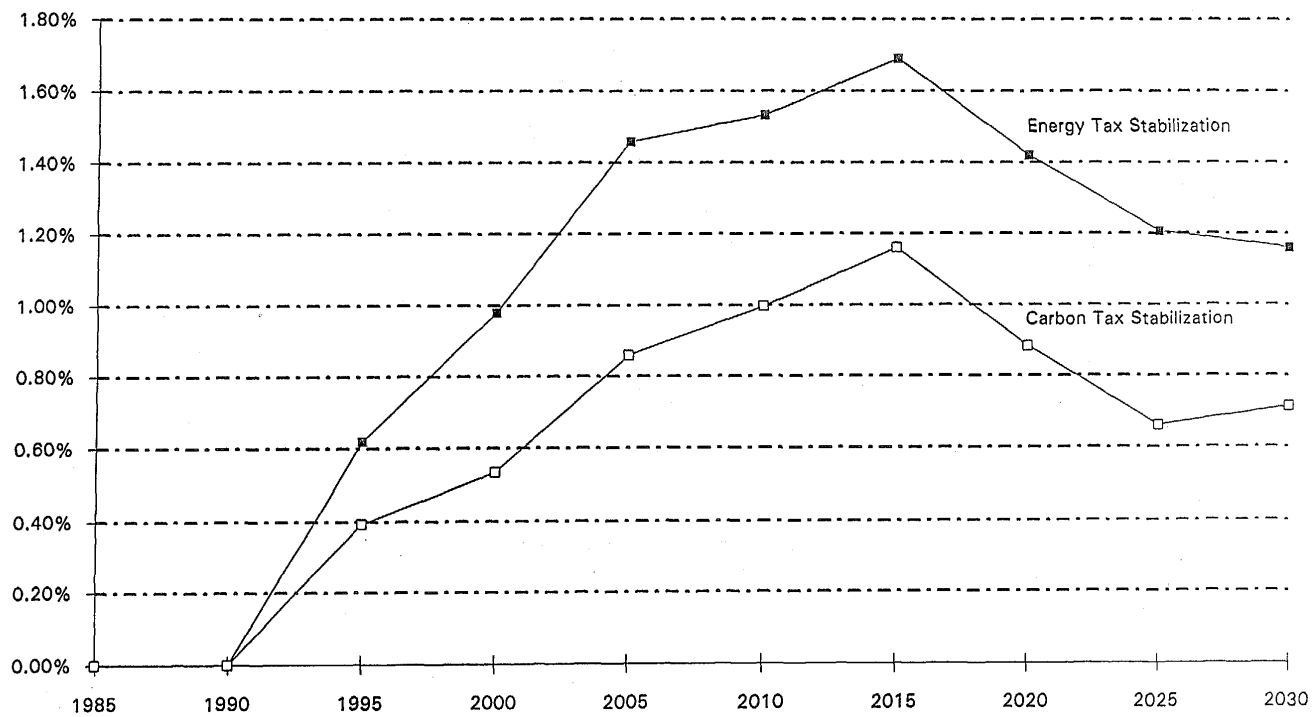
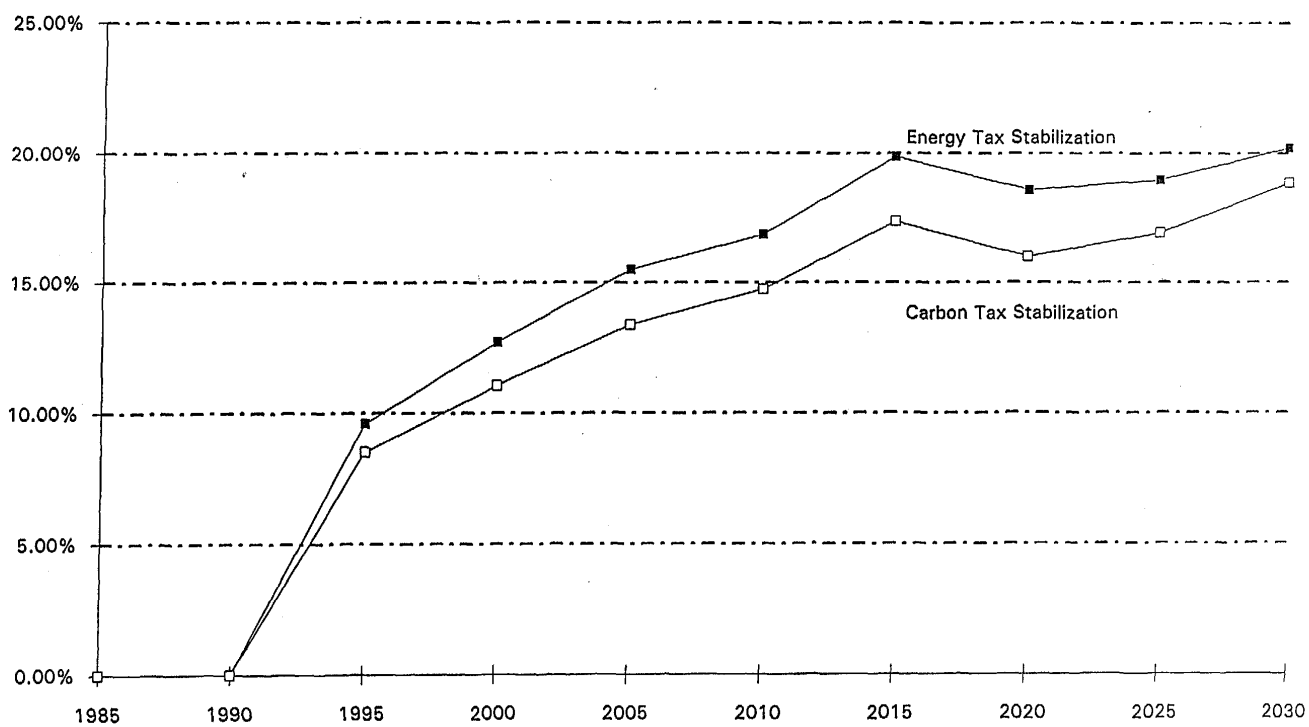


FIGURE 17: % Reduction in Primary Energy Consumption from Reference Case



References

- Edmonds, J., and Barns, D.W., 1990. *Estimating the Marginal Cost of Reducing Global Fossil Fuel CO₂ Emissions*. PNL-SA-18361. Pacific Northwest Laboratory. Washington D.C.
- Edmonds, J., and Barns, D.W., 1991. *Factors Affecting the Long-term Cost of Global Fossil Fuel CO₂ Emission Reductions*. PNL Global Studies Program Working Paper. Washington D.C.
- Edmonds, J., and Reilly, J.M., 1985. *Global Energy: Assessing the Future*. Oxford University Press. New York.
- Edmonds, J., Reilly, J.M., Gardner, R.H., and Brenkert, A., 1986. *Uncertainty in Future Global Energy Use and Fossil Fuel CO₂ Emissions 1975 to 2075*. TR036, DO3/NBB-0081Dist. Category UC-11. National Technical Information Service. U.S. Department of Commerce. Springfield, Virginia 22161.
- Edmonds, J., H.M. Pitcher, D. Barns, R. Baron, and M.A. Wise. 1993. *Modeling Future Greenhouse Gas Emissions: The Second Generation Model Description*. Proceedings of the United Nations University Conference on "Global Change and Modeling." Tokyo, JAPAN. October, 1991.
- Elliott, Douglas C. et al. 1990. "A Technical and Economic Analysis of Direct Biomass Liquefaction," in *Energy from Biomass and Wastes XIII*. Donald L. Klass ed. Chicago: Institute of Gas Technology.
- Fisher-Vanden, K. et al. 1993. *Calibration of the Second Generation Model*. Pacific Northwest Laboratory. Washington, D.C.
- IPCC (Intergovernmental Panel on Climate Change). 1990.
- Jorgenson, Dale W. 1990. "Productivity and Economic Growth" in *Fifty Years of Economic Measurement*, edited by Ernst R. Berndt and Jack E. Triplett. National Bureau of Economic Research, published by the University of Chicago Press.
- OECD (Organisation for Economic Co-operation and Development). 1992. International Energy Agency. *Electricity Supply in the OECD*. Paris, France.
- U.S. Department of Agriculture, 1992. *Conversations with E. Ball*.
- U.S. Department of Commerce, Bureau of Economic Analysis, 1987. *Fixed Reproducible Tangible Wealth in the United States, 1925-85*. Washington D.C., Government Printing Office.
- U.S. Department of Commerce, Bureau of Economic Analysis, 1990. "Annual Input-Output Accounts for the U.S. Economy, 1985" in *Survey of Current Business*. Washington D.C.
- U.S. Department of Commerce, Bureau of Economic Analysis, 1991. *Benchmark Input-Output Accounts for the U.S. Economy, 1982*. Washington D.C., Government Printing Office.
- U.S. Department of Commerce, Bureau of Census, 1992. *Statistical Abstract of the United States*. Washington, D.C., Government Printing Office.
- U.S. Department of Energy (DOE), 1991a. *Limiting Net Greenhouse Gas Emissions*. Report to the Congress. Bradley, R.A., E.C. Watts, and E.R. Williams (eds.). Volumes 1 & 2. DOE/PE-0101. Washington, D.C.
- U.S. Department of Energy (DOE), 1991b. *Integrated Analysis Supporting the National Energy Strategy: Methodology, Assumptions and Results, Technical Annex 2*. DOE/S-0086P. National Technical Information Service. Springfield, Virginia.

U.S. Department of Energy, Energy Information Administration (EIA), June 1991. *Financial Statistics of Selected Investor-Owned Electric Utilities, 1989*. DOE/EIA-0437(89). Washington, D.C.

U.S. Department of Energy, Energy Information Administration (EIA), June 1992. *Annual Energy Review 1991*. DOE/EIA-0384(91). Washington, D.C.

Wan, Edward I. and Malcolm D. Fraser, 1990. "Economic Assessment of Advanced Biomass Gasification Systems," in *Energy from Biomass and Wastes XIII*. Donald L. Klass ed. Chicago: Institute of Gas Technology.

World Bank, 1990. *World Tables*. 1989-1990 Edition. Johns Hopkins University Press. Baltimore, New York.

Acknowledgements

The authors would like to express appreciation to the United States Department of Energy, Office of Energy Research and the Office of Policy, Planning, and Analysis, and to the United States Environmental Protection Agency, for support in the conduct of this research.

The Millennium Project and Global Issues of Population and Environment

Jerome C. Glenn, United Nations University

Abstract

This paper describes the United Nations University's Millennium Project, its feasibility study, and the results from a 2-round international Delphi study on population and environment.

Introduction

The pace, complexity, and globalization of change requires attention to the future. The change of millennia - the coming of the year 2001 - provides an opportunity, a psychological focus, for a global review of past achievements and problems and a unique chance to assess and reflect on future issues and opportunities.

Unfortunately, there is as yet no mechanism or information utility to support a worldwide effort to study past achievements and problems or future issues and opportunities. Although there are many individual, isolated, special purpose, and one-time study efforts underway, there is no international system that can provide coherence or continuity to these studies, including feedback and sharing of information, and, in particular, the systematic exploration of future possibilities and policy alternatives. With growing interest in the future, the spread of instantaneous and global communications, the advent of powerful new non-deterministic modeling techniques, the ability to evoke, capture and share information and perceptions with systematic questioning techniques and software, the proliferation of data bases, and knowledge visualization, it is now possible for futurists, scholars and others around the world to interact globally and take a fresh look at the future possibilities and policies in ways not previously possible.

As the World Bank provides an ongoing system for research and feedback to improve economic policy, so too the United Nations University could provide an on-going system for the improvement of futures research and its application to the policy process. According to a series of interviews, questionnaires, and meetings with leading futurists and scholars around the world, the proposed "Millennium Project" has the potential to become such a system.

The Concept of the Millennium Project

The Millennium Project was originally conceived as a major landmark in studies dealing with the future. Although its time horizon was set as the next century, the next two decades were to be given analytic emphasis and the next hundred years would be given more normative emphasis. The project was designed to synthesize judgments on a global basis about future developments and issues in a way that would enrich and illuminate policy discussions.

Subsequent discussions have served to sharpen the central goal of the full-scale Millennium Project. In essence it is to:

Organize a program of futures research and a facility for global research that will continuously update and improve humanity's thinking about the future, and to make that thinking available through feedback systems in a variety of media for consideration in public policy, advanced training, and public education.

The pre-feasibility study identified the following ten objectives for the Millennium Project:

1. Link futurists, scholars, creative thinkers, and institutions around the world
2. Create an international information system of forecasts, key questions, issues, lessons from history, and potential futures research agendas, that allows for feedback to improve the state of research and thought in these areas

3. Evaluate futures research methodology and potential for setting standards
4. Forecast important technological, social, and scientific achievements and their likely effects on the future human condition
5. Integrate forecasts to describe and evaluate, to the extent possible, the likely future conditions of people in the world over the next century
6. Explore possible strategies in coping with these issues - the strategies that relate to actions that might be taken by governments, other institutions, or individuals
7. Increase awareness of the future and its promise, while inculcating the view that the future could be shaped in the interests of society through thoughtful policies
8. Develop a set of educational materials useful at all levels, dealing with the history of achievement of the last century and social issues of the next century, critical choices, and policy options that relate to these issues, and other future-related topics
9. Conduct advanced training
10. Provide a legacy to the people of the next century: what was thought about the future and how we, at the turn of the century had hoped it would evolve.

The Millennium Project Feasibility Study has three phases; this report describes Phase I.

PHASE I, from which this paper is drawn, began in November 1992 with support from the U.S. EPA to identify and link futurists and scholars around the world. Two Delphi inquiries were involved: a two round Delphi on the process of organizing the Millennium Project and a two round Delphi on population & environment (which composes most of this paper). The full 151 page report of this phase is available on Internet via Gopher futures.wic.epa.gov or on diskette from the Millennium Project, 4421 Garrison Street, N.W., Washington, DC 20016 (202)-686-5179.

PHASE II began in August 1993 with support from the UNDP/African Futures to produce a set of booklets and computer diskettes on futures methods and issues via an environmental scanning system. Sixteen methods will be considered for review in Phase II and for presentation in individual booklets; these methods include:

1. Introduction & Overview
2. Participatory Methods
3. Delphi
4. Systems and Modeling
5. Simulation Modelling
6. Decision Modelling
7. Scenario Construction
8. Trend Impact Analysis
9. Technology Sequence Analysis
10. Cross Impact Analysis
11. Futures Wheels
12. Structural Analysis & Actors' Strategies
13. Normative Forecasting
14. Relevance Trees and Morphological Analysis
15. Simulation Gaming as Applied to Forecasting
16. Visioning, Genius Forecasting, Intuitive Approaches

17. Methods Frontiers, Integration, and Strategic Planning

Each methods booklet would have the following sections or chapters: 1) History of the Method; 2) What the technique IS; 3) How to DO IT; 4) Strengths and weaknesses of the method; and 5) Alternative uses of the method AND how to use it in combinations with other methods. An Appendix will also include several actual examples of the method and an annotated bibliography

Phase II will also include identification and analysis of futures issues of importance to Africa. These include:

1. Technological Capacity
2. International Economic Policy and International Trade
3. Agriculture and Food Security Trends
4. Global Life Support Systems and Sustainable Development
5. Population, Education and Human Welfare (special attention to AIDS)
6. Peace, Governance, and Culture

Each issue scan and analysis includes: 1) Historical forces; 2) How these forces are changing today and in the foreseeable future; 3) Alternative scenarios/impacts out to 2025 in Africa - futures chapter; 4) How to measure change on this issue; 5) How this scan was done and how to keep it going; 6) Template for data base ; 7) Key sources, individuals, institutions, computer network groups, etc. plus an appendix of relevant forecasts for The Futures Group scan in hard copy and on diskette

PHASE III will consist of a consultative meeting to focus on recommendations for the feasibility study report. The location and design of such a meeting will depend on the responses to this report and the results of Phase II. Use of advanced telecommunications systems could augment the proposed face-to-face consultative meeting. The primary product of this phase will be the final report on the feasibility of the Millennium Project.

As described in the full Phase I report, two 2-round Delphi inquiries were conducted to gather judgments from the international panel consisting of futurists, scholars, and interested participants. The first of these two-round Delphis was concerned with the process of organizing the full-scale Millennium Project. The second two-round Delphi focused on the leading issues identified by the international panel in the first Delphi: population and environmental change. The conclusions of this second Delphi study is below.

Results of Population and Environment Delphi Rounds I and II

In the first round, the panelists were asked to rate some forces that led to the reduction of the world population growth rate from 2.06% in the late 1960s to 1.7% currently and to assess how these forces might change over the next 25 years. They were invited to add forces, which were rated in the second round by the same Scale A below. Table 1 orders these forces by their historic influence; table 2 orders them by their future influence.

Scale A

<i>Historic Influence</i>	<i>Future Influence</i>
1 = Very Important	1 = Greatly increasing in importance
2 = Important	2 = Increasing in importance
3 = Marginally important	3 = Remaining the same in importance
4 = Unimportant	4 = Decreasing in importance
5 = Counter impact	5 = No longer a factor, or mixed

Table 1 - The importance of some HISTORIC factors on global population growth

- 2.0 Availability of inexpensive, simple effective contraceptives;
- 2.0 Family planning and public health programs
- 2.0 China's population policy
- 2.0 Legitimization of contraception

- 2.0 Decreases in infant and child mortality
- 2.1 Government policies in developing countries that encourage smaller families
- 2.1 Increasing number of years that women attend school
- 2.1 Rising incomes and the spread of middle class values
- 2.2 Move away from agricultural society/primary sector
- 2.3 Spread of new communications media (Television, etc)
- 2.3 War, famine, disease, and pestilence
- 2.4 Improved literacy by improved children's schooling
- 2.4 Decreasing Catholic Church's social influence
- 2.5 Rise of "woman's power."
- 2.9 Availability of male contraceptives
- 2.9 Environmental deterioration
- 3.1 Education about relation of environment and population
- 3.1 Increasing futuristic orientation
- 3.2 Demonstration by the North that fewer children can mean more wealth

Table 2 - The importance of some forces affecting population over the NEXT 25 years

- 1.8 Spread of new communications media (Television, etc)
- 2.0 Increasing number of years that women attend school
- 2.0 Environmental deterioration
- 2.1 Government policies in developing countries that encourage smaller families
- 2.1 Family planning and public health programs
- 2.1 Improved literacy by improved children's schooling
- 2.2 Availability of inexpensive, simple effective contraceptives
- 2.2 Legitimization of contraception
- 2.2 Education about relation of environment and population
- 2.2 Decreases in infant and child mortality
- 2.3 Rising incomes and the spread of middle class values
- 2.3 Availability of male contraceptives
- 2.4 Rise of "woman's power"
- 2.4 Increasing futuristic orientation
- 2.6 Move away from agricultural society/primary sector
- 2.6 Decreasing Catholic Church's social influence
- 2.7 War, famine, disease, and pestilence
- 2.8 China's population policy
- 2.8 Demonstration by the North that fewer children can mean more wealth

In the first round, the panelists were asked to rate some forces the could explain why populations rates have remained relatively high in the developing countries and to assess how these forces might change over the next 25 years. They were invited to add forces, which were rated in the second round by the same Scale A above. Table 3 orders these forces by their historic influence; table 4 orders them by their future influence.

Table 3 - Some reasons for high population growth in developing countries

- 1.6 Need of children for social security, to support parents in their old age.
- 1.7 Discrimination against women (little autonomy, education, and lack of social power)
- 1.8 Poverty
- 1.8 Low levels of literacy and lack of understanding ecological view of planet earth
- 2.0 Family-based, labor-intensive economies needing children's input
- 2.0 Low levels of literacy
- 2.1 Close adherence to religious tenets that lead to avoidance of contraceptives
- 2.1 Rural areas receive less attention from population programs than urban areas

- 2.2 Lack of information and access to contraceptives
- 2.2 High infant mortality
- 2.3 Masculinity associated with increasing numbers of children
- 2.4 Belief by governments that larger populations mean greater political strength
- 2.8 Government policies supporting large families
- 3.1 Liberal immigration policies in richer countries

Table 4 - How important will these forces be over the NEXT 25 years?

- 2.2 Low levels of literacy and lack of understanding ecological view of planet earth
- 2.4 Poverty
- 2.6 Rural areas receive less attention from population programs than urban areas
- 2.7 Low levels of literacy
- 2.8 Discrimination against women (little autonomy, education, and lack of social power)
- 2.8 Lack of information and access to contraceptives
- 3.0 Need of children for social security, to support parents in their old age.
- 3.1 Family-based, labor-intensive economies needing children's input
- 3.1 High infant mortality
- 3.2 Belief by governments that larger populations mean greater political strength
- 3.4 Close adherence to religious tenets that lead to avoidance of contraceptives
- 3.4 Government policies supporting large families
- 3.4 Liberal immigration policies in richer countries
- 3.5 Masculinity associated with increasing numbers of children

The panelists were asked to assess new forces and unprecedented events that might influence population growth in the future and suggest and assess additional such forces as to their likelihood of occurrence and impacts over the next 25 years using Scale B. Table 5 shows the average of the panel's responses about their likelihood; Table 6 shows their impacts over the next 25 years.

Scale B

<i>Likelihood of occurrence</i>	<i>Eventual impact on population within the next 25 years of growth</i>
1 = almost certain	1 = reduces growth rate by 30% or more
2 = likely	2 = reduces growth rate by 5-30%.
3 = even or 50/50 chance	3 = no impact.
4 = unlikely	4 = increases growth rate by 5-30%.
5 = almost impossible	5 = increases growth rate by 30% or more

Table 5 Likelihood of occurrence of new forces or unprecedented events that might influence future population growth

- 1.8 Simple test for identifying the sex of unborn children
- 1.9 Increasing survival in middle age and early old age due to curing or improved therapy for heart disease, cancer and stroke
- 1.9 Public health programs decrease mortality of infant and young adults by 5-10%
- 2.1 Simple, safe, effective male birth control pill
- 2.3 Changes in death rates due to spread of AIDS in developing countries
- 2.4 Massive starvation of the scale of Somalia at least once every three years
- 2.5 Long term (at least one year) contraceptives widely used and accepted as birth control pills are today
- 2.5 Important negative changes in the environment (e.g., accumulation of toxic wastes, failures of mono-agriculture crops, contamination of drinking water) resulting in increased mortality
- 2.6 New deadly viruses, including AIDS mutations

- 2.6 Wide spread use of relatively cheap and easy ways to affect the aging process, resulting in diminished mortality and extension of the life span by about 5 years
- 2.6 3% of births via new methods of impregnation and prenatal development ("Outside womb" fertility, artificial inseminating, surrogate motherhood, other such techniques)
- 2.8 Simple method for selecting sex at conception
- 2.9 Rising incomes in most developing countries
- 2.9 Increasing sterility, worldwide, by 10% due to environmental degradation, higher level of stress, and other factors
- 3.3 Change in the Vatican's position: use of contraceptives becomes available without limit
- 3.4 Two children per family becomes the social norm in the majority of the developing world
- 3.4 Increasing impotency by 10% due to environmental degradation, higher level of stress, and other factors
- 3.4 Doubling, worldwide, of today's level of male and female homosexuality
- 3.5 20% increase of people active in religions that encourage higher fertility
- 3.6 Successful new prototype habitats in oceans, cold regions, or in earth orbit gives stimulates popular frontier spirit and alternatives to previous urbanization patterns
- 3.9 Basic economic needs met for 90% of global population (minimum acceptable health care, food, and shelter)

Table 6 Eventual impact of these new forces and unprecedented events on population growth over the next 25 years.

- 2.0 Long term (at least one year) contraceptives widely used and accepted as birth control pills are today
- 2.1 Rising incomes in most developing countries
- 2.1 Two children per family becomes the social norm in the majority of the developing world
- 2.3 Simple, safe, effective male birth control pill
- 2.3 Change in the Vatican's position: use of contraceptives becomes available without limit
- 2.3 Increasing sterility, worldwide, by 10% due to environmental degradation, higher level of stress, and other factors
- 2.4 Changes in death rates due to spread of AIDS in developing countries
- 2.4 New deadly viruses, including AIDS mutations
- 2.5 Important negative changes in the environment (e.g., accumulation of toxic wastes, failures of mono-agriculture crops, contamination of drinking water) resulting in increased mortality
- 2.6 Massive starvation of the scale of Somalia at least once every three years
- 2.6 Simple method for selecting sex at conception
- 2.7 Simple test for identifying the sex of unborn children
- 2.7 Increasing impotency by 10% due to environmental degradation, higher level of stress, and other factors
- 2.7 Doubling, worldwide, of today's level of male and female homosexuality
- 2.9 3% of births via new methods of impregnation and prenatal development ("Outside womb" fertility, artificial inseminating, surrogate motherhood, other such techniques)
- 2.9 Basic economic needs met for 90% of global population (minimum acceptable health care, food, and shelter)
- 3.0 Successful new prototype habitats in oceans, cold regions, or in earth orbit gives stimulates popular frontier spirit and alternatives to previous urbanization patterns
- 3.2 Public health programs decrease mortality of infant and young adults by 5-10%
- 3.4 Increasing survival in middle age and early old age due to curing or improved therapy for heart disease, cancer and stroke
- 3.4 Wide spread use of relatively cheap and easy ways to affect the aging process, resulting in diminished mortality and extension of the life span by about 5 years
- 3.5 20% increase of people active in religions that encourage higher fertility

With the information provided by the panel in round I, we identified two possible future events that were judged to be unlikely, but nevertheless potentially effective in reducing population growth rates, if they were to occur. Both have been considered before and one is the focus of large international programs. Without repeating what has been said and tried many times, we asked for suggestions about novel policy approaches that

might be practical, and if implemented, improve the probability of these developments. A distillation of those suggestions follow:

1.4.1 Novel policies that could lead to a social norm of two children per family throughout the majority of the developing world:

- Contraceptives added to the water supply, to conceive potential parents obtain counter active agent.
- Family-size tax that increased substantially for every child over two.
- Prohibition of child labor.
- Universal and compulsory schooling, for all to age 15-16, with mandatory immunizations from birth onward, and free available family planning support. Clean water is also a necessity.
- Tie all government subsidies to number of children per family in an inverse relationship.
- Why not encourage acceptability of no children for some couples in all countries? Why must we all reproduce? New paradigm needed
- Policies should focus on women's access to work and education.
- With enough effort and resources, the present policies (national, bilateral, and multi-lateral) that try to make family planning services universally available and to promote information, education and communication about family planning, hence the two child norm should be virtually universal within 25 years.
- Study what would motivate those who have more than two to limit to two. Different programs for different groups, China's program seems to work; Indians would need to design one for India.
- Free circulation of contraceptives in public health programs supported by churches and other religious organizations.
- International information utilities in education, health, and training in information society.
- Adopt Chinese incentives
- Bi- and multi-lateral aid requires it.
- Without significant reductions in poverty and infant mortality, and increase in women's education and empowerment, these policies would need to be coercive in nature or at least provide strong economic Disincentives to having many children.

1.4.2 Novel policies that could lead to the Vatican's acceptance of contraception without limit:

- Remove Holy See from the United Nations on grounds that it is not really a country and give it the same status as the World Council of Churches.
- Promotion of contraceptive that is accompanied with strong dissemination of moral values.
- Separate personal religious beliefs from personal choices as to the use of contraceptives.
- Theological doctrine developed by U.S. Catholic Bishops in support of sustainable development (of Government Statements at UNCED 1992).
- New international convention of religious leaders - Vatican III - with supporting papal encyclical letter - 1995.
- Elect a non-European Pope.
- Schism
- It is not practical.
- Create philosophical shift among powerful within the Roman Catholic Church.
- Policies that focus on responsabilization of woman to make choices.
- Only decline of organized religion is likely to affect the situation.
- Allow priests and nuns to marry and pay for the raising of children.
- Reconsideration of the theology of St. Thomas Aquinas - especially the discarding of the Thomastic view of "natural law" as it applied to human sexuality. Christ said nothing whatsoever about human sexuality. Separate the notion of procreation as the only "natural end" from that of enjoyment.
- Try to influence Vatican that some kinds of contraceptions could possibly be acceptable-not "chemical" and for women contraception but "mechanical" and for men (condoms) contraception.

1.4.3. Other policy areas of your selection and how you would address it:

- Research for long-term implantable ovulation suppression device.
- Global televised debate on population policy, environmental protection, and social ethics.
- Sex education in schools
- All nations will have to eventually adopt policies that clearly state the freedom of individual choice.
- Policies that emphasizes people to take charge of their lives and reduce dependence on governments.

Environment Section

Panelists were presented several reasons why population growth may led to environmental degradation and were asked to suggest additional reasons. They assessed each using Scale C below. Table 8 lists the average of the panel's rating.

Scale C

- 1 = of paramount importance
- 2 = of great importance
- 3 = of modest importance
- 4 = of little importance
- 5 = no effect

Table 8 Some reasons why population may lead to environmental degradation

- 1.7 The drive to improve "standard of living" through consumption
- 1.8 Sheer size
- 2.0 Lack of understanding about the environment
- 2.0 The generation of waste
- 2.0 Diminished carrying capacity of the environment
- 2.0 Increased energy consumption per capita
- 2.2 Increased material use per capita leading to increased global demand for local resources
- 2.9 Increased social stress, breakdown of community values

Panelists were asked to rate the forces that led to changes in environmental quality over the previous five decades and to asses how these forces might change over the next 25 years. There were also asked to suggest and rate new such forces. Table 9 lists the averages of the panel's judgments on their historic influence over the past five decades; Table 10 lists the averages of the panel's judgement on their future influence using Scale D below.

Scale D

Historic Influence

- 1 = Very Important
- 2 = Important
- 3 = Marginally important
- 4 = Unimportant
- 5 = Counter impact

Future Influence

- 1 = Greatly increasing in importance
- 2 = Increasing in importance
- 3 = Remaining the same in importance
- 4 = Decreasing in importance
- 5 = No longer a factor, or mixed

Table 9 - Some historic influences on Environmental Quality in order of importance

- 1.4 World population growth
- 1.7 Lack of availability or use of clean energy generation systems compared to coal, oil, wood, or dung burning
- 1.7 Economic systems that treat the environment as a free good
- 1.8 Popularization of the automobile

- 1.8 Excessive use of industrial processes that have by-products of toxic wastes, without appropriate attention to the disposal of such by-products
- 1.9 Lack of political leadership, public support or will to address environmental problems
- 1.9 National accounting systems that do not reflect natural resources and environmental damage
- 1.9 Pricing of energy in a way that does not account for its full environmental costs
- 1.9 Lack of adequate waste management in most places in the world
- 1.9 Aggressive forestry
- 2.0 Short term, reductionist, anti-generalist, and reactive policy making with a lack of concern for future generations in most developed countries
- 2.0 Lack of economic incentives for corporation and individuals to be environmentally responsible.
- 2.0 Rise of a "throw-away society" in developing nations
- 2.0 Excessive consumption and contamination of water aquifers
- 2.1 Lack of understanding of environmental interrelationships by average person
- 2.1 Use of chemicals in agriculture
- 2.1 Lack of integrated action among land use planning, family planning, environmental protection, and sustainable development in most places of the world
- 2.1 Excessive farming on marginal lands
- 2.2 The desire by most people to be "modern," and by most nations to "industrialize," and value systems to be anthropocentric
- 2.2 Pricing of transportation, including public subsidies for transportation in a way that does not account for it full environmental costs
- 2.2 Military activity (also in the sense of diverting expenditures from other activities)
- 2.2 Mining and the search for and utilization of raw materials
- 2.3 Urbanization and suburban sprawl
- 2.4 Lack of funds for adequate environmental research, development and related environmental education and enforcement programs
- 2.5 Lack of effective communication among experts and with the public about environmental issues
- 2.6 Lack of academic attention, particularly from economists and social scientists
- 2.6 Fishing practices
- 2.9 Aspects of globalism and or nationalism that lead to loss of local culture and sense of individual responsibility
- 3.1 Lack of ecological ethic in religious values
- 3.1 Natural phenomena (e.g. volcano eruptions, etc.)
- 3.7 Failure to expand human community or habitation into oceans and space

Table 10 - Future influence of these forces over the NEXT 25 years in order of importance

- 2.0 World population growth
- 2.2 Lack of adequate waste management in most places in the world
- 2.2 Excessive consumption and contamination of water aquifers
- 2.3 Urbanization and suburban sprawl
- 2.3 Excessive farming on marginal lands
- 2.4 Lack of integrated action among land use planning, family planning, environmental protection, and sustainable development in most places of the world
- 2.4 National accounting systems that do not reflect natural resources and environmental damage
- 2.4 Lack of funds for adequate environmental research, development and related environmental education and enforcement programs
- 2.5 Popularization of the automobile
- 2.5 The desire by most people to be "modern," and by most nations to "industrialize," and value systems to be anthropocentric
- 2.5 Lack of availability or use of clean energy generation systems compared to coal, oil, wood, or dung burning
- 2.5 Lack of understanding of environmental interrelationships by average person
- 2.6 Lack of economic incentives for corporation and individuals to be environmentally responsible

- 2.6 Pricing of energy in a way that does not account for its full environmental costs
- 2.6 Excessive use of industrial processes that have by-products of toxic wastes, without appropriate attention to the disposal of such by-products
- 2.6 Aspects of globalism and or nationalism that lead to loss of local culture and sense of individual responsibility
- 2.6 Lack of political leadership, public support, or will to address environmental problems
- 2.6 Short term, reductionist, anti-generalist, and reactive policy making with a lack of concern for future generations in most developed countries
- 2.7 Aggressive forestry
- 2.7 Pricing of transportation, including public subsidies for transportation in a way that does not account for its full environmental costs
- 2.8 Economic systems that treat the environment as a free good.
- 2.8 Use of chemicals in agriculture
- 2.8 Lack of academic attention, particularly from economists and social scientists
- 2.8 Natural phenomena (e.g. volcano eruptions, etc.)
- 2.8 Lack of effective communication among experts and with the public about environmental issues
- 2.9 Rise of a "throw-away society" in developing nations
- 2.9 Mining and the search for and utilization of raw materials
- 3.0 Fishing practices
- 3.0 Lack of ecological ethic in religious values
- 3.1 Failure to expand human community or habitation into oceans and space
- 3.3 Military activity (also in the sense of diverting expenditures from other activities)

Panelists were asked to assess new forces and unprecedented events that might influence environmental quality in the future and to suggest additional such forces and events. They were asked for their judgments about the likelihood of occurrence and impacts of these forces and events over the next 25 years using Scale E below. Table 11 lists the averages of the panel's responses about the likelihood of occurrence of these events within the next 25 years; Table 12 lists these in terms of the future impact (only the responses on future impact are given from the first round, because the second round had a typographical error in Scale E asking impact on population instead of on environment).

Scale E

<i>Likelihood of occurrence within the next 25 years.</i>	<i>Eventual impact on environmental quality</i>
a = almost certain	1 = very positive impact.
b = likely	2 = positive impact.
c = even or 50/50 chance	3 = no impact.
d = unlikely	4 = negative impact.
e = almost impossible	5 = very negative impact.

Table 11 Likelihood of some new forces and unprecedented events that might influence environmental quality

- 1.9 Wide use of new clean energy generating technologies, producing 10% of total output
- 2.0 Economic competition for certain raw materials, jobs, and markets becomes much more intense; protectionism through tariffs and non-tariff barriers increases
- 2.1 Actual demonstration [irrefutable evidence] of greenhouse warming
- 2.2 Cessation in production/consumption of ozone depleting gases
- 2.3 World per capita GDP grows by a total of 25% from 1990 level
- 2.3 Doubling 1990 levels (as a percentage of GDP) of per capita financing for environmental work
- 2.3 10% reduction, through substitution or other means, in the use of the polluting sources of energy such as coal, wood, and dung in developing countries
- 2.4 Development and wide spread use of plant varieties that thrive in salt or brackish water

- 2.4 National CO2 emission goals established by most nations
- 2.4 Energy prices increase 100 % by additional taxation in most developed countries
- 2.5 Observation of a rise in sea level as predicted by Intergovernmental Panel on Climate Change
- 2.5 Transportation prices (public and private) increased 200% by additional taxation in most developed countries
- 2.6 GNP/GDP accounting revised in most countries to account for environmental costs
- 2.6 "Sustainable agriculture" practiced by 50% of farms in western countries
- 2.6 Global depression including 20% unemployment in OECD countries for more than one year
- 2.7 Environmental education essentially everywhere
- 2.7 "Congestion tolls" established on most major highways in developed countries
- 2.7 Creation and implementation in essentially every country of economic incentives for corporations and individuals to act in an environmentally responsible way, through mechanisms such as pollution "rights," including prices that reflect environmental costs, etc
- 2.7 Detonation of one or more nuclear weapons in war, terrorist incident, or by accident
- 2.8 Global environmental fund established with payment based on CO2 emissions (Al Gore proposal)
- 2.8 Environmental ethics and sustainability thinking dominates policy of most governments
- 2.9 Reduction of annual greenhouse gases emissions by 25%
- 2.9 Sales of electric cars account for 10% of all new cars sold world-wide
- 2.9 Wide spread planting of genetically engineered trees that mature in 3-5 years
- 2.9 50% of humanity views waste as "immoral"
- 2.9 Primary raw materials prices increased 200% by additional taxation in most developed countries
- 2.9 Pesticide use reduced 50% worldwide
- 3.0 U.N. enforcement against a government, corporation, or institution for violating environment treaty
- 3.0 Most LDC governments institute helpful environmental policies
- 3.0 75% of all waste is recycled in OECD countries
- 3.0 General rise of conservative governments deny most environmental deterioration, exaggerate extent of progress in protection, and roll back many environmental initiatives
- 3.1 25% decline in per capita meat consumption worldwide
- 3.1 The rise of political power of "green" parties in essentially all countries
- 3.1 50% reduction in waste generation (e.g. Canadian green plan goal)
- 3.2 Commercial demonstration of production of electricity through fusion power process
- 3.2 Revival of spiritual values for most people in the world which includes treating plants and animals as partners
- 3.3 Widespread war occurs, at least as extensive as World War II
- 3.4 Tropical rain forests stop shrinking
- 3.5 Major volcanic explosion leads to widespread global cooling and crop failure
- 3.9 Eating beef considered immoral in western countries

Table 12 Some new forces and unprecedented events rated by Scale E for their eventual impact on the Environment (average responses from first round only)

- 1.7 Doubling 1990 levels (as a percentage of GDP) of per capita financing for environmental work
- 1.7 Most LDC governments institute helpful environmental policies
- 1.7 75% of all waste is recycled in OECD countries
- 1.7 Environmental education essentially everywhere
- 1.8 National CO2 emission goals established by most nations
- 1.8 GNP/GDP accounting revised in most countries to account for environmental costs
- 1.8 50% reduction in waste generation (e.g. Canadian green plan goal)
- 1.9 50% of humanity views waste as "immoral"
- 1.9 Tropical rain forests stop shrinking
- 2.0 Reduction of annual greenhouse gases emissions by 25%
- 2.1 U.N. enforcement against a government, corporation, or institution for violating environment treaty
- 2.1 The rise of political power of "green" parties in essentially all countries

- 2.2 "Congestion tolls" established on most major highways in developed countries
- 2.2 Wide spread planting of genetically engineered trees that mature in 3-5 years
- 2.3 Sales of electric cars account for 10% of all new cars sold world-wide
- 2.3 Development and wide spread use of plant varieties that thrive in salt or brackish water
- 2.6 Actual demonstration [irrefutable evidence] of greenhouse warming
- 2.9 Eating beef considered immoral in western countries
- 3.1 World per capita GDP grows by a total of 25% from 1990 level

With the information provided by the panel in round I, some possible future events were judged to be unlikely, but nevertheless potentially effective in reducing environmental degradation. All of these have been considered before and some have been subject to large scale international programs. Without repeating what has been said and tried many times, we asked for novel suggestions about policy approaches that might be practical, and if implemented, improve the probability of these developments. Below is a distillation of the panel's responses:

2.4.1 Novel policies that could lead to the end of tropical rain forests shrinkage:

- Total ban on tropical hardwood imports.
- Manufacture pharmaceuticals from cultivated "tropical" plants.
- Ocean farming
- National accounts including environmental costs.
- Restriction on use of unsustainably produced tropical woods.
- Strict laws for replanting for each tree cut.
- Provide technical support, funding, and environmental facilities to countries with tropical rain forests.
- International supervision and national trade of CO2 discharge right.
- Global fund to substitute governments for lost income from reduced logging.
- Rain Forest Guard Program that pays those who destroy rain forests to protect them from others who destroy
- Educate to value these forests as natural treasure house of medicines
- Proper valuation of living forests. This can be achieved in part by incorporating natural resources into the national accounting system. A key factor will be the inclusion of non-market values.
- Policies that adopt new incentives not based on the economic calculus.
- Adopt a forest incentive connect local communities to rain forests to sister communities around the world.
- Free plants and trees to those who recycle
- Credit for carbon dioxide sequestering. In Canada and many other countries, there are programs to reduce the emissions of "greenhouse gases". Certain industries that emit large quantities of these gases, such as fossil-fuel-burning utilities, are facing new restrictions on production and expansion. -- They have argued that they should receive credit for initiatives they undertake in other parts of the world. For example, a utility in North American could purchase or secure an area of tropical forest that will sequester carbon dioxide to balance new emissions from the utility.
- Debt for Nature swaps. I believe the World Bank and other lending organizations have recognized the value of living forests by retiring portions of a country's debt in exchange for preservation of tropical rain forest.
- Increase the commercial benefits of the living forest.
- Tropical rain forests produce unique biological materials for pharmaceuticals, chemicals, foods and other commercial products. The benefits of these materials accrue to the company that develops them commercially and establishes patents. Some mechanism should direct benefits to the originating country.
- Agricultural policies that favor harvest of endemic foods (nuts, berries, fruits, etc) over clearing and cultivation of introduced crops and livestock. This could include policies to support research into rain forest foods as well as marketing and consumer education.
- Support for farm co-operatives This may encourage smaller-scale farming and forestry and direct the benefits to the local people.

2.4.2 Novel policies that could lead to 75% of all waste being recycled in OECD countries:

- Fines for not recycling; tax relief for those who do.
- Don't collect/dispose of unrecycled garbage.
- Taxes on raw materials and subsidy on recycled materials could achieve this, although it would make no "economic" sense.
- International supervision and fine of pollution trade.
- Improvements in design of household items utilizing recycled material
- Copy Swedish example of deposit/reimbursement for cans.
- Policies in which waste recycling efforts are shared with both producers and consumers.
- Tax companies and households for waste generated/collected.
- Introduce simple ways for companies and households to separate their waste for recycling.
- ...the best policy for future is to eliminate waste rather than to recycle the accumulated waste. The recent experience in Germany shows that its recycling program has produced 400,000 tons of waste plastic for recycling instead of the expected 100,000 tons. Germany can recycle 80,000 tons (40%) of plastic a year at present, and will have to ship the rest abroad - much of it to Eastern Europe and the Third World. News reports tell of contractors filling false recycling reports, then tipping their loads in dumps. Nowadays recycling costs in Germany are as much as around \$9 per pound. By 1997 the German government says it will be able to recycle 800,000 tons of plastic. Even this one example shows limited capacities of 'developed' countries to do recycling within their countries. Even they recycle 75% of their waste, because of its cumulative growth, if unchecked, the remaining 25% will severely affect the environment. Dumping waste, especially toxic, hazardous, nuclear, abroad following the not-in-my-backyard approach philosophy which is fueled by present day economics is a real future issue. I wonder how other respondents see it.
- Full-cost accounting. Policies that support full-cost accounting of waste disposal will generally increase the cost of disposal and encourage waste reduction and recycling.
- Ban the export of waste beyond national borders. This includes solid waste and sewage, but not materials destined for recycling.
- Tradable permits or quotas for waste disposal in international waters.
- Tradable permits are being implemented for companies emitting air contaminants. A similar system based on an international agreement could be applied to companies disposing of waste into oceans and international waterways. A company that recycles waste destined for ocean-dumping would be able to sell or trade its excess quota.
- more opportunities at community level to achieve higher levels of recycling

2.4.3 Novel policies that could lead to sustainable agriculture being practiced by 50% of farms in more economically advanced countries:

- Compulsory and clear labelling of all food products on whether they have been produced in a sustainable manner ("Eco-labelling") coupled with strong consumer education.
- Reduction in area of land set-aside under European Community common agricultural policy.
- Proportionate reduction of allowable agricultural imports.
- Heavy taxation on petro-chemical fertilizers and pesticides, even soil loss.
- Global sustainable agricultural awards for new breakthroughs
- Requirement that every farm should generate a 5-year conservation plan subject to regular review
- New scientific methods including bioresources would (in the long term) contribute to sustainable agricultural practices.
- subsidy for soil conservation, integrated pest management, polyculture, organic fertilizers.
- Popularization of the way of agricultural planning by macro-coordination science.
- Educate about value of healthier foods.
- Reduce subsidies and let them compete.
- Consumer education favoring produce from sustainable agriculture. This should include a credible system for consumers to identify produce that meets "sustainable agriculture standards".
- Increase monitoring of toxic chemicals in our environment. There is growing public concern about the threat of the environment to our health and a sensitivity to chemicals in our food, water and air.

- By monitoring and publishing data on toxic chemicals, pressure will grow to eliminate their sources, including agricultural chemicals.

2.4.5 Other policy areas:

- All environmental and population policy areas have to be re-invented. Policies that are currently in effect are completely inadequate for a world that is interdependent and has at its disposal safe and useful technologies.
- Encourage transition from competitive economy to synergetic economy.
- Establish the law of the Life Cycle Assessment in each country.
- Each country publish a forward looking statement every 5 or 10 years.
- Each country should hold an annual "search" conference to assess priorities and review public expenditure
- Massive increase in financial support in status of scientific research in tropical biology (NOT just agronomy or forestry).
- ecological economies (considering nature as a scarce natural capital, properly valuing it, allocating natural commons similar to allocating electromagnetic spectrum through the ITU, and ecological taxes, make environment a global security issue, off-set investments that allow firms to remedy environmental damage in one country by cheaper countervailing measures in another, tradeable pollution permits that fix global emission limits for countries or industrial sectors, reduce import tariffs on environmentally sound technologies, good and equipment, tax breaks for more environmentally sound practices),
- more flexible repatriation limits for income made from these technologies which could provide firms with necessary financial break to enable investments in more costly, green technologies in less developed countries,
- higher tariffs or taxes on polluting products or technologies, with the revenues collected to be used to subsidize the acquisition of environmentally safe technologies,
- bulk purchase agreements for a region,
- purchase guarantees by bilateral, multilateral or regional funding agencies which could underwrite less developed country purchases of sound technology,
- an international technology bank, funded by country pledges, could acquire the rights to innovative green technologies so as to make them easily available to environmentally less advantageous countries,
- an international center to settle investment disputes could curb restrictive business practices that block environmentally less advantageous country access to sound technologies, such as restrictive licensing arrangements and prohibitively high prices.
- debt-for-nature swaps.
- development assistance programs that could also provide additional impetus to green technology transfers. Some of these instruments have already been used. Their efficiency varies.

FFC/93 Papers & Proceedings

Demographics of the Population and Environment Delphi Panel (round I and or round II):

SECTOR/INSTITUTIONS	NUMBER RECEIVED	PERCENT OF TOTAL
---------------------	-----------------	------------------

Private Sector	13	26 %
Universities	12	24 %
International Organizations	10	20 %
Governments	8	16 %
NGOs	7	14 %
TOTALS	50	100 %

Number Sent	Percent of Total	Country/ Region	Percent responded of number sent by country/region
29	38 %	United States	38 %
19	25 %	Europe	53 %
13	17 %	Asia	77 %
7	9 %	Africa	71 %
5	7 %	Latin America	60 %
3	4 %	Canada	66 %
76	100 %	TOTALS	

EXPERIMENTAL ECONOMICS

SESSION I

MODERATOR Doug Hale, Energy Information Administration

PRESENTATION 1 Charissa Wellford, Federal Trade Commission

SESSION II

MODERATOR Andy Kydes, Energy Information Administration

PRESENTATION 1 Steven Elliott, Oak Ridge National Laboratory

PRESENTATION 2 Jamie Kruse, University of Colorado

PRESENTATION 3 David Bjornstad, Oak Ridge National Laboratory

Experimental Economics

Antitrust: Results from the Laboratory

Charissa Wellford, Federal Trade Commission

I. Introduction

Antitrust practitioners are well acquainted with the challenges of applying economic theory. Industrial organization issues have been studied in the laboratory for over three decades, and the results of this research offer useful insights for applied work in antitrust.¹ Experimental methods provide a means for enhancing our understanding of markets, broadening the base of knowledge that is appropriate to draw from as an expert witness. Some have suggested that laboratory findings could be admissible as evidence in court.² In this paper I present results from the experimental literature that address competition issues.³

In short, laboratory markets are created by defining the structure and rules of the market, and by paying volunteer market participants in cash according to how successful they are at achieving their economic goals within this setting. In contrast to theory (or computer simulations of theory), in which the nature of behavior is assumed, real human behavior directly impacts market performance.

Landes recalls that Ronald Coase's reason for growing tired of antitrust was

because when the prices went up the judges said it was monopoly, when the prices went down, they said it was predatory pricing, and when they stayed the same, they said it was tacit collusion.⁴

In the laboratory, the experimenter knows the true competitive equilibrium and monopoly equilibrium, for the experimenter sets the costs, demand, and other market parameters that underlie these theoretical predictions. Thus, when a firm cuts price in a laboratory market, it can be unambiguously determined whether it is behaving in a competitive or a predatory manner.

The market performance resulting from behavior observed in the laboratory can be measured relative to the theoretical predictions. For instance, exploiting the experimenter's ability to control the underlying cost structure so as to create conditions that are theoretically amenable to predatory pricing, enables tests of the theory to be examined in a market setting. The behavioral deviations from the theoretical competitive, monopoly, and predatory predictions, which are known, can be measured directly. Similarly, the market performance in laboratory monopoly and oligopoly markets can be compared directly to monopoly and oligopoly theory. Measuring the

¹ Chamberlin (1948) typically receives credit for conducting the first laboratory market experiment, although his procedures do not satisfy current generally accepted methods of experimentation. Fascinated with the potential of Chamberlin's experiment in which he was a participant (see Smith (1991b)), Smith (1962) began running several series of experiments in 1955 that demonstrated the competitive equilibrium as predicted by the standard textbook intersection of supply and demand schedules. Smith implemented his test using the double-auction institution, which he developed based upon the basic trading rules of the New York Stock Exchange. Other important early research includes Hoggatt (1959) on oligopoly and Fouraker and Siegel (1963) on bargaining and oligopoly. The use of laboratory methods has expanded significantly in the past few decades, with over 40 universities worldwide installing and maintaining computerized laboratories. Many other sites conduct experiments without the use of sophisticated equipment.

² See, for instance, Kirkwood (1981).

³ This brief survey of experimental results relating to antitrust is not a comprehensive compendium, but rather an introduction. Readers interested in further reading on industrial organization experiments with implications for antitrust are referred to Plott (1982, 1989) and Holt (1989, 1991). Isaac (1983) considers the role of experimental economics as a tool in public policy analysis and provides several examples. Smith (1981b) discusses the relationship of theory, experiments, and antitrust policy. The *Handbook of Experimental Economics*, edited by Kagel and Roth (1991), surveys a broad array of laboratory results. Smith (1990) summarizes the major findings from the first 30 years of experimental research, and Smith (1991a) contains reprints of many of the prominent papers in the field.

⁴ See Kitch (1983), p. 191.

distance of the observed behavior from the competitive, Cournot, and the monopoly equilibria under different conditions is possible.

Control and the ability to replicate are the premier advantages of laboratory analysis. Each time that a design parameter (control variable) is altered, a different experimental *treatment* is created. Some examples of different laboratory treatments that relate to competition issues include: allowing sellers to openly discuss their choices, prohibiting direct communication except through direct market choices (*e.g.*, price, quantity), providing more information to some parties than others, creating market power, restricting or permitting entry, and changing the trading institution (rules of contracting). Changing one control variable at a time, the treatment effect can be measured directly. Replication of laboratory markets (repeating structurally identical markets with different groups of participants) at other laboratories by other scientists increases the confidence with which we can accept the results.

The ability to control many features of laboratory markets and to replicate experiments under the same conditions allows tests that discriminate among theories. This provides an important tool to antitrust practitioners, who are frequently faced with this very task of identifying the relevant theory. While laboratory results may not identify which theory holds for a specific antitrust case, they can influence how an antitrust practitioner interprets the relevance of a theory. Once a theory is shown to be robust in laboratory markets under the conditions that the theory itself suggests, it can be "stress tested" to determine under which other conditions the theory continues to predict behavior and those under which its predictive ability vanishes. This also provides insight as to how markets work and which theories are most useful in particular circumstances.

Those who advocate the policy application of a particular theory bear the burden of proving why the theory is appropriate as a basis for policy. One way to examine a theory is in the laboratory. When a theory fails to predict well in a simple laboratory setting under conditions the theory itself suggests, it is difficult to believe that it would predict better in a more complex market in the naturally-occurring economy.⁵

Experiments enhance what we learn from traditional empirical analysis of field markets. In general, laboratory analysis is well suited for some sorts of antitrust issues and not for others. For example, experiments are not practical for estimating some case specifics such as the true industry cost structure, demand elasticity, or proper definition of the market. Traditional econometric analysis is better suited for that task, as well as the analysis of marginal changes in specific field markets. In contrast, laboratory experiments are better suited for examining regularities under general industry structural characteristics, characteristics often influenced by both theory and empirical studies of naturally-occurring markets. As Smith (1992, p. 2) notes, laboratory analysis is less a substitute for traditional econometric/empirical methods than it is a supplement which seeks appropriate data for dealing with structural questions that are beyond the scope of aggregate observational data.

Both methods of empirical analysis share the same goal: coherent observation.⁶ The laboratory enables the formation of data that address fundamental changes which often are difficult, if not impossible (or illegal), to address with field data. For instance, consider comparing the effects on competition of different information conditions, institutions, market concentration levels, or communication conditions (*e.g.*, permitting sellers to talk about price), while holding everything else constant. Moreover, as mentioned previously, the laboratory permits the observations to be evaluated against the true theoretical benchmarks.⁷

The prerequisite for effectively interpreting and applying experimental findings is understanding how a microeconomic system is created in the laboratory. The next section gives a succinct overview of methodology, followed by a discussion of laboratory results of relating competition issues.

⁵ Plott (1987) and Smith (1992 and 1993) summarize several laboratory experiments designed to inform policy in field markets.

⁶ See Smith (1985).

⁷ The laboratory also facilitates the analysis of econometric estimators. See Cox and Oaxaca (1990).

II. Experimental Methodology

Every microeconomic system investigated in the laboratory consists of (1) an *environment*, (2) an *institution*, and (3) *behavior*.⁸ An environment is made up of the utility functions and technology (knowledge and skill endowments) that each participant brings to the experiment, as well as the commodities initially endowed upon each individual by the experimenter. The utility functions and technology are inherently unobservable; however, their consequences are observable. In a market experiment, the reduced form of the environment is given by the supply and demand schedules, which summarize the incentives to exchange.

Elements of an institution include a language (comprised of messages sent by each participant), allocation rules, cost imputation rules, and adjustment process rules.⁹ All components of an institution, which define each agent's property rights in communication and in exchange, are observable. An institution is given in the laboratory by the experimental instructions, which describe the rules of exchange that each participant receives. One example of a familiar institution is the English (or progressive) auction, commonly used to sell objects (*e.g.*, art or livestock) in art/antique auction houses or at county fairs. Bidders in an English auction must improve upon (*i.e.*, bid higher than) the standing bid in order for their bid (message) to be recognized by the seller. The last bidder, the one who bids the highest, is allocated the object and pays the seller the amount of the final bid.

Once an environment and institution are established in the laboratory, the third major element of a microeconomic system, behavior, can be observed. The resulting behavior of participants in an experiment is a function of both the environment and the institution. *Behavior is not induced*, the incentives that participants face are.¹⁰

A laboratory microeconomic system is created by *inducing value*, which involves assigning monetary value to units of experimental assets.¹¹ For instance, each seller in a laboratory market is assigned a cost for each unit of a commodity that he might choose to sell. The difference between the price at which he sells a unit and its cost, determines the seller's profit from the sale of that unit. Likewise, each buyer is given the redemption value, guaranteed by the experimenter, associated with each unit of the commodity that he might choose to purchase. The difference between the resale (redemption) value of a unit and the price he pays for the unit, determines the profit that the buyer earns on the unit. In this manner, the monetary value of a decision becomes well defined for each market participant.

Control of preferences in a laboratory microeconomic system requires that four precepts are satisfied:

- (1) *Nonsatiation*. Utility increases in cash rewards (*i.e.*, each participant prefers more money to less).
- (2) *Saliency*. Rewards properly reflect incentives, rewards increase (decrease) in the favorable (unfavorable) outcomes.
- (3) *Dominance*. Rewards sufficiently compensate each market participant for any subjective costs associated with participation in the experiment.¹²

⁸ For a more complete and rigorous presentation on the creation of microeconomic systems in the laboratory, see Smith (1982b). For a condensed version, see Smith (1987). My summary of methodology utilizes his terminology.

⁹ Messages, for example, might include a bid or an offer, or acceptance of a bid or an offer. The allocation rule is a function of the messages of all participants, and specifies the allocation of commodities. The cost imputation rule gives the monetary payment each agent receives as a result of the messages. Adjustment process rules simply define the starting, transition, and stopping rules of the experiment.

¹⁰ If behavior were induced, experimentation would be synonymous with simulation, in which behavior is assumed and some element of randomness is injected into the process. For instance, participants in a Cournot (quantity-setting market) institution are not constrained to choose the Cournot quantities. Rather, each participant selects a quantity from his choice set, which includes the Cournot quantity among several other production alternatives.

¹¹ See Smith (1976b) for the theory of induced value.

¹² Failing to satisfy dominance could lead to participants making their decisions in a purely random fashion, because the monetary compensation does not sufficiently compensate them for their decision costs. Many experiments have been run in which payoffs have been doubled, tripled, or even more. In general, these increases do not significantly affect the mean of observed participant behavior. As long as participants earn at least the opportunity cost of their time, the results are not significantly affected. Smith and Walker (1993) report that increasing the payouts usually does not change the mean of the data, although it generally reduces the variance of the observations about their mean.

(4) *Privacy*. Each participant receives information on his own reward schedule, but not the reward schedules of the other participants.

Precepts (1) through (4) permit tests of theory. However, an additional precept is necessary for addressing whether replicable laboratory results shed light on the behavior in and performance of naturally-occurring markets: (5) *Parallelism*. Hypotheses that are supported in one laboratory or field market experiment also hold in other laboratory or field markets when corresponding conditions are held constant.

Like theory, laboratory markets may or may not parallel the forces that drive naturally-occurring markets. Of course, if specific theories or laboratory markets are used to inform policy, it is desirable that they provide a credible parallel. The use of careful experimental design, sometimes enhanced by computerized laboratories, permits the study of fairly sophisticated markets.¹³

The above methodology is usually applied in university laboratories, and laboratory market participants are generally college undergraduates or graduate students, who likely are not as sophisticated as the agents in markets of interest. Hence, the methodological issue arises that laboratory market outcomes might vary with the pool of participants. Most economic theories are based on the concept of rationality, agents maximizing their utility given constraints. When market outcomes differ with the subject pool, this suggests either that rationality is not a generalizable trait (*e.g.*, students are not rational, but OPEC members are), or that experience with a parallel institution in the naturally-occurring economy is important. Participants usually receive training in a specific laboratory institution, especially for the more complex institutions. Laboratory markets of different types have been run with GE executives,¹⁴ Chicago Mercantile Exchange over-the-counter stock traders,¹⁵ Energy Administration officials, natural gas corporation executives,¹⁶ Eastern European reformers,¹⁷ and other professionals. In general, these groups do *not* exhibit behavior that is significantly different than that of student participants in laboratory markets.¹⁸ Note that if subject pool selection becomes an issue, the laboratory accommodates its analysis.^{19,20} Still, in the interest of satisfying payoff dominance while optimizing the research output subject to budget constraints (the opportunity cost of a student participant likely is lower than that of a Chicago Mercantile Exchange trader),

¹³ See, for instance, the Rassenti, Reynolds, and Smith (1989) study of cotenancy in laboratory networks which are designed to parallel natural gas networks. See also Hong and Plott's (1982) telephone market study of rate filing policies for inland water transportation; and Smith (1993), which discusses the laboratory market foundation of the recently formed Arizona Stock Exchange (AZX).

¹⁴ Fouraker and Siegel (1963).

¹⁵ King, Smith, Williams, and Van Boening (*forthcoming*).

¹⁶ Smith presented the results of running experiments with both regulators and then with some of the regulated parties when summarizing his research on natural gas networks in Rassenti, Reynolds, and Smith (1989) at the American Enterprise Institute Conference on Policy Approaches to the Deregulation of Network Industries, October 10-11, 1990.

¹⁷ This data was collected by Michael Block and Vernon Smith, both of the University of Arizona, during a World Bank sponsored conference on market reform, Vienna, Austria (September 1992 and July 1993).

¹⁸ See Ball and Cech (1993) for a review of subject pool selection in experimental economics. Many psychology experiments involve different subject pools, however, psychology experiments frequently violate the precepts of nonsatiation, saliency, and dominance (usually by not monetarily rewarding laboratory participants for their decisions). Hence, the results from such experiments should be interpreted with caution. See Cox and Isaac (1986) and Smith (1991) for a comparison of the laboratory methodologies utilized in economics and psychology. Wilcox (1989 and 1992) and Smith and Walker (1993) evaluate the use of monetary rewards in laboratory markets.

¹⁹ See, for instance, Gerety (1987) which finds that undergraduates and prisoners do not exhibit significantly different propensities to conspire when presented with opportunities to do so in sealed offer auctions (markets in which each seller submits an offer to supply a good or service—similar to government procurement) without the threat of antitrust enforcement. However, as the Gerety hypothesized *a priori*, when the threat of enforcement was introduced, by means of a random monetary penalty for a violation, the subject pools viewed the threat differently. Namely, the students exhibited risk-aversion, concentrating excessively on the size of the penalty; while the prisoners were risk-seeking, focusing more on the low probability of detection.

²⁰ Dyer and Kagel (1992) test common value auction theory in the laboratory using students and professional construction contractors as participants. The use of the professionals highlighted the failure of current theory to parallel their industry. Several contractors requested information about the architect, suggesting that reputation, though not part of the theory, is an important factor in their decisionmaking.

and because most economics laboratories are located at universities, college students have been and likely will continue to be the preferred subject pool for most cases.²¹

In laboratory studies, as in econometric studies of naturally-occurring markets, it is possible to incorrectly implement the methodology and consequently draw inappropriate inferences. Beyond the basic precepts of good experimental design, discussed above, many other elements must be considered when designing an experiment.²² A few other design matters will be discussed below in presenting laboratory findings that relate to competition issues.

III. Some Experimental Results

A. Competitive Behavior and Institutional Effects

Laboratories are well-suited to examining straight-forward (and falsifiable) theoretical predictions. For instance, it is often hypothesized that the competitive outcome, where total surplus is maximized at the intersection of the demand and supply schedules, requires many buyers and sellers, as well as perfect information. The competitive equilibrium prediction provides a relatively accurate description of market performance across many laboratory institutions and environments. However, the speed with which competitive equilibrium is reached, the proximity of observations to competitive predictions, and the propensity for cooperative behavior under various conditions varies with the institution. As Smith (1989) highlights, until the 1960s, economic theories remained largely institution free, with outcomes derived from the environment (market structure and agent knowledge) and *ad hoc* assumptions about demand (e.g., perfect demand revelation or "price-taking" by buyers). For instance, the standard competitive equilibrium model says nothing about *how* equilibrium is achieved. Without specifying the mechanics of the underlying institution, such as the trading rules, the competitive model merely assumes that convergence will occur.²³ Important early developments in institution-specific theory include Shubik's introduction (1959) of the extensive form game to industrial organization theory; Hurwicz's theory (1960) of mechanisms; and Vickrey's work (1961) on (first-price, Dutch, second-price, and English) auctions. As Shubik notes (1959, p. 183):

It is foolish to entertain the delusion that the Cournot, Edgeworth, Chamberlinian, or other 'reaction-curve' assumptions are good approximations of non-co-operative behavior in the market *and* that we can leave all aspects of asset and corporate structure safely assumed away in a *ceteris paribus* condition and still come out with a useful theory of oligopoly.

Before we are in a position to examine the detailed meaning of any type of behavioristic assumption or reaction-curve we must be able to describe with reasonable verisimilitude the salient observable features of a market in a dynamic setting. [*Emphasis in original*]

Institution-specific theories, and laboratory tests of institution-free theories across different institutions, suggest that institutions matter.²⁴

²¹ In most laboratory markets, each participant has the same degree of familiarity or *experience* with the institution. This, of course, can be controlled if it is important that the market participants have varying degrees of experience with the market.

²² Methodological issues are discussed throughout the literature. See, for example, volumes of *Research in Experimental Economics*, Greenwich, CT: JAI Press, and Davis and Holt (1992).

²³ The concept of the Walrasian auctioneer comes to mind. The Walrasian model provides an institutional framework to describe how equilibrium is achieved. However, the abstract concept of employing a fictitious auctioneer raises the issue of whether the model parallels observable features of any institution in the naturally-occurring economy.

²⁴ The behavioral results from laboratory tests of auctions led to extensions of Vickrey's auction theories. For example, see Cox, Roberson, and Smith (1982) and Cox, Dinkins, and Smith (1993). Bargaining theory has also been altered to reflect laboratory results; see Bolton (1991). Other examples of theoretical work prompted by experimental results include Friedman (1984, 1991), Wilson (1987), and Easley and Ledyard (1988).

Several institutions have been studied in the laboratory. Two of the more important from an antitrust perspective are the *double-auction* and the *posted-offer* institutions.²⁵

1. The Double-Auction Institution

The double-auction, formulated by Smith (1962, 1964), involves buyers who are free to make bids and accept offers, and sellers who are free to make offers and accept bids for units of a homogenous commodity. Trading takes place over several market periods of specified duration. As Smith (1976a) points out, these markets are similar to real estate and over-the-counter securities markets, as no unaccepted bid or offer is binding in a later period unless it is resubmitted. Another version of the double-auction institution imposes an improvement rule, permitting only those bids (offers) that are higher (lower) than the outstanding bid (offer). Under the improvement rule, the double-auction more closely parallels organized exchanges such as the American or New York Stock Exchange, as the traders cannot cancel a bid or offer that they have made that is currently standing as the lowest offer to sell or the highest bid to buy. With or without the improvement rule, the observed outcomes are roughly the same.

During the past decade, the majority of antitrust cases that were challenged or settled by the FTC involved nonretail markets.²⁶ The double-auction institution also parallels important features of wholesale or nonretail markets, which frequently involve negotiation of contracts.²⁷ Hong and Plott (1982) find that when an institution involves direct negotiations (telephone-based negotiations are a key feature of some of the markets that they study in the laboratory) the performance of the market appears similar to that of double-auction markets in the laboratory.

In the laboratory, double-auction markets have very strong *convergence* properties and exhibit particularly high *market efficiency*, as measured by the percentage of the total potential surplus captured by the traders (or alternatively, the absence of dead weight loss). The tendency for contract prices to converge from the side of the market that has the greater theoretical surplus is well documented.²⁸

Figure 1 presents the cost and value assignments for a double-auction experiment and the resulting market performance with four buyers and four sellers inexperienced in laboratory markets.²⁹ Each seller (buyer) is assigned costs (redemption values) for three units of the experimental commodity. For instance, Buyer 1's redemption value (denoted B1 on Figure 1) for the first unit purchased is \$5.65. The experimenter guarantees to redeem Buyer 1's first unit for \$5.65. Thus, Buyer 1 earns the difference between the redemption value of the unit and the price he pays for it. As indicated in Figure 1, the redemption values for Buyer 1's second unit is \$4.60, and the third \$4.45. Similarly, Buyer 2 is assigned redemption values of \$4.95, \$4.75, and \$4.50 for the first, second, and third units purchased. On the supply-side, units of the commodity are produced to order in this market, so that sellers do not bear the cost of production or inventory for unsold units.³⁰ Seller 1's cost (denoted by S1 in Figure 1) of producing the first unit is \$3.75. Seller one earns the difference between the price at which he sells the unit and his cost. His production cost for the second unit is \$4.80, and the third, \$4.95. Redemption value and cost assignments did not change with trading period, each seller (buyer) faced the same costs (values) for three units each period. By the fourth round of trading, prices have stabilized at the competitive level. The competitive

²⁵ Thousands of double-auctions and posted-offer experiments have been conducted. See, for instance, Smith (1962, 1964, 1967, 1976b, 1981a); Plott and Smith (1978); Smith, Williams, Bratton, and Vannoni (1982); Friedman (1984); Isaac, Ramey, and Williams (1984); Ketchum, Smith, and Williams (1984); and Cox and Oaxaca (1990).

²⁶ I am indebted to Malcolm Coate, Andrew Kleit, and Rene Bustamante (1992) for use of their data set.

²⁷ I am indebted to John Morris, former FTC Assistant to the Director of Antitrust, for suggesting the parallel to me.

²⁸ For example, if the surplus is distributed asymmetrically such that the consumers surplus exceeds producers surplus at the static competitive equilibrium, prices will tend to converge to the competitive equilibrium from above. See Smith and Williams (1982).

²⁹ The market design and data presented in Figure 1 are reported in Ketcham, Smith, and Williams (1984).

³⁰ See Mestelman, Welland, and Welland (1987) and Mestelman and Welland (1988) for analysis of inventory effects in laboratory markets.

exchange quantity of seven units is a Nash equilibrium in this market.³¹ Appendix 1 provides sample instructions for a seller in a hand-run double auction market.³²

Other experiments have also shown that four buyers and four sellers can be sufficient to assure rapid and consistent achievement of the competitive price in double-auction markets. These results are robust with nonstationary demand and supply parameters. In other words, when the market demand and supply schedules change over time, the competitive prediction continues to describe the observed market performance.³³ Laboratory research in markets with nonstationary demand and supply schedules suggests the importance of *not* enforcing price discrimination laws. In a dynamic setting with nonstationary demand and supply schedules, the convergence to a new competitive equilibrium *requires* that different prices be charged to different customers. Thus, enforcement (or threat of enforcement) of price discrimination laws, such as the Robinson-Patman Act, in a dynamic setting likely would perversely encourage anticompetitive outcomes. Charging different prices to different customers in a dynamic market is consistent with competition.

In double-auction *duopoly* markets the outcomes remain competitive.³⁴ In fact, even with *one* seller, the hypothesis of monopoly pricing often is rejected in the double-auction institution.³⁵ Figure 2 shows data from a single seller, five buyer double-auction market.³⁶ The buyers in this market session signal one another by bidding low, strategically *withholding demand* so as to mitigate the power of the monopolist, who does not know the true demand curve. Withholding demand occurs when at least one buyer refuses to purchase units of the laboratory commodity, even when his marginal valuation for a unit exceeds the posted price. The seller(s) relies on the observed buyer behavior to estimate demand. Thus, if buyers under-reveal their true willingness-to-pay then the demand curve is altered in effect. The buyers withhold demand with the intention of influencing the future pricing decisions of the seller(s). The laboratory data on withholding show that the true competitive equilibrium is often preserved. In fact, the misrepresentations of the buyers encourages the competitive outcome. In the market presented in Figure 2, the buyers effectively exert downward pressure on price until the competitive outcome is achieved. Some double-auction monopolists are able to sustain prices above the competitive equilibrium; however, they usually are unable to sustain prices and profits at the level predicted by monopoly theory.

The data from double-auction experiments have profound implications for traditional competitive price theory. The laboratory results show that the competitive equilibrium is attainable with less stringent informational assumptions than thought necessary. The data indicate that each double-auction trader needs only know his own valuations; extensive knowledge of other traders' demand or cost valuations is not necessary to generate a stable competitive outcome. Moreover, the results show that the market need not involve a "large" number of buyers and sellers to achieve competitive results.³⁷ Trading experience is also not imperative to achieve competitive market performance, although experience may speed convergence. Moreover, the results do not require the standard assumption that agents must be "price takers" for the competitive equilibrium to be attained. In the double-auction

³¹ This is to say that if all but one seller (buyer) traded each profitable unit at the competitive equilibrium price, the remaining participant could not do better than to offer (bid for) each profitable unit at the competitive equilibrium price.

³² Double auction and other experiments either are hand-run, without the use of a computer network, or computer-run. Using the latter method involves assigning each market participant to a computer terminal, which is networked to the other participant's terminals, at which they can enter decisions and receive market feedback.

³³ See Miller, Plott, and Smith (1977); Williams (1979); Smith (1982a); Cox and Oaxaca (1990); Davis and Williams (1990a); and Davis, Harrison, and Williams (1993). Hoffman and Plott (1981) also supports this finding when speculators or "middlemen" are present.

³⁴ Smith and Williams (1989).

³⁵ See, for example, Smith (1981a); Smith and Williams (1981); and Smith and Williams (1989).

³⁶ The market design and data presented in Figure 2 are from Smith and Williams (1989).

³⁷ This is evident even in the early double-auction experiments (see, for instance Smith (1962, 1964). Even though the "large numbers" condition remains in many textbook explanations of competition (see, for example, Carlton and Perloff (1989) p. 67), theorists also recognize that the number of buyers and sellers required to achieve competitive outcomes in an industry is analytically unimportant. See, for instance, McGee (1971) or Fama and Laffer (1972). Demsetz (1973, p. 26) points out that "there has been an irresistible inclination among economists to identify real world monopoly power with the structure postulate of the monopoly model -- the one-firm industry. It is but a short step from this to the conviction that market concentration is an index of monopoly power."

institution, every trader is a price maker.³⁸ Thus, in the double-auction institution, the competitive prediction is more robust than theory indicates.

Thus, the results from double-auction experiments suggest that there should be limited, if any, antitrust concern in markets with institutional features that resemble those of the double-auction.³⁹ The strength of competitive tendencies in the double-auction institution arise from the ability of both sides of the market, buyers and sellers, to have a voice in the market. Buyers make bids to buy and sellers make offers to sell, both can accept or reject contracts. To the extent that an antitrust market exhibits such institutional features, experimental evidence offers little support for anticompetitive concerns.

2. The Posted-Offer Institution

Frank Williams (1973) was the first to examine experimentally the posted-offer institution, in which each seller sets production capacity (maximum output) and posts a "take-it-or-leave-it" unit price.⁴⁰ Sellers produce to order, so that no inventories are carried from one period to the next, and no production costs are incurred for unordered units. Each seller chooses his price before knowing his rivals' prices. Buyers are selected in random order each period and take turns completing their purchases. A trading period consists of a round of posting prices followed by shopping. When the lowest price producer exhausts his production capacity for the period, buyers move on to the next lowest price producer, and so on. All buyers may not be satisfied.

The posted-offer institution parallels several features of naturally-occurring retail markets. Sellers in retail markets set production capacities and post prices that cannot be easily changed for some interval. For instance, catalog marketers are unlikely to change their prices until the next issue. Airlines post special rates that are good for travel within certain dates. Stock-outs (hitting capacity constraints within a trading period) are also observed in both retail and posted-offer markets. Sometimes retailers post prices with disclaimers such as "while supplies last" or "subject to available seating." In the laboratory posted-offer institution sellers generally do not incur the cost of units that are not sold. Thus, subject to capacity constraints, production meets demand.

Plott and Smith (1978) and Ketcham, Smith, and Williams (1984) demonstrate that changing the institution from a double-auction to a posted-offer auction can affect outcomes. Prices tend to be higher on average over time in the posted-offer institution than in the double-auction institution. However, and more importantly from an antitrust perspective, prices in both market institutions converge to the competitive price. The reason that posted-offer prices tend to be higher on average than double-auction prices is that in posted-offer markets prices usually converge from above and take longer to adjust to the competitive level.⁴¹ Market efficiency is higher in double-auction markets than in comparable posted-offer markets.⁴²

³⁸ Each trader can actively participate in the pricing process: buyers make bids and sellers make offers. See Smith (1962, 1964, 1976a, and 1982a); Ketcham, Smith, and Williams (1984); and Plott and Smith (1978).

³⁹ Further evidence, presented below, enhances this conclusion.

⁴⁰ This analysis of the posted-offer institution was actually an attempt to compare the single unit per agent endowment feature of the Smith (1962 and 1964) double-auction experiments to a double-auction in which traders are permitted to buy or sell more than one unit per trading period. However, when altering the number of units per trader, Williams also unintentionally modified several of the trading rules. These modifications change the institution from a double-auction to what is commonly known today as a posted-offer auction. Williams reports behavior that is somewhat different than the behavior observed in the Smith single unit per trader double-auctions. Notice that the hypothesis Williams is testing is compound, contaminating the pure effect of moving from a single to a multiple unit assignment of costs and values in the same institution. Plott and Smith (1978) recognize this problem. They also classify and then test the difference between the double-auction and posted-offer institutions.

⁴¹ Davis and Williams (1985) find that, unlike in double-auction markets, posted-offer contract prices converge to competitive equilibrium from above, irrespective of the asymmetry between producers' and consumers' surplus. The institutional effect also interacts with other design parameters. In posted-offer markets with no Nash equilibria, prices converge towards the competitive equilibrium; while in cases in which there exists a nonNash competitive equilibrium and a Nash equilibrium, prices lie closer to the Nash equilibrium than to the competitive or the limit price equilibrium. See Ketcham, Smith, and Williams (1984).

⁴² This comparison of institutions is conducted holding all else equal. Thus, while the posted-offer markets exhibit lower levels of market efficiency than double-auction markets, in practice there remains the issue of institution efficiency. For example, consider grocery shopping under the two alternative institutions. The transactions costs of negotiating the price of each of many items under the double-auction trading

When sellers produce in advance, rather than to demand, and there is no carry over of inventory from one period to the next, the mean price across the two institutions no longer differs significantly. However, the double-auction remains more efficient, because more unsold units are produced in the posted-offer institution. Thus, the production characteristics of an institution may have an impact on market performance.⁴³

In contrast to double-auction settings, buyers do not possess the ability to signal in posted-offer markets. While buyers can reject posted offers, these decisions are private, eliminating a means to signal to other buyers. Sellers, on the other hand, maintain the ability to signal to other sellers via pricing decisions, which are posted for all market participants to see. Figure 3 presents data from a posted-offer market with experienced sellers.⁴⁴ The underlying market structure is the same as that of the double-auction experiment presented in Figure 1.⁴⁵ However, for the posted-offer trial displayed in Figure 3, attempts are made to create conditions under which tacit collusion would have a good chance of being observed, by recruiting the sellers on the basis of their prior propensity to collude tacitly.⁴⁶ Although some signaling efforts (attempts to encourage supracompetitive pricing) arise in the early periods, they are eventually aborted. The trial can hardly be classified as a case of successful collusion. Prices stabilize at the competitive level. Ketcham, Smith, and Williams attribute the slight increase in price during the later periods to "end effect."⁴⁷

Recall that in the double-auction institution, the monopoly model has little predictive power. The monopoly prediction does a better job of describing behavior in posted-offer markets with single sellers. Yet, while the prices lie above the competitive equilibrium in most of the posted-offer markets studied, the index of monopoly effectiveness (the percentage of theoretical monopoly profits actually captured by the single seller) is seldom 100 percent. For instance, across the posted-offer monopoly sessions reported in Isaac, Ramey, and Williams (1983), the average index of monopoly effectiveness is 35 percent.⁴⁸ The Isaac, Ramey, and Williams study uses *human* buyers. Brown-Kruse (1986) shows that when demand is fully revealed (generally the case when human buyers are replaced by *simulated buyers*) sellers extract more rents from posted-offer markets. This suggests that when sellers lack complete information on demand, human buyers can strategically influence the market outcome.⁴⁹

B. Oligopoly

1. Institutions and Information

Theory presents several alternative models of oligopoly. The laboratory permits the examination different institutions suggested by the theories, as well as the impact of various information conditions upon market performance. Many oligopoly experiments use either double-auction or posted-offer institutions, but other institutions

rules could be substantial and avoided by trading according to posted-offer rules. In general, the gain from avoiding such transactions costs in retail settings by employing a posted-offer institution likely will dominate the loss in market efficiency from not adopting the double-auction trading rules.

⁴³ Mestelman, Welland and Welland (1987) and Mestelman and Welland (1988).

⁴⁴ The sellers have participated previously in posted-offer laboratory markets.

⁴⁵ The demand and cost parameters across the two experiments differ by a constant, \$0.55, but the underlying structure is the same.

⁴⁶ See Ketcham, Smith, and Williams.

⁴⁷ Because subjects were experienced, they probably had a reasonable idea that the experiment would end around period 25. Generally, participants are not informed of the number of periods to be conducted in a given market session.

⁴⁸ The per-period measures range from -222% to 100%.

⁴⁹ Monopoly theory also encounters difficulties in a *posted-bid* institution. Smith (1982) shows that in markets with *one* seller who can accept or reject bids, and five buyers who post bids (a *posted-bid* as opposed to a *posted-offer* institution), monopoly power is not exercised. This results because the buyers successfully signal other buyers via their bids to withhold their true willingness-to-pay. Thus, through misrepresenting their willingness-to-pay, the buyers neutralize the seller's monopoly power, because the seller does not know the true demand curve. Market price and production converge to the competitive prediction in this institution and environment.

such as Cournot⁵⁰ and Bertrand⁵¹ are also studied. In both Cournot and Bertrand settings, the market demand curve is provided to all sellers. In Cournot markets, sellers set quantities; in Bertrand markets, they set price.

As first shown by Carlson (1967) for large markets (more than 20 sellers), and later by Wellford (1989) for oligopolies (five firms), quantity-setting markets that are theoretically unstable in the cobweb or Cournot sense exhibit strong convergence properties in laboratory markets. Subjects behave "as if" the markets are stable. The hypothesis that subjects form expectations consistent with the cobweb model, *i.e.*, that price next period will be the same as price in the current period, is not supported by laboratory tests. Further, rational expectations also fails to predict well. Of the traditional models considered, adaptive and extrapolative expectations best describe the data.⁵² However, the structure, not just the parameterization, of expectations seems to be adaptive. Market participants alter their expectation process when they realize that their myopic responses lead to relatively unfavorable outcomes.⁵³

Information plays an important role in laboratory markets. Most oligopoly theories require *complete* and *common information* on preferences to achieve noncooperative equilibria.⁵⁴ This raises issues for applied work in antitrust and other areas, as it is difficult to believe that each market participant in the naturally-occurring economy actually knows the utility values for all other participants. In the laboratory competitive and noncooperative equilibrium outcomes are more likely to arise behaviorally however, under the conditions of *private information* than under complete information.⁵⁵ If *sellers* actually had complete information, they often would prefer to achieve a different equilibrium.⁵⁶

Fouraker and Siegel (1963) document the effect of complete versus private information in their study of Cournot and Bertrand duopolies and triopolies. Tables 1 and 2 summarize Fouraker and Siegel's results, which have been confirmed in many other environments and institutions. The tables provide the percentage of sessions in which the market quantity (price) in a Cournot (Bertrand) market converges to (or near) a particular equilibrium concept after a specified number of periods.⁵⁷ Sellers in complete information treatments know the profit functions of their rivals. In particular, sellers know how each rival's quantity decision affects the rival's profit. In the private information treatments, each seller knows only his own profit function. Referring to Table 1, duopolies in the reported quantity choice markets under private information show no support for the shared monopoly outcome (*i.e.*,

⁵⁰ See, for instance, Fouraker and Siegel, *ibid*; Binger, Hoffman, Libecap, and Schachat (1990); and Wellford (1990).

⁵¹ See, for example, Friedman (1967) and Friedman and Hoggatt (1980); and Hoggatt, Friedman, and Gill (1967).

⁵² It is interesting to note that macro theorists are reverting from rational to adaptive expectations. See, for instance, Thomas Sargent's address in honor of Milton Friedman's eightieth birthday at the 1992 Western Economic Association meetings.

⁵³ Wellford, *ibid*. Schmalensee (1976) finds in a stark (nonmarket) setting that the speed of response in the adaptive structure retards during turning points. See Grether (1980 and 1992) and Camerer (1987) for tests of Bayes rule.

⁵⁴ Complete information holds when each market agent knows the preferences of all the other agents. Common information holds when all of the agents have information, and they know that all of the other agents possess this information and so forth.

⁵⁵ Complete information in the laboratory generally refers to information on the payout functions and previous decisions of other market participants. Under conditions of private information in the laboratory, market participants know only how their own decisions affect their payout, and can observe the market outcomes (*i.e.*, the market price and market production level). For other discussions on the effect of complete information in laboratory markets, see Smith (1989) and (1990), and McCabe, Rassenti, and Smith (1989). Holt (1985) conducts a quantity-choice duopoly game under complete information, and notes that the subjects want to earn more than the other firm, yet show no tendency to maximize the difference between their earnings. He conjectures that there "...probably would have been less variability in the data if the subjects in these experiments had not been given the complete information necessary to compute the other seller's profits," (p. 323).

⁵⁶ In general, a symmetric Cournot duopolist can earn more by choosing the joint profit maximization equilibrium rather than playing the noncooperative (Cournot) strategy.

⁵⁷ Fouraker and Siegel report data for period 21 in the Cournot markets, and period 14 in the Bertrand markets. A period is one round during which sellers set quantity (price) in each Cournot (Bertrand) market session.

the joint maximum), and the strongest support (87.5%) for the Cournot equilibrium.⁵⁸ By contrast, the joint maximum equilibrium concept received support in 31.25 % of the Cournot markets conducted under conditions of complete information. While the Cournot prediction continues to describe behavior in 46.87 % of the markets, it has diminished predictive power relative to the otherwise equivalent private information treatments. The ability of the noncooperative equilibria (*i.e.*, Cournot and Bertrand) to predict also improves under conditions of private information for the both Cournot triopoly and Bertrand duopoly markets. Table 2 suggests that the incentives in the Bertrand markets are sufficient to overcome the information effect as the number of sellers increases, so that the Bertrand equilibrium predicts well under both information treatments.

The ability of markets to economize information (*i.e.*, without intending to, market agents achieve efficient outcomes) was recognized by Adam Smith (1776) in his discussion of the invisible hand and described in more detail by Hayek (1945). However, until the advent of laboratory methods, it was difficult to test *directly* whether this hypothesis would hold. At least through the mid-fifties, the economics profession at large did not accept the relevance of the claim.⁵⁹ In fact, before he started conducting laboratory experiments, Vernon Smith (1993), who some call the "father" of experimental methods in economics, also doubted its applicability until his experimental subjects convinced him otherwise.

In many experimental markets, poorly informed, error-prone, and uncomprehending human agents interact through the trading rules to produce social algorithms which demonstrably approximate the wealth maximizing outcomes traditionally thought to require complete information and cognitively rational actors.

While the outcomes of Smith's early experiments supported the notion that markets economize information,⁶⁰ Smith (1982a) provides a more stringent test of the Hayek Hypothesis, which he specifies as (p. 167):

Strict privacy together with the trading rules of a market institution are sufficient to produce competitive market outcomes at or near 100% efficiency.

Using the double auction institution, three treatments are considered: (1) stationary, (2) cyclical, and (3) irregular shifts in demand and supply. Figure 4 presents the data from one laboratory session conducted under the most rigorous design reported. The data are representative of other sessions conducted under the same design. In this market, after the first week of trading (five periods), the demand and supply schedules shift from D_1 and S_1 to D_2 and S_2 . The four sellers have identical unit costs of \$5.70, with an 11 unit aggregate production capacity during the first week and a 16 unit one in the second. The four buyers have the same unit value of \$6.80, with the aggregate capacity to demand no more than 16 units during the first week, and no more than 11 units the second.⁶¹ The competitive equilibrium during the first (second) week occurs at the price of \$6.80 (\$5.70) with an output of 11 (11) units. This market design provides a rigorous "stress test" of the Hayek Hypothesis. The competitive equilibrium is not the unique price and production mix which satisfies 100% market efficiency. Complete efficiency occurs when 11 units are sold at *any* price ranging from \$5.70 to \$6.80. This provides a stringent test of the ability of price to converge to the competitive equilibrium in this design, because at the competitive equilibrium during week one (week two) all of the \$12.10 per period exchange surplus accrues to the sellers (buyers). Thus, at the competitive equilibrium during the first week (second week), buyers (sellers) receive only a 10¢ trading commission per unit exchanged. The extreme asymmetry of the exchange surplus between buyers and sellers creates tremendous incentives for the market agents that are cut out of the exchange surplus at the

⁵⁸ This is not to say that the Cournot equilibrium was observed in 87.5% of the markets during period 21, rather that the observations lie closer to the Cournot prediction than to the alternative equilibria of interest.

⁵⁹ As Smith (1993) notes, the Chicago and Austrian schools of thought were the exceptions. See also Kitch (1983) for a remembrance of the days at Chicago prior to the general professional acceptance of these ideas.

⁶⁰ See Smith (1962 and 1964).

⁶¹ Market participants are not informed in advance that market conditions might change. Smith (1982a, p. 172) notes that "from the point of view of the participants this change is rather subtle in that their individual value and cost assignments remain constant." Only the aggregate demand and supply capacities change.

competitive equilibrium to exert more resistance in the bargaining process. Such resistance could curtail, if not prevent, convergence to the competitive equilibrium.

The results presented in Figure 4 reveal an slow but steady convergence of price to the competitive solution by periods four and five in week one. It appears as though sellers' expectations of high prices and profits initially retard the rate of convergence after the shift in the demand and cost schedules. However, the new competitive equilibrium is achieved by periods nine and ten. These data provide impressive support for the Hayek Hypothesis in an environment with stationary demand and supply curves in the double-auction institution. Smith (1984b) also finds strong support for the hypothesis in experiments with dynamic environments. Van Boening and Wilcox (1992) have constructed markets with avoidable fixed costs in a double-auction institution. Their markets provide an even greater challenge to the ability of the market to achieve the competitive outcome, and have spawned research efforts to explore and design institutions that overcome more effectively the lumpiness inherent in such cost functions.⁶²

Another interesting informational finding relates to situations of asymmetric information. Camerer, Lowenstein, and Weber (1989) find evidence that supports a popular theory in psychology, that better informed agents often fall victim to the "curse of knowledge." The curse occurs when agents fail to ignore information that they have that others do not, even when it is unnecessary information. For instance, investment bankers, who possess more information than the public about a stock to be offered publicly, must anticipate the demand for the securities by the relatively uninformed public when pricing the shares.⁶³

The curse runs contrary to the conventional assumption in agency theory that agents with better information can correctly forecast the decisions of less informed agents.⁶⁴ More information may not always be an advantage. This is not to say that no information is necessarily better than some. In fact, the anticompetitive distortions of asymmetric information outcomes can be overcome in markets.⁶⁵ Market forces in the double-auction institution decrease the magnitude of the "curse" by about 50%, this decrease is largely due to the more rational traders disproportionate share of actions in the market. Thus, the disciplining forces of the market promote the paradox that individual irrationality can encourage collective rationality.

The curse is otherwise difficult to erode, even with very large incentives and extensive training, learning does not improve much. The curse remains. Camerer *et al.* note that the curse could affect strategic behavior in oligopoly settings, as firms erroneously decide to enter or exit. However, this remains to be tested. The curse has also yet to be analyzed in a systematic fashion in other laboratory institutions. Such study may enhance our knowledge of the breadth of the curse and the ability of markets to overcome it.

2. Market Power

The role of information, the form of institution, and other factors come into play in laboratory studies of market power. *Market power*, classified as a "unilateral competitive effect" by the Horizontal Merger Guidelines (1992, §2.2), is the ability of a firm to maintain price above the competitive level unilaterally and profitably. *Equilibrium Market Power*, a general game-theoretic concept, refers to a noncooperative equilibrium that is consistent with supracompetitive prices.

The laboratory facilitates the analysis of market power, in that it can be directly built into the structure of the market. Thus, the behavioral consequence of the existence of market power can be directly compared to markets that are similar except for the lack of theoretical market power.

Figure 5 presents a market design in which unilateral market power exists in a noncooperative game.⁶⁶ At the competitive equilibrium in this market with five sellers, 16 units are produced at the price of \$2.60. In order

⁶² See, for instance, Durham, Rassenti, Smith, Van Boening, and Wilcox (1993).

⁶³ Other examples include fashion, wine, and movies. Industry experts have informed opinions about the value of the product, but if they are subject to the curse of knowledge their prices will tend to be biased upward (downward) for high (low) quality items relative to the uninformed general public's willingness to pay, and thus also relative to the profit maximizing level.

⁶⁴ See Conn, Young, and Bishop (1991) for tests of principal-agent theory.

⁶⁵ See, for instance, Isaac, Ramey, and Williams (1984).

⁶⁶ This design originates in Holt, Langan, and Villamil (1986) and is implemented in several subsequent laboratory studies on market power.

to encourage the purchase and sale of the marginal units, each market participant receives a 5¢ commission per unit exchanged.⁶⁷ Either of the low cost sellers, seller one (S1) or seller two (S2), could unilaterally and profitably increase their price from \$2.60 to \$2.85. At the competitive equilibrium, the two low-cost sellers would each earn \$2.60 per unit for five units of production, less the \$11.50 cost of producing the five units, plus \$.25 in commissions, for earnings of \$1.75.⁶⁸ However, assuming that the other sellers behave competitively, if either seller one or seller two withholds its two zero-profit units from sale, the *effective* supply curve shifts back so that the *effective* competitive equilibrium, where the demand and (effective) supply schedules cross, occurs at the price of \$2.85. At this price, the seller who suppresses production by two units would earn \$2.40, including commissions. Hence, it is profitable for either seller one or seller two to unilaterally increase price above the fundamental competitive equilibrium. In this example, the other low-cost seller, who withholds no units of production from the market would earn \$3.00. Hence in this example, the two low-cost sellers increase their earnings by either \$0.65 or \$1.25. Any of the high-cost sellers (sellers three, four, and five) could also withhold a unit, and unilaterally and profitably increase price above the competitive level, but not as dramatically as the low-cost sellers. Unilaterally withholding one unit from the market would lead to a price between \$2.60 and \$2.85, thus a high-cost seller would increase his earnings by no more than 20¢ and could decrease it by less than 5¢.⁶⁹

Table 3 presents a terse summary of several laboratory studies of market power.⁷⁰ I briefly discuss the overall findings of the research and some of the issues that they raises.

a). Supracompetitive Pricing, Efficiency, and the Hayek Hypothesis

While not consistently observed across all design treatments, supracompetitive pricing does occur in some of the laboratory markets. Probably the most striking result across the studies is that *even when market power is built into the structure of the market, the observed market efficiency remains relatively high*. Thus, the deadweight loss associated with the presence of theoretical market power seems to be small in laboratory markets.

What do the results of supracompetitive pricing and high market efficiency imply when evaluated in the context of the Hayek Hypothesis? One, interpretation of the Hayek Hypothesis, the *price interpretation*, implies that prices should converge to the competitive level. Another interpretation, the *efficiency interpretation*, suggests that the surplus captured should approach 100%. Davis and Williams (1991) find that when sellers recognize they possess market power, the price interpretation is often violated in double-auction and posted-offer markets. However, the efficiency interpretation of the hypothesis cannot be rejected, even when the price interpretation fails to hold. Across the studies reported in Table 3, the average market efficiencies consistently exceed 90%.⁷¹

Several of the studies note that sellers with theoretical market power rarely withhold units in the manner that static market power theory predicts. For instance, Davis and Williams (1991, p. 270) report that across their eight posted-offer markets the two low-cost sellers

offered less than nine of the ten units collectively available to them in only eight out of one hundred and forty six periods. Five of these instances occurred in the first two periods of trading in four different experiments, while the remaining three instances were isolated instances in three different experiments.

The authors also report that one of four of their double-auction trials exhibits evidence of strategic withholding by sellers one and two. Yet, these low-cost sellers fail to achieve prices that are significantly higher than those in the other double-auction trials.

⁶⁷ Plott and Smith (1978) find that 5¢ commissions are sufficient to motivate marginal trades.

⁶⁸ As Figure 4 indicates, seller one (or two) incurs production costs of \$1.60 for the first unit, \$2.35 per unit for the second and third units, and \$2.60 per unit for the fourth and fifth units.

⁶⁹ Withholding two units would unilaterally make the effective competitive equilibrium occur at the price \$2.85, however any high-cost seller doing so would forego earnings altogether.

⁷⁰ Several of the studies (*see*, for instance, Holt and Davis (1990)) include a baseline treatment which are similar to the other treatments as possible except that no market power exists. These baseline market trials are not included in Table 3.

⁷¹ A few studies do not report efficiencies.

These unexpected results could be due to the fact that sellers in the laboratory markets do not recognize that they have potential market power. Only in a few cases does experience with the institution seem to increase the sellers' awareness of their potential market influence to the point that they attempt to exercise their power.⁷² The general absence of withholding in these markets suggests that the observed supracompetitive pricing is due to tacit collusion rather than the exercise of market power. When market power exists, the likelihood of tacit collusion increases; yet, the achievement of perfect collusion is infrequently observed.⁷³

b). Institution Matters

When controlling for market power, larger deviations from the competitive equilibrium are observed in posted-offer markets than in double-auction markets.⁷⁴ The differential impact is most obvious when the *buyers*, rather than the sellers, have market power. When buyers have market power in double-auction markets, they are able to elicit subcompetitive prices,⁷⁵ however in posted-offer markets they are unable to overcome the institutional asymmetry (*i.e.*, that buyers cannot post bids, but sellers can post offers - - so that buyers have a more indirect influence on price) and prices converge to the competitive equilibrium.

Friedman and Ostroy (1991) argue that in most recent laboratory studies of market power, market agents have the opportunity to trade only a few units of an indivisible commodity. This indivisibility significantly reduces the profitable opportunities to raise price by withholding units. As the authors point out, in some markets the only way a trader can influence price is by exiting from the market and therefore giving up all profit for the trading period. When divisibility of sales is incorporated, it affects the ability of sellers to exercise market power in the double-auction institution. Most notably, when divisibility is allowed, supracompetitive pricing is not observed. Friedman and Ostroy describe the behavior of the market participants (p. 3),

Through their misrepresentations, the subjects spontaneously imposed a perfectly competitive environment on each other...[the institution with divisibility] encourages individuals with potential monopoly power to exercise that power in such a way that it is neutralized.

Another institutional feature that affects the exercise of market power involves whether the message space of the buyers is limited to include only their true reservation values, as is often the case when buyers are simulated rather than human. As found in other environments under the posted-offer institution, prices are significantly lower when human buyers are present than when buyers are simulated to behave as the theory predicts.⁷⁶ Davis and Williams (1991) find that when human buyers are present and sellers have market power, price converges to the competitive equilibrium in posted-offer markets.

While the data from experiments reveal that institutions matter, the data also suggest that a feature of the environment, the nature of aggregate production capacity at the competitive equilibrium, might determine whether market power will in fact be exercised or whether tacit collusion is more effectively achieved. When aggregate

⁷² Holt, Langan, and Villamil (1986) find one experienced seller who figures out the advantages of withholding. However, other studies (*e.g.*, see Davis and Williams (1991)) find no evidence of increased incidence of withholding of production by experienced participants. Another way to increase the sellers' familiarity with their potential market power would involve establishing the condition of complete information, so that each seller knows the cost and payoff functions of his rivals. It could also be important to require the sellers to "earn the right" to be in a position of potential power. Hoffman and Spitzer (1982) find that in laboratory tests of bargaining that if the preferential property rights are randomly assigned rather than "earned" (through a trivial game) that participants are less likely to fully exercise those rights. This would make it clear to all sellers and buyers (when human buyers are used) that there exist fundamental differences in cost structures across sellers. Barring these two design changes, complete information and "earned" assignment of property rights, it is unlikely that the opportunities to the sellers could be made clearer.

⁷³ See, for instance, Davis, Holt, and Villamil (1990).

⁷⁴ See, for instance, Holt (1989) and Davis and Williams (1991).

⁷⁵ This is found in one of two buyer-power market trials (not summarized in Table 3) from the Holt, Langan, and Villamil (1986) study. This trial, which has a symmetric distribution of exchange surplus between buyers and sellers, establishes that contracts occurring at prices below the competitive equilibrium are not due solely to the asymmetric distribution of exchange surplus (*see supra* note 24).

⁷⁶ See Brown-Kruse (1986 and 1991).

production capacity of the market is constrained at the competitive equilibrium, the likelihood of supracompetitive pricing might be different than when this aggregate capacity is nonbinding. The column in Table 3 marked "excess capacity at CE" indicates whether or not the production capacity is nonbinding. This issue is discussed in greater detail in the next section, as it also arises in research on the effect of communication upon collusion.

3. Tacit and Explicit Conspiracies in Restraint of Trade

Many laboratory studies of oligopoly focus on tacit or explicit conspiracy.⁷⁷ The laboratory provides a useful mechanism for studying conspiracies. By controlling directly the communication of market participants (*e.g.*, whether or what kind of communication is allowed), the impact of explicit versus tacit conspiracy can be compared to theoretical predictions. Collusion is studied in several laboratory institutions, including the Cournot, double-auction, and posted-offer auction.⁷⁸

a). Opportunities for Tacit Collusion

Permitting communication only through the choice variables specified by the theory and institution allows tests of the ability of sellers to collude tacitly. In other words, sellers have no means of communicating directly with each other except through their market decisions.

In Cournot industries with *more than two sellers*, prices tend to lie between the static Cournot and competitive predictions, the joint profit maximum (*i.e.*, the monopoly outcome) is not achieved.⁷⁹ Theoretically, the Cournot solution approaches the competitive level as the number of sellers increases. Behaviorally, the fewer the number of sellers, the higher the observed variance in prices in laboratory quantity-setting markets.⁸⁰

In laboratory Cournot *duopolies*, the behavioral results are less clear. Binger, Hoffman, Libecap, and Schachat (1990) observe prices converging to the collusive outcome from above the Cournot-Nash equilibrium under conditions of complete information, while Morrison and Kamarei (1990) find successful collusion in only one of their eight laboratory duopolies conducted under conditions of incomplete information. Under complete information in quantity-choice duopoly games, the Cournot-Nash equilibrium predicts better than the consistent conjectures and the collusive equilibria.⁸¹

Similar mixed results arise in price-choice institutions with two sellers. Dolbear, Lave, and Bowman (1966) mention more variability in outcomes with duopolies than with triopolies in price-setting markets. Davis, Holt, and Villamil (1990) also find that some duopolies exhibit rivalistic behavior in posted-offer markets. Benson and Faminow (1988) show that subject experience with the institution increases the incidence of tacit collusion in posted-offer duopolies, but they too observe experienced participants who compete fiercely. So, while tacit collusion between two sellers appears to be easier to coordinate than when there are more sellers in laboratory markets, head-on competition is also observed in duopolies.

Posner (1976) points out that the concentration level at which concerns of collusive pricing arise depends on the economist. For instance, Scherer (1970 p. 185) would be concerned when the four largest firms hold a forty percent industry share. Bork (1978 pp. 221-222), on the other hand, argues:

⁷⁷ R. M. Isaac (1990) provides a useful overview.

⁷⁸ See Isaac and Walker (1985) and Gerety (1987) for worthwhile laboratory studies of collusion in sealed-bid and sealed-offer auctions. Examples of these institutions in field markets include the application of sealed-bid auctions to assign off-shore oil leases, and sealed-offer auctions to select procurement contractors.

⁷⁹ This is found in many studies. See, for instance, Fouraker and Siegel, *ibid*; Binger, Hoffman, Libecap, and Schachat (1990); and Morrison and Kamarei (1990).

⁸⁰ See, for instance, Fouraker and Siegel (1963) and Wellford (1990).

⁸¹ See Holt (1985). At the consistent conjectures equilibrium, each firm knows the other's output choice and, given the conjectured reaction of the other firm, any deviation from their own output level would be unprofitable.

we are in the area of uncertainty when we ask whether mergers that would concentrate a market to only two firms of roughly equal size should be prohibited. My guess is that they should not and, therefore, that mergers up to 60 or 70 percent of the market should be permitted - a figure, curiously enough, that resembles the old Sherman Act..."

Bork also suggests that horizontal mergers that would create no fewer than three significant firms should be considered presumptively lawful. In general, applying these alternative rules of thumb for identifying presumptively collusive markets, the Bork rule outperforms the Scherer rule in laboratory markets.

Over time, average prices in laboratory posted-offer oligopolies are higher than in otherwise identical double-auctions. With more than two sellers however, prices in both the double-auction and posted-offer institutions eventually converge to competitive levels. Alger (1986, 1987) presents some evidence that longer experiments, which allow participants more time to coordinate a collusive strategy, can lead to more anticompetitive outcomes in posted-offer markets. At present it is not clear how to interpret his results in relation to other studies, as he made numerous design changes.⁸²

As discussed in the previous subsection, the structural existence of market power in double-auction and posted-offer institutions sometimes leads to its exercise.⁸³ More often, it appears to increase the likelihood of tacit collusion. However, even when supracompetitive pricing occurs in the laboratory markets studied, the efficiency loss is generally small.⁸⁴ The posted-offer markets adjust more slowly than their double-auction counterparts to capture the gains from exchange. However, Smith (1982c) and Davis and Williams (1990b) find that the posted-offer institution mitigates market power and the ability to tacitly collude when the power is on the nonposting (*i.e.*, the buyers') side. Similarly, in *posted-bid* markets, in which buyers post bids and sellers accept or reject them, seller market power and tacit collusion is ineffective. Thus, the absence of a voice (*i.e.*, the ability to post either a bid or an offer) in the market appears to explain many of the behavioral differences observed in posted-offer and double-auction monopolies and oligopolies. Laboratory results suggest that the effect of the institution can dominate the effect of market structure, and thus, influence the likelihood of successful tacit collusion.

b) Opportunities for Explicit Collusion

Several laboratory studies research the impact of allowing open communication among sellers regarding price and other matters. Explicit communication among sellers about price, the "smoking gun" of antitrust, constitutes a *per se* violation of the Sherman Act in naturally-occurring markets.⁸⁵ The experiments with explicit opportunities to collude permit situations that would blatantly violate the antitrust laws in field markets. Thus, the laboratory markets provide benchmarks of behavior in a setting without the threat of antitrust enforcement.⁸⁶ Different types of permissible communication, such as communications akin to trade press announcements, can also be considered in laboratory markets.

When communication among sellers is permitted between trading periods in laboratory markets, participants quickly decide that price is an interesting topic to discuss.⁸⁷ In laboratory markets with open communication, buyers generally are (1) not informed of the sellers' deliberations, and (2) are physically isolated from the sellers. For

⁸² Some of the changes include the number of sellers, the information available to sellers (*e.g.* quantity sold by rivals), simulated buyer behavior, the number of periods, the length of periods, the instructions, and the cost and demand configurations (*e.g.* a Walrasian excess supply at the competitive equilibrium). Taking one of the changes for instance, as shown by Brown-Kruse (1986 and 1991), altering the buyer rationing process is not a trivial modification. Thus, one cannot attribute the result of making numerous changes simultaneously to any individual change.

⁸³ See Davis and Williams (1990b, 1991); Holt, Langan, and Villamil (1986); Davis, Holt, and Villamil (1990); and Holt (1989).

⁸⁴ See Table 3.

⁸⁵ See, for instance, *U.S. v. Socony-Vacuum Oil Co. et al.*, 310 U.S. 150 (1940), which remains the precedent of *per se* illegality of explicit horizontal price fixing.

⁸⁶ The threat of antitrust enforcement can be incorporated in laboratory tests, although it is not commonly implemented. See Gerety (1987).

⁸⁷ Often the participants are prohibited from revealing their individual unit costs and from making side-payments. See, for instance, Isaac, Ramey, and Williams (1984).

example, Isaac, Ramey, and Williams (1984) geographically separate the buyer and sellers. The sellers, located in Arizona, transact with the buyers, who are situated in Indiana, via a computer network.

Even when open communication is permitted in double-auction markets in the laboratory with at least two sellers, prices still converge to the competitive equilibrium. By contrast, in posted-offer markets with at least two sellers, prices tend toward the higher Cournot-Nash equilibrium.⁸⁸ Binger *et al.* find that communication facilitates collusion in laboratory Cournot markets, in which sellers set quantity. When communication is permitted in these Cournot experiments, markets with two or five sellers are statistically indistinguishable from the monopoly outcome, *i.e.*, the joint profit maximum is maintained throughout. Brown-Kruse, Cronshaw, and Schenk (1993) show that communication can lead to collusive outcomes in otherwise competitive spatial markets.

Holt and Davis (1990) observe an absence of stable collusive outcomes in posted-offer markets under communications (nonbinding price announcements) of the type traditionally considered in the grey area of antitrust law. This result holds even when market power exists. In other words, in the Holt and Davis experiments, when (1) sellers are provided an opportunity to increase prices unilaterally and profitably by strategically withholding quantity and (2) nonbinding price announcements are made, the sellers do *not* exercise market power.⁸⁹

Davis and Holt (1991) present four experiments showing that market power in posted-offer markets of three and five sellers leads to supracompetitive pricing. This result conflicts with the competitive results Holt and Davis (1990) find in similar markets of three sellers with nonbinding price communication (which would presumably only enhance the erosion of competition). The authors do not compare the results in their 1991 paper or the significant design change that seems to be driving the outcomes. The change has to do with the nature of aggregate production capacity. In the 1990 design, aggregate production capacity exceeded the competitive equilibrium level (as in Figure 5); while in the 1991 study, capacity is constrained at the competitive solution. Thus, it appears that the aggregate capacity constraint promotes the exercise of market power, when it exists. Further research is necessary to complete the analysis of the effect of capacity on unilateral market power, with and without nonbinding communication.

In markets not explicitly designed to investigate unilateral market power, Isaac and Reynolds (1990) find that the nature of aggregate capacity at the competitive equilibrium has a significant impact on market performance in posted-offer markets, even more so than a pure numbers effect of moving from four to two sellers. High excess aggregate capacity at the competitive equilibrium encourages more competitive behavior than the low excess aggregate capacity market designs. This supports the examples from the naturally-occurring economy presented in Scherer (1980, pages 209-212),⁹⁰ and contests the Davidson and Deneckere (1990) model of excess capacity as a device (a credible threat to punish defectors) to facilitate collusion.⁹¹

C. Mergers

Wellford (1990) examines horizontal mergers in laboratory Cournot markets. The study considers the ability of the critical levels of concentration demarcated in the *Horizontal Merger Guidelines* (1992) to screen for anticompetitive mergers in controlled markets.⁹² Production efficiencies resulting from merger and methods of measuring concentration are considered in the study.

⁸⁸ Note that while prices may be higher in posted-offer duopolies when communication is permitted, the prices do not reach the perfectly collusive (monopoly) level. See Isaac, Ramey, and Williams (1984), Isaac and Plott (1981), and Smith (1981a).

⁸⁹ Before each period, one seller (selected in an announced sequence) is permitted to communicate via a computer network by completing the statement "___ is an appropriate price for the market this period." Other sellers can respond 'A' if they agree, and 'L' ('H') if they think it should be lower (higher). The announcements and responses are displayed to each seller every period.

⁹⁰ See also the experiments in Sherman (1971); and Brown-Kruse *et al.* (1993), which also shows that when sellers have capacity constraints, market behavior is best described by the Bertrand-Edgeworth "cycling" model. However, while the Bertrand-Edgeworth model predicts better than the mixed-strategy Nash equilibrium, the joint monopoly equilibrium, and the competitive equilibrium, each has some predictive power. None of the models is entirely consistent with the laboratory data.

⁹¹ Moreover, the data from the field markets upon which the Isaac and Reynolds (1990) design is based, (see Isaac, Oaxaca, and Reynolds (1988)) are consistent with their laboratory results.

⁹² The critical levels of concentration did not change in the revision of the 1984 Guidelines.

Table 4 presents the basic experimental design. The Cournot markets studied have either five symmetric firms (each with 20% of industry capital -- see column I) or eleven firms (three large, three medium, and five small firms, each with 15%, 10%, and 5% of industry capital, respectively -- see column II) before the merger. The initial market supply and demand conditions are identical for both of the industry structures. In the five-firm industries described in column I, merger creates a new entity with 40% of the industry's capital and increases the Herfindahl-Hirshman Index (HHI) of market concentration from 2000 to 2800. Based on the HHI levels, the Guidelines classify these industries as "highly concentrated" and mergers that "potentially raise significant competitive concern" (§1.51(c)). In the eleven-firm industries characterized in column II, merger occurs between two small firms, increasing the HHI from 1100 to 1150. This corresponds to a merger that the Guidelines state would be unlikely to have adverse competitive consequences and ordinarily require no further analysis" (§1.51(b)). The performance of the laboratory markets with eleven firms is indeed competitive. Yet, some competitive improvement is apparent when the merged firm enjoys economies of scale.

In light of the laboratory findings relating to complete versus private information conditions, the experiments are conducted under conditions of private information.⁹³ As in other laboratory investigations of Cournot markets, each seller chooses the quantity to produce and knows the true market demand curve. The experimental design prohibits entry into the laboratory industries.

Figure 6 presents the series of average prices by period from five Cournot laboratory markets with five symmetric firms premerger and four postmerger. No efficiency gains result due to merger in this experimental treatment. All market participants have experience in similar Cournot markets (without merger), so that they are familiar with choosing production quantities and calculating their profits. Figure 7 provides the average price by period from markets that have the same underlying conditions as those shown in Figure 6, except that the merged firm in each of five market sessions enjoys economies of scale in production.

As Fouraker and Siegel found in their study of Cournot private information triopolies,⁹⁴ the data in both Figures 6 and 7 tend to lie between the static competitive and Cournot predictions.⁹⁵ No tacitly collusive outcomes are observed in *any* individual market session. The market quickly disciplines sellers who try to cut output. Those who reduce production in one period generally increase it in the next.

When the merged firm experiences no economies from merging, the mean postmerger price across all market sessions in this treatment is not significantly greater than the premerger price. The data suggests that the critical HHI levels established by the Guidelines may be overly strict, especially considering that entry is precluded during, and the threat of antitrust enforcement is absent from, the experiments.

Economies of scale due to merger have a statistically significant impact on industry performance in the markets studied; efficiencies make the markets more competitive. Figures 6 and 7 illustrate the downward pressure on mean price by period when the merged firm experiences increasing returns to scale in production. Moreover, a statistically significant increase in mean total surplus and mean consumer surplus is observed. The efficiency result may have implications for mergers in declining industries, as it is generally believed that mergers in such industries exhibit increasing returns to scale cost structures.

The Guidelines state that concentration may be measured by either capacity or sales. Capacity is a control variable set by the experimenter, sales are observable *ex post*. All of the industries studied exhibit significantly higher levels of concentration as measured by sales than when measured based on capacity or on capital. In fact, when measured by sales, the eleven-firm industry HHI increases from 1900 premerger to 2000 postmerger (compared to 1100 to 1150 based on capacity). This measurement discrepancy is enough to move the behaviorally competitive eleven-firm industries into the range of concentration that the Guidelines classify as raising "significant competitive concern." For the mergers studied, relying on the HHI as measured by sales instead of capacity leads to inappropriately increasing the number of cases to be challenged. The laboratory data indicate that basing the HHI on capacity and altering the policy demarcation line would improve measurement of anticompetitive effects.

⁹³ See §III B(1), *supra*.

⁹⁴ See §III B(1), *supra*.

⁹⁵ It should be noted that these markets structurally differ from the Cournot markets reported in the previous section. Symmetric firm size is not maintained in these merger experiments, but is standard in the other market studies.

D. The Illusive Giant: Predation

Isaac and Smith (1985) set out to find predatory pricing under conditions that theory suggests are amenable to predation, noting that (p. 320), "if such behavior is a human trait we ought to be able to observe it in the laboratory." Their design involves two firms: one large firm (firm A) with "deep pockets" and a cost advantage over the other firm, a small firm (firm B). Both firms are given an up front capital endowment to help offset possible future losses. Firm A receives twice the capital endowment of firm B. Figure 8 gives the demand and supply conditions, specifying seller costs and buyer values. The market is structured so that firm A can exclude B and still experience a positive cash flow. The design sufficiently separates the alternative equilibria, including joint monopoly, dominant firm, Edgeworth cycling, predation, and competition.⁹⁶ Predation is not observed in this design with or without sunk entry costs to discourage entry. Satisfying the theoretical assumption of complete information on competitors' costs, in addition to sunk entry costs again proves fruitless, as does attempting to induce rivalistic incentives.⁹⁷ The dominant firm equilibrium best describes the data from all the treatments.⁹⁸ After unsuccessful attempts at creating conditions suitable for observing predation, Isaac and Smith plant a confederate instructed to consistently price at predatory levels, to verify that firm B can indeed be forced to exit in their design. Exit occurs.

Isaac and Smith also study the effect of imposing various antitrust remedies for predation aimed at markets where predation is theoretically likely.⁹⁹ In their experimental design, application of a remedy corresponds to a situation of type two regulatory error, as no predation is observed in these markets. Market performance is more anticompetitive (*i.e.*, higher prices and lower market efficiency) when regulatory sanctions are in place, compared to identical markets without the supposed remedies.

Harrison (1988) runs one experiment using the Isaac and Smith parameters, but makes several design changes. The first relaxes the deep pocket advantage of the incumbent. This likely biases the design against observing predation. The more substantive changes involve several levels of prior subject experience in laboratory monopoly and contestable markets, as well as a multi-market feature. The opportunity cost of participating in one market becomes the foregone profits of not participating in the some other market, rather than a simple constant cost factor as in the Isaac and Smith design. Harrison reports evidence of predation. Although, Rutstrom (1985) initiates a careful attempts to reconcile the design changes between the two studies, she fails to replicate the Harrison result in the several treatments considered.¹⁰⁰ The search for predation continues.

Jung, Kagel, and Levin (1991) explore for predatory pricing in an abstract game theoretic setting. They examine the reputation arguments of Kreps and Wilson (1982) and Milgrom and Roberts (1982). The experimental session each use three monopolists and four entrants in a round of decisions. Each monopolist faces a series of entry threats from the four entrants. The entrants do not know *a priori* whether a monopolist is 'strong' (will punish the entrant by predating) or 'weak.' However, over the course of the round, information on a monopolist's reputation can be collected. Weak monopolists may strategically choose to fight the entrant, to build a reputation

⁹⁶ This is to say that from each seller's perspective his profit in one equilibrium is sufficiently different than his profit in another (*i.e.* dominance in payoffs is satisfied not only over the entire experiment, but also across relevant alternatives at a given point in time). For instance, if each seller's earnings only differ by a cent in one equilibrium versus another, it would be unlikely to observe a clean behavioral separation in market predictions when an experiment is replicated.

⁹⁷ Inducing rivalistic incentives raises methodological issues, as it move away from the issue of whether predatory behavior is innately a human trait.

⁹⁸ The dominant firm equilibrium (*dfe*) entails firm B following firm A's lead in setting price. For the Isaac and Smith design, both firms set the same price of \$2.84 in the *dfe* and firm A (firm B) earns \$1.99 (\$0.51) each period. Firm A earns \$3.42 per period at the uncontested monopoly equilibrium, from \$1.10 to \$1.80 in the competitive range, and no more than \$0.88 per period with a predation strategy.

⁹⁹ The "remedy" applied prohibits firm A from expanding production for beyond the maximum quantity ever set by the firm for two periods whenever firm B enters. Firm A also faces a semipermanent price reduction regulation, requiring that all its price reductions last for at least five periods.

¹⁰⁰ The Rutstrom paper is an unpublished manuscript which pre-dates the published version of the Harrison note. The paper does not test all of the design changes or the possible combinations of treatments implemented by Harrison. In some sense this is the case of cold fission in the economics laboratory. Until other researchers can replicate the data point, it remains shrouded in uncertainty.

of being a strong monopolist, in hopes that other potential entrants decide not to enter. Some predatory behavior is documented.

Unfortunately, it is altogether unclear how the abstract game theory results map into behavior in well defined market institutions. Further work building the bridge between analysis in such abstract settings and in well defined institutions is warranted.

Smith and Isaac make a point that is worth repeating. Current theories and experiments of predation fail to acknowledge the importance of capital. Assets of exiting firms will likely be sold, perhaps at a discount, to new or other potential entrants. This severely dilutes the benefits from a predatory strategy.

E. Entry: A Look at the Contestable Markets Hypothesis

Entry prevails as a key variable in antitrust analysis. The laboratory facilitates the study of entry. Most laboratory studies of entry involve tests of the contestable market hypothesis under various conditions in posted-offer markets.¹⁰¹ The basic structure of the markets examined in these studies includes two firms with identical cost structures that exhibit declining average cost, with each firm owning capacity sufficient to satisfy the entire market. Thus, the conditions for natural monopoly are satisfied. Both strong and weak forms of the contestable markets hypothesis are considered. The former conjectures that the threat of entry is sufficient to assure that competitive equilibrium outcomes are achieved, while the weaker version predicts that market outcomes will converge toward competitive levels, and lie closer to the competitive solution than to the monopoly prediction.

Coursey, Isaac, and Smith (1984) find that when sunk costs associated with entry are zero, economies of scale in and of themselves are an *ineffective* barrier to entry. In fact, the contestable market hypothesis provides a reasonable prediction of behavior. Four of their six markets quickly converge to the sustainable competitive equilibrium, and the other two gradually approach, but do not reach the competitive range. Over all, the Coursey, Isaac, and Smith data support the weak version of the contestable market hypothesis, yet several of the markets also appear to be consistent with the strong interpretation.

The study includes a *baseline* treatment of markets with a single seller who has the same cost structure as either of the two duopolists in the contestability experimental treatments.¹⁰² Even *without* the threat of entry in the baseline sessions, the unregulated monopolists have difficulty in achieving the monopoly prediction due to buyers strategically withholding demand. Holding back purchases in this design hits the seller at his most profitable units. However, the disciplining effect diminishes over time. Strategic withholding appears at the rate of roughly 9% in the Coursey, Isaac, and Smith monopoly markets, and 1% in the duopoly markets. Thus, the competitive results observed in the duopoly markets arise largely from the contesting discipline of sellers rather than the withholding behavior of buyers. Of course the mere existence of human rather than computerized buyers creates the a credible threat that withholding could occur. Brown-Kruse (1991) examines this using the Coursey, Isaac, and Smith (1984) design, and finds that markets with human, rather than robot, buyers converge faster.

When the sunk costs of entry are positive, Coursey, Isaac, Luke, and Smith (1984, p. 69) find that support for the strong version of the contestable markets hypothesis weakens, "yet the disciplining power of contestability remains impressive," as sustained monopoly pricing is not observed. Data from all twelve of the markets reported are consistent with the weak form of the contestable markets hypothesis. Although, in contrast to the zero entry cost treatment presented in Coursey, Isaac, and Smith (1984), the dynamics exhibit less of a monotone convergence path in seven of the twelve laboratory markets. No evidence of natural monopoly outcomes, market collapse, or sustainable tacit collusion surface.

Harrison and McKee take issue with (1) the use of human buyers in the Coursey *et al.* (1984) studies instead of computerized perfectly demand revealing buyers that are in closer line with what contestable market theory, as presented in Bamoul, Panzar, and Willig (1982), assumes and (2) the simultaneous nature of price choice when the theory involves a sequencing so that potential entrants can make their entry decision based on the current price set by the incumbent. With respect to the first criticism, Brown-Kruse (1991) points out, the fact that human buyers do not fully reveal their willingness to pay is less of a shortcoming of the experiments than in fact a

¹⁰¹ See also section D. on predation above.

¹⁰² A baseline set of market trials is a treatment against which all other treatments are compared. In the Coursey, Isaac, and Smith study, the benchmark against which contestability must be compared is the monopoly case. While the *theoretical* monopoly benchmark is known without conducting any laboratory markets, the *behavioral* benchmark is not.

weakness of the theory to adequately include strategic behavior on the demand side of the market. Taking both points (1) and (2) into account, the Coursey *et al.* experiments can be interpreted as boundary experiments of the contestable market hypothesis, and the theory is in fact fairly robust to the deviations in conditions reported.

When Harrison and McKee examine contestability under the behavioral demand conditions assumed in the theory, and with the Bertrand-Nash sequential choice condition in affect, they find support for the strong version of the contestability hypothesis. However, they fail to compare the two demand conditions *directly* (human versus computerized) or the Bertrand-Nash treatments in isolation, hence compounding the influence of the effects of the conjectures that they advance. In a later study, Harrison, McKee, and Rutstrom (1989) argue that simulation of buyers rather than using human buyers does not significantly affect the results. In fact, both Harrison *et al.* studies digress substantially from the Coursey, *et al.* design.¹⁰³ Brown-Kruse (1991) explicitly analyzes the demand conditions with data from experiments designed to address the buyer issue. Her data reveal that *prices are significantly lower when human buyers are present*. The sheer possibility that strategic withholding might occur sufficiently disciplines the market to assure competitive market performance. This supports the findings of Coursey, Isaac, Luke, and Smith (1984), which raised the possibility that buyers' strategic withholding behavior enhances the convergence to the competitive solution. Using the Coursey *et al.* market parameters, Brown-Kruse also tests the contestable markets hypothesis when each duopolist can opt to receive a certain rate of return in an alternate market instead of entering the contestable market. The data reveal no significant difference in adjusted price relative to markets in which the opportunity cost of entry is zero. The effect of sequencing remains to be isolated by holding all other treatment variables constant.

Millner, Pratt, and Reilly (1990) study contestability using a new institution, a real-time posted-offer flow market. In their computerized duopoly markets, sellers may change their price at any point in time, setting a price for a specified *flow* of output per second. The seller with the lowest price for the longest interval makes proportionately more sales the other seller, and costs depend on the rate of production per second. Sellers in this design have identical decreasing marginal cost schedules. Each firm can opt to accept a per second alternate rate of return instead of posting an offer in the flow market. The sunk costs of entry and exit are zero.

Buyers in the flow market are computer simulated to reveal demand fully, purchasing units of the laboratory commodity at their true willingness-to-pay for a given quantity per second. One experimental treatment allows only one seller, who is fully protected from potential entry, and provides a benchmark against which to measure the behavior observed in the theoretically contestable laboratory markets.

Under the real-time continuous setting, contestability ostensibly weakens. However, it is difficult to know to what degree the institution actually affects the outcome, for no direct comparison of institutions is made under the same industry parameters. Millner, Pratt, and Reilly find that their data in the theoretically contestable markets fail to converge to any stable equilibrium, including the sustainable Ramsey-efficient competitive equilibrium. Yet, the market performance is significantly more competitive than in the protected monopoly laboratory markets. Contestability increases welfare. Several of the markets exhibit behavior consistent with the unstable price hypothesis, as introduced in Coursey, Isaac, Luke, and Smith: Prices decline while two firms serve the market, until they fall sufficiently so that one firm exits, after which prices increase until entry occurs and the cycle repeats. In the Millner, Pratt, and Reilly experiments, for the periods in which units were traded, the market was contested by both sellers over 70% of the time. The impact of using human buyers in the flow market remains for later study.

All of the reported contestability experiments involve firms with decreasing marginal cost functions, a common condition of natural monopoly. While many antitrust markets that practitioners evaluate do not involve such specialized cost conditions, it is instructive to observe that the threat of entry indeed imposes competitive pressure on behavior, even in the most potentially problematic types of markets. The pressure increases when human buyers replace the perfectly demand-revealing robot buyers that theory assumes. Little laboratory work exists on entry under conditions of constant and increasing marginal cost, and remains an area for future study of interest to the antitrust community. However, preliminary results of (noncontestable) monopoly posted-offer markets under different cost scenarios and simulated buyers in Harrison, McKee, and Rutstrom (1990) suggest that monopolists exert monopoly power most effectively in constant cost industries, followed by decreasing cost and then

¹⁰³ Other treatment variables introduced include: different cost and buyer values, a new computerization of a posted-offer market, altered experience levels. The participants in Harrison and McKee are also prescreened for exhibiting risk neutral behavior using a variation of the method employed in Becker, DeGroot, and Marshak (1964).

increasing cost industries. The underlying behavioral difference may stem from the nature of equilibrium in the three price settings. This remains for later research, but may impact the manner in which theorists and antitrust practitioners think about monopoly, even when entry is perfectly suppressed.

F. Facilitating Practices: Elements of the Ethyl Case in the Laboratory

Grether and Plott (1984) designed experiments to parallel the industry structure and practices in the Ethyl case.¹⁰⁴ Both sides of the case agreed on the general structural form of the lead-based antiknock compound industry. Thus, the design of the laboratory market structure was relatively straight-forward. Features of the market included a homogeneous product, inelastic demand, no entry,¹⁰⁵ two large firms, two small firms, low transportation costs, and large buyers.

While the defense and complaint counsels agreed on industry structure, they argued about the conduct that resulted from the following seller practices: (1) most favored nation clauses (guarantees the buyer that no other customer will obtain a lower price for a similar quantity of the good); (2) advanced price announcements; and (3) delivered pricing (price is not contingent on location). Holding the economic parameters constant across experiments, the practices can be introduced, or omitted, so that the effect of the change can be studied. Delivered pricing is present in all of the Grether and Plott laboratory markets.

The FTC counsel contended that uncertainty about rivals firms' actions or reactions encourages competition, and therefore the most favored nation clauses and the advanced price announcements promote collusion. Grether and Plott find that without the practices, the observed prices lie closer to the competitive level than to any other prediction. Introducing both the most favored nation clauses and advanced price announcements leads to a decrease in competition, however the prices lie approximately halfway between the predictions of the competitive and Cournot models. The joint maximum, the perfectly collusive outcome, is rejected. Additionally, the most favored nation clause in isolation exerts a *negative* impact on price, the opposite result as argued. However, the price announcements do generate the hypothesized effect of increasing prices.

The conclusions of the Grether and Plott research are consistent with other experimental findings, several of which are discussed above. The application of laboratory analysis to the Ethyl case illustrates how experimental analysis can be useful for increasing our understanding of the prominent elements of complex antitrust situations.

IV. Concluding Remarks

The economics laboratory is an important tool which can enhance understanding of how markets work. The first three decades of laboratory research expands the general knowledge of theoretical industrial organization and applied field analysis relating to antitrust. From this literature, several significant themes suggest directions for policy or topics for further study:

- Institutional form matters more than theory and policy imply. In fact, institutional effects often dominate the effects of structural characteristics of the market. (For instance, the competitive outcome is more robust than economic theory suggests, markets with single sellers do not necessarily lead to monopoly outcomes in certain institutional settings.)
- Standard theories of noncooperative equilibria assume that all market participants have complete information on the payout functions of others, an assumption that is likely violated in most field markets. However, these theories actually predict better in laboratory markets under conditions of private rather than complete information.
- Communications of the form the antitrust laws prohibit sometimes lead to collusive outcomes. No support has been found to date that communications in the form of trade press announcements facilitate stable collusion in laboratory markets. Tacit collusion is difficult to coordinate in most

¹⁰⁴ *Ethyl Corp. v. F.T.C.*, 729 F.2d 128 (1984).

¹⁰⁵ Entry was unlikely, as the EPA was in the process of phasing out the use of lead in gasoline.

settings. When explicit collusion is permitted, the successful implementation of a cartel depends on the form of trading institution: (1) in double-auction markets, collusion usually falls apart; (2) in posted-offer markets, the sellers raise price to the Cournot-Nash level; and (3) in Cournot markets, the sellers sometimes achieve the perfectly collusive outcome.

- Predation is difficult to generate behaviorally in a market context, even under market conditions designed to give the theory its "best shot." Another lesson to be gleaned from research on this topic is that simple game-theoretic models may perform well when tested under the rigid assumptions they dictate however, when the models are imbedded in laboratory market institutions, the models often fail to describe behavior.
- Economies of scale resulting from merger in Cournot laboratory markets are passed on to consumers. The HHI critical values set in the Guidelines are overly restrictive for the markets studied, even when the merging parties do not enjoy economies.
- Market power, when it exists, may be exercised. The effect of institutional form, the nature of capacity, the divisibility of production, and the presence of human buyers are important in determining whether or not it will be. When market power exists and prices are supracompetitive, it is often the case that market power is not exercised fully (in the sense that sellers with power do not withhold production to the extent that theory predicts). Rather, the presence of market power seems to facilitate tacit collusion. Further study is warranted to better understand the relationship of the above variables on the effective exercise of market power, as well as on the incidence of tacit collusion.

Almost all of the results surveyed above involve industries with concentration levels sufficient to "raise significant anticompetitive concern" under the Guidelines (§1.51(b)), when using the Hirfindahl-Hirshman Index (HHI) of concentration to screen for potential horizontal merger cases. Most of the laboratory markets have HHIs of at least 2500 (*e.g.*, four symmetric firms) and many have values of at least 5000 (*e.g.*, duopoly or monopoly). In merger cases litigated by the Federal Trade Commission or the Department of Justice from 1982-1991, roughly 35% of the cases involved industries with HHIs of 2500 or less, 65% of 3333 or less, and 84% of 5000 or less.¹⁰⁶ Thus, the concentration of the markets studied in the laboratory is on average higher than the market concentration in litigated field cases.¹⁰⁷

It is worthwhile to note that most of the laboratory analysis summarized above omits two features that are critical for antitrust applications: (1) the possibility of entry¹⁰⁸ and (2) the threat of antitrust enforcement.¹⁰⁹ The absence of these factors should bias the data in favor of observing collusive outcomes.¹¹⁰ Thus, to the extent that collusion is not observed, or when it is, it is unstable, presents a strong result: Even *without* the threat of entry or antitrust penalty, many of the markets are still competitive. It is possible to incorporate both of these factors in

¹⁰⁶ See Coate (1992).

¹⁰⁷ This may not necessarily be representative of market concentration in cases that fold or settle prior to litigation.

¹⁰⁸ The most notable exceptions include the literature on predation and contestable natural monopolies in the laboratory, *see* §III (D) and (E) *supra*. In particular, little is known currently about entry in laboratory markets with constant or increasing marginal cost environments.

¹⁰⁹ Gerety (1987) incorporates the threat, *see* note 23 *supra*.

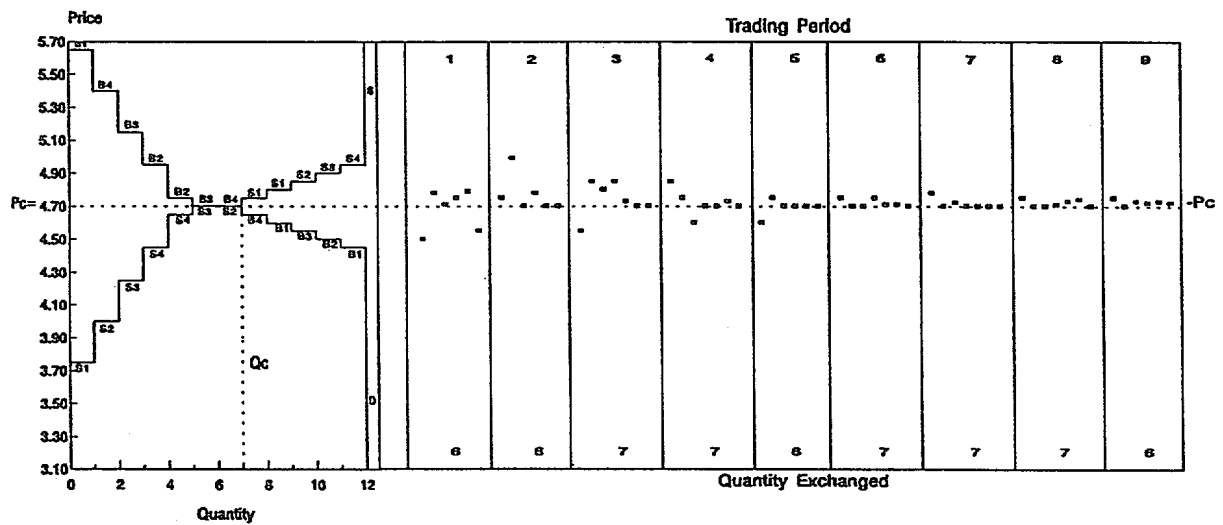
¹¹⁰ By and large, existing laboratory studies and antitrust analysis assume that the structural parameters (*e.g.* costs and demand) of an industry are constant over time. In this regard, both methodologies likely overestimate the ease of tacit collusion if structural shocks are relevant, for the shocks would likely make coordination more difficult. It is possible to test directly whether the incidence of supracompetitive pricing decreases when the parameters are not constant.

future laboratory studies.¹¹¹ Doing so would likely decrease the incidence of anticompetitive market outcomes, although, it remains to be tested.

Undoubtedly, future laboratory research will bring further progress and perhaps help to reach a consensus on how oligopoly markets work. As laboratory evidence accumulates, it will build an empirical foundation for our understanding of markets.

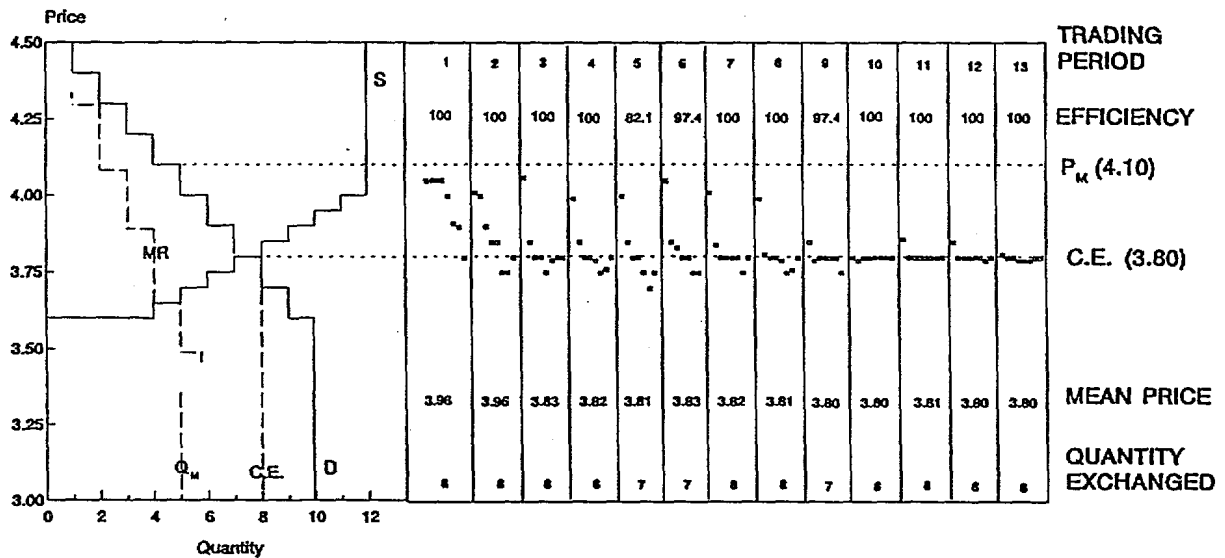
¹¹¹ Another feature that could be important is agency theory. Conn, Young, and Bishop (1991) consider principal-agent theory, but not imbedded in an oligopoly framework.

Figure 1*
Double Auction 2DA 24



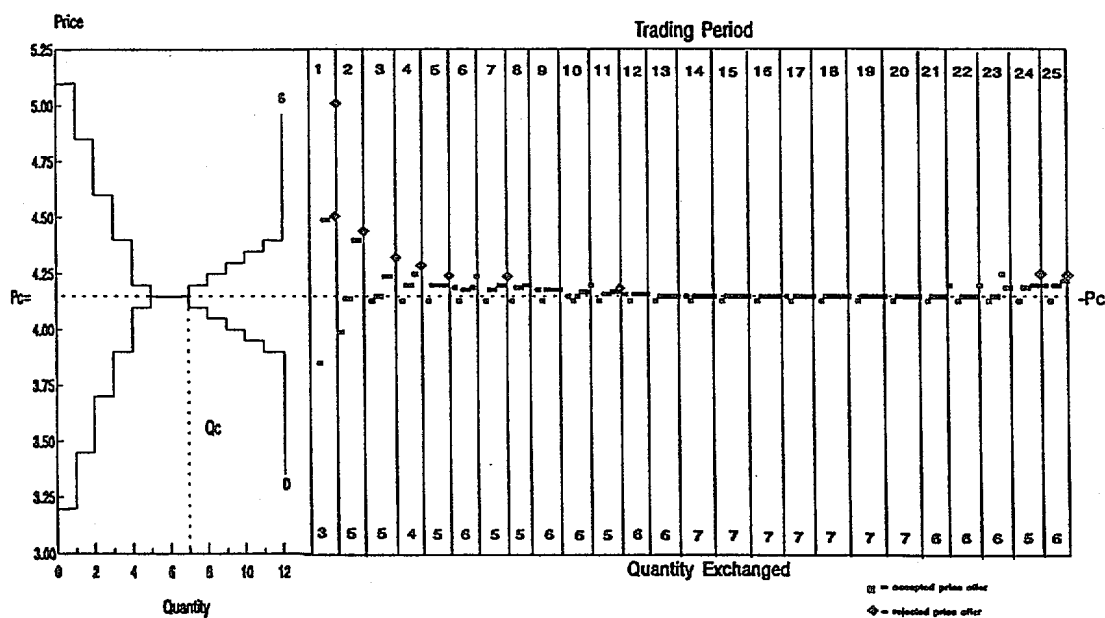
*Source: Ketcham, Smith, and Williams
(1984), page 603.

Figure 2*
Double Auction Monopoly M1x



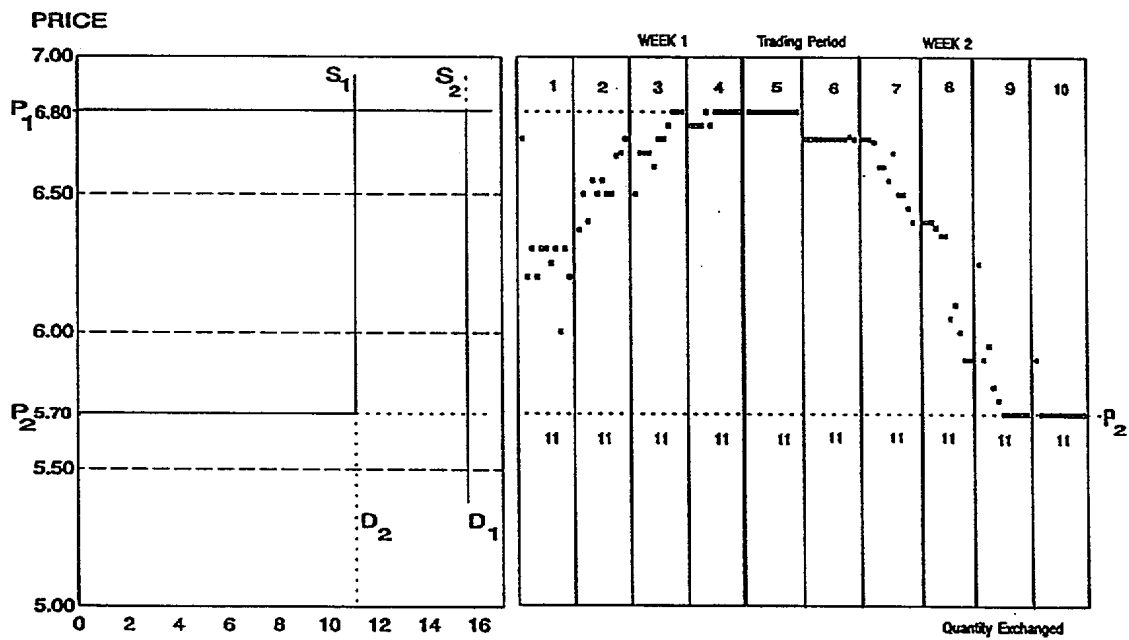
* Source: Smith and Williams (1989), page 38.

Figure 3*
Posted Offer PO17ixs



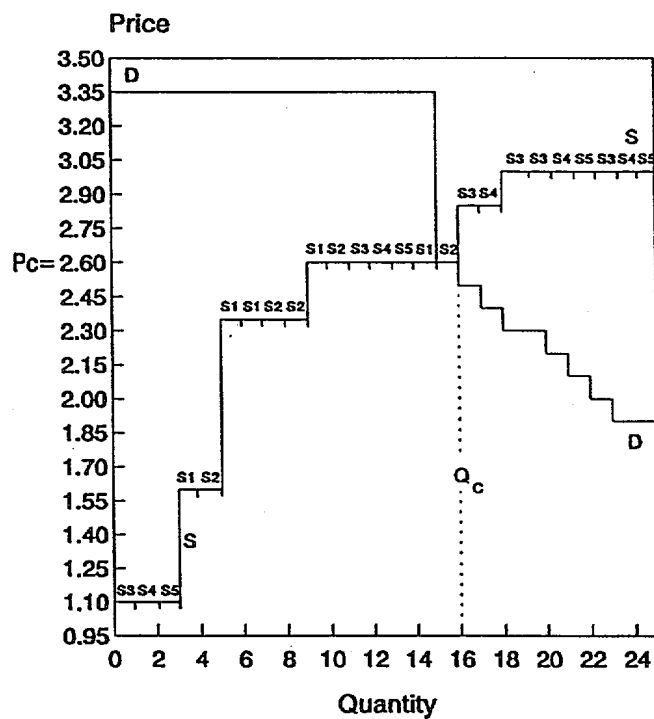
*Source: Ketcham, Smith, and Williams (1984), page 605.

Figure 4*
Double Auction Ilpda43



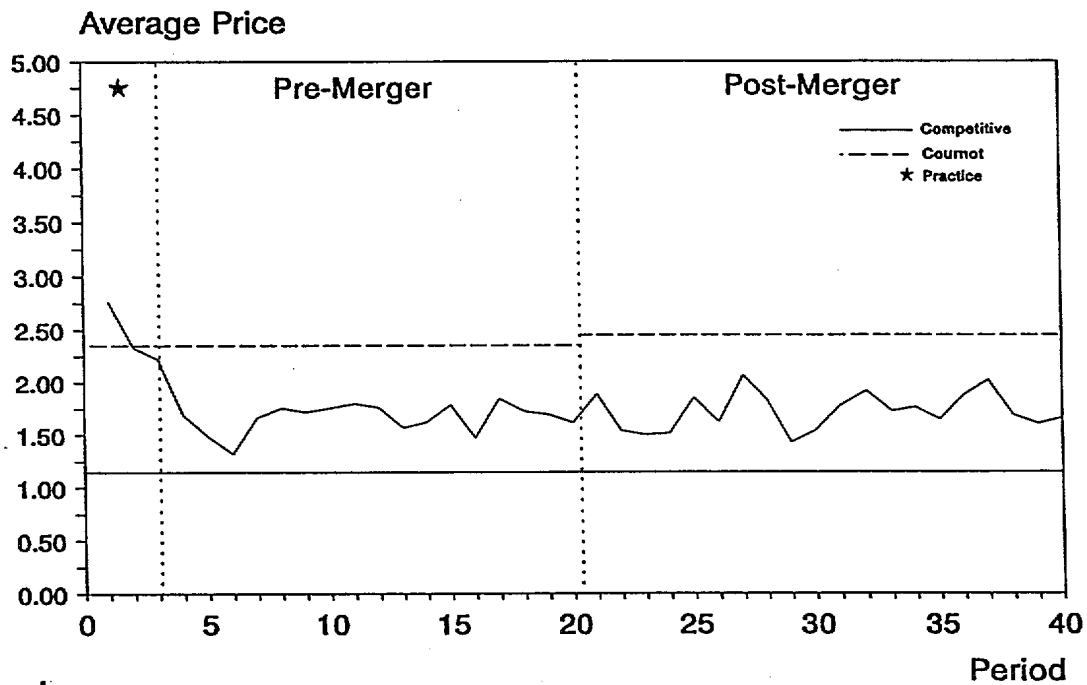
*Source: Smith (1982a)

Figure 5*
Market Power Design



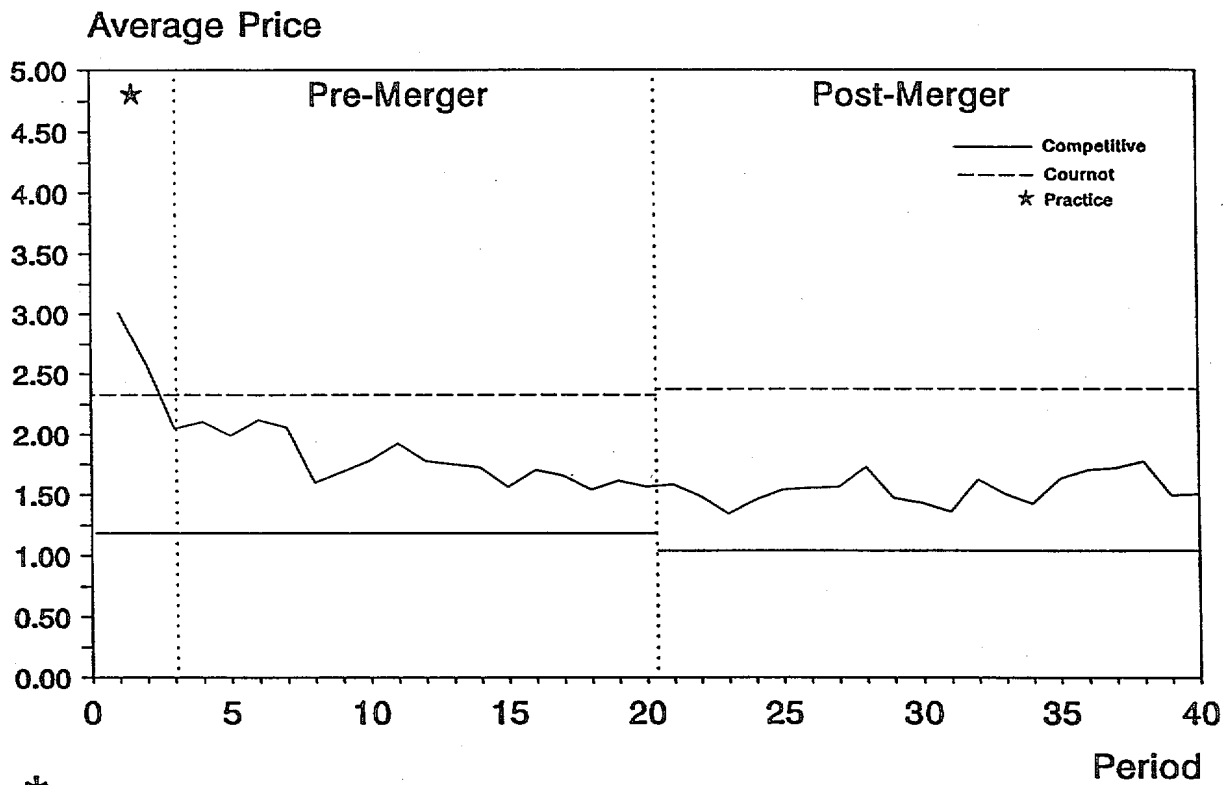
*Source: Holt, Langan, and Villamil (1986)

Figure 6*
Experiments 501 - 505
(No Economies)



*Source: Wellford, 1992.

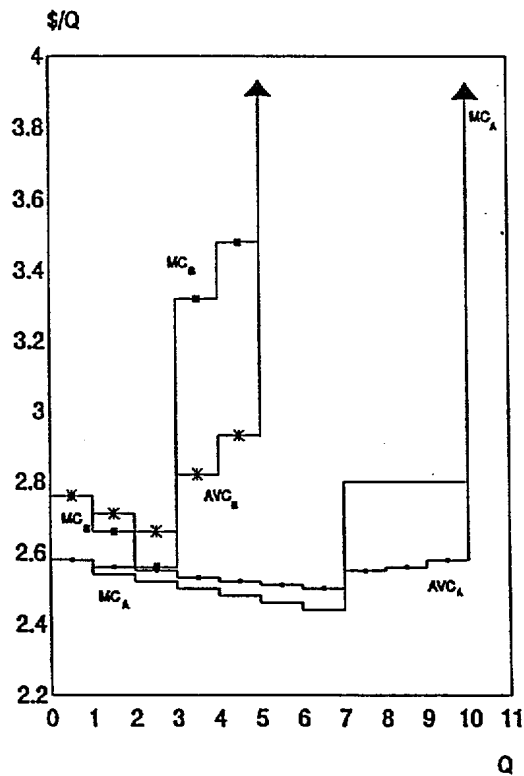
Figure 7*
Experiments 506 - 510
(With Economies)



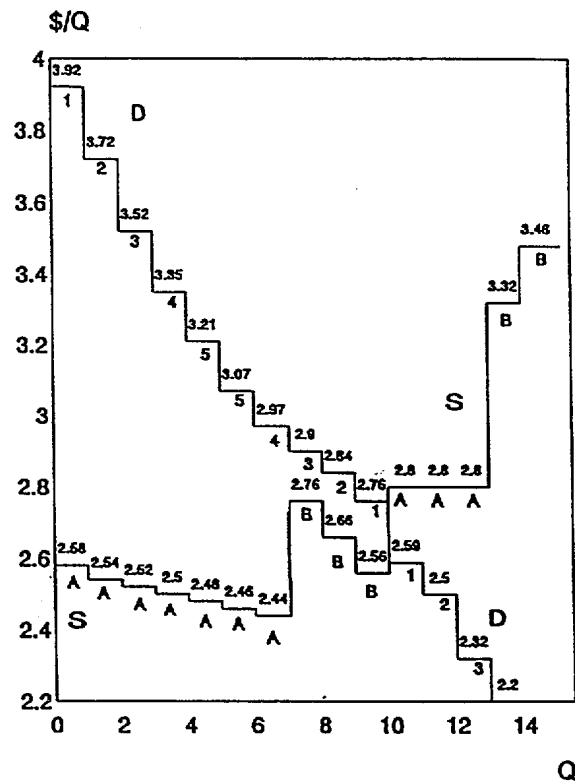
*Source: Wellford, 1992.

Figure 8^{*}

Cost and Demand Parameters Predatory Pricing Experiments



(a)



(b)

* Source: Isaac and Smith (1985),
page 325.

Table 1

Percentage of Period 21 Market Quantity Observations
Supporting a Particular Equilibrium Concept
Under Private Information Conditions
in Fouraker and Siegel's Cournot Markets*
(Complete Information Treatments in Parentheses)

Equilibrium Concept	Duopoly	Triopoly
Joint Max	0.00	0.00
	(31.25)	(0.00)
Cournot	87.50	81.82
	(46.87) [†]	(45.45)
CE	12.50	18.18
	(21.88) [†]	(54.55)

* Compiled from pages 131, 134, 140, and 141.

[†] One observation is equidistant between the Cournot and the Competitive Equilibrium (CE). Each of the two equilibria is given equal weight of 0.5 for this observation in the calculation of the percentages.

Table 2

Percentage of Period 14 Market Price Observations
Supporting a Particular Equilibrium Concept
Under Private Information Conditions
in Fouraker and Siegel's Bertrand Markets*
(Complete Information Treatments in Parentheses)

Equilibrium Concept	Duopoly	Triopoly
Joint Max	0.00	0.00
	(35.29)	(0.00)
Bertrand	100.00	100.00
	(64.71)	(100.00)

* Compiled from pages 174, 177, 181, and 183.

Table 3
Laboratory Studies of Market Power

Study	Institution	Excess Capacity at CE	Sellers	Buyers	Trials	Market Performance	Average Market Efficiency
Holt, Langan, and Villamil (1986)	DA	Yes	5 [†]	5	3, 2x	Two of the markets converge to the CE, the others converge to a price slightly above the CE price. On average, the deviations from the CE are observed with higher frequency when the sellers have previous experience in the PO institution.	100 % (in period 8)
Davis and Williams (1991)	DA	Yes	5 [†]	5	2, 2x	Supracompetitive prices are observed, but cannot be explained by the strategic withholding of quantity by sellers with theoretical market power.	98 %
Friedman and Ostroy (1991)	DA with Divisible Units	In Some Markets	4	4	5x	Strategic withholding is negligible. By strategically misrepresenting their true valuations, market agents unintentionally achieve the competitive outcome.	> 95 %
Davis, Holt, and Villamil (1990)	PO	Yes	either 2 or 3	*	2, 2x 2, 2x	Static market power leads to observed supracompetitive prices. Median price is less than (exceeds) the mixed strategy Nash equilibrium prediction in the duopoly (triopoly) markets. Some evidence of tacit collusion exists, but it is imperfect, with price lower than the shared monopoly outcome.	NR
Davis and Williams (1990b)	PO	Yes	5 [†]	5	4x ^{††}	Markets with <i>buyer</i> market power are unable to overcome the institutional asymmetry, prices converge to the competitive levels.	99 %
Holt and Davis (1990)	PO with Non-Binding Price Notices	Yes	3	*	2x	After early success at raising price above the pre-notice level (but not to the shared monopoly level), coordination breaks down and price converges to the pre-notice level.	NR
Davis and Williams (1991)	PO	Yes	5 [†]	5(*)	4, 4x	Prices exceed CE price in all of the market trials with market power and simulated buyers. However, it is not clear whether market power is exercised, for the firms that structurally possess market power do not withhold units from the market. When human buyers are introduced into the analysis, price converges to the CE.	91%* 92 %
Davis and Holt (1992)	PO	No	either 3 or 5	*	12x	Supracompetitive prices are sustained throughout all the market power trials. However, the elevated prices seem to be due to tacit collusion rather than the exercise of market power.	NR

CE Competitive Equilibrium; DA Double-Auction; PO Posted-Off; NR Not Reported; * Simulated buyers, and sellers have full information on the demand curve; x Trials in which participants have prior experience with the institution; † This uses the Holt, Langan, and Villamil (1986) design in which two sellers have substantially greater power than the other three sellers; and †† Four market power trials are summarized with the Davis and Williams (1991) experiments.

Table 4

Experimental Design

Cournot Homogeneous Product Mergers

	I	II
Premerger	<p>5 firms (each with 20% of industry capital)</p> <p>$HHI = 2000$</p>	<p>11 firms (3 with 15%, 3 with 10%, and 5 with 5% of industry capital)</p> <p>$HHI = 1100$</p>
Postmerger	<p>4 firms (1 with 40%, and 3 with 20% of industry capital)</p> <p>$HHI = 2800$</p>	<p>10 firms (3 with 15%, 4 with 10%, and 3 with 5% of industry capital)</p> <p>$HHI = 1150$</p>

Appendix

SAMPLE DOUBLE-AUCTION INSTRUCTIONS

General

This is an experiment in the economics of market decisionmaking. The instructions are simple and if you follow them carefully and make good decisions you can earn money which will be paid to you in cash at the end of the experiment.

In this experiment we are going to conduct a market in which some of you will be buyers and some of you will be sellers in a sequence of market days or trading periods. Attached to the instructions you will find a sheet labeled Buyer or Seller, which describes to you the value of any decisions you might make. **You are not to reveal this information to anyone.** It is your own private information.

Specific Instructions to Buyers

During each market period you are free to purchase units from any seller or sellers. The first unit that you buy *during a trading period* you will receive the amount listed under unit 1 marked redemption value; if you buy a second unit you will receive the additional amount marked unit 2 redemption value. The profits from each purchase are computed by taking the difference between the redemption value and purchase price of the unit bought. That is,

$$\text{your earnings} = (\text{redemption value}) - (\text{purchase price}).$$

Suppose, for example, that you buy two units and that your redemption value for the first unit is 200 and for the second unit is 180. If you pay 150 for your first unit and 160 for your second unit, your earnings are:

$$\text{earnings from first unit} = 200 - 150 = 50$$

$$\text{earnings from second unit} = 180 - 160 = 20$$

$$\text{total earnings} = 50 + 20 = 70.$$

The blanks on the table will help you record your profits. The purchase price of the first unit that you buy during the first period should be recorded on your sheet *at the time of purchase*. You should then record your profit for that unit. At the end of the period record your total of profits for the period in the space provided. Subsequent periods should be recorded similarly.

FFC/93 Papers & Proceedings

Specific Instructions to Sellers

During each market period you are free to sell units to any buyer or buyers. The first unit that you sell *during a trading period* you obtain at a cost of the amount listed under unit 1 marked cost; if you sell a second unit you incur the cost listed for unit 2. Your profits from each sale are computed by taking the difference between the sales price of the unit and its cost. That is,

$$\text{your earnings} = (\text{sales price of unit}) - (\text{cost of unit}).$$

Suppose, for example, that you sell two units and that your cost of the first unit is 140 and of the second unit is 160. If you sell the first unit at 200 and the second unit at 190, your earnings are:

$$\text{earnings from first unit} = 200 - 140 = 60$$

$$\text{earnings from second unit} = 190 - 160 = 30$$

$$\text{total earnings} = 60 + 30 = 90.$$

The blanks on the table will help you record your profits. The sales price of the first unit you sell during the first period should be recorded on your sheet *at the time of purchase*. You should then record your profit for that unit. At the end of the period record the total of profits in the space provided. Subsequent periods should be recorded similarly.

Market Organization

The market for units is organized as follows. The trading period is open for 5 minutes. Any buyer is free to bid and any seller is free to offer at any time that recognition is gained from the auctioneer. The bid (offer) is tendered by giving the sequence: name and bid price (offer price). For example, "buyer 7 bids 120" or "seller 6 offers 210." Recognized bids and offers are written on the board and will remain there until accepted, canceled or replaced by a higher (lower) bid (offer). Anyone is free to accept a standing bid (offer). For example, "seller 6 accepts buyer 7's bid of 120" or "buyer 7 accepts seller 6's offer of 210." Once a bid (offer) is accepted, a binding contract has been closed and both parties record the transaction. Completed transactions will be circled on the board.

Things to note

Always buy (sell) units in order of unit number.

You can only buy (sell) as many units as you have values (costs).

There is no carryover of untraded units across market periods.

FFC/93 Papers & Proceedings

SAMPLE SELLERS RECORD TABLE
SELLER # 2

PERIOD 1		UNIT 1	UNIT 2	UNIT 3	
	(A) Cost	4.00	4.70	4.75	
	(B) Sales Price				
	(C) Profit per unit [(B) - (A)]				
	(D) Profit this period				
	(E) Cumulative Profit				

PERIOD 2		UNIT 1	UNIT 2	UNIT 3	
	(A) Cost	4.00	4.70	4.85	
	(B) Sales Price				
	(C) Profit per unit [(B) - (A)]				
	(D) Profit this period				
	(E) Cumulative Profit				

PERIOD 3		UNIT 1	UNIT 2	UNIT 3	
	(A) Cost	4.00	4.70	4.85	
	(B) Sales Price				
	(C) Profit per unit [(B) - (A)]				
	(D) Profit this period				
	(E) Cumulative Profit				

PERIOD 4		UNIT 1	UNIT 2	UNIT 3	
	(A) Cost	4.00	4.70	4.85	
	(B) Sales Price				
	(C) Profit per unit [(B) - (A)]				
	(D) Profit this period				
	(E) Cumulative Profit				

References

- Alger, D. (1986), *Investigating Oligopolies Within the Laboratory*, FTC Bureau of Economics Report.
- Alger, D. (1987), "Laboratory Tests of Equilibrium Predictions with Disequilibrium Data," *Review of Economic Studies* 54, 105-145.
- Ball, S. and P. Cech (1991), "Subject Pool Selection," Kellogg Graduate School of Management Working Paper 91-42, Northwestern University.
- Bamoul, W. J., J. C. Panzar, and R. D. Willig (1982), *Contestable Markets and the Theory of Industry Structure*, San Diego: Harcourt-Brace-Jovanovich.
- Benson, B. L. and M. D. Faminow (June 1988), "The Impact of Experience on Prices and Profits in Experimental Duopoly Markets," *Journal of Economic Behavior and Organization*, 9, 345-365.
- Binger B. R., E. Hoffman, G. D. Libecap, and K. M. Schachat (1992), "An Experimetric Study of the Cournot Model," University of Arizona Working Paper 92-13.
- Bolton, G. E. (December 1991), "A Comparative Model of Bargaining: Theory and Evidence," *American Economic Review*, 81(5), 1096-1136.
- Bork, R. H. (1978), *The Antitrust Paradox*, New York: Free Press.
- Brown-Kruse, J. L. (March 1986), "Laboratory Tests of Buyer Rationing Rules in Bertrand-Edgeworth Duopolies," Revised Manuscript, June 1990, University of Colorado.
- Brown-Kruse, J. L. (Spring 1991), "Contestability in the Presence of an Alternative Market: An Experimental Examination," *RAND Journal of Economics*, 22(1), 136-147.
- Brown-Kruse, J. L., M. B. Cronshaw, and D. J. Schenk (January 1993), "Theory and Experiments on Spatial Competition," *Economic Inquiry*, 31(1), 139-165.
- Brown-Kruse, J., S. Rassenti, S. S. Reynolds, and V. L. Smith (1987), "Bertrand-Edgeworth Competition in Experimental Markets," University of Arizona Discussion Paper 87-12, forthcoming in *Econometrica*.
- Carlson, J. A. (January 1967), "The Stability of An Experimental Market With a Supply-Response Lag," *Southern Economic Journal*, 33(3), 305-321.
- Carlton, D. W. and J. M. Perloff (1989), *Modern Industrial Organization*, Glenview, IL: Scott, Foresman/Little, Brown Higher Education.
- Chamberlin, E. (1948), "An Experimental Imperfect Market," *Journal of Political Economy*, 56(2), 95-108.
- Camerer, C. F. (December 1987), "Do Biases in Probability Judgement Matter in Markets? Experimental Evidence," *American Economic Review*, 77(5), 981-997.
- Camerer, C. F., G. Lowenstein, and M. Weber (1989), "The Curse of Knowledge in Economic Settings: An Experimental Analysis," *Journal of Political Economy*, 97(5), 1232-1254.
- Coate, M. B. (1992), "Economics, the Guidelines and the Evolution of Merger Policy," *Antitrust Bulletin*, Winter, 997-1024.

Coate, M. B., A. N. Kleit, and R. Bustamante (1993), "Fight, Fold, or Settle?: Modeling the Reaction to FTC Merger Challenges," FTC Working Paper 200.

Conn, D., S. M. Young, and R. A. Bishop (1991), "Testing Principal-Agent Theory: A Framework for Analysis and Results From Three Experiments in a Non-Stochastic Setting," University of Southern California School of Business.

Coursey, D., R. M. Isaac, R. M. Luke and V. L. Smith (Spring 1984), "Market Contestability in the Presence of Sunk Costs," *RAND Journal of Economics*, 69-84.

Coursey, D., R. M. Isaac, and V. L. Smith (April 1984), "Natural Monopoly and Contested Markets," *Journal of Law and Economics*, 27, 91-113.

Cox, J. C., Dinkins, and V. L. Smith (1993), "Endogenous Entry and Exit in Common Value Auctions," University of Arizona Department of Economics Working Paper.

Cox, J. C. and R. M. Isaac (1986), "Experimental Economics and Experimental Psychology: Ever the Twain Shall Meet," *Economic Psychology: Intersections in Theory and Application*, A. J. and H. W. MacFadyen, eds., Elsevier Science Publishers B. V. (North-Holland).

Cox, J. C. and R. L. Oaxaca (1990), Using Laboratory Market Experiments to Evaluate Econometric Estimators of Structural Models," Manuscript, University of Arizona.

Cox, J. C., B. Roberson, and V. L. Smith (1982), "Theory and Behavior of Single Object Auctions," *Research in Experimental Economics, Volume 2*, Greenwich, CT: JAI Press, 1-43.

Davidson, C. and R. Deneckere (1990), "Excess Capacity and Collusion," *International Economic Review*, 31(3), 521-541.

Davis, D. D., G. W. Harrison, and A. W. Williams (1993), "Convergence to Nonstationary Competitive Equilibria: An Experimental Analysis," *Journal of Economic Behavior and Organization*, forthcoming.

Davis, D. D. and C. A. Holt (1991), "Capacity Asymmetries, Market Power, and Mergers in Laboratory Markets with Posted Prices," Manuscript, Virginia Commonwealth University.

Davis, D. D., C. A. Holt, and A. P. Villamil (1990), "Supra-Competitive Prices and Market Power in Posted-Offer Experiments," BEBR Working Paper 90-1648, University of Illinois Urbana-Champaign.

Davis, D. D. and A. Williams (1985), "The Effects of Rent Asymmetries in Posted-offer Markets," *Journal of Economic Behavior and Organization*, 7, 303-316.

Davis, D. D. and A. Williams (April 1990a), "Convergence to Nonstationary Competitive Equilibria: An Experimental Analysis," Manuscript, Indiana University.

Davis, D. D. and A. Williams (1990b), "Market Power and the Institutional Asymmetry of the Posted-offer Trading Institution," *Economic Letters*, 34(3), 211-214.

Davis, D. D. and A. Williams (1991), "The Hayek Hypothesis in Experimental Auctions: Institutional Effects and Market Power," *Economic Inquiry*, 29, 261-274.

Demsetz, H. (1973), *The Market Concentration Doctrine*, Washington, D.C.: American Enterprise Institute.

Deneckere, R. and D. Kovenock (1989) "Capacity-Constrained Price Competition when Unit Costs Differ," CMSEMS Discussion Paper 861, Northwestern University.

FFC/93 Papers & Proceedings

Dolbear, F. T., L. B. Lave, and G. Bowman (1966), "Collusion in Oligopoly: An Experiment of Numbers and Information," *Quarterly Journal of Economics*, 82(2), 240-259.

Durham, Y., S. Rassenti, V. L. Smith, M. Van Boening, and N. T. Wilcox (1993), "Can Core Allocations Be Achieved in Avoidable Fixed Cost Environments Using Two-Part Price Competition?" presented at the Western Economic Association Meetings, Lake Tahoe, NV, June 22.

Dyer, D. and J. H. Kagel (1992), "Experienced Bidders in Common Value Auctions: Behavior in the Laboratory versus the Natural Habitat," Manuscript, Department of Economics, University of Pittsburgh, presented in the FTC Bureau of Economics Seminar Series, Spring 1992.

Easley, D. and J. Ledyard (1988), "Theories of Price Formation and Exchange in Oral Auctions," Manuscript, California Institute of Technology.

Fama, E. F., and A. B. Laffer (1972), "The Number of Firms and Competition," *American Economic Review*, 62, 670-674.

Fouraker, L. and S. Siegel (1963), *Bargaining Behavior*, New York, McGraw-Hill.

Friedman, D. (March 1984), "On the Efficiency of Double-Auction Markets," *American Economic Review*, 74(1), 60-72.

Friedman, D. (1991), "A Simple Testable Model of Double-Auction Markets," *Journal of Economic Behavior and Organization*, 15, 47-70.

Friedman, D. and J. Ostroy (September 1991), "Competitiveness in Auction Markets: An Experimental and Theoretical Investigation," Manuscript, University of California, Santa Cruz.

Friedman, J. W. (1967), "An Experimental Study of Cooperative Duopoly," *Econometrica*, 379-397.

Friedman, J. W. and A. C. Hoggatt (1980), "An Experiment in Noncooperative Oligopoly," *Research in Experimental Economics, Supplement 1*, Greenwich, CT: JAI Press.

Garvin, S. and J. Kagel (December 1990), "An Experimental Investigation of Limit Entry Pricing," Manuscript, University of Pittsburgh.

Gerety, V. E. (1987), *Formation and Control of Collusion in a Sealed Offer Market: An Experimental Examination*, University of Arizona Ph.D. Dissertation.

Grether, D. (1980), "Bayes Rule as a Descriptive Model: The Representativeness Heuristic," *Quarterly Journal of Economics*, 95, 537-557.

Grether, D. (1992), "Testing Bayes Rule and the Representativeness Heuristic: Some Experimental Evidence," *Journal of Behavior and Organization*, 17(1), 31-57.

Grether, D. and C. Plott (October 1984), "The Effects of Market Practices in Oligopolistic Markets: An Experimental Examination of the Ethyl Case," *Economic Inquiry*, 22, 479-507.

Harrison, G. (1987), "Experimental Evaluation of the Contestable Markets Hypothesis," *Public Regulation: New Perspectives on Institutions and Policies*, E. E. Bailey, ed., Cambridge: MIT Press, 191-225.

Harrison, G. (1988), "Predatory Pricing in a Multiple Market Experiment," *Journal of Economic Behavior and Organization*, 9, 405-417.

FFC/93 Papers & Proceedings

- Harrison G. and M. McKee (Spring 1985), "Monopoly Behavior, Decentralized Regulation, and Contestable Markets: An Experimental Evaluation," *RAND Journal of Economics*, 16(1), 51-69.
- Harrison, G., M. McKee, and E. E. Rutstrom (1990), "Experimental Evaluation of Institutions of Monopoly Restraint," *Advances in Behavioral Economics, Volume II*, Green and Kagel, eds., Norwood, N.J.: Ablex, 54-94.
- Hayek, F. A. (September 1945), "The Use of Knowledge in Society," *American Economic Review*, 35(4), 519-530.
- Hoffman, E. and C. R. Plott (1981), "The Effect of Intertemporal Speculation on the Outcomes in Seller Posted Offer Auction Markets," *Quarterly Journal of Economics*, 96, 223-241.
- Hoffman, E. and M. Spitzer (April 1982), "The Coase Theorem: Some Experimental Tests," *Journal of Law and Economics*, 25, 73-98.
- Hoggatt, A. C. (July 1959), "An Experimental Business Game," *Behavioral Science*, 4, 192-203.
- Hoggatt, A. C., J. W. Friedman, and S. Gill (1967), "Price Signaling in Experimental Oligopoly," *American Economic Review*, 66(2), 261-266.
- Holt, C. A. (June 1985), "An Experimental Test of the Consistent-Conjectures," *American Economic Review*, 75, 314-325.
- Holt, C. A. (October 1989), "The Exercise of Market Power in Laboratory Experiments," *Journal of Law and Economics*, 32, S107-S130.
- Holt, C. A. (1991), "Industrial Organization: A Survey of Laboratory Research," *Handbook of Experimental Economics*, Princeton University Press.
- Holt, C. A. and D. D. Davis (December 1990), "The Effects of Non-Binding Price Announcements on Posted-Offer Markets," *Economic Letters*, 34, 307-310.
- Holt, C. A., L. W. Langan and A. P. Villamil (1986), "Market Power in Oral Double-auctions," *Economic Inquiry*, 24, 107-123.
- Hong, J. and Plott, C. (1982), "Rate Filing Policies for Inland Water Transportation," *Bell Journal of Economics*, 13(1), 1-19.
- Hurwicz, L. (1972), "The Design of Mechanisms for Resource Allocation," *American Economic Review, Papers and Proceedings*, 63, 1-30.
- Isaac, R. M. (July 1983), "Laboratory Experimental Economics as a Tool in Public Policy Analysis," *The Social Science Journal*, 20(3), 45-58.
- Isaac, R. M. (1990), "Laboratory Studies of Monopoly and Conspiracy," Manuscript, University of Arizona.
- Isaac, R. M. and C. Plott (1981), "The Opportunity for Conspiracy in Restraint of Trade: An Experimental Study," *Journal of Economic Behavior and Organization*, 2, 1-30.
- Isaac, R. M., R. L. Oaxaca, and S. S. Reynolds (1988), "Competition and Pricing in the Arizona Gasoline Market," University of Arizona Discussion Paper 88-16.
- Isaac, R. M., V. Ramey, and A. Williams (1984), "The Effects of Market Organization on Conspiracies in Restraint of Trade," *Journal of Economic Behavior and Organization*, 5, 191-222.

- Isaac, R. M. and S. S. Reynolds (1989), "Appropriability and Market Structure in a Stochastic Invention Model," *Quarterly Journal of Economics*, 103, 647-672.
- Isaac, R. M. and S. S. Reynolds (1990), "Two or Four Firms: Does it Matter?" Manuscript, University of Arizona.
- Isaac, R. M. and V. L. Smith (1985), "In Search of Predatory Pricing," *Journal of Political Economy*, 93, 320-345.
- Isaac, R. M. and J. Walker (1985), "Information and Conspiracy in Sealed Bid Auctions," *Journal of Economic Behavior and Organization*, 6, 139-159.
- Jung, Y. J., J. H. Kagel, and D. Levin (1991), "On the Existence of Predatory Pricing: An Experimental Study of Reputation and Entry Deterrence in the Chain-Store Game," Manuscript, University of Pittsburgh.
- Kagel, J. and A. Roth, eds. (1991), *Handbook of Experimental Economics*, Princeton University Press.
- Ketcham, J., V. L. Smith, and A. Williams (1984), "A Comparison of Posted-Offer and Double-Auction Pricing Institutions," *Review of Economic Studies*, 51, 595-614.
- Kirkwood, J. (1981), "Antitrust Implications of Recent Experimental Literature on Collusion," in *Strategy, Predation, and Antitrust Analysis*, S. Salop, ed., FTC Bureau of Economics Report.
- Kreps, D. and J. Scheinkman (1983), "Quantity Precommitment and Bertrand Competition Yield Cournot Outcomes," *Bell Journal of Economics*, 14, 326-337.
- Kreps, D. and R. Wilson (1982), "Reputation and Imperfect Information," *Journal of Economic Theory*, 27, 253-279.
- Ketcham, J., V. L. Smith, and A. Williams (1984), "A Comparison of Posted-Offer and Double-auction Pricing Institutions," *Review of Economic Studies*, 51(4), 595-614.
- King, R., V. L. Smith, A. Williams, and M. Van Boening (forthcoming), "The Robustness of Bubbles and Crashes in Experimental Stock Markets," *Nonlinear Dynamics and Evolutionary Economics*, I. Prigogine, R. H. Day, and P. Chen, eds., Oxford: Oxford University Press.
- Kitch, E. W., ed. (1983), "The Fire of Truth: A Remembrance of Law and Economics at Chicago, 1932-1970," *Journal of Law and Economics* 26, 163-234.
- Levin, D. (April 1988), "Stackelberg, Cournot, and Collusive Monopoly: Performance and Welfare Comparison," *Economic Inquiry*, 26, 317-330.
- McCabe, K., S. Rassenti, and V. L. Smith (1989), "Lakatos and Experimental Economics," University of Arizona Discussion Paper 89-24.
- McGee, J. S. (1971), *In Defense of Industrial Concentration*, New York: Praeger Publishers Incorporated.
- Mestelman, S., D. Welland, and D. J. Welland (June 1987), "Advance Production in Posted-Offer Markets," *Review of Economic Studies*, 55(4), 641-654.
- Mestelman, S. and D. J. Welland (October 1988), "Advance Production in Experimental Markets," *Journal of Economic Behavior and Organization*, 8, 249-264.
- Milgrom, P. and J. Roberts (1982), "Predation, Reputation, and Entry Deterrence," *Journal of Economic Theory*, 27, 280-312.

FFC/93 Papers & Proceedings

- Miller, R., C. Plott, and V. L. Smith (November 1977), "Intertemporal Competitive Equilibrium: An Empirical Study of Speculation," *Quarterly Journal of Economics*, 91, 599-624.
- Morrison, C. C. and H. Kamarei (1990), "Some Experimental Testing of the Cournot-Nash Hypothesis in Small Group Rivalry Situations," *Journal of Economic Behavior and Organization*, 13(2), 213-231.
- Plott, C. (1982), "Industrial Organization Theory and Experimental Economics," *Journal of Economic Literature*, 20(4), 1485-1527.
- Plott, C. (1987) "Some Policy Applications of Experimental Methods," *Laboratory Experimentation in Economics*, A. E. Roth, ed. Cambridge: Cambridge University Press, 193-219.
- Plott, C. (1989), "An Updated Review of Industrial Organization: Applications of Experimental Methods," *Handbook of Industrial Organization, Volume II*, eds. R. Schmalensee and R. D. Willig, Elsevier Science Publishers.
- Plott, C. and V. L. Smith (1978), "Experimental Examination of Two Exchange Institutions," *Review of Economic Studies*, 45(1), 133-153.
- Posner, R. A. (1976), *Antitrust Law*, Chicago: The University of Chicago Press.
- Rassenti, S. J., S. S. Reynolds, and V. L. Smith (October 1989), "Cotenancy and Competition in an Experimental Auction Market for Natural Gas Pipeline Networks," Manuscript, University of Arizona.
- Rutstrom, E. E. (December 1985), "In Search of a Reconciliation of Results in Predatory Pricing Experiments," Manuscript, University of South Carolina.
- Sargent, T. J. (1992) "Milton, Money, and Mischief," presented at the Western Economic Association Meetings, San Francisco, CA, Forthcoming, *Economic Inquiry*.
- Scherer, F. M. (1980), *Industrial Market Structure and Economic Performance*, Second Edition, Boston: Houghton Mifflin Company.
- Schmalensee, R. (January 1976), "An Experimental Study of Expectation Formation," *Econometrica*, 44, 17-41.
- Sherman, R. (April 1976), "An Experiment on the Persistence of Price Collusion," *Southern Economic Journal*, 26, 489-495.
- Smith, A. (1937), originally 1776, *The Wealth of Nations*, New York: Random House, Modern Library Edition.
- Smith, V. L. (1962), "An Experimental Study of Competitive Market Behavior," *Journal of Political Economy*, 70(2), 111-137.
- Smith, V. L. (1964), Effect of Market Organization on Competitive Equilibrium," *Quarterly Journal of Economics*, 78(2), 181-201.
- Smith, V. L. (1967), Experimental Studies of Discrimination Versus Competition in Sealed-Bid Auction Markets," *Journal of Business*, 40, 56-84.
- Smith, V. L. (1976a), "Bidding and Auctioning Institutions: Experimental Results," *Bidding and Auctioning for Procurement and Allocation*, Y. Amihaud, ed., New York: New York University Press, 43-64.
- Smith, V. L. (1976b), "Induced Value Theory," *American Economic Review*, 66(2), 274-279.

- Smith, V. L. (1981a), "An Empirical Study of Decentralized Institutions of Monopoly Restraint," *Essays in Contemporary Fields of Economics in Honor of Emanuel T. Weiler (1914-1979)*, G. Horwich and J. Quirk, eds., Purdue University Press, 83-106.
- Smith, V. L. (1981b), "Theory, Experiment, and Antitrust Policy," *Strategy, Predation, and Antitrust Analysis*, ed. S. Salop, FTC Bureau of Economics Report.
- Smith, V. L. (April 1982a), "Markets as Economizers of Information: Experimental Examination of the 'Hayek Hypothesis'," *Economic Inquiry*, 20, 165-179.
- Smith, V. L. (December 1982b), "Microeconomic Systems as an Experimental Science," *American Economic Review*, 72, 923-955.
- Smith, V. L. (1982c), "Reflections on Some Experimental Market Mechanisms for Classical Environments," *Choice Models for Buyer Behavior: Research in Marketing*, Supplement 1, Greenwich, CT: JAI Press, 13-47.
- Smith, V. L. (1987), "Experimental Methods in Economics," *The New Palgrave*, volume 1, J. Eatwell, M. Milgate, and P. Newman, eds., New York, NY: W. W. Norton and Company, 94-111.
- Smith, V. L. (1988/1989), "New Directions for Economics," *Journal of Business Administration, Special Issue: Future Directions for Economics*, 18(1,2), 41-52 (an earlier version was presented on February 28, 1986 at the first official meeting of the Economic Science Association, Tucson, Arizona).
- Smith, V. L. (Winter 1989), "Theory, Experiment and Economics," *Journal of Economic Perspectives*, 3(1), 151-169.
- Smith, V. L. (1990), "Experimental Economics: Behavioral Lessons for Microeconomic Theory and Policy," Nancy L. Schwartz Memorial Lecture, Kellogg Graduate School of Management, Northwestern University.
- Smith, V. L. (1991a), *Papers in Experimental Economics*, Cambridge: Cambridge University Press.
- Smith, V. L. (1991b), "Game Theory and Experimental Economics: Beginning and Early Influences," University of Arizona Working Paper 91-5.
- Smith, V. L. (1991c), "Rational Choice: The Contrast between Economics and Psychology," *Journal of Political Economy*, 99, 877-897.
- Smith, V. L. (July 1992), "Experimental Economics: Evaluating Microeconomic Policies and Institutional Designs," Manuscript, presented in the Economic Analysis Group Seminar Series, Department of Justice Antitrust Division.
- Smith, V. L. (November 1993), "Economics in the Laboratory," *Journal of Economic Perspectives*, forthcoming.
- Smith, V. L. and J. Walker (April 1993), "Monetary Rewards and Decision Cost in Experimental Economics," *Economic Inquiry*, 31, 245-261.
- Smith, V. L. and A. Williams (June 1981), "On Nonbinding Price Controls in a Competitive Market," *American Economic Review*, 71(3), 467-474.
- Smith, V. L. and A. Williams (September 1982), "The Effects of Rent Asymmetries in Experimental Auction Markets," *Journal of Economic Behavior and Organization*.
- Smith, V. L. and A. Williams (September 1989), "The Boundaries of Competitive Price Theory: Convergence, Expectations, and Transaction Costs," *Advances in Behavioral Economics, Volume I*, L. Green and J. Kagel, eds., Norwood, N.J.: Ablex, 31-53.

- Smith, V. L., A. Williams, K. Bratton, and M. Vannoni (1982), "Competitive Market Institutions: Double-Auctions versus Sealed Bid-Offer Auctions," *American Economic Review*, 72(1), 58-77.
- U.S. Department of Justice (1984), "Merger Guidelines," *Federal Register*, 49, 26284.
- U.S. Department of Justice and Federal Trade Commission (April 2, 1992), "Horizontal Merger Guidelines," *Antitrust Trade and Regulation Report*, 1559.
- Van Boening, M. V. and N. Wilcox (1992), "A Fixed Cost Exception to the Hayek Hypothesis," presented at the Economic Science Association Meetings, October 24.
- Vickrey, W. (1961), "Counterspeculation, Auctions, and Competitive Sealed Tenders," *Journal of Finance*, 16(1), 3-37.
- Wellford, C. P. (1989), "Price Dynamics and Expectations in the Cobweb Model: An Experimental Analysis," University of Arizona Discussion Paper 89-15, Revised Manuscript, 1991, Federal Trade Commission.
- Wellford, C. P. (1990), "Horizontal Mergers: Concentration and Performance," Chapter 2 of University of Arizona Ph.D. Dissertation.
- Wilcox, N. T. (1989), "Well-Defined Loss Metrics and the Situations that Demand Them," presented at the Economic Science Association Meetings, Tucson, AZ, October 28-29.
- Wilcox, N. T. (1992), "Incentives, Complexity, and Time Allocation in a Decision-Making Environment," presented at the Public Choice/Economic Science Association Meetings, New Orleans, LA, March 27-29.
- Williams, F. (1973), "The Effect of Market Organization on Competitive Equilibrium: The Multiunit Case," *Review of Economic Studies*, 40(1), 97-113.
- Williams, A. (1979), "Intertemporal Competitive Equilibrium: On Further Experimental Results," *Research in Experimental Economics*, vol. 1, V. L. Smith, ed., Greenwich, CT: JAI Press.
- Wilson, R. (1987), "On Equilibria of Bid-Ask Markets," *Arrow and the Ascent of Economic Theory: Essays in honor of Kenneth J. Arrow*, G. Feiwel, ed., London: McMillan Press, 375-414.

The Marketable Emissions Allowance

Steven R. Elliott, Oak Ridge National Laboratory

Abstract

This paper examines the results of experiments designed to study various aspects of tradable emissions permits as set forth by the Clean Air Act Amendments of 1991. In particular, this paper develops an understanding of some of the results that we attained in our experimental work in this area at the Universities of Arizona and Colorado. In these experiments subjects acted as managers of firms that must trade emissions permits. After an initial allocation of permits, subjects were allowed to trade them in both an "informal" double auction and a "formal" revenue neutral auction or in some sessions just the "formal" auction. Further, some sessions allowed subjects to hold permits across rounds for later use (intertemporal banking of permits), while others did not.

The results show, when faced with only one arena in which to trade permits, subjects are able to lock on the efficient trading price and come close to the efficient intertemporal allocation of permits. However, when we introduce the double auction mechanism to a setting with banking, results change dramatically. Now there are arbitrage possibilities across markets and permits become an asset. Price becomes very unstable across markets and periods, and it may show signs of bubbling and crashing as other assets have done in experimental markets. Further research in this area is outlined to understand more of the dynamics of these influences on the efficient solution predicted by theory.

I. Introduction

The 1991 clean Air Act Amendments (CAAA) created a system of tradeable emissions permits for the reduction of SO₂. The participants in this new system are large electricity generation facilities in the U.S.. By initiating such a system, proponents argue that reductions in emissions can be achieved at the least cost to society as a whole. The idea is to introduce market based solutions to what had traditionally been relegated to systems of command-and-control. This is the first test on such a scale of this type of incentive-based regulatory system. Its success or failure will have a profound influence on the adaptation of such systems in the future.

The trading system that has been mandated by the U.S. Environmental Protection Agency (USEPA) has three separate opportunities for permit transactions (see Federal Register 1991a and 1991b). The first mechanism is a formal auction process using a revenue-neutral auction mechanism. This system allows firms to submit both bids to buy and offers to sell permits. The bids are arrayed from highest price to lowest to form a demand curve for permits while the offers are arrayed from lowest price to highest to form a supply curve. In standard fashion, all those whose bids and offers are to the left of the intersection of the supply and demand curves trade. Those to the right are shut out of that particular trading opportunity. The auction is discriminative price- and revenue-neutral so that each purchaser of a permit pays exactly what he or she bids, and the auctioneer (in this case USEPA) does not receive any of the auction's proceeds. In the mechanism specified, the highest bidder purchase the permit from the agent making the lowest offer, and pays exactly what he or she bids. Therefore, the lowest offer receives the amount paid by the highest bidder and so on up and down the respective schedules.¹ This formal market is to occur once a year and will be the arena in which the permits withheld at the annual allocation are traded².

The second trading opportunity is an informal process. In this case, agents are allowed to negotiate individual deals for permits at any time. Such transactions have occurred already with the Tennessee Valley authority as principle purchaser of permits (Charlier, 1992).

The final opportunity for permit transactions is direct purchase of permits from USEPA. In this situation, agents contact USEPA to purchase permits at some posted price. The price is designed to be significantly higher

¹As might be expected this type of incentive structure is not incentive compatible, and may create opportunity for profitable strategic bidding and offering. The reader is referred to Cason, 1992 for a full theoretical development of this auction's strategic characteristics.

²The units withheld from the allocation enter as permits with a zero offer price. Their revenue is distributed equal on the bases of the number of units actually withheld.

than the price established in the other trading markets. Agents will see this as an option of last resort that is used very infrequently.

Permits have a life of one year but can be banked. That is, once a permit is used to cover SO₂ emissions at a plant, it cannot be used again. However, permits need not be used in the year they are issued. Permits can be held or "banked" until such time in the future that the firm wishes to use or sell them. This introduces an intertemporal aspect to this emission right system.

While there has been much theoretical work in the past describing, and analyzing a system of marketable permits (see Montgomery, 1972 and Dales, 1968), very little literature has addressed the hybrid system that is currently in place. Instead this new program is being analyzed within the controlled laboratory environment of experimental economics.

In an initial study before the legislation was finalized, the discriminative price revenue-neutral auction mechanism was tested. That work focused on the size of the mandatory contribution to this market, and the effects of using a discriminative price versus a more well understood uniform price. The results of that work (Franciosi et al., 1992) added to our understanding of some of the fundamentals of this novel system.

This paper reports the results of two experimental economic investigation into the workings of the mandated pollution permit market. One study focuses on the formal revenue-neutral auction. While other contains both this formal auction and the opportunity for informal trades. As is explained in grater detail below, the informal market takes the form of a double auction. Each study allows for banking of permits.

The results show that one mechanism alone can be effective in realizing much of the gains from permit trade. Yet, by adding the second trading opportunity, the results begin to diverge from our equilibrium forecasts, and gains from trade fall substantially. Further, effective use of banking may be hindered by this complication.

It is important to note at this time that the comparison of these results is qualitative only. While both studies used the same experimental parameters, they used different procedures, so a rigorous statistical procedure is not possible. A study already underway will attempt to replicate these results using a standard procedure and parameter set. In this way we will be able to investigate formally the results reported in the present paper.

The remainder of the paper is organized in three sections. Section II briefly describes the design and procedures used in these experiments. The next section will turn to the results. Here we will focus on the issues of permit price fluctuation and system efficiency. The final section draws some tentative conclusions and outlines future analysis in this area to help understand the dynamics at work within this complicated institution.

II. Experimental Designs

These experiments were conducted in two different locations using two different procedures. The University of Colorado's Laboratory for Economics and Psychology (LEAP) undertook experiments that focused on the revenue-neutral auction. The University of Arizona's Economic science Laboratory (ESL) created an environment that contained the informal market and a simplified revenue-neutral auction. In the remainder of this section I focus on a short description of the procedures used by the specific experiments at each location. One important feature that will be discussed is the training procedure used at both locations, as these may have some influence on the results. For a detailed discussion on the exact procedures used see Franciosi et al. 1992 and Brown-Kruse and Cronshaw 1992.

Revenue-Neutral Auction Only

In these experiments run at LEAP, subjects acted as the managers of revenue-maximizing-rate-of-return regulated firms. That is, at the beginning of each round each firm was given a fixed revenue and its profit depended on its cost reduction behavior. Also at the beginning of each round, firms were given an allocation of emission permits. The allocation in these series of experiments was the same for all firms (6 permits to start) with the allocation being cut in half in round 5 to simulate the change that occurs in the late 1990's in the USEPA's regulation. Subjects knew in advance the length of the experiment and each period's allocation.

For these experiments firms would produce, as a matter of course, some emissions that had to be cleaned up. Firms could chose to use permits to cover the emissions produced or pay to clean them up. There were five firm types which varied in the amount of emissions and cost to clean them up. This gave incentive for firms with low emissions to sell permits to high emission firms. Table 1 contains the parameter set used in these experiments, including the value of permits to firms and their initial allocation.

FFC/93 Papers & Proceedings

Once the initial allocation of money and permits was made, subjects made decisions about use of permits. This involved deciding how many permits to buy or sell, and how many to use or bank. At the end of the round new balances were calculated and added to the subject's total. At the end of the experiments, subjects were paid in cash.

For this series of experiments subjects first participated in a "banking only" experiment. In this case, the initial allocation of money and permits was made, and the only choice subjects faced was how many permits to bank or use; there was no trading between subjects. The purpose was to introduce the subject to the somewhat complex concept of banking without including the further complexity of the auction. All subjects participating in this training session were paid as usual, but only those who made over a specified amount of the optimal profit were invited back for the full session.

The experienced subjects who were brought back faced the same experiment as their trainer expect it now featured both banking and the revenue-neutral auction. Then when the initial allocation of permits was made, some were withheld (see Table 1 for initial allocation and withholding quantity). Subject were then asked to entered their bids for permits and then were asked to enter their offers to sell. In this way, subjects could operate on both sides of the market and operate as true traders in permits.

Once all bids and offers had been entered, the computer ranked them as described above, and notified subjects as to the number of permits they bought and sold and of their net profit. Subject were given time to study this information before they had to decide on the number of permits to use or bank. Once all subjects had made and entered their banking decisions, a new allocation was made, and the auction was opened up. At the end of the experiment all permits had to be used or fell to a zero value. Subjects again were paid in cash at the conclusion.

Table 1 Experimental Parameters					
Firm No.					
Unit No.	1,6	2,7	3,8	4,9	5,10
1	3.80	15.20	7.60	19.00	22.00
2	3.40	13.60	6.80	17.00	20.40
3	3.00	12.00	6.00	15.00	18.00
4	2.60	10.40	5.20	13.00	15.60
5	2.20	8.80	4.40	11.00	13.20
6	1.80	7.20	3.60	9.00	10.80
7	1.40	5.60	2.80	7.00	8.40
8	1.00	4.00	2.00	5.00	6.00
9	.60	2.40	1.20	3.00	3.60
10	.20	.80	.40	1.00	1.20
Allotment (forced)					
Prd 1-4	6(2)	6(2)	6(2)	6(2)	6(2)
Prd 5-12	3(1)	3(1)	3(1)	3(1)	3(1)

Two Permit Markets

As noted, the procedure used at the ESL varied from that used at LEAP. This set of experiments utilized software designed by ESL for the Department of Energy for their continuing research effort in the area of marketable emissions permits. Subjects were told that the permits could be redeemed for cash. That is, instead of acting as firm managers, these subjects were simply utility-maximizing individuals. As in the other experiments, permits were allocated at the beginning of each round. Subjects knew the allocation for the entire experiment, and the redemption value for permits.

Training sessions for this series of experiments were also conducted independently of the sessions reported here. In these sessions, subjects went step by step through the computerized, on-screen instructions. At the conclusion of the instructions, subjects were allowed to participate in three rounds of the experiment to see how the procedures worked first hand. After these practice rounds had been completed a written examination about the procedures in the experiment was given to the subjects. Those subjects who scored greater than 60% were asked to be in the regular sessions. All subjects participating in this training session were paid a flat fee for their time.

The sessions using these experienced subjects again began with the computerized instructions. After subjects had reviewed the instructions the sessions started with an allocation of permits. Again some permits were withheld for mandatory participation in the auction. Once the allocation was made, an informal trading opportunity was opened. In these sessions this informal market was a double auction. This institution was chosen for its well-known properties of conversion to the competitive equilibrium. This market was open for 3 minutes, allowing all subjects to buy and sell permits as they liked.

At the close of the market, subjects were asked by the computer if they would like to voluntarily contribute more units to the auction that was coming up. In this set of experiments, there was no opportunity for subjects to put a reservation price on contributed units as in the LEAP experiments. After all voluntary contributions were registered by the computer the auction was opened. Here subjects were asked to indicate to the computer their bids to buy permits. Once all bids had been entered, the computer ranked bids from highest to lowest, and distributed them to those with winning bids. It also reduced their balance by the amount of their bid. Because the auction was still revenue-neutral, the proceeds were distributed on a per unit basis to all those who had units forced into, or voluntarily contributed to, the auction. As noted, this was a more simpler procedure than either the one used in LEAP or mandated by the USEPA rules. However, it was agreed that the complexity of the experiment necessitated this compromise.

After the results of the auction were posted, subjects were asked to consult their individual redemption values to consider how many units they wished to redeem for cash and how many they wished to bank for the future; balances are adjusted accordingly. At the end of the experiment subjects are again paid in cash.

Thus we have results from one set of experiments that utilize only the revenue-neutral auction, and one that combines this mechanism with an opportunity to trade within a double market institution. By including this new mechanism we have introduced the opportunity to arbitrage across markets. However, as will become obvious this new feature may create problems for the efficient operation of this system of marketable permits.

III. Results

As noted in the introduction, the focus of this paper is on the prices of the traded permits and the overall efficiency of the system. There are of course other important points that should be considered such as the optimal banking pattern, the number of permits used (or redeemed) each period, and the bid and offer pattern. These topics can and will be addressed in later papers and are discussed in some detail in both the Franciosi, and Brown-Kruse and Cronshaw papers.

Figures 1-3 are representative of the results of the auction-only experiments. The lowest trading price is the last accepted bid in the revenue-neutral auction, and the mean trading price is the average price bid on, and paid for, permits. The adapted CE price is a measure of the expected competitive equilibrium given the banking pattern within the given experiment. That is, instead of comparing these price measures with a competitive equilibrium based on optimal banking patterns (\$71), the adaptive CE is based on the previous banking patterns observed in the particular experiment and is a dynamic measure across the life of the experiment. It was hypothesized that subjects made price decisions based not on some expectation of optimal banking established in period one, but on their current behavior and that of the other subjects.

As can be observed from the figures, prices seem to converge to the adaptive CE price. Further, as the

FIGURE 1

Market Prices PSPMF1

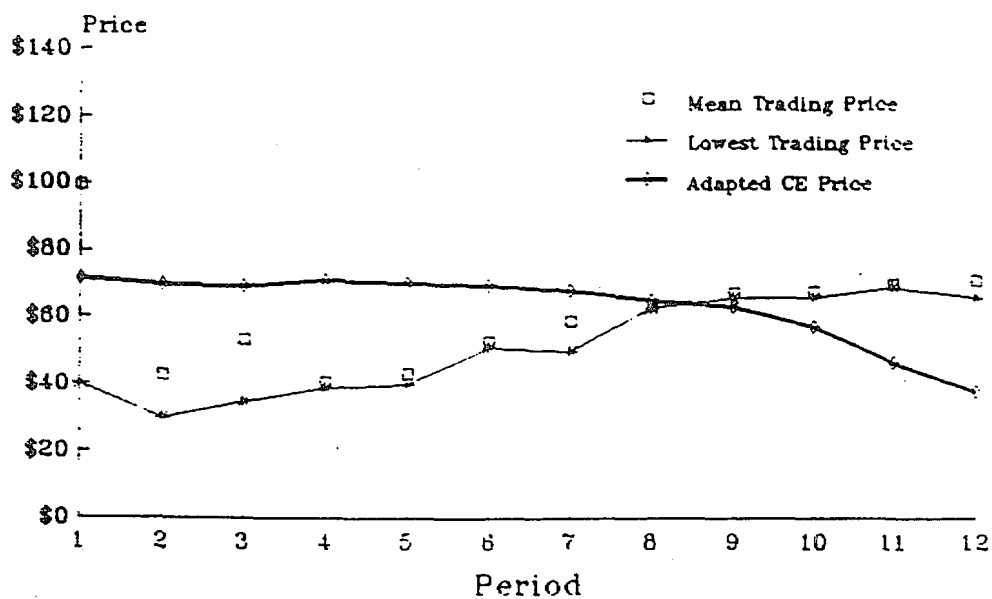


FIGURE 2

Permit Market Prices PSPMF3

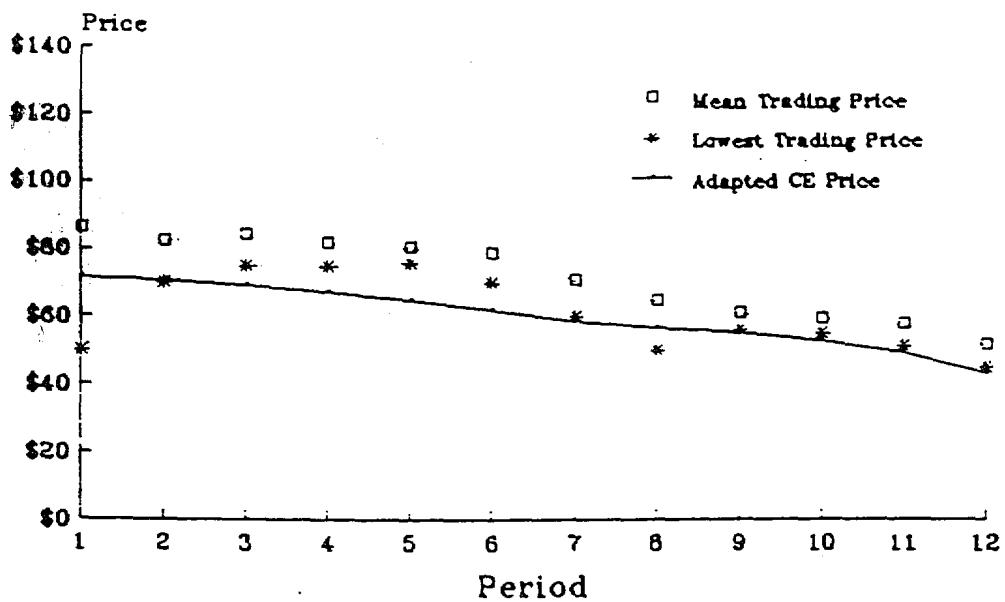
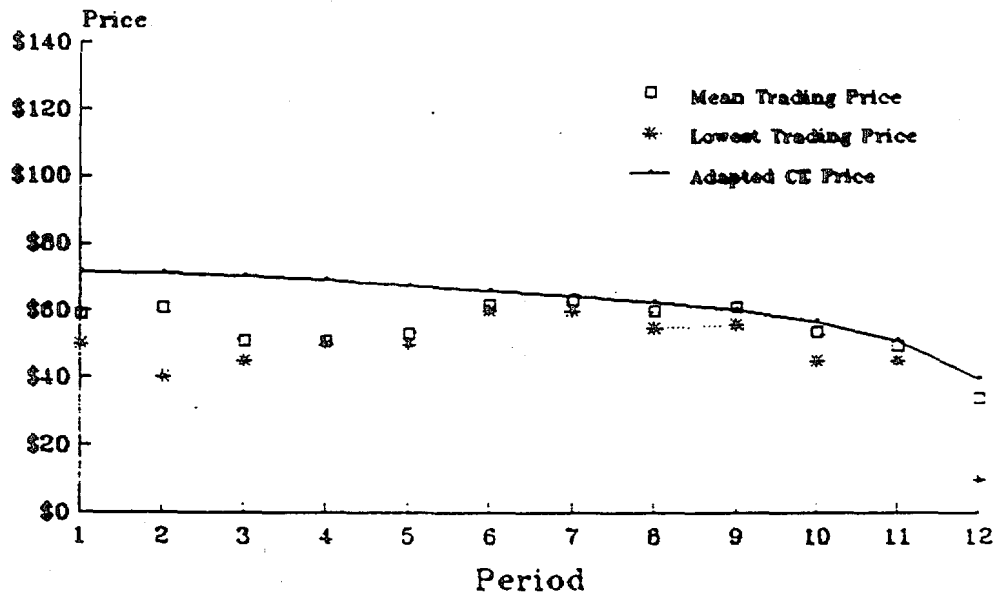


FIGURE 3

Permit Market Prices PSPMF6



experiment progresses, the bids tighten. The variation across winning bids becomes very small; the mean and the lowest trading prices converge. The falling CE price is the result of the excessive number of permits banked (thereby making subsequent purchases less important). Even if we ignore this CE price, we can see that the price of traded permits does not move too far from the original expected price of \$71. The system seems to operate effectively and efficiently.

Another important observation is that there is no price fluctuation when the allocation of permits is cut in half at the start of period five. This would indicate that banking is serving its purpose by helping to stabilize the prices in this market. This is consistent across all the sessions. Thus, prices in under this revenue-neutral auction seem stable across time and tend to converge to the adaptive CE price.

However, when we examine the results from the experiments with both the auction and the market we see a completely different story. Figures 4-6 show much of the same data as the previous figure. Again we can see the mean trading price, and in this case, the highest reject bid (the one right below the lowest accepted bid). Further, the solid connected points represent the contract price established in the double auctions. The dashed line at \$7.10 represents the experiment's equilibrium permit price with banking, while the solid line is the equilibrium price with each round³.

As can be observed, prices are unstable relative to the earlier results. We no longer see the convergence of the mean bid price and the lowest rejected bid. In fact there is quite a pronounced divergence. Also there is a definite effect when the allocation is reduced in period five. Here, at least in the Figures 4 and 5, we see a sharp increase in both the prices in the auction and in the market. It is possible that this pattern could be a "bubble and crash" phenomenon that has been reported in other experimental asset markets (Smith et al. 1988).

Further there is a significant divergence in the prices between the auction and the market. This would mean that it would be possible to buy a permit under one institution and sell in another at a profit. This is especially pronounced in Figure 6. Even though the market price seems to converge early to the banking equilibrium, it is well above the auction prices. It seems that the arbitrage is not working. By adding the second trading opportunity, the system seems to inherit a number of non-optimal characteristics.

This is reinforced when we examine the efficiencies within the experiments. Here we define efficiency as a percentage of the optimal profit realized across the entire experiment. As would be expected, the average efficiency in the auction-only experiments was quite high, at just over 80% across all the experimental sessions. This compares with efficiencies ranging from 7.5% to 47% in the auction and market experiments. Somehow, the introduction of this new market drastically affects the operation of the system in a negative way.

IV. Conclusions

It would seem that the introduction of a new trading institution dramatically effects the outcome in this permit market. Prices become wildly unstable, not only across time, but across markets in a period; we have no indication of an optimal social price for permits. Realizations of gains from trade are lost as banking becomes very erratic. In effect a system that operates near its expected optimum in the one-market case seems to fly apart when we introduce this double auction.

What is it about the introduction of this second trading opportunity that so adversely affects these results? The answer is not easy or obvious. It is important, before postulating too extensively to remember that these two sets of experiments used two different procedures -- not only to run the experiments but to create the experienced subject pool used. It will be important to study this in a more systematic approach by using one procedure both for data collection and training.

Also, it will be important to use a parameter set that is designed for this type of study. If we look at other results from the two-market experiments using a different data set, we see much higher efficiencies, similar to those seen in the one-market sessions. Of course these experiments still yielded prices that were similar to those in figures 4-6. But the radical and consistent fall in efficiencies seen with the introduction of the double auction may be caused by some quirk in the parameters.

Experiments are being designed to look at these issues. They will employ the software used in the two market experiments. A completely new parameter set also will be created. Further, subjects not from either location will be used to control the unlikely event there is some kind of subject pool effect. There experiments will

³The prices in the two experimental designs are stated in different units, but are comparable.

FIGURE 4
MARKET AND AUCTION PRICES
R10_12DC

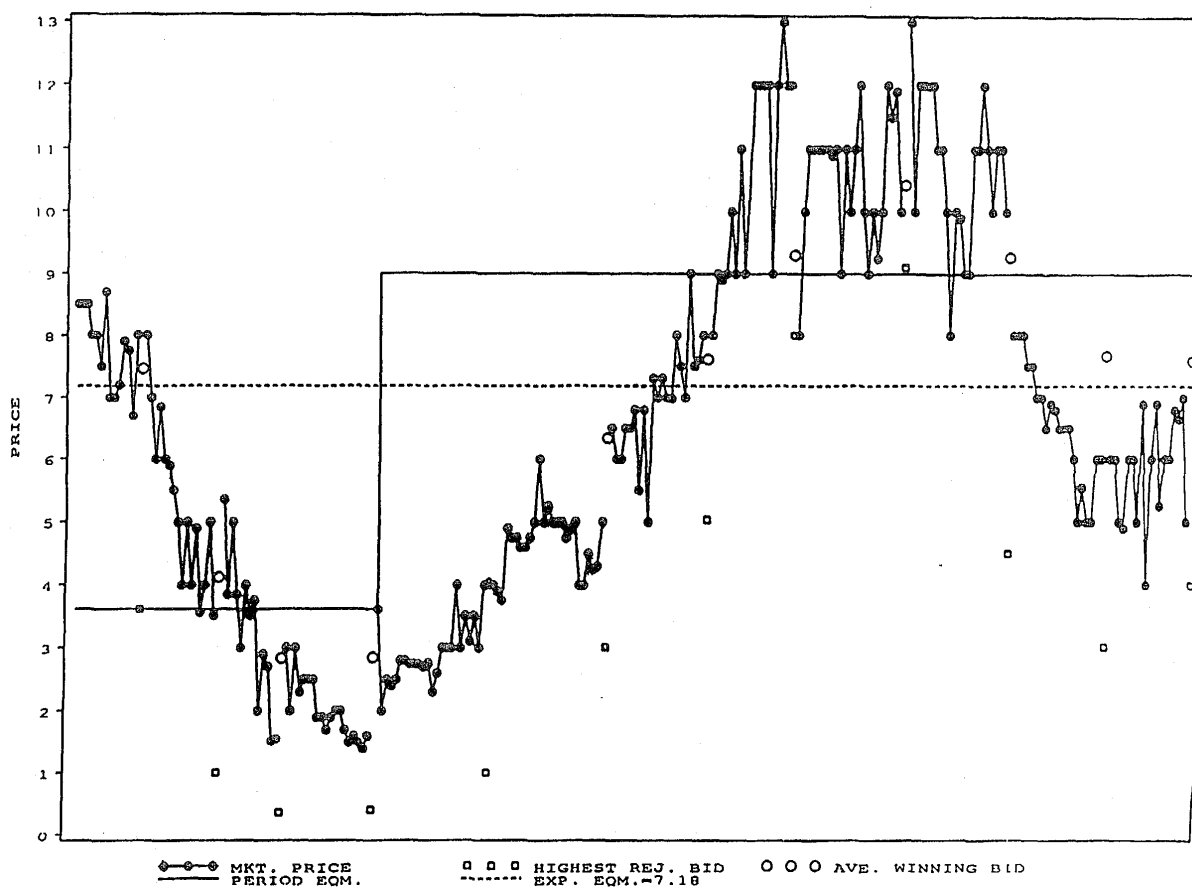


FIGURE 5
MARKET AND AUCTION PRICES
R11_12DC

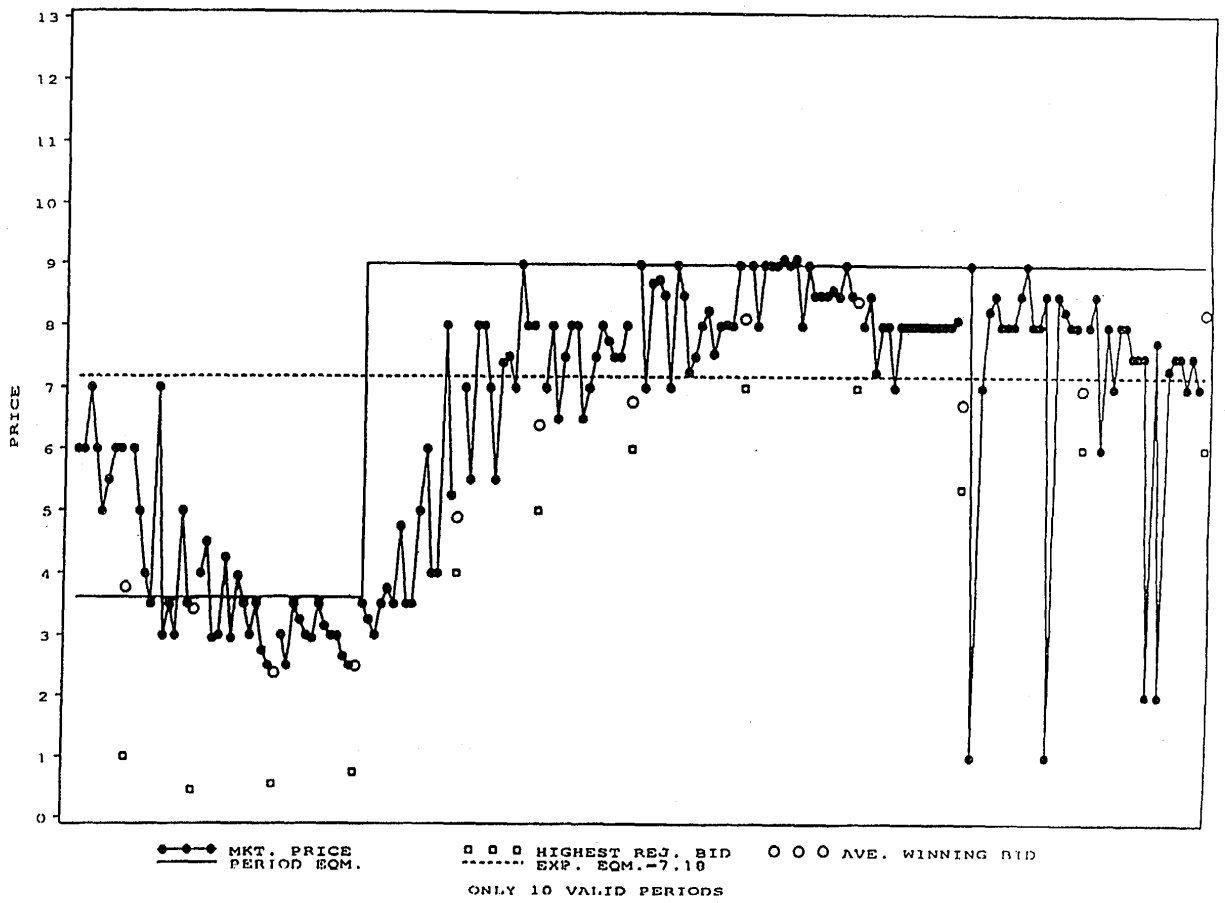
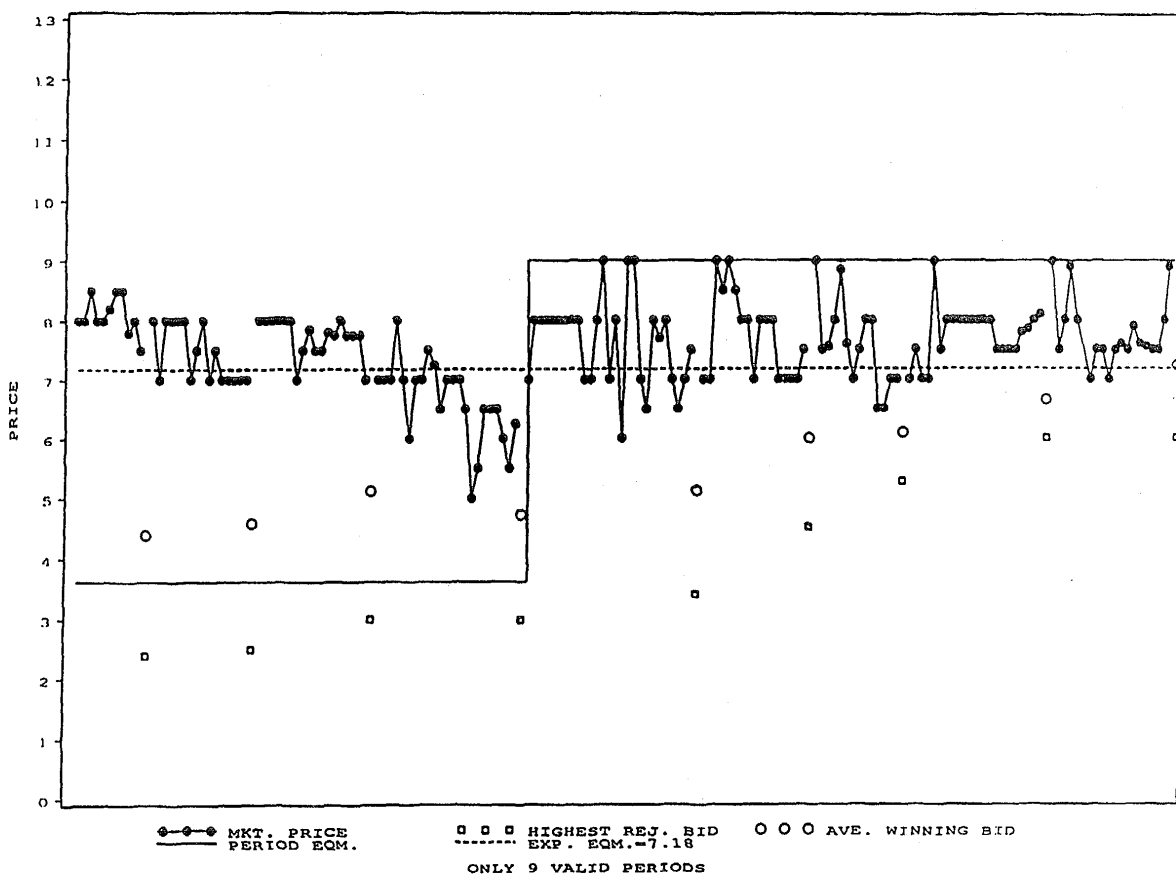


FIGURE 6
MARKET AND AUCTION PRICES
R12_10DC



be run both with and without the double auction. Also the training procedure will be varied systematically in order to control for training effects.

It is important that we understand what is going on in this market. With one experiment giving a thumbs up and another a thumbs down to this incentive-based system, we are sending radically different signals to the policy makers who are closely watching. If we can create a system that controls emissions while reducing government "involvement," we may see widespread acceptance and use of incentive-based regulation. If we create a system that blows up (and we cannot explain why), this grand experiment may be for naught and we will continue with command-and-control.

Experimental economics offers a testing bed for this type of policy, but only if we can show its consistency, reliability and usefulness in identifying, explaining and mitigating problems such as those we have found here. We must carefully examine anomalies like those seen in this paper and methodically work to explain them. by doing so we will demonstrate the power of experimental economic methods as a policy analysis tool.

References and Selected Readings

- Baumol, W.J.; Oates, Wallace E. (1988) *The Theory of Environmental Policy*. 2nd ed. Cambridge U.K: Cambridge University Press.
- Brown-Kruse, Jamie; Cronshaw, Mark. (1991). Permit Market Experiments Preliminary Report. University of Colorado, Department of Economics. 92DOCS 016 - ET
- Brown-Kruse, Jamie; Cronshaw, Mark; Elliott, Steven (1992-03). Experimental Investigation into Marketable Emissions Permits: Some Results and Questions. University of Colorado, Department of Economics.
- Cason, Timothy N. (1992). Seller Incentive Properties of EPS's Emissions Trading Auction. University of Southern California, Los Angeles, CA. Dept of Economics. 92DOCS 028 - et
- Cronshaw, Mark B.; Brown-Kruse, Jamie. (1991-3-8). Temporal Properties of a Market for Emission Permits. University of Colorado, Department of Economics. 92DOCS 018 ET
- Cronshaw, Mark B.; Brown-Kruse, Jamie. (1992-05-26). An Experimental Analysis of Emission Permits with Banking and the Clean Air Act Amendments of 1990. University of Colorado, Department of Economics. 92DOCS 030
- Cropper, Maureen L.; Oates, Wallace E. *Environmental Economics: A Survey*. *Journal of Economic Literature* 30(2):675-740.
- Dales, John H. (1988) *Pollution, Property and Prices*. Toronto: University of Toronto Press.
- Franciosi, R.; Isaac, R. Mark; Pingry, David E.; Reynolds, Stanley S. (1992-5). Extension of Research into Marketable Emissions Permits. 92DOCS 029
- Hahn, Robert W. (1988). Promoting Efficiency and Equity through institutional design. *Policy Sciences* 21:41-66. 91DOCS 031
- Hahn, Robert W. (1989) "Economic Prescriptions for Environmental Problems: How the Patient Followed the Doctor's Orders" *Journal of Economic Perspectives* 3(2): 95-114.
- Hahn, Robert; Noll, Roger G. (1982) "Designing a Market for Tradable Emission Permits" in *Reform of Environmental Regulation*. Wesley Magat (ed.) Cambridge, MA: Ballinger.
- Hale, Douglas R.; Bjornstad, David J. (1991). Marketable Permits for Controlling Sulphur Dioxide Emissions. Oak Ridge National Laboratory, Energy Division. 92DOCS 019 - ET
- Heggelund, Merete. (January, 1991). Emissions Permit Trading: A Policy Tool to Reduce the Atmospheric Concentration of Greenhouse Gases. Calgary: Canadian Energy Research Institute. 92DOCS 006
- Montgomery, W. David. (1972) Markets in Licenses and Efficient Pollution Control Programs. *Journal of Economic Theory* 5(3): 395-418.
- Plott, Charles R. (1983). Externalities and Corrective Policies in Experimental Markets. *Economic Journal* 93: 106-127. 91DOCS 043 ET Sheffield, Arthur (1992) Canada, Department of the Environment, Personal Communication.
- Tietenberg, Tom (1988) *Environmental and Natural Resource Economics*. 2nd ed.

FFC/93 Papers & Proceedings

Smith, Vernon L., Svchanek, Gerry L., and Williams, Arlington W. (1988) "Bubbles, Crashes, and Endogenous Expectations in Experimental Spot Asset Markets," Econometrica, Vol 56, NO 5, September, 1119-1152.

Voluntary Selection of Real Time Pricing: Laboratory Evidence

Jamie Kruse, University of Colorado

Abstract

This paper reports an experimental study whose goal is to capture behavior in the laboratory that mimics observed field behavior. A laboratory model is created in which the access fee portion of a two part tariff can lead to acceptance or rejection of a real time pricing program. Further, changes in the laboratory conditions which generate derived demand for an input result in optimal reversal of the decision to participate in the program. This behavior mirrors that observed when such a program was implemented on a trial basis in the field. Further, two proposed adjustments are tested and both judged successful in increasing participation rate.

I. Introduction

In 1989, Niagara Mohawk Power Corporation (NMPC) offered on an experimental basis, real-time pricing service. The experimental service called Hourly Integrated Pricing Program (HIPP), is a two-part tariff. A customer specific access fee plus marginal cost-based variable charge comprise the HIPP tariff. The societal benefit of real-time marginal cost-based pricing is well documented (For example, see R. Willig, 1978; R. Schmalensee, 1982; H. Varian, 1987; and K. Roberts, 1979). Under HIPP the goal of customer specific access fee is to preserve revenue neutrality--the customer is at least as well off under the two part-tariff as he would have been under conventional pricing. In order to preserve incentive compatibility, the current or future access fee cannot be affected by the customer's current decisions. The access fee is based on the customer baseline load (CBL) which was determined prior to creation of the program. Specifically the access fee equals the CBL multiplied by the nonnegative difference between the conventional and real-time price. Under stable conditions, this access fee plus real-time price yields exactly the same revenue to the utility and profit to the customer as conventional pricing if the customer maintains the same load pattern.¹ Figure 1 presents a simple example. In case a, P_c is the conventional price and x the quantity the customer would use based on derived demand D . Suppose x becomes the customer's CBL and the real time price P_R is realized. The customer can shift load to x_R to take advantage of the low input price. The customer's gain from the transaction at P_R is equal to the area of acP_R minus the access fee (area $P_c b d P_R$). However in its experimental offering of HIPP, NMPC found that customers would enroll in HIPP and then retire from it later when economic conditions changed.

As long as the customer's economic conditions do not worsen, the benefits to the customer and society are both positive. A problem arises with the fixed access fee when economic conditions change. When conditions change so that customers' consumption under conventional pricing would differ from their established CBL, HIPP is no longer revenue neutral. Figure 1 case b illustrates an example that would prompt the customer to retire from the real time pricing program. If economic conditions deteriorate for the firm, then the relevant derived demand is D' as depicted in case b. In this case, with the favorable real time price, P_{R_1} , the customer will use x_{R_1} . The customer's gain from the transaction is the area $a'eP_{R_1}$ minus $P_c b d P_{R_1}$ --which makes him worse off than with the conventional price. At such a time, this customer would retire from the HIPP program only to reenter when an upturn in the business cycle makes HIPP preferable again. When customers leave the HIPP program, the societal benefit of marginal cost-based real-time pricing is lost for the time being.

The primary goal of this study is to determine whether laboratory conditions similar to those described above prompt a behavioral response similar to the customer reaction observed in the field. Our secondary goal is to compare two adjustments to the original program which changes the incentives to remain in the real time pricing program.

The laboratory environment was created so that subjects would self-select into two groups: New Program customers and "conventional" customers. From this base case, we then introduce a series of changes in the laboratory environment. Subjects experience a "recession" of known length. A subset of the New Program

¹ By stable, we mean that the economic conditions that determine customer profitability do not change. Product prices may be stochastic, but the underlying distribution is constant.

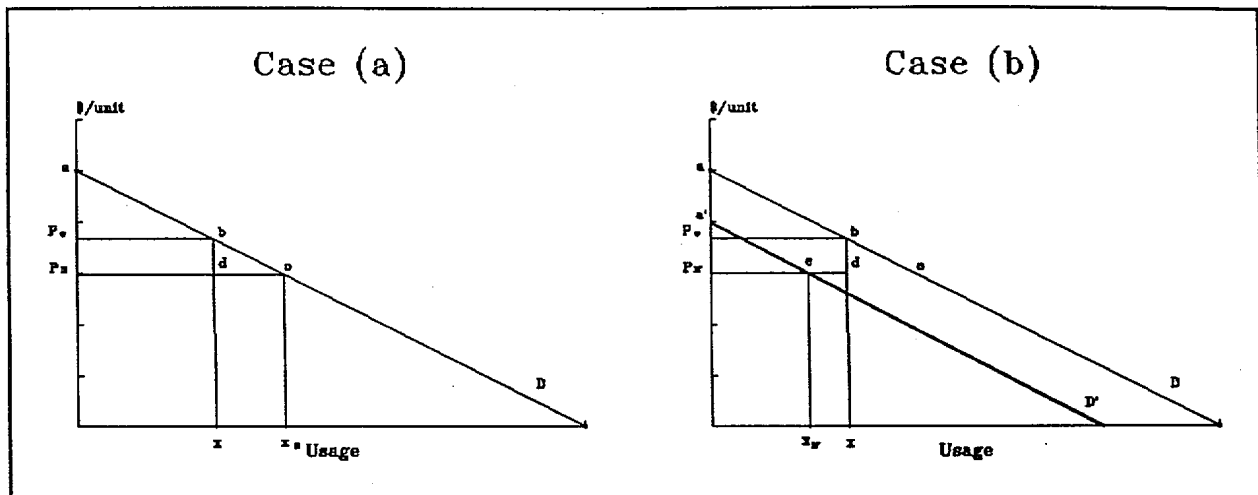


Figure 1 The acceptance of a two part tariff under changing demand conditions.

participants find their CBL-based access fee so high that the conventional program is more profitable than the New Program during the recession. A second recession is initiated with either the additional stipulation that retiring from the New Program means that a mandatory waiting period must pass before reentry is allowed or a reduced access fee for the New Program accompanies the recession. In the next section we will elaborate on the experimental design and discuss expected profit maximization based incentives. This is followed by our experimental results and discussion in the final section.

II. Experimental Design and Theoretical Predictions

The experiment is divided into **Parts**, **Cycles**, and **Rounds**. Each of the four parts of the experiment places the subject under a different set of rules which govern his/her decision. Figure 2 demonstrates the relationship between parts, cycles and rounds. In each **part** there are several cycles. At the beginning of a **cycle**, the subject chooses either Conventional pricing or the New Program. In general, the subject does not know how many cycles are in each **part**. Each cycle contains four rounds (a, b, c, & d) in which the subject makes production decisions.²

In **Part I**, subjects are told that they are making production decisions for a firm which uses a single input to produce a fictitious product. All firms face the "conventional" input price in Part I. The product price for each firm is determined by the toss of a six sided die. The numbers on the die map into a set of six product prices. There are six different mapping. After the die toss is entered, the computer calculates the subject-specific profit opportunities for all possible input decisions. Figure 3 shows a typical screen display that would be presented to a subject. The profit maximizing production decision within a round is transparent—costs, revenue, and profit for every possible input choice are calculated for the subject. The subject can simply choose the number of input units that yield the highest profit for the round.

Part I familiarizes subjects with the general layout and timing of each cycle and round. There are one or two practice cycles containing four rounds each. One additional cycle in Part I establishes the subject's historical baseline usage which is used to calculate the access fee for the New Program.

Part II introduces the "new" real time pricing program. Subjects are informed that at the beginning of each cycle, they can choose either the constant (conventional) input price or a "new program" in which they face a lower average input price plus an access fee. The access fee is calculated as follows:

$$\text{Access Fee} = (\text{old input price} - \text{new program input price}) \times \text{Historic Usage}^3$$

² A complete set of instructions, protocol and profit calculations is available from the author on request.

³ The minimum access fee is zero. If the HIPP input price exceeds the old input price, then the access fee is set to zero.

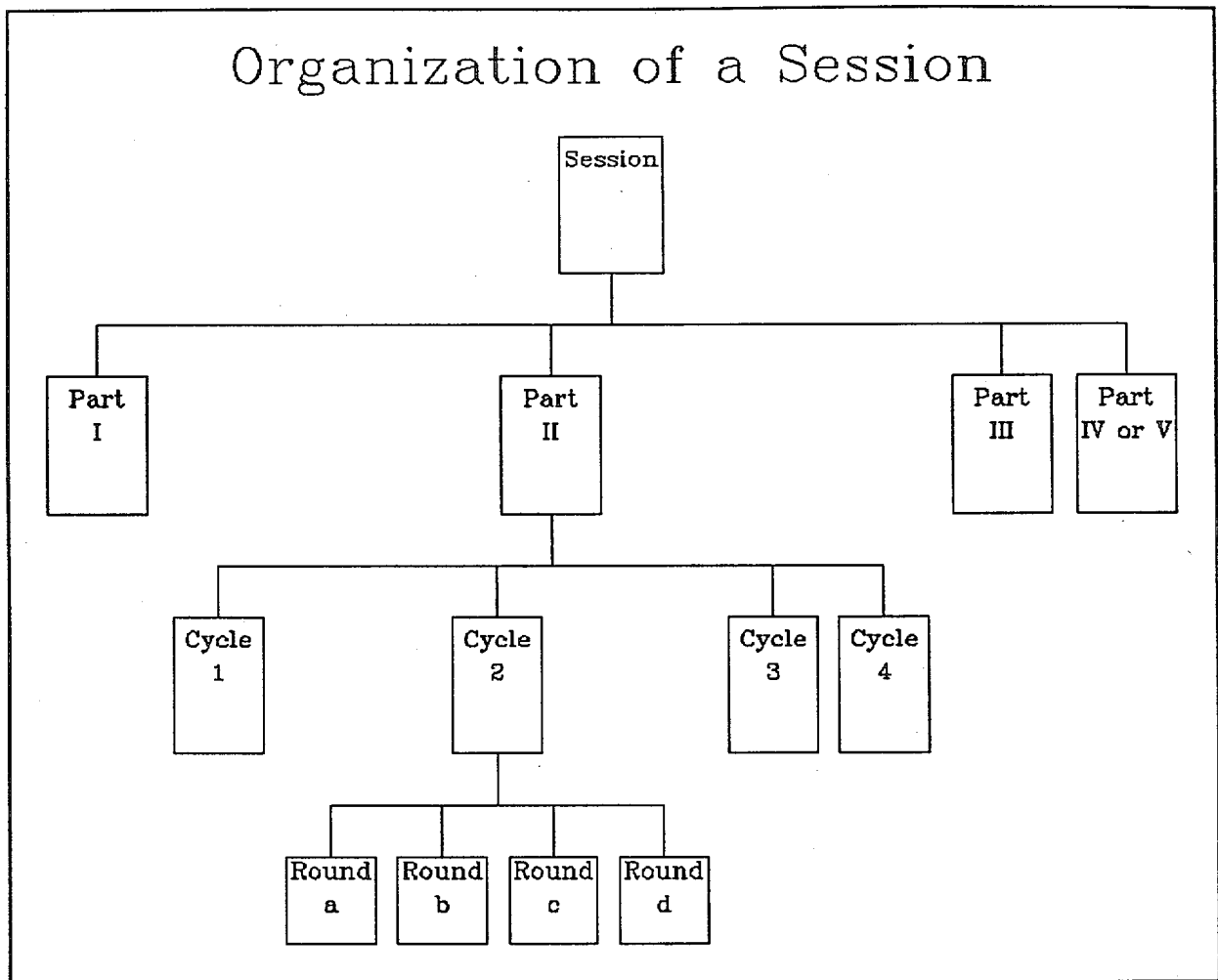


Figure 2

Figure 3.

PRODUCTION TABLE				
Input Units	Input Cost	Product Units	Product Revenue	Profit
1	2.40	3.95	3.16	0.76
2	4.80	7.80	6.24	1.44
3	7.20	11.55	9.24	2.04
4	9.60	15.20	12.16	2.56
5	12.00	18.75	15.00	3.00
6	14.40	22.20	17.76	3.36
7	16.80	25.55	20.44	3.64
8	19.20	28.80	23.04	3.84
9	21.60	31.95	25.56	3.96
10	24.00	35.00	28.00	4.00
11	26.40	37.95	30.36	3.96
12	28.80	40.80	32.64	3.84
13	31.20	43.55	34.84	3.64
14	33.60	46.20	36.96	3.36
15	36.00	48.75	39.00	3.00
16	38.40	51.20	40.96	2.56
17	40.80	53.55	42.84	2.04
18	43.20	55.80	44.64	1.44
19	45.60	57.95	46.36	0.76
20	48.00	60.00	48.00	0.00

INFORMATION		
Part 1 Cycle 1 Round B		
Input Price		2.40
Product Price		0.80
Input Units	Product Revenue	Profit
----	----	----
Balance	4.00	

INPUT REQUEST
Enter Input Units
---->

Historic usage was established in Part I. The New Program input price is determined by the draw of a poker chip from a bag each round. There are twenty chips of each color in the bag and one hundred chips in all. A poker chip is drawn each round to determine the New Program input price and the die is tossed at the beginning of each cycle to determine the product price.

The first two cycles of Part III represent an economic downturn which hits all firms. The lower Product price Schedule B is used for the first two cycles. The two "recession" cycles are followed by three "normal" (Price Schedule A) cycles for a total of five cycles in Part III.

Part III is followed by either Part IV or V. In Part IV, if a subject does not choose the New Program during a recession cycle, then he/she must wait four cycles before being given the New Program input price option. In Part V, the access fee is scaled down during recession cycles. Both Parts IV and V last a total of six cycles.

All sessions were conducted in the Laboratory for Economics and Psychology (LEAP) at the University of Colorado. Subjects were volunteers from undergraduate classes at the University of Colorado at Boulder. Subjects were allowed to participate once. Typically, the duration of the experiment was two hours with subjects earning from \$10 to \$46. Table 1. summarizes the number of subjects and treatments used in each session.

Table 1. Summary of Sessions and Treatments

ID#	# of Subjects	I	II	III	IV	V
NMUCB1	11	X	X	X	X	
NMUCB2	11	X	X	X		X
NMUCB3	8	X	X	X	X	
NMUCB4	11	X	X	X	X	
NMUCB5	6	X	X	X		X
NMUCB6	9	X	X	X	X	
NMUCB7	6	X	X	X		X
NMUCB8	8	X	X	X		X
NMUCB11	9	X	X	X	X	
NMUCB12	6	X	X	X		X
NMUCB13	7	X	X	X	X	

III. Parameters

The experiments reported here were designed and parameters chosen so that under the assumption of risk neutral profit maximization subject/firms self select into three groups. Roughly one third of the subjects would maximize expected profit by always choosing the New program when it is offered. Another one third of the subjects maximize expected profit if they choose the old program. The remaining one third of the subjects would maximize expected profit if they choose the New Program during nonrecession cycles and choose the old program during the recession cycles.

All subjects have a production function of the form, $Q = ax^2 + bx$. Q is the quantity of output which is sold on the product market. x is the quantity of the input used. In general, $a < 0$ and $b > 0$. Specifically, the production function all firms use is $Q = -0.05x^2 + 4x$. Firms face a product price that is drawn from a known distribution. The input price and product price are known by the subject before he/she must make a production decision. The old program input price is \$2.40 per unit. The set of product prices and their associated probabilities (in parentheses) are shown below. In addition, the optimal quantity of input is listed below the price and probability.

FFC/93 Papers & Proceedings

Product					
Price	\$0.65	\$0.70	\$0.80	\$0.90	\$0.95
(probability)	(1/6)	(1/6)	(1/3)	(1/6)	(1/6)
Optimal usage	3	6	10	13	15

The experiment is designed so that all subjects face the same distribution of product prices. However, the random draw signifies different prices for different subjects. In other words, when one subject sees the draw and receives a product price of \$0.90, another subject receives \$0.80, another receives \$0.65 and so on. In part I, the historic usage is recorded and used to calculate the New Program access fee for subsequent parts of the experiment. Our design generates subject decisions which should yield historic usage which ranges from 3 to 15 units of input and thus divides the subject pool. The proportion of the subject pool with a specific historic usage is roughly equal to the probability of the price associated with that optimal usage. So about one sixth of the subjects should have historic usage equal to three, one third of the subjects will have a ten unit history, etc.

The "New Program" is offered for the first time in Part II. Subjects can choose the old program with fixed input price or the new program with access fee and input prices that are drawn from a known distribution. The access fee equals the historic usage times the difference between \$2.40 and the new program input price and is nonnegative. If the new program input price exceeds the old program price of \$2.40, the access fee is zero. The subject must decide whether to subscribe to the new program before he/she knows the current cycle's product price. The New Program input prices, their associated probabilities and the Access Fee are shown below:

New Program					
Price	\$1.90	\$2.00	\$2.05	\$2.30	\$2.50
probability	(1/5)	(1/5)	(1/5)	(1/5)	(1/5)
Access Fee	H*0.50	H*0.40	H*0.35	H*0.10	0.00

H=historic usage

Restricting our attention to the historic usage predicted by simple profit maximization in Part I, we can look at the difference in *expected* profit between the Old Program and the New Program. Recall that profit maximizing firms in Part I will establish historic usage of 3, 6, 10, 13, or 15 units. Table 2. below compares their expected profit given their participation in the Old or New Program.

Using expected profit maximization as the decision criteria, subjects with historic usage of 3, 6, or 10 should choose the New Program and subjects with histories of 13 or 15 units should choose the Old Program.

Parts III, IV, and V have a two-cycle recession. The product prices during a recession are shown below.

RECESSION

Product					
Price	\$0.65	\$0.67	\$0.70	\$0.73	\$0.75
(probability)	(1/6)	(1/6)	(1/3)	(1/6)	(1/6)

Table 2. Profit Expectations from the Old (conventional) program and New Program.

Historic Usage	Old Program	New Program	difference in expected profit/round
3	\$4.63	\$6.89	\$2.26
6	\$4.63	\$6.08	\$1.45
10	\$4.63	\$5.00	\$0.37
13	\$4.63	\$4.19	-\$0.44
15	\$4.63	\$3.65	-\$0.98

Table 3 compares the expected profits during a recession. Expected profit maximizing firms with histories of 3 or 6 should would choose the New program. Firms with a 10 unit history would choose the Old Program as would firms with historic usage of 13 or 15 if the decision criteria is expected profit maximization. Contrast the

risk neutral prediction for recession v. nonrecession cycles for the firms with a 10 unit history. During a "normal" cycle they would choose the New Program and during a "recession" cycle, they would choose the Old Program.

Parts IV and V change the incentive to participate in the New Program during recession cycles. Comparison of the last two parts is analogous to comparing a "carrot or stick" approach. Part IV penalizes subjects if they leave HIPP during a recession by not allowing re-entry for four cycles. Under the penalty system, firms with histories of 3, 6, and 10 should choose the New Program, whereas the Old Program is the expected profit maximizing choice of firms with historic usage of 13 or 15 units.

In Part V we reduce the access fee so that the New Program dominates the Old Program in terms of expected profit for the firms with a ten unit history. The magnitude of the difference is about the same as for the nonrecession case.

The experimental design and the parameters set the stage so that the choices of Old Program always, New Program always, and New Program during normal cycles and Old Program during recession cycles could be observed for some subset of the subjects. Using risk neutral profit maximization as the decision criteria, we can partition the subjects precisely as shown in Table 4. However, this partition is not necessarily a prediction. Risk attitudes can shift the classifications. The risk neutral partition provides a baseline for comparison with observed choices.

IV. Results and Discussion

One major concern in conducting this study was whether or not we could create conditions in the laboratory that would give rise to behavior that mimics field behavior. We wanted to create a set of conditions that would induce some subjects to participate in the New program during "normal economic periods" but retire from the program during economic downturns. All subjects had identical production functions, faced the same product price distribution and potentially had the same per unit input costs. The only differences induced by the lab environment were differences in historical usage which affected subjects' access fees. The parameters were chosen so that the array of subject decisions would cover all possibilities. The next step is to assess whether the variety in subject decisions were due to purely random behavior or to the laboratory stimuli we used.

Figure 4. summarizes the participation rate in the New Program for all subjects. Normal cycles were unrestricted cycles in which schedule A product prices were used. Recession cycles were unrestricted cycles in which the lower schedule B product prices were used. Penalty cycles were also recession cycles with reentry allowed into the New Program after four cycles had passed since retirement. Finally, Lower Fee cycles were recession cycles with a reduced access fee. The New Program participation rate went from 59.2% in normal cycles to 49.8% in unrestricted recession cycles. In addition, both the penalty and lower fee resulted in higher participation than in unrestricted recession cycles.

The incentives for different decisions by otherwise identical firms were created by setting the individual decision conditions so that subjects would have different historical usage and thus different access fees. Firms with a low CBL would find the New Program very attractive because of a relatively low access fee. On the other hand, high CBL firms would find the access fee prohibitively high. If we can separate decisions on the basis of their CBL, then we have evidence that subjects are responding to the stimuli we created. Figure 5 provides graphical evidence that subject CBL can be used to classify subject decisions. High CBL firms had the lowest New Program participation rate on average for all cycle types. Low CBL firms had the highest New Program participation rates for all cycle types. A 2x3 ANOVA was used to test for differences in participation rate in unrestricted normal and recession cycles. We can reject the hypothesis that subjects with different CBL's responded similarly. In addition for alpha greater than .061, we can reject the hypothesis that participation in the New Program was the same for recession and normal cycles. These results are summarized in Table 5.

We compared subject participation in the New Program during the three types of recession cycles: unrestricted recession cycles, penalty recession cycles and lower fee recession cycles. The difference in participation rate according to CBL class and recession cycle type is highly significant in both cases. Table 6 provides a summary of the results of the ANOVA described above.

Analysis of Covariance was used to determine whether subjects responded differently to the penalty v. the lower fee. Subject participation during unrestricted recession cycles was the covariate. There is no significant difference in the rate of subject participation in the two adjustments to the New Program. Apparently, subjects reacted to the two incentive programs in a similar fashion. Table 7 presents the results of the ANCOVA.

FFC/93 Papers & Proceedings

Table 3. Profit Expectations from the Old (conventional) and New Program during recession cycles.

Historic Usage	RECESSION		difference in expected profit/round
	Old Program	New Program	
3	\$1.24	\$2.62	\$1.39
6	\$1.24	\$1.81	\$0.58
10	\$1.24	\$0.73	-\$0.50
13	\$1.24	-\$0.08	-\$1.31
15	\$1.24	-\$0.62	-\$1.85

Table 4. Partition of sample based on expected profit maximization.

Historic Usage Part I	Part II	Part III cycle 1-2/3-5	Part IV cycle 1-2/3-6	Part V cycle 1-2/3-6	Proportion of Sample
3 or 6	New	N/N	N/N	N/N	1/3
10	New	O/N	N/N	N/N	1/3
13 or 15	Old	O/O	O/O	O/O	1/3

Figure 4.

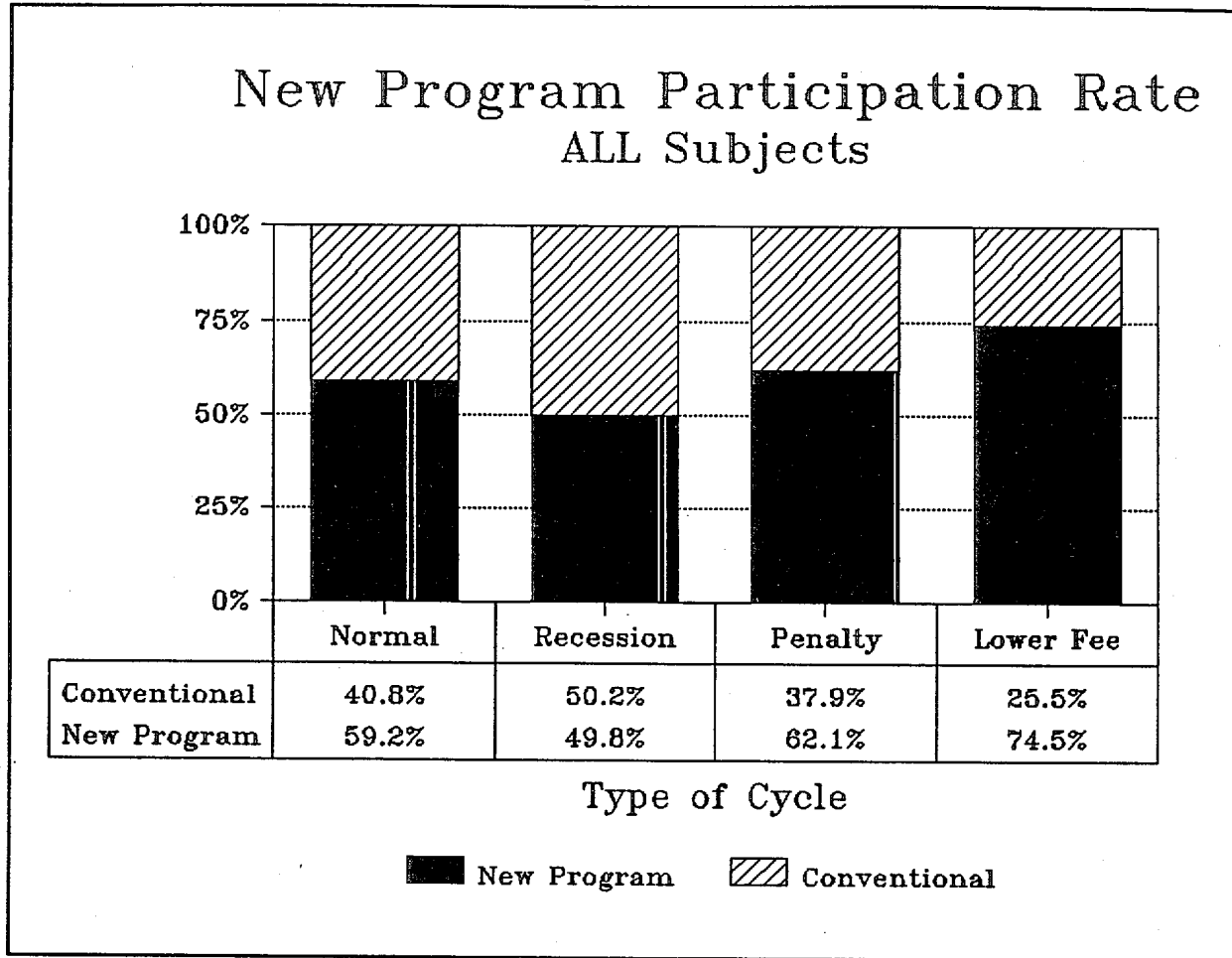
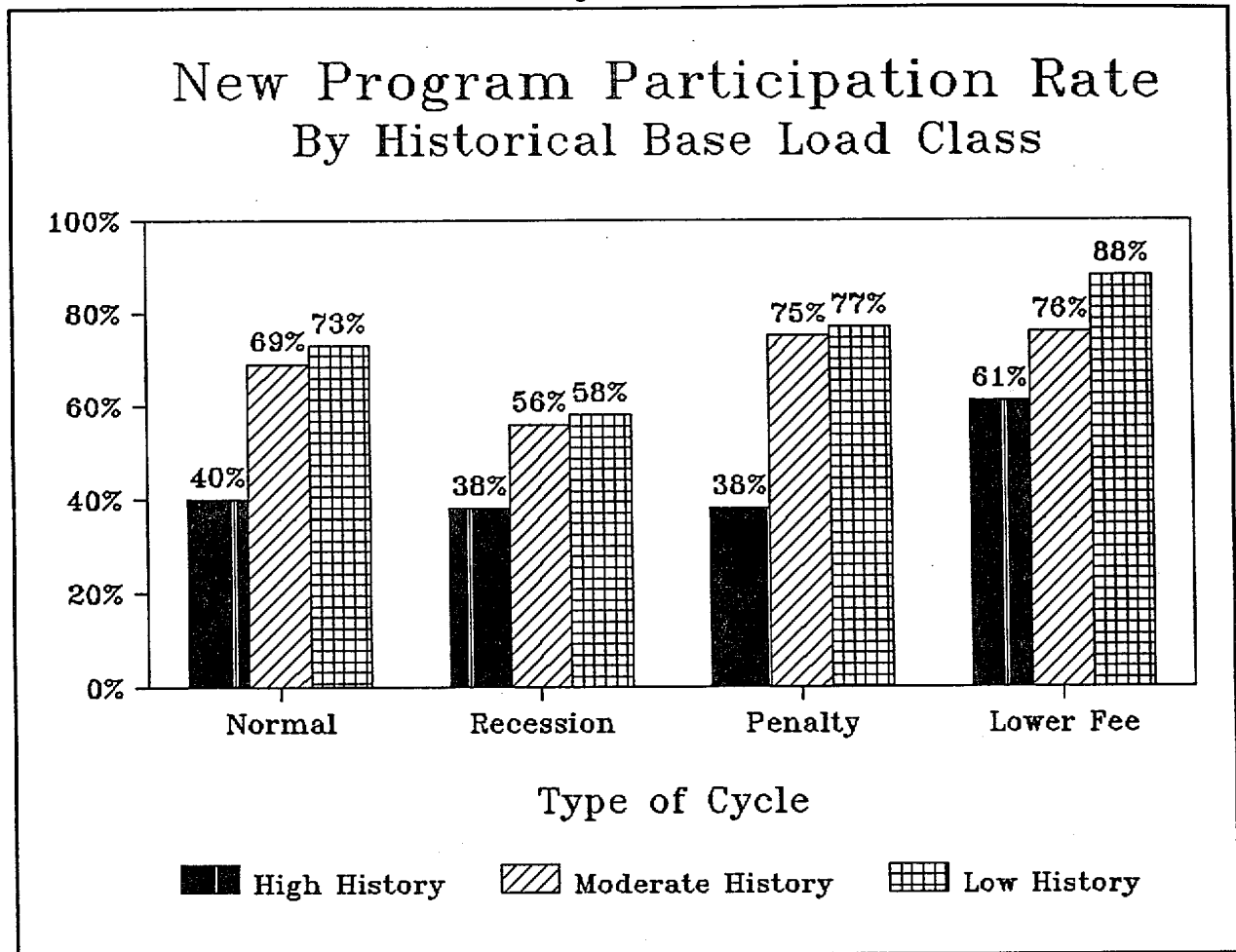


Figure 5.



FFC/93 Papers & Proceedings

Table 5. ANOVA Summary Table for Unrestricted Cycles and CBL Class
2 (normal and recession cycles) x 3 (high, moderate, low)

Source of Variation	sum of squares	d.f.	mean-square	F-ratio	P ⁴
CBL class	2.813	2	1.407	13.362	0.000
Unrestricted cycle type	0.376	1	0.376	3.569	0.061
CBL x cycle type interaction	0.223	2	0.111	1.059	0.349
error	18.634	177	0.105		

Table 6. ANOVA Summary Table for Recession Cycles and CBL Class
3 (recession, penalty, lower fee) x 3 (high, moderate, low)

Source of Variation	sum of squares	d.f.	mean-square	F-ratio	P ⁵
CBL class	4.169	2	2.084	14.049	0.000
Recession cycle type	2.504	2	1.252	8.437	0.000
CBL x cycle type interaction	0.529	4	0.132	0.891	0.470
error	35.758	241	0.148		

⁴ P is the calculated probability of a type I error. If $P < \alpha$ then we can reject the null hypothesis that the means are equal.

⁵P is the calculated probability of a type I error. If $P < \alpha$ then we can reject the null hypothesis that the means are equal.

V. Conclusions

The impetus for this study came from observed field behavior. Customers of a real time pricing program instituted through a two-part tariff would "bail out" during recessions and return when business improved. A model based on risk neutral profit maximization generates implications that give rise to the same behavior. Theoretically and behaviorally, the magnitude of the access fee is a determinant of the decision to subscribe to the New Program. Subjects did respond to adjustments in the New program to retain customers. Both of these adjustments work equally well from a statistical viewpoint although customers probably prefer the reduced access fee to the penalty.

The point of departure for this study is what makes it unique. Smith(1986) contrasts nomothetical experiments that compare theory and observation with nomoempirical that compare the effect of different institutions and/or environments as a means of documenting replicable empirical "laws". A subset of the nomoempirical are heuristic experiments that provide empirical probes into new topics. This study was not theory driven--although the theory plays an important part in the design and parameterization process. We do test different institutions--the alternative designs of the new program, but with a theoretical basis for the hypotheses we test. The experiment should not be classified as heuristic. The theoretical basis, once established, plays an important role. This study represents an attempt to capture the motives which give rise to a pattern of behavior that has been observed in the field--to generate similar behavior in the lab--and to test changes in the lab environment intended to modify the behavior. One remarkable consequence of this effort is that we had some degree of success.

Table 7. ANCOVA Summary Table for Penalty and Reduced Fee Recession Cycles using the Unrestricted Recession Rate as a Covariate

Source of Variation	sum of squares	d.f.	mean-square	F-ratio	p ⁶
Recession cycle penalty or reduced fee	0.015	1	0.015	0.0849	0.772
Covariate: Unrestricted recession cycle participation rate	1.899	1	1.899	10.626	0.002
error	15.910	89	0.179		

⁶ P is the calculated probability of a type I error. If $P < \alpha$ then we can reject the null hypothesis that the means are equal.

References

- Katz, M. 1983. "Nonuniform Pricing Output and Welfare Under a Monopoly." *Review of Economic Studies*, 37-56.
- Oren, S., S. Smith and R. Wilson. 1983. "Competitive Nonlinear Tariffs." *Journal of Economic Theory*, 49-71.
- Roberts, K. 1979. "Welfare Considerations of Nonlinear Pricing." *Economic Journal*, 66-83.
- Smith, V. 1986. "Experimental Methods in Economics." in The New Palgrave: A Dictionary of Economic Theory and Doctrine. Macmillan Press, Ltd. New York.
- Spence, M. 1977. "Nonlinear Prices and Welfare." *Journal of Public Economics*, 1-18.
- Willig, R. 1978. "Pareto-superior nonlinear outlay schedules." *Bell Journal of Economics*, 9(1): 167-207.

Investment, Uncertainty, and Risk

David J. Bjornstad and Steven Elliott, Oak Ridge National Laboratory

Abstract

Traditional economic theory suggests that investment will occur when the expected return on the investment is at least equal to the cost of capital (the interest rate). However, observations of investment behavior suggest that the predictions from this theory overstate the penetration of new technologies into markets. A new literature has developed to explain why investment seems to require a rate of return greater than the traditional cost of capital. We propose to use controlled laboratory experimental methods test the implications of this new literature, and thereby develop an understand of what drives investments in such things as energy saving technologies.

Introduction

The level of a nation's technology base is closely related to the nature of its capital stock. Changes in that base are therefore affected by decisions to make changes in the capital stock. Stated differently, technological change is determined by investment behavior, whether by firms in the production of goods for the marketplace or by households in the production of household services. In the aggregate, the summation of individual decisions regarding durable goods purchases determines the introduction of technology into the economy. Many of the nation's perceived problems rely to a great extent on technological change as a solution. Environmental protection, reduced CFCs, and reduction of greenhouse gas emissions have all been linked to new, innovative technologies. Hence understanding the underlying basis of investment is key to developing efficient policy responses to these issues.

The traditional theory of investment is based in fundamental Marshallian analysis, in which factor productivities are valued in product terms and related to opportunity costs. For investment, the opportunity cost is viewed as the cost of capital obtained through retained earnings or through capital markets. The firm is expected to invest when presented with opportunities to earn a rate of return that exceeds its opportunity costs. Technical innovations that lower costs or otherwise present profit opportunities are often forecasted to penetrate markets promptly. Conversely, factors whose value of marginal product falls below opportunity costs are forecasted to be discarded promptly.

This theory has proven a sound base for approaching the fundamental behavior of the firm as it trades off the purchases of various factor inputs. Yet it has been a poor device for forecasting the timing or the scope of individual investment decisions. In an effort to sharpen the forecasting properties of the theory, a new body of literature has developed that takes into account the irreversibility (sunk cost) attributes of the decision to invest and the role played by uncertainty. In simple terms, this literature postulates a theory of "optimal inertia" by creating a cost component that represents the option value placed by the firm on the opportunity to wait until more information becomes available and uncertainty is reduced.

This theory is linked to technological change by the assertion that the larger the difference between the attributes of an innovative durable good and the traditional durable good it seeks to replace, the larger is the uncertainty surrounding the decision. One theory (Dixit [1992]) argues that the value of the option to wait is directly and positively related to uncertainty, from which it follows that firms will be more reluctant to adopt technologically innovative products or processes than more traditional ones. Stated differently, the "hurdle rate" the firm applies to investment decision making (opportunity cost of capital adjusted for uncertainty) is systematically higher for technologically innovative goods than for traditional ones. Moreover, specific hurdle rates will be conditional on the market structure, the cost of information to the decision making unit, and the ability of the decision making unit to diversify, among other factors.

In this paper we outline a plan to investigate the relationship between investment behavior and choices among durable good opportunities characterized by different degrees of performance uncertainty. The approach to be taken will apply the methodology of experimental economics which offers the ability to control such institutional concerns as market structure, ability to diversify, etc., while systematically varying parameters of interest, such as

risk. It therefore offers an intermediate option for gathering information between pure theory and expensive and relatively complicated gathering of field data.

This work is of particular interest because of its potential ability to shed light on behavior that must be understood clearly in order to develop policies that influence market acceptance of new technologies. Past research has indicated that new technologies typically do not achieve performance levels anticipated by the bench-testing. Past research has also established that technologies penetrate at far slower rates than would be predicted by the simple Marshallian model. Despite this, Marshallian postulates remain at the heart of much of investment analysis particularly at the federal level. This proposed research would provide a foundation for analysis that could lead to an improved understanding of investment behavior controlling for confounding influences.

Background

The theory of investment has been codified in a rich and diversified literature, which for present purposes can be best understood through two seminal pieces. Pindyck [1991] summarizes this literature and introduces the notion that because decisions to purchase capital goods are characterized by irreversibility, uncertainty, and the ability to delay the decision until a future time period, investment opportunities can be arrayed and analyzed in a manner similar to financial options. In another paper Dixit [1992] retraces many of the steps taken by Pindyck while emphasizing to a greater extent the notion of "optimal inertia." Taken together, these two pieces essentially formulate the theory of investment as a problem of optimal timing rather than the more traditional approach of choosing an optimal level.

Pindyck takes as a point of departure the notion that a firm's value lies in part in the option value of the alternatives open to it. Once exercised, this option value is translated from a potential activity to an actual one and evaluated in real terms rather than financial ones. However, the firm cannot, in general, retrace its steps, exchanging the real asset for a financial one at full value. Hence, the firm waits to invest while it gathers information to reduce uncertainty. As we note below, there is an implied value of information problem embedded in the approach.

Dixit emphasizes the symmetry between the decision to invest and the decision to abandon (cease operations). Casting his work (upon which Pindyck draws heavily) as a theory of optimal inertia, he relates uncertainty to the size of the hurdle rate needed to undertake investment. He then demonstrates that firms have a symmetric incentive to maintain unprofitable operations to degrees that exceed the point dictated by the payment of variable costs. Thus, the firm both invests and disinvests more slowly than under the traditional model.

Taken together, this literature yields three primary conclusions:

1. When there is uncertainty, irreversibility and opportunity to wait, firms delay decisions to modify durable good holdings longer than traditional analysis would forecast.
2. The scope of delay is directly related to the uncertainty associated with the performance of the durable good. By inference, the greater the departure from current practices, i.e., the more innovative the technology, the greater will be the inertia.
3. From the perspective of the firm, the cost of this inertia is less than is evident from the cash flow implications of a simple investment analysis, because the market attaches values to the options available to the firm.

Lind [1992] seizes upon this literature as a means to place a value on R&D related to global warming issues. He argues that the opportunity to wait, coupled with uncertainty and irreversibility, is analogous to segmenting a larger decision such as immediate implementation of a carbon tax into a sequence of smaller steps, and permits comparison of the costs and benefits of gathering additional information to the costs and benefits of some irreversible action.

Thus there exists an alternative theoretic foundation for examining investment decisions. However, the data that exist make direct testing of such theories difficult. For this reason we turn to the arena of experimental economic methods to conduct empirical tests of this complex investment process.

The use of direct experimental testing of economic principles and theories began in the 1950's and grew the 1960's and 1970's as economists such as Vernon Smith and Charles Plott began to demonstrate the benefits of

studying economic phenomenon under controlled conditions. The field has taken off in the 1980's and 1990's and is now recognized through numerous publications in prestigious economic journals.

A summary of this approach is contained in Smith [1992] that was prepared for the Secretary of Energy Advisory Board Workshop on Innovative Tools for Economic Analysis. This paper reviews the influence of experimental methods in the analysis of many economic issues such as energy policy. In this paper, Smith points out the difficulty of obtaining a complete picture of economic transactions in market settings. In most cases only aggregate data are available, with many compounding and confounding components that make analysis difficult. Regarding standard econometric data he points out that

... [data] limitations are spotlighted in standard econometric textbooks which begin with an ideal world in which the observations for econometric analysis have been derived from an experiment in which the independent variables are controlled treatments with randomization used where appropriate. Necessarily the econometric exercise has been largely defined in terms of what to do when your data do not come from such an experiment. Experimental economists have sought to open up the constraints on traditional empirical methodology in economics by means of a direct investment effort in experimental methods (Smith [1992] p.1).

Economic experiments that study the operation of markets have a core of three elements: (1) the value/cost environment; (2) an institution that defines the communication of information within the market; and (3) behavior of the market participants. The environment defines the conditions and the motivations for exchange. Examples include the preference structure imposed on subjects, their initial endowments, and the technology or costs that they must face in the experiment. The institution defines the rules under which subjects make their decisions. One easily recognizable institution is an auction. As an example, under the English Auction, such those for paintings or antiques, participants raise the bid on an object until only one bidder remains. This person must then purchase the item up for auction, and pay the price she or he last bid. Institutions may be quite formal or casual depending on the issue under examination. The final element is behavior; the actions that the subjects reveal in the experiment. The experimentalist controls the environment and the institutions and then observes the behavior to draw inferences about the market, theory or policy being tested.

It should be noted that experimental economic methods are one of the strongest links between the "science" of economics and other "hard" sciences such as physics and biology (Smith [1992]). The rigorous, controlled settings employed by the experimental economist have appeal to other disciplines that demand this kind of controlled method to investigate the validity of their hypotheses. Thus, experimental methods offer a strong tool for gaining more widespread acceptance of many of the theories within the field of economics. Therefore, we will investigate issues of investment decisions as expounded by Dixit and Pindyck in the controlled environment of the laboratory as described by Smith.

Experimental Methods and the Study of Investment Decisions.

This section will paint, in broad strokes, our protocol or the investigation of investment decisions by various economic agents using experimental economic methodology. We begin by restating that in its simplest form experimental economics places real people in a controlled economic environment to test various economic theories. For example, in a basic situation, human subjects are placed in a room with instructions to buy and sell a "good" to test predictions implied by the Law of Supply and Demand. In more complicated experiments, subjects may act as managers of firms who decide whether it is best to enter a particular market and compete for customers, or to stay on the sidelines. Subjects are provided positive incentives to behave rationally through cash payments based on the profits earned through their decisions.

The tasks described below outline a multi-year program of research in the area of investment behavior. Each focuses on an aspect of the investment process described by the theory above. The first task outlines baseline experiments that strip the problem to its most salient elements. The results of these experiments will be used to judge the performance of the more complicated ones in the following tasks. In Task 2, the information about investment payoffs is manipulated. Portfolio effects and trade offs are examined in Task 3 by introducing multiple investment opportunities. In Task 4, we change the basic structure of the baseline experiments to understand better the effects of investment decisions across time. Finally in the last task, we examine various market structures and the influences that they exert on investment decisions.

Task 1: Baseline Experiments.

In these experiments, as in all the other experiments, subjects will act as managers of firms. In the baseline experiments, these firms have the potential to invest in a new technology for their production process. They will be given an endowment of cash, and will be asked if they would like to put some or all of this in a fund (the investment) that has some type of uncertain return. This distribution of possible returns on their investment represents the present value of the future stream of the revenues. For these experiments, the distribution will be known to all subjects. That is, subjects will know all the possible values and respective frequencies for the return on their investments. Any and all money not put in the investment is placed in a safe haven. These funds will earn some rate of return that is less than the expected return from the investment (though it is possible that the realized return from the investment will be below the safe haven).

When all subjects have made their investment decisions, each will draw their return from the distribution (with replacement). Subjects will be paid their return, and the experiment will move to another investment decision. In subsequent rounds, the experimental procedure works much the same with the distribution of investment returns varying in order to discover evidence of the "hurdle" rate and other phenomenon described above.

In this experiment, Marshallian investment theory predicts that as long as the new investment's certainty equivalent is greater than the return from the safe haven, subjects should invest. However, for the models of Dixit and Pindyck to be theoretical improvements, this rate of return must be systematically and predictably greater than the certainty equivalent rate. For example, if Marshallian theory best describes investment behavior, then if the expected rate of return in the investment is even marginally higher than the safe haven funds should be placed in the investment. However, if subjects do systematically put some funds into the safe haven, we can hypothesize that investment behavior follows another paradigm such as those discussed above. If the traditional Marshallian approach is shown to predict behavior in the laboratory, we will begin to manipulate parameters of this environment to test its limits. If the data support another theory, we will again begin to manipulate the experimental environment in order to understand more about this behavior. The following tasks are predicated on the assumption that Marshallian theory will fail in these baseline experiments; however, the tasks can be easily modified to investigate the bounds of this traditional theory.

Task 2: Investments Under Uncertainty.

An important variation will be making the distribution of payoffs more uncertain. This can be done by changing the way the payoff distribution is presented to the subjects. In this way, we will be able to capture many important aspects of the uncertainty of making investments in new technology. In a case of less certainty, subjects will be given the mean, the variance, and the bounds of the distribution of returns. In decreasing information from this point, subjects will receive only the mean and variance. Removing more certainty, subjects will receive only the mean of the distribution with the variance and the bounds left unknown. In a final case, it is possible to imagine a case where the distribution of returns is completely unknown. Here, subjects can pay to receive informative draws from the distribution.

Task 3: Multiple Investment Opportunities.

Another line of investigation is the area of multiple investment opportunities. In this case, instead of subjects deciding between a single investment or the safe haven fund, there are two or more investment opportunities. In some cases, these options will be mutually exclusive (i.e., subjects can invest all or nothing in one or another), or may have some divisibility (i.e., some funds may be invested across a number of different options). In the first case we examine competing technologies, and in the second, we look at ways that firms can diversify their portfolio of technologies and mitigate the risk they face.

An important aspect to this line of inquiry is the manipulation of information, as was described above. That is, while at first it may seem obvious that a subject would invest all her/his funds in the investment with the highest expected return, this is predicated on the assumption that all the distributions are the same. Therefore, we need to vary the distributions around the various investment options. This means that while two distributions may have the same expected payoff, we may find that one distribution that does not have a negative payoff but has no large high payoffs may be preferred to one with high potential gains matched by high potential losses.

This is only one example of how the information variables above can be combined with the multiplicity of investment opportunities. Each environment captures a different aspect of the investment process in the field, and each gives us a deeper insight into how investments are made and what affects their outcome.

Task 4: The Time Element in Investment Decision.

To this point our investigation intentionally has ignored time. Investment decisions have been made in a single time periods and the returns have been paid out in the same period. This task will lift this restriction and make time a part of the decision process. The theory cited above suggests that time allows firms to postpone an investment today to wait for more or better information in the future. As noted, this is a rational strategy if the payoff to waiting is greater than the income foregone by not investing (and therefore earning some return) today.

To introduce time into our experiments, we will have payoffs for an investment drawn every period throughout the experiment (or life of the investment). That is, if a subject chooses to invest in the first period, he or she will draw an income from some distribution in subsequent rounds. By adding this dimension, it is possible for subjects to postpone investing at present, and wait until more information is available or until a more favorable distribution is presented. It may be explained that more information will be made available as the experiment progress, or that different distributions will be attached to investments in different rounds to point out that there is some difference between investing right away and waiting and postponing the investment.

It should be noted that the information or change of distribution need not necessarily be a positive change. In other words, the future information may be that the investment will only pay off for two more rounds and then it will return nothing. In this case the subject has already foregone some number of rounds of income. Likewise, a change in information may indicate the introduction of a distribution with a higher probability of a negative income from the investment. Because all of these possibilities exist, it is not always true that waiting is an optimal strategy. Another important issue brought out in the theory is that it might be optimal to spend a little today to insure the option to invest more at a later date. This can be used to examine some issues of sunk cost.

In terms of our experiments we can give subjects the chance to place a little of their cash in a pot that will allow them to take advantage of investment opportunity in the future should they so desire. They are told that whatever they spend now cannot be returned to them if they choose to invest, but some part may be applicable to the price of the investment. In this way we can begin to examine directly the issues of option value and the sequencing of investment.

Another facet of investment behavior characterized by moving to this more realistic representation of time is the option of abandoning or mothballing projects if they become unprofitable. In some treatments, subjects will be given the option of setting the investment aside (abandoning it) and not drawing a return on it. In other treatments, subjects can set aside the investment but reserve the right to pay some price in the future to bring it back and draw income from it again.

All of these treatments can be combined with salient variables that are identified in Task 2 or 3. Interaction of time and various information conditions as noted will be quite important in developing our understanding of investment decisions in the field. This is also true of the interactions with various uncertainty conditions. Multiple investments will open the door to how subjects react to shifting portfolios of investment opportunities.

Task 5: The Effect of Market Structure on Investment.

The experiments described thus far have assumed that each subject is making investment decisions in a vacuum; the decisions of one subject have no impact on the investment opportunities or payoffs of the other subjects. In market terms, the subjects are monopolists in their individual markets with no threat of entry of a competitor. Yet with a few notable exceptions, this is not the structure of most markets.

In a purely competitive market, if one firm is able to make greater returns than the competition, others will enter (or adapt that firm's technology), which will bring profits back into equilibrium. While this is also an extreme case, it points out that interactions between firms are important to consider in the analysis. If one firm invests in a new technology, others may be forced to take action just to keep up.

There are many ways to introduce this type of interaction into the laboratory setting. It is possible that the investment opportunity will be limited (there is only a discrete number of funds that can be divided among a number of subjects). In this case, an auction may be necessary to achieve allocative efficiency amongst the subjects. In another treatment, subjects have to invest in order to keep unwanted competition from entering the market and

running the established firms out of the experiment. Pay-offs may also be affected by the number of subjects investing in a given opportunity or fund. In a simple example, the draws of return from the distribution may be made without replacement, thus complicating the investment decisions and their impacts.

As with the experiments in the other tasks, the market structure experiments can be combined with aspects of the other scenarios that we more fully understand. This will also allow for the relaxation of some of the controls of the laboratory and bring the actions of subjects closer to that which is observed in the field. While this task list is not meant to be an exhaustive description of the possible impacts and interactions that affect investment decisions, it seems to capture many of the important ones.

Conclusion

The tasks outlined above form the basis of a multi-year research program in the area of investment decisions. Because of the time involved in the careful planning and execution of experiments, and the attention that must be paid to the analysis of the results, we will undertake no more than Task 1 and 2 in the first year of this project. Once we have this framework in place, we can more easily address the other tasks and the issues they embody.

It is also important to note that this entire research program is but a first step in itself. It does not address the important area of consumer investment. As pointed out above, a large part of the problem with technological innovation lies in acceptance by consumers. Yet it can be argued that individuals do not necessarily make investment decisions the same way firms do. That is, we do not view the purchase of a refrigerator as having a return on our investment. Instead, we see many of these as things that we need to exist in the modern world, or live the life we would like.

But even at the end of a research program on consumer investment behavior, we would only have an understanding of two investment sectors in isolation. It will be important to put both parts together and observe how they interact. We can create a situation where the return to a firm's investment decision is not determined by a random draw from a distribution, but by the investment decision of a consumer also in the experiment. Thus, as we begin to understand and to draw all the pieces together, we can more clearly understand the entire process.

References and Selected Bibliography

Davis D.D. and Holt, C.A. [1993], *Experimental Economics*, Princeton, Princeton University Press.

Dixit, A. [1992], "Investment and Hysteresis," *Journal of Economic Perspectives*, Vol 6, No 1, p 107-132.

Lind R.C. [1992], "Global Warming Policy: The Need to Consider Real Options and Take a Sequential Approach," Final Report submitted to the United States Environmental Protection Agency, May.

Lind, R.C. [1992a], "The Analysis of Intertemporal Decisions Under Uncertainty in Energy Policy," Prepared for Workshop by the Task Force on Economic Analysis and Modeling Research Agenda, Secretary of Energy Advisory Board.

Pindyck, R.S. [1991], "Irreversibility, Uncertainty, and Investment," *Journal of Economic Literature*, Vol 29, No 3, p 1110-1148.

Smith V.L. [1992], "Experimental Economics: Evaluating Microeconomic Policies and Institutional Designs," Prepared for Workshop by the Task Force on Economic Analysis and Modeling Research Agenda, Secretary of Energy Advisory Board.

Thaler, R.H. [1992], *The Winner's Curse: Paradoxes and Anomalies of Economic Life*, New York, Free Press.

FORECASTING FUTURES MARKETS

MODERATOR **Ethan Smith, U.S. Geological Service**

PRESENTATION 1 **Ronald Babula, Phil Colling, Gregory Gajewski, Economic Research Service**

PRESENTATION 2 **William Melick, Charles Thomas, Federal Reserve Board**

PRESENTATION 3 **David Torgerson, Economic Research Service**

Forecasting Futures Markets

Employing Historical Price Dynamics to Aid in Forecasting Price Effects: The Case of Lumber Price Increases on Housing-Related Prices

Ronald A. Babula, Phil L. Colling, and Gregory R. Gajewski, Economic Research Service, U.S. Department of Agriculture

The increased volatility of the U.S. lumber market in late-1992 and early-1993 was evident from the large swings in the prices of lumber and lumber futures (hereafter, futures price) chronicled by Babula, Gajewski, and Colling (1993) and Gorte (1993). Lumber futures prices doubled between October 1992 and March 1993, and then fell 25 percent in April. Prices of lumber have also been moving widely, though by less than futures prices. Lumber prices were nearly 34 percent higher over the year ending April 1993, and then fell 4.7 percent in May. Since May 1993, lumber and futures prices have remained volatile, while levels have receded somewhat from early-1993 highs. Such swings are the largest since November 1978-August 1979 when lumber futures price rose 34 percent, and since November 1982-July 1983 when lumber price rose 22 percent with heightened housing activity during the economic recovery of the early 1980's (Gorte 1993; Babula, Gajewski, and Colling 1993).

Reasons for the increased volatility in the prices of lumber and lumber futures include: a recurring lumber demand increase during recovery from such recessions as that of the early-1990's; a reduction in timber harvests from the Pacific Northwest because of legal and environmental constraints; and perhaps because of some increased Japanese purchases, as that nation emerges from its recession (Gorte 1993). Whatever the cause, the movements (increases and decreases) over the last year should have impacts on lumber-related prices in the construction and housing sectors of the economy. Our aim is to use historical dynamic patterns of interactions among lumber and lumber-related housing and construction prices to discern the future effects and effect patterns elicited by the recent wide swings in the prices of lumber and lumber futures.

We use data-oriented statistical methods to ascertain how much and with what dynamic patterns related construction and housing prices should rise (fall) due to rises (declines) in lumber and lumber futures prices. For reasons stated below, we use vector autoregression (VAR) methods, in two separate experiments, to map the historical dynamic effects of:

- (1) a one-time 10 percent rise in lumber price on futures price, the price of construction materials, the consumer price of housing, and the consumer price of shelter (hereafter called the lumber price experiment).
- (2) a one-time 10-percent rise in lumber futures price on lumber price, the price of construction materials, the consumer price of housing, and the consumer price of shelter (hereafter, the futures price experiment).

Specifically, we set out to answer six questions concerning how shocks in lumber and lumber futures prices dynamically affect the remaining four respondent prices in each experiment: (a) What are the reaction times required for the prices to begin responding to each shock?; (b) What dynamic patterns do the monthly responses take?; (c) To what degree do the prices ultimately respond to each shock?; (e) What are the strengths of relationships among the five prices?; and (f) What are the differences in (a) through (e) elicited by a shock in lumber prices as opposed to a shock in lumber futures price (hereafter futures price)? Question (e) includes whether futures price responds to a greater (lesser) degree to lumber price movements than lumber price responds to futures price movements. Question (e) also includes whether lumber price movements elicit more (less) pronounced materials and housing-related price effects than effects elicited by futures price movements.

While not forecasts, such dynamics reveal how, on average, construction- and housing-related prices have historically responded to changes in the prices of lumber and lumber futures. These dynamics therefore provide insight concerning patterns, timing, and size for specific monthly 1993-1994 forecasts in the wake of the recent volatility in lumber and futures prices. As with all models, actual event-specific 1993-94 levels and movement patterns of lumber-related prices may differ from what the model's captured long run dynamic patterns would

predict and what the model's impulse response function would simulate. Yet the average historical dynamic patterns reflected by the model constitute an evidential point of reference for determining which 1993-94 point-forecasts (and forecasted price patterns) from competing models are "reasonable" by historical standards to characterize what may happen in the wake of the recent movements in lumber and lumber futures prices.

Methods, Model, and Data

Common sense, observed history, and economic theory suggest that large increases in the price of such a major construction/housing input as lumber should elicit increases in the prices of related lumber, construction materials, and housing/shelter prices. And our results presented below confirm as much. Answering questions (a) through (f) implies focusing not so much on if respondent prices react, as on how they react to shocks in lumber and futures prices. Conventional econometric models that intensively use static economic theory are equipped to handle questions concerning what happens at the static equilibria before and after the shock (Bessler 1984a, b). Such "structural" models often have little or nothing to say about what occurs dynamically between equilibria to the observed choice variables -- that is in answering the above six dynamic questions about each experiment's price responses (Bessler 1984a, b). Such time series models as Sims' (1980) vector autoregression (VAR) and Johansen and Juselius' (1990, 1992) maximum likelihood methods of vector error correction (VEC) models for cointegrated systems are equipped to address dynamic issues (a) through (f).

We chose to model the five-price system with a vector autoregression (VAR) model in logged levels over a vector error correction (VEC) model. For a VEC to be appropriate, the variables in levels must be individually nonstationary (integrated of order d , $d > 0$), and exhibit stationary behavior as a system. That is, for a VEC to be appropriate, the five prices must be individually nonstationary, but form at least one stationary long run (cointegrating) relationship such that the prices move tandemly through time (Granger 1986; Hendry 1986; and Johansen and Juselius 1990, 1992). Yet results from Dickey-Fuller (1979) tests performed on the logged series suggest that evidence at the five-percent significance level is sufficient to reject the null hypotheses that each levels-variable is nonstationary.¹ The variables are stationary, thereby permitting us to model the system of lumber, futures, materials, housing, and shelter prices as a system (Johansen and Juselius 1990, 1992).

VAR econometric methods are equipped to address the dynamic inter-equilibria issues (a) through (f). The technique is data-oriented and imposes as few a priori theoretical restrictions as possible, so as to permit the dynamic regularities in the time-ordered data to reveal themselves (Bessler 1984a, b).

The literature is replete with detailed summaries and derivations of VAR methods. Those interested should consult Sims (1980), Bessler (1984a, b), and VanTassel and Bessler (1988). Our five-equation VAR model takes the following form:

$$\begin{aligned}
 x_t = & a_{0,x} + a_{x,T} * \text{TREND} \\
 & + a_{x,1} * \text{LUMBER}_{t-1} + \dots + a_{x,10} * \text{LUMBER}_{t-10} \\
 & + a_{x,11} * \text{FUTURES}_{t-1} + \dots + a_{x,20} * \text{FUTURES}_{t-10} \\
 & + a_{x,21} * \text{MATLS}_{t-1} + \dots + a_{x,30} * \text{MATLS}_{t-10} \\
 & + a_{x,31} * \text{HSG}_{t-1} + \dots + a_{x,40} * \text{HSG}_{t-10} \\
 & + a_{x,41} * \text{SHELT}_{t-1} + \dots + a_{x,50} * \text{SHELT}_{t-10} + R_{x,t} \quad (1)
 \end{aligned}$$

The subscript t denotes the current value, while subscript $(t-i)$ refers to the i th lag from the period- t value. The upper-cased subscript T represents the coefficient on time trend or TREND. On the left hand side, x = LUMBER, FUTURES, MATLS, HSG, and SHELT. The latter variable labels reflect, respectively, the prices of lumber, lumber futures, construction materials, consumer housing and consumer shelter. The coefficient with a nought subscript represents the intercept. $R_{x,t}$ represents white noise residuals.

Monthly producer and consumer price indices (PPIs, CPIs) obtained from the U.S. Bureau of Labor Statistics (BLS) served for all but the lumber futures price. The BLS PPI for lumber serves as lumber price

¹The procedures for Dickey and Fuller's unit root test are described in Fuller (1976) and Dickey and Fuller (1979). The "pseudo- t " values for the five prices ranged from -3.35 to -6.92. Since all five pseudo- t values were negative, and of an absolute value exceeding that of the τ_α critical value of -2.89, then evidence at the five percent significance level was sufficient to reject the null hypotheses that each variable in levels was nonstationary. The variables in levels are integrated of order zero ($I(0)$). Hence were the five prices modeled as a VEC in Johansen and Juselius' equation 1.2 (1990, p. 170), then π would be of full rank.

(LUMBER). The wholesale price of construction materials (MATLS) is represented by the PPI for construction materials. The CPI of all urban consumers for housing services represents the consumer price of residential housing services (HSG). What was desired was a broader price than that of newly-built home-owned residential units. The CPI for housing includes prices of owned and rented shelter services, as well as prices of upkeep, household expenses, and furnishings that are to some extent, lumber-dependent. The CPI of all urban consumers for shelter (SHELT) is a more narrowly defined price than the housing CPI. The CPI for shelter reflects the consumer price of rented and owned residential shelter.

The lumber futures price used is the closing price of the nearby futures contract (that contract nearest to expiration). To avoid potential problems associated with contract delivery, we did not use the nearby contract during its delivery month. If the nearby contract was into its last month, which is the delivery month, the next nearby contract was used. To stay consistent with the timing of the BLS index data, we used the closing price on the Tuesday of the week containing the 13th of that month.

Following VanTassel and Bessler (1988) and Bessler (1984a, b), the VAR model's lag structure was chosen using Tiao and Box's (1981) likelihood ratio test procedure (also see Lutkepohl 1981). Results (not reported here) suggest a 10-order lag. Each equation includes a constant, a time trend to account for time-dependent influences not of direct interest to this study, and a series of 11 indicator variables to account for seasonal influences.

The five VAR equations may have contemporaneously correlated innovations. Failure to correct for contemporaneously correlated current errors will produce impulse responses not representative of historical patterns (Sims 1980). A Choleski decomposition was imposed on the VAR for each experiment to orthogonalize the current innovation matrix, such that the variance/covariance matrix was identity in each experiment. Choleski decompositions resolve the problem of contemporaneous feedback.

Each decomposition requires an arbitrary imposition of a Wold causal ordering among the current values of the dependent variables (see Bessler 1984a, b). In the lumber price experiment, the chosen ordering was LUMBER, FUTURES, MATLS, HSG, and SHELT. The chosen ordering in the futures price experiment was the same except that the ordering of the lumber and lumber futures prices was reversed. The choices of these orderings were based on a number of considerations. First, common sense, observed history, and economic theory all suggest that there is a valid line of causality from lumber and lumber futures price movements to the lumber-dependent prices of construction materials, housing, and shelter. Second, futures and lumber prices move tandemly. Third, Sims (1989) and Bessler (1984a, b) suggest that when there is a valid line of causality, as from lumber and lumber futures prices to the other prices, then the variable shocked is placed atop the ordering. Fourth, the bottom three prices in each experiment were ordered as MATLS, HSG, and SHELT. Lumber is a prime residential housing input, so a shock in lumber or futures price could be expected to influence materials price before housing or shelter prices, suggesting MATLS' third place in each ordering.

VAR Model Diagnostics

Data for all five prices was available from January 1974 through December 1992 or 1974:1-92:12 (hereafter, months are denoted numerically with 1 reflecting January and 12 reflecting December). A 24-month period, 1974:1-75:12, was chosen as a period over which we applied the Tiao-Box lag search procedures. The 24-month period, 1991:1-92:12, was reserved as the validation period, over which the model's out-of-sample forecasts were evaluated. This left the period 1976:1-90:12 as an estimation period over which we initially estimated the model. We provide in-sample and out-of sample diagnostics for the initially estimated model. Because this model met the accepted diagnostic standards of in-sample fit and out-of-sample predictive accuracy, we then re-estimated the model over the "simulation" period of 1976:1-92:12, which includes the estimation and validation periods. The model estimated over the simulation period generated the two experiments' impulse response results and the forecast error variance (FEV) decompositions that we analyzed. The re-estimation was necessary to incorporate the maximum information possible for simulation of the near-future (1993-94) conditions (see Babula and Bessler 1990). All estimations and simulations were generated by Doan's (1990) package, Regression Analysis of Time Series (RATS).

The diagnostic test results for the model are reported. Ljung-Box portmanteau or Q tests and Dickey-Fuller tests on VAR equation residuals are two in-sample tests concerning model adequacy. Theil-U values are calculated for the 24 step-ahead forecasts permitted by the validation period, and provide insight on how reliably the estimated VAR model predicts events yet "unseen" beyond its own sample (Bessler and Kling 1984; Doan).

The Ljung-Box Q values, calculated for an equation's residuals, tests the null hypothesis that the equation has been adequately specified (Harvey 1990; Granger and Newbold 1986, p. 100). Evidence at the five percent significance level was not sufficient to reject the null hypotheses that each VAR equation was adequately specified. The Q values ranged from 32.4 to 46.4, and were less than the 58.1 critical chi-square value.

Stationarity (white noise residuals) of the estimated equations is required. We therefore tested for the stationarity of the innovations or residuals of each equation using the Dickey-Fuller (DF) tests. The pseudo- or t-like values on the nondifferenced lagged regressors in the τ_μ and τ_τ tests ranged from -14.04 to -14.46. Not only were these values negative, they also had absolute values which exceeded the critical values of -2.89 for the τ_μ test and -3.45 for the τ_τ test (Fuller 1976, p. 373). Evidence at the five percent significance level was sufficient to reject the null hypotheses that each VAR equation's residuals were nonstationary. Combined with the Ljung-Box evidence of model adequacy, these ADF tests on VAR equation residuals point to model adequacy.

Each equation (initially estimated over the 1976:1-90:12 period) generated as many "step-ahead" forecasts as the validation period would allow. The forecasts were run through a Kalman filter. Thus the 24-month validation period permitted 24 one-step-ahead forecasts; 23 two-step-ahead forecasts, etc. Theil-U values were provided for each of the 24 forecast horizons (i.e., 24 Theil-U's for each equation). A Theil-U value of less than unity suggests a superior (more accurate) performance at a horizon than the "naive" or random walk model forecasts. A naive model forecast equals the previous period's actual value. Further, a sub-unity Theil-U suggests that there were gains, in terms of forecast accuracy, from modeling the VAR equations as opposed to expending relatively no modeling effort through naively forecasting (Babula and Bessler 1990). Gains to modeling were apparent in the following three equations with the following proportions of Theil-U values below unity: 15/24 for materials price and 24/24 for housing and shelter prices.

Lumber and futures price forecasts confirm the FEV decomposition results (below) suggesting that these two prices are highly exogenous in each experiment. Typically, highly exogenous variables do not predict well beyond the sample, although they can (and do) add much to the prediction and simulation of relevant endogenous variables (MATLS, HSG, and SHEL). No more than five of the lumber and futures price Theil-U values were sub-unity. Lumber price is affected by other influences beyond the price influences of housing-related markets. Lumber price does not forecast well, as expected, and two reasons are cited. First, only one lumber futures price contract is traded on the Chicago Mercantile Exchange: the contract for Spruce-Pine-Fir, 2x4, Standard and Better Grade (Babula, Gajewski, and Colling 1993). This lumber is harvested primarily in Canada. Hence the lumber futures price, while perhaps a good "surface signal" of lumber price movements and anticipations, is too narrowly defined to serve as an instrumental variable for, and directly reflect, the supply and demand forces in all of the different U.S. lumber markets (Babula, Gajewski, and Colling 1993). And second, the anticipatory nature of the futures price renders its prediction hampered by expectations, and sometimes even emotions, which are difficult to capture in a model.

We conclude that, on balance, all five equations are adequately specified. Most (three: MATLS, HSG, and SHEL) of the five equations "passed" all three diagnostic checks, while all five equations passed most (at least two) of the three sets of diagnostic checks. While lumber and futures prices failed to generate Theil-U's that suggest gains to modeling, these equations did generate evidence of adequate model specification by accepted standards.

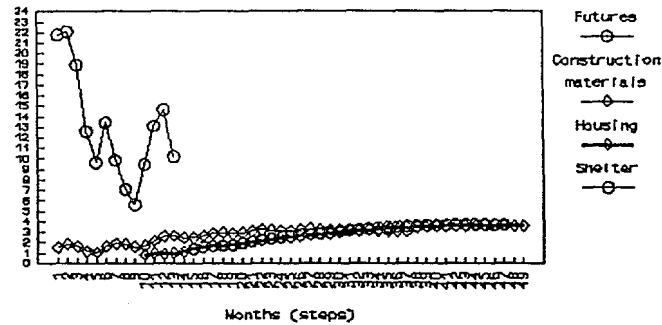
Impulse Responses in Related Prices to Shocks in Lumber and Futures Prices

The impulse response function simulates, over time, the effect of a one-time shock in one of a VAR's series on itself and on other series in the system (Bessler 1984a, b). Increases of 10 percent were chosen because one does not currently know at this writing to what point lumber and lumber futures prices will ultimately climb or recede. A 10-percent shock is conveniently sized because of the VAR model's linearity. The shapes of both experiments' impulse patterns in figures 1 and 2 would remain the same, with only the scales of the vertical axes varying with differently sized shocks.²

The impulse responses are reported for the lumber price experiment in figure 1 and the futures price experiment in figure 2. Dynamic aspects obtained from the impulse response results are summarized in table 1.

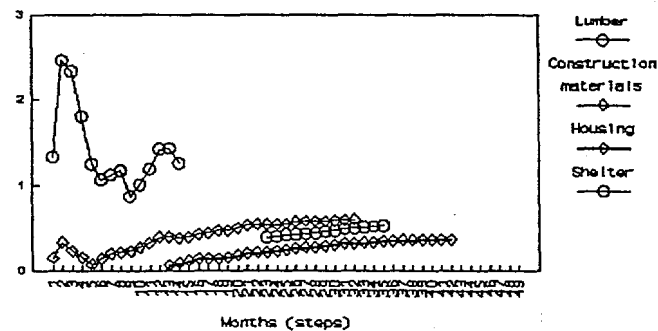
²For example, one can, by the model's linearity, characterize the impulse response simulations to 20-percent shocks by simply multiplying the impulses from the 10-percent shock experiments by a scalar of 2.0.

Figure 1. Impulse responses in the prices of
lumber futures, construction materials, and
housing services from a 10% rise in lumber price
Percent change in price



Impulse responses are statistically non-zero at
the 5-percent significance level. Impulses in
the lumber futures price are included.

Figure 2. Impulse responses in the prices of
lumber, construction materials, and housing
services from a 10% rise in lumber futures price
Percent change in price



Impulse responses are statistically non-zero at
the 5-percent significance level.

FFC/93 Papers & Proceedings

Table 1 -- Dynamic aspects of price response patterns of the lumber and futures price experiments.

Dynamic Aspect:	Lumber Price	Futures Price	Materials Price	Housing Price	Shelter Price
<u>Reaction times (months):</u>					
Lumber price exp.	---	0	0	9	14
Futures price exp.	0	---	0	12	22
<u>Response directions:</u>					
Lumber price exp.	---	rise	rise	rise	rise
Futures price exp.	rise	---	rise	rise	rise
<u>Response patterns:</u>					
Lumber price exp.	---	sharp, then decay	shallow, then accelerate	shallow, then accelerate	shallow, then accelerate
Futures price exp.	sharp, then decay	---	shallow, then accelerate	shallow, then accelerate	shallow, then accelerate
<u>Response durations (months):</u>					
Lumber price exp.	---	13	37	40	33
Futures price exp.	14	---	32	31	13
<u>Multipliers:</u>					
Lumber price exp.	---	1.6	0.48	0.54	0.50
Futures price exp.	0.38	---	0.22	0.26	0.10

These results reflect patterns averaged over all of the sample's interactions. The results reflecting such average historical dynamics are valid in characterizing the dynamic effects of the recent rises in lumber and lumber futures prices insofar as recent/current conditions are similar to history's average conditions captured by the model. These long run average dynamics can provide standards by which to judge the "reasonableness" of monthly 1993-94 point-forecasts of competing models, as well as provide some indication on how forecasted price levels will be dynamically achieved.

Impulse responses are approximate changes in the non-logged prices, and are not price levels. Klok and Van Dijk's (1981) Monte Carlo methods generated t-values for each impulse response. These values test the null hypothesis that each impulse is zero-valued. We focus on those impulses that were statistically nonzero at the five percent significance level.

Impulse responses: the lumber price experiment

Reaction times required for futures, materials, housing, and shelter prices to respond to shocks in the lumber prices have varied historically. As perhaps expected because of its speculative and anticipatory nature, futures price has begun reacting during the same month as (within 29 days of) the lumber price movement. And perhaps because of the short times required for construction materials to be manufactured, construction materials prices have also begun reacting to the lumber price shocks during the same month as the shock. On average, reaction times for the remaining prices have been longer: nine months for housing price and 13 months for shelter price. Reaction times of nine to 13 months may reflect the lags inherent in planning, constructing, and marketing (selling or renting) residential units. In addition to shelter price, housing price includes such other housing-related prices as furnishing and upkeep service prices that may be lumber-dependent, and that may respond to lumber shocks sooner than shelter prices. Hence, the housing price's reaction time is less than that required of the shelter price's responses.

Generally, history has had rises and falls in lumber price elicit similarly-directioned movements in futures, materials, housing, and shelter prices. While there may be event-specific examples in the past where this was not true, the model captures patterns that are averages over all interactions. Hence, these average patterns suggest that lumber price and the other prices generally moved up and down in a generally tandem manner.

The impulse patterns of the materials, housing, and shelter prices have been more enduring than those of futures price. Dynamic history has had futures price respond rapidly, sharply, and then take-on a gradually decaying pattern, while lasting just over a year, 13 months. This is consistent with recent media reports of sharply escalating and volatile lumber futures prices (see Taylor 1993). The impulse response patterns of the materials, housing, and shelter prices have differed from those of futures prices in being generally more muted in magnitude; in lasting for longer periods of time; and in taking-on patterns where responses gradually accelerate, rather than sharply decay, over the response period. These initially shallow patterns begin at low magnitudes and then accelerate to higher magnitudes that are still muted when compared in magnitude to the futures price impulses. The three patterns have endured from month-one through month-37 or about three years for materials price; from month-10 through 49 or just over three years for housing price; and from month-15 through month-47, or almost three years, for shelter price.

Multipliers of price response to lumber price shocks are calculated, and suggest the degrees to which prices ultimately respond to the shocks.³ The multipliers provide history's average percentage point increase per point rise in lumber price. The futures price's multiplier for the lumber price experiment is 1.6 suggesting that, on average, each 10-point rise (fall) in lumber price has historically elicited a larger 16 percent rise (fall) in futures

³By a VAR model's very definition, each variable is posited as a function of a specified number, here 10, lags of each endogenously modeled variable in the system. Hence a one-time shock to the system places all five variables into cycles of monthly pulsation, including the shock variable. Insofar as the data levels are modeled in natural logarithms, then shocks to and impulse responses in, the logged variables constitute proportional changes in the non-logged variables, and percent changes in the non-logged variables when multiplied by 100. As an example, consider the materials price's multiplier. One first adds up the 37 statistically nonzero materials price impulses to obtain a cumulative percent change in MATLS response. One secondly sums the corresponding shock variable impulses into a cumulative change in lumber price. Finally, one then divides the percent change of the shock variable into the percent change in the response variable to obtain, here, the materials price's multiplier of percentage point response to a point change in lumber price. These are history's average responses. One calculates such an elasticity for the four respondent prices in each of the two experiments. A positive multiplier suggests that rises/falls in the shock variable have generally elicited the same in the respondent variable for which the multiplier is calculated.

price over a 13 month period.⁴ Likewise, each 10-percent increase in lumber price has elicited, over periods ranging from 33 to 40 months, increases of 4.8 percent in materials price; 5.4 percent in housing price; and 5.0 percent in shelter price. The effects of the lumber price shock have been less than one-for-one on all prices except futures.

Impulse responses: the futures price experiment

The dynamic aspects with which lumber, materials, housing, and shelter prices have historically responded to futures price shocks are reported in table 1 and figure 2. Lumber and materials prices have historically reacted during the same month as the futures price shock. The housing and shelter prices have historically taken from 12 to 22 months to start reacting to the futures price shock. Generally, history has had lumber futures price and the four other prices move in similar directions.

Futures price increases have elicited lumber price increases that were pronounced early in the response cycle, that have decayed through the cycle's remainder, and that have endured, on average, for just over a year (14 months). The impulse responses of the materials, housing, and shelter prices have differed from those of lumber price in being generally more muted in magnitude; in lasting for longer periods of time; and in taking-on patterns that gradually accelerate, rather than sharply decay, over the response cycle. Construction materials price responses began in month-one and lasted for 32 months. Housing price impulses endured for a 31-month period beginning in month-13. Shelter price responses have lasted for 13 months after requiring nearly two years (22 months) to begin.

Multipliers for the futures price experiment are reported in table 1. These multipliers suggest that responses to futures price movements have been less than one-for-one. On average historically, a 10-percent increase has elicited an increase of 3.8 percent in lumber price over a 14 month period; an increase of 2.2 percent in materials price over a 32-month period; an increase of 2.6 percent in housing price over a 31-month period; and an increase of one percent in shelter price over a 13-month period.

Strength of Relationships: Analyses of Forecast Error Variance Decompositions

Analysis of decompositions of forecast error variance (FEV) is another tool of VAR econometrics for discerning relationships among the modeled system's time series. FEV is, at alternative horizons or steps, attributed to shocks in each of the dynamic system's series, such that a measurement of relative "strength" of relationships emerges (Bessler 1984a, b). Error decompositions attribute within-sample variance to alternative series and thus give measures which are useful in applied work. In table 2, the top portion contains the FEV decompositions of the lumber price experiment, and the bottom portion contains the FEV decompositions of the futures price experiment. Recall that the lumber price experiment was conducted under the following Choleski causal ordering: LUMBER, FUTURES, MATLS, HOUSING, and SHELTER. The ordering for the futures experiment was the latter one with the reversal of the first two prices such that futures price is first and lumber price second.

Ordering influences the results, but perhaps not to as great an extent as is sometimes thought. Certainly, the materials, housing, and shelter prices, being lumber-related, are plausibly situated at the bottom of the orderings of the lumber and futures price experiments. Given this, FEV decompositions for these three variables, the sub-ordering of which is constant across experiments, take on the same patterns of FEV decompositions in both experiments. One can verify this by simply comparing FEV decompositions down a particular column across experiments. Further, the combined FEV decompositions take on the same patterns across experiments. For example, the combined explanation of materials price FEV from futures and lumber prices at month-24 is about 71 percent in both experiments. What differs across experiments is the breakdown of the combined FEV decompositions for these two prices. Hence, the degree to which lumber or futures price influences or explains uncertainty in the other prices depends on the ordering, that is on which of the two prices is placed causally first in the ordering.

⁴The PPI for lumber is an index of a number of lumber product prices at many pricing points, whereas the futures price is for a certain grade of lumber at one point. Yet the futures price can serve, and has often served, as a representative price for a wider array of lumber products. Therefore, we can not conclude that the futures price response is 1.6 times the "lumber" price response. However, the figures and table 1's results do indicate that futures price tends to overreact in the short-term. This is particularly evident from the futures price impulses of figure 1.

FFC/93 Papers & Proceedings

Table 2 -- Decompositions of forecast error variance (FEV) for the lumber and futures price experiments.

Percent Explanation of Forecast Error Variance from:						
Variable:	Step	Lumber Price	Futures Price	Materials Price	Housing Price	Shelter Price
<u>FEV results from the lumber price experiment:</u>						
Lumber price:	1	95.36	3.96	0.03	0.23	0.42
	3	84.88	6.06	4.64	4.10	0.31
	6	70.37	4.49	10.19	12.26	2.69
	12	65.08	3.62	11.32	16.09	3.88
	18	58.79	3.35	20.53	13.17	4.15
	24	53.30	3.39	28.58	10.67	4.26
	30	47.20	3.63	35.21	9.47	4.48
	36	42.78	3.92	39.04	9.42	4.83
Futures price:	1	35.35	64.17	0.12	0.03	0.33
	3	35.22	54.26	5.88	4.33	0.31
	6	34.49	47.51	10.66	6.77	0.57
	12	37.04	38.83	13.61	9.77	0.76
	18	36.13	36.49	15.94	10.64	0.80
	24	35.67	35.09	18.21	10.25	0.78
	30	34.51	33.93	20.59	10.09	0.88
	36	33.59	32.91	22.05	10.26	1.19
Materials price:	1	42.65	1.35	55.26	0.71	0.03
	3	39.24	0.73	52.25	7.52	0.26
	6	42.45	1.20	42.24	13.01	1.09
	12	53.67	0.85	24.59	20.22	0.67
	18	65.05	1.04	16.07	17.20	0.64
	24	69.00	1.59	16.36	12.17	0.88
	30	67.41	2.16	20.99	8.45	1.00
	36	63.59	2.64	26.57	6.29	0.91
Housing price:	1	0.10	0.18	0.09	99.16	0.47
	3	0.29	0.21	0.09	97.95	1.47
	6	0.72	1.39	0.44	95.52	1.93
	12	9.68	2.82	0.55	82.63	4.32
	18	27.67	5.50	1.10	58.59	7.13
	24	46.15	5.71	0.92	37.81	9.41
	30	58.23	4.82	0.91	25.73	10.30
	36	64.84	3.97	2.78	18.42	9.99
Shelter price:	1	0.23	0.24	0.19	74.61	24.73
	3	0.51	0.16	0.37	73.79	25.17
	6	0.66	1.88	1.42	69.45	26.59
	12	5.30	2.10	1.60	57.03	33.97
	18	16.04	2.64	2.76	41.75	36.81
	24	30.04	2.35	2.79	30.28	34.54
	30	42.78	1.94	2.12	23.33	29.83
	36	52.14	1.66	3.44	18.21	24.55

FFC/93 Papers & Proceedings

FEV results from the futures price experiment:

Lumber price:	1	54.95	44.37	0.03	0.23	0.42
	3	44.01	46.93	4.64	4.10	0.31
	6	37.01	37.84	10.49	12.26	2.69
	12	34.83	33.87	11.32	16.09	3.88
	18	31.13	31.01	20.53	13.17	4.15
	24	27.63	29.01	28.38	10.67	4.26
	30	23.97	26.86	35.21	9.47	4.48
	36	21.46	25.24	39.04	9.42	4.83
Futures price:	1	1.21	98.31	0.12	0.03	0.33
	3	1.89	87.58	5.88	4.33	0.31
	6	2.71	79.28	10.66	6.77	0.57
	12	6.48	69.39	13.61	9.77	0.76
	18	6.78	65.84	15.94	10.64	0.80
	24	7.12	63.64	18.21	10.25	0.78
	30	6.90	61.54	20.59	10.09	0.88
	36	6.83	59.67	22.05	10.26	1.19
Materials price:	1	29.37	14.63	55.26	0.71	0.03
	3	27.12	12.85	52.25	7.52	0.26
	6	33.11	10.54	42.24	13.01	1.09
	12	37.71	16.82	24.59	20.22	0.67
	18	42.46	23.62	16.07	17.20	0.64
	24	42.37	28.23	16.36	12.17	0.88
	30	39.49	30.08	20.99	8.45	1.00
	36	35.86	30.37	26.57	6.29	0.91
Housing price:	1	0.00	0.28	0.09	99.16	0.47
	3	0.06	0.44	0.09	97.95	1.47
	6	0.08	2.03	0.44	95.52	1.93
	12	3.50	8.99	0.55	82.63	4.31
	18	10.57	22.60	1.10	58.59	7.13
	24	20.51	31.35	0.92	37.81	9.41
	30	28.65	34.41	0.91	25.73	10.30
	36	33.83	34.98	2.78	18.42	9.99
Shelter price:	1	0.07	0.40	0.19	76.61	24.73
	3	0.31	0.36	0.37	73.79	25.17
	6	0.29	2.25	1.42	69.45	26.59
	12	2.09	5.31	1.60	57.03	33.97
	18	7.21	11.46	2.77	41.75	36.81
	24	15.82	16.57	2.79	30.28	34.54
	30	24.42	20.29	2.12	23.33	29.83
	36	30.72	23.08	3.44	18.21	24.55

Lumber price is highly exogenous in the lumber price experiment, especially at horizons of 18 months or less, with 59 to 95 percent of lumber price's FEV self-attributed. After 18 months, this percentage drops below half and suggests lumber price's increasing exogeneity at the longer horizons. After 18 months, materials price takes on increasing importance in explaining lumber price FEV -- steadily up to 39 percent by the 36th month horizon. These FEV results coincide with the same experiment's impulse response results: lumber price shocks have required from two to three years before related price effects have played themselves out.

In the futures price experiment, lumber price has been more endogenous than in the lumber price experiment. The lumber price FEV has been from 21.5 to 55 percent self-attributed, and from 25 to 47 percent attributed to futures price variation.

Futures price becomes more endogenous in the lumber price experiment beyond the month-3 horizon, when no more than about 47 percent of its FEV attributed to own-variation. From 34 to 37 percent of futures price FEV is attributed to the shock variable, lumber price. Futures price is more exogenous in the futures price experiment, with from 60 to 98 percent of its FEV having been self-attributed.

Whether lumber price affects futures price or futures price affects lumber price depends on which of the two prices is placed causally first in the ordering as the shock variable, that is on which variable which moves first. Each shock variable exhibits high degrees of exogeneity, and explains comparable FEV proportions of each other, in the two experiments. So across experiments, the patterns with which lumber price accounts for futures price variability is similar to the patterns with which futures price accounts for lumber price variability. But considering these FEV results in combination with the two experiments' impulse response results adds further insight. While lumber and futures explain similar percentages of each other's FEV across experiments, the impulses suggest that the sizes of effects vary noticeably across experiments. Compared with the futures price experiment, lumber price accounts for similar percentages of larger price effects.

Materials price's FEV becomes increasingly dependent on the combined variation of lumber and futures price after the one-year horizon. After a year, no less than 66 percent of materials price FEV is attributed to variation in lumber and futures prices. This coincides roughly with the materials price impulses (figures 1 and 2) which require 12 to 18 months to achieve or approach peak strength levels in response to the shocks.

Housing price is highly exogenous early on, with no less than 59 percent of its FEV being self-attributed through the 18-month horizon. Thereafter, the combined variation of lumber and futures prices accounts for most (52-69 percent) of housing price's FEV. These results reinforce the housing price impulse responses which take from nine to 12 months to activate, and from 2.5 to three years after the shock to peak in strength (figures 1 and 2).

Shelter price remains highly endogenous in both experiments, with from 25 to 37 percent of its FEV being self-attributed at all reported horizons. The price's exogeneity peaks at month-18, after which contributions of own-variation to FEV steadily and moderately decline. These results are not inconsistent with the shelter price impulses of figures 1 and 2, where from 14 to 22 months are required for responses to begin, where responses remain within a narrow band of impulse magnitude over the response cycle, and where responses require from 2.5 to three years to approach peak strength.

Comparative Dynamics of the Lumber and Futures Price Experiments

Compared with a futures price increase, a 10-percent rise in lumber price has elicited price increases that form similarly shaped patterns, but which have engaged more rapidly; have endured for longer periods; and have achieved generally greater magnitudes. This is evident from table 1. When reaction times for a price across experiments have differed, those of the lumber price experiment have been shorter. A lumber price shock's effects were generally more enduring than price effects elicited by a futures price shock (see table 1, response durations). Figures 1 and 2 reveal, however, that the impulses of the two experiments take on similar monthly patterns -- reaction times and durations notwithstanding.

The most evident difference in the price effects from shocks in lumber and futures prices lies in the effect magnitudes. Differences are apparent from comparing the scales of figures 1 and 2 which contain plotted impulses, and from table 1's response multipliers. Lumber and futures price movements elicit statistically significant responses in each other, as well as in materials, shelter, and housing prices. Yet a lumber price shock has elicited far greater impulse magnitudes than a similarly sized shock in futures price.

Another interesting comparison involves futures price responses to lumber price shocks as opposed to lumber price responses to futures price shocks. Figures 1 and 2 reveal that lumber and futures prices

respond to each other with similar shapes and for comparable periods. Yet figures 1 and 2, as well as table 1's multipliers, suggest that futures price "over-reacts" relative to lumber price. The multipliers suggest that historically, percentage futures price response to lumber price movement has been four times the percentage lumber price response to futures price movement.

There were similarities and dissimilarities in FEV decomposition patterns across experimental orderings. The percentages of FEV of all prices attributed to variations in the three prices not influenced by the experimental orderings -- materials, housing, and shelter prices -- were the same across experiments. The combined influence of lumber and futures price uncertainty on the FEV's of the system is the same across experiments, although the distribution of this combined influence on the prices' FEV's varies according to which of the two prices is placed first in the ordering and engages the system's shock. The lumber price percentage of this combined influence is greatest in the lumber price experiment, while the futures price's percentage is greatest in the futures price experiment. But taken together with the impulse response results, another result emerges: the FEV decompositions are similar across experiments, but these similar percentage patterns hold for differently sized effects.

Summary and Conclusions

The average monthly dynamic patterns captured by the model can characterize what may happen to the five prices from the recent volatile swings in lumber and lumber futures prices.

Lumber price increases should elicit materials price increases that begin within a month, that gradually accelerate over about a year before peaking in magnitude, and that endure for about three years. Response should be less than one-for-one, with materials price rising by about half of the percent increase in lumber price over this period. A futures price increase should similarly influence materials price, but by less and for shorter time periods.

Some time (up to 14 months) would elapse before lumber price increases noticeably influence housing and shelter prices. Patterns of housing and shelter prices increases would begin at low levels and then gradually gain strength, while enduring at least 2.5 years. Responses would be less than one for one, with housing and shelter price rising by about half of the percent increase in lumber price. Housing and shelter price impulses from futures price shocks would be more delayed, weaker, and shorter-lived than those elicited by lumber price movements.

Futures price should respond to a far greater degree to lumber price than lumber price would respond to futures price. Each percent rise in lumber price should elicit a greater 1.6 percent increase in futures price over about a year, while lumber price should respond by about a quarter of that percentage to similar increases in futures price.

Therefore, each increase, say of 10 percent, in futures price should elicit responses in construction and housing related prices that are often more delayed, that are weaker in strength, and that are less enduring than similar movements in lumber price. Compared with lumber price responses to futures price movements, futures prices should over-react to changes in lumber price. So to elicit given changes in materials, housing, and shelter prices, futures price would have to swing more widely than futures price.

References

- Babula, R.A., G. Gajewski, and P. Colling. "Lumber Prices Linked to Housing Markets." Agricultural Outlook AO-197(June, 1993):27-29.
- Babula, R.A. and D.A. Bessler. "The Corn-Egg Price Transmission Mechanism." Southern Journal of Agricultural Economics 22(no. 2, Dec. 1990):79-86.
- Bessler, D.A. "An Analysis of Dynamic Economic Relationships: An Application to the U.S. Hog Market." Canadian Journal of Agricultural Economics 32(1984a):109-24.
- Bessler, D.A. "Relative Prices and Money: A Vector Autoregression on Brazilian Data." American Journal of Agricultural Economics 66(1984b):25-30.
- Bessler, D.A. and J.L. Kling. "A Note on Tests of Granger Causality." Applied Economics 16(1984):335-42.
- Dickey, D. and W. Fuller. "Distribution of the Estimates for Autoregressive Time Series with a Unit Root." Journal of the American Statistical Association 74(1979):427-31.
- Doan, T. Regression Analysis of Time Series, Users' Manual, Version 3.0. Evanston, Illinois:VAR Econometrics, 1990.
- Ericsson, N. "Cointegration, Exogeneity, and Policy Analysis: An Overview." Journal of Policy Modeling 14(no. 3, June 1992):251-80.
- Fuller, W. Introduction to Statistical Time Series. New York:John Wiley and Sons, 1976.
- Gorte, R.W. "Lumber Prices." Memorandum, Environment and Natural Resources Policy Division, Congressional Research Service, Library of Congress, Washington DC, March 10, 1993.
- Granger, C.W.J. "Developments in the Study of Cointegrated Economic Variables." Oxford Bulletin of Economics and Statistics 48(1986):213-28.
- Granger, C.W.J. and P. Newbold. Forecasting Economic Time Series. New York:Academic Press, 1986.
- Harvey, A. An Econometric Analysis of Time Series. Cambridge, MA: The MIT Press, 1990.
- Hendry, D. "Econometric Modelling with Cointegrated Variables: An Overview." Oxford Bulletin of Economics and Statistics 48(1986):201-12.
- Johansen, S. and K. Juselius. "Testing Structural Hypotheses in a Multivariate Cointegration Analysis of the PPP and UIP for the UK." Journal of Econometrics 53(1992):211-44.
- Johansen, S. and K. Juselius. "Maximum Likelihood and Inferences on Cointegration: With Applications to the Demand for Money." Oxford Bulletin of Economics and Statistics 52(1990, no. 2):169-210.
- Kloek, T. and H.K. VanDijk. "Bayesian Estimates of Equation System Parameters: An Application of Monte Carlo." Econometrica 46(1978):1-20.
- Larue, B. "Farm Input, Farm Output, and Retail Food Prices: A Cointegration Analysis." Canadian Journal of Agricultural Economics 39(1991):335-53.
- Lutkepohl, H. "Comparison of Criteria for Estimating the Order of a Vector Autoregression Process." Journal of Time Series Analysis 6(1985):35-52.

Sims, C. "Macroeconomics and Reality." Econometrica. 48(1980):1-48.

Sims, C. "Models and Their Uses." American Journal of Agricultural Economics 71(1989):489-94.

Taylor, J. "Bull Market in Lumber May be Over." Wall Street Journal
Section C, April 2, 1993, pp. C1 & C12.

Tiao, G. and G.E.P. Box. "Modeling Multiple Time Series: With Applications." Journal of the American Statistical Association 76(1981):802-16.

VanTassell, L.W. and D.A. Bessler. "Dynamic Price Adjustments Between Commercial and Purebred Cattle Markets." Southern Journal of Agricultural Economics 20(no. 2, Dec. 1988):137-43.

Reading the Market's Mind: Using Crude Oil Options to Recover the PDF

William Melick and Charles Thomas, Federal Reserve Board*

Energy economics is often concerned with decision making under uncertainty. Derivative markets, used to offset risk, can help quantify this uncertainty. Option markets in particular, given their contingent nature, have been used by economists in order to better understand the uncertainty in a give market. (One such example is Overdahl and Matthews (1988), who used options to construct confidence intervals for the future spot price of oil.)

Naturally, the standard tools developed in finance have been brought to bear on these questions. However, these tools are not well designed to handle some of the situations encountered in the energy markets. During unsettled times (that is, most of the time in energy markets) the assumption that prices are drawn from a lognormal distribution, which lies behind the Black-Scholes (1973) option pricing model, does not seem appropriate. Yet the usual alternative from the finance literature, that prices follow a jump diffusion process (e.g. Bates (1990)), seems overly restrictive.

An alternative method to quantify uncertainty has been developed by Melick and Thomas (1993) (hereafter MT). They use option prices, imposing as little structure as possible, to recover the market's implied distribution for the underlying commodity. As an application, they consider options on crude oil futures during the Persian Gulf crisis. This paper presents an overview of their technique, using data from the time of the Persian Gulf crisis and the summer of 1993 for illustrative purposes. The paper is organized as follows. The first section presents the intuition behind the MT procedure, using the straightforward case of European-style options as a starting point. The second section examines several episodes from the Persian Gulf crisis, while the third section used illustrations from the most recent period. Conclusions are presented in the final section.

I. Option Prices and PDFs

A call option allows its holder to buy the underlying commodity at the agreed upon strike price. Therefore, the call option has value at expiration (is in the money) whenever the price of the commodity is above the strike price. Unlike an American option, European options can be exercised at the time the contract expires. Therefore, the value of a European call option, with a strike price denoted by X , can be expressed as the product of two terms

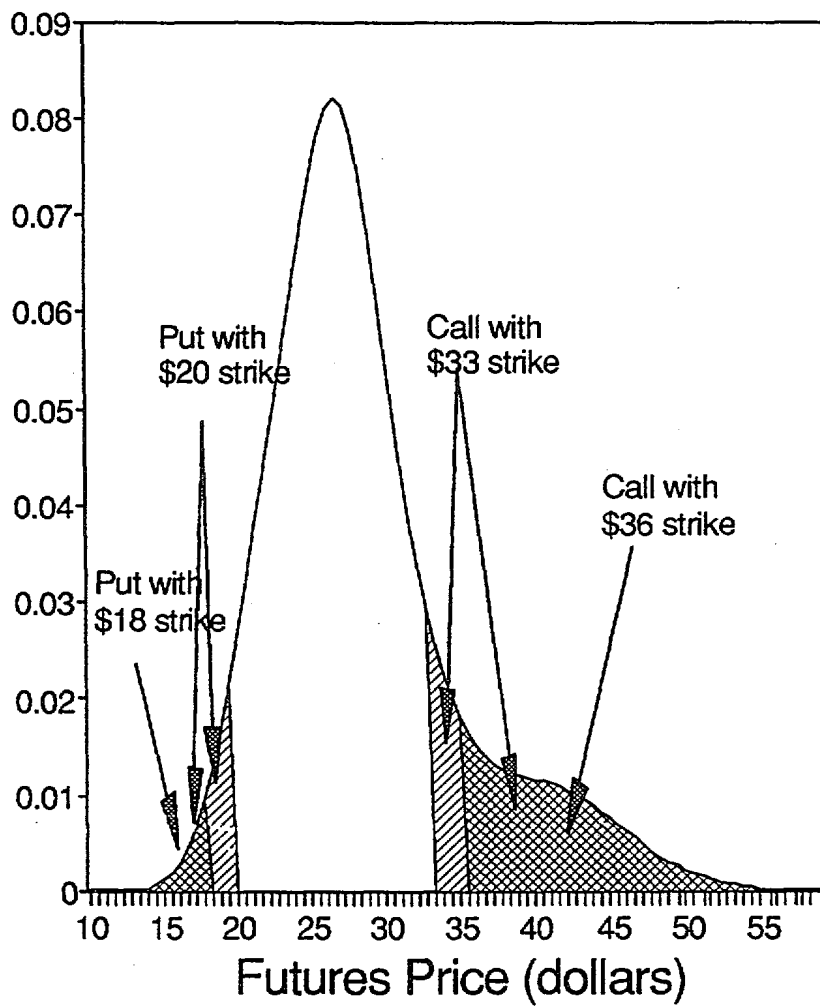
$$C(X) = \int_X^{\infty} f(p) dp \cdot \left(\int_X^{\infty} p \cdot f(p) dp / \int_X^{\infty} f(p) dp - X \right)$$

where p is the price of the commodity, and $f(p)$ is its PDF. The first term represents the probability that the price of the commodity, at expiration, is above the strike price. The second term represents the difference between the expectation for the price of the commodity, given that it is above the strike price, and the strike price. Both of these terms are integral of the probability density function (PDF) for the underlying commodity's price. Given prices for European options with different strike prices, a functional form for the PDF, and a nonlinear solver, it is relatively straightforward, to "invert" a set of equations like (1) and recover the parameters of the PDF. The technique involves searching for the parameter values that minimizes the pricing errors generated by the set of equations.

These concepts are illustrated in Chart 1 which considers what portions of a PDF could be recovered with four options: two puts with strikes of \$18 and \$20, and two calls with strikes of \$33 and \$36. The \$36 call would only have value if the price of the underlying commodity were above \$36, therefore, it only offers information about the portion of the PDF above \$36. The \$33 call also provides information about the PDF above \$36, as well as filling in the portion of the PDF between \$33 and \$36. The story for the puts is similar. The \$18 put helps to

* This paper represents the views of the authors and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System or other members of its staff. The authors may be reached by writing to William Melick, Mail Stop #42, Federal Reserve Board, Washington, D.C. 20551, or by calling (202) 452-2296.

Chart 1
Options Values and the
Density Function



recover the tail of the PDF below \$18. The \$20 put also helps to recover this lower tail, as well as filling in the part of the PDF between \$18 and \$20.

It is useful to note how data limitations and the assumed functional form for the distribution interact. The fact that strikes are at discrete intervals and, more importantly, that they do not span the entire possible range of the price of the commodity, places an important limitation on what the option prices can reveal about the distribution. The recorded option prices only contain information about the conditional expectation and probability mass in the following segments of the range of the price: 1) the segment below the lowest strike, 2) the segments between each strike, and 3) the segment above the highest strike. Any number of distributions could generate the same results for these conditional expectations and probabilities. For example, for any given distribution we can construct a second distribution out of a series of non-overlapping uniform densities which will be observational equivalent to the given distribution. Any estimated distribution requires careful interpretation, especially in the regions below the lowest strike and above the highest strike. For crude oil, strikes are almost always \$1.00 apart (in a few instances \$5.00), allowing a fine demarkation of the distribution within the range of strikes. However, the shape of the distribution in the tails will depend importantly on the functional form assumed for the distribution.

The recovery of the PDF is more complicated for American options. These options can be exercised at any time prior to expiration, adding to their value relative to European options. Thus, the value of an American call option, in terms of the PDF, cannot be written as compactly as equation (1). The idea of MT is to bound the value of the American option using two equations (similar in spirit to equation (1)). As before, parameter values are found that minimize the pricing errors generated by a weighing of the pairs of bounds in the set of equations.

The method of MT requires one important condition to derive the bounds; namely, that the price of the underlying commodity martingales. This is not unreasonable for options on futures, as there is no cost in holding a futures contract. In brief, the martingale assumption means that today's price for the commodity equals the expectation for the price of the commodity at expiration. (For evidence that crude oil features martingale see Dominguez (1989), Kumar (1992), and Deaves and Krinsky (1992)).

The intuition behind the bounds is fairly straightforward. Given that we know the value of a European option, the upper bound will maximize the premium associated with early exercise, subject to the martingale assumption. The early exercise premium will be maximized if prices moved tomorrow and stayed fixed thereafter through expiration. Conversely, the early exercise premium would be minimized if the price did not move until just before expiration, so that the option holder would only exercise today or at expiration. In this situation there would be little value associated with the right of early exercise. Both the upper and lower bound, given the martingale assumption, can be written in terms of the PDF for the price of the underlying commodity. A detailed derivation and further intuition is presented in MT.

II. Application to Crude Oil: Persian Gulf Crisis

Section I showed that the MT method allows a researcher to recover the PDF for the price of the commodity underlying an American option given: 1) observed options prices, 2) an assumed functional form for the PDF, and 3) a nonlinear optimization routine. Throughout the Persian Gulf crisis, market commentary focused on three distinct outcomes: 1) a return to pre-Crisis conditions (e.g. Iraq would peacefully withdraw from Kuwait), 2) a severe disruption to Persian Gulf oil supplies (e.g. damage to Saudi Arabian facilities during a war), and 3) a continuation of unsettled conditions over the relevant horizon (e.g. a prolonged stalemate in which outcome 1 or 2 might eventually occur). Given these three possibilities, we chose a mixture of three lognormals (MLN) as the form of the distribution to be estimated, written as

$$g = \pi_1 \cdot g_1[u_1, \sigma_1] + \pi_2 \cdot g_2[u_2, \sigma_2] + \pi_3 \cdot g_3[u_3, \sigma_3].$$

If in fact market participants felt that prices were likely to be drawn from a trimodal distribution this could be easily captured by the mixture. Ex ante, we expected that as news hit the market, the relative weighing of the three lognormals might change, as well as the parameters of each of the three lognormals. For example, news of an Iraqi rocket attack on a Saudi Arabian oil field might increase the weighing on the lognormal distribution with the highest mode, as well as increase the relevant range encompassed by this lognormal distribution.

As an alternative, we also followed Overdahl and Matthews and recovered a single lognormal density (roughly the Black-Scholes model) using the Barone-Adesi and Whaley (1987) approximation to account for the American style options. These estimations correspond to a single lognormal (SNL), and would be roughly equal to those from MLN in which $\pi_1 = \pi_2 = 0$.

Estimated distributions for selected events during the Persian Gulf crisis (both SLN and MLN) are found in Charts 2 and 3. Comparing the estimated PDFs from the two models immediately before and after the receipt of "news" allows us to infer how the market interpreted the news and highlights the differences between the NLN and SLN models.

On Thursday, October 25, 1990, the London Financial Times carried a report that Iraqi forces had attached explosives to 300 of Kuwait's 1000 oil wells, quoting a senior Kuwaiti engineer who had left Kuwait one week earlier. This revelation pushed oil prices up sharply, with the futures contract nearest to expiration (December) rising \$3.17 per barrel. Chart 2 plots the PDFs from MLN and SLN for October 22 (top panel) and October 25 (bottom panel) using the January contract. On October 22, market expectations for future prices were centered quite tightly around \$24 per barrel. The news of the mining widened each model's distribution significantly, with the MLN allowing for a sizeable probability mass between \$60 and \$70 per barrel.

The largest one-day change in oil prices in NYMEX history occurred in Thursday January 17, 1991 when 1) several governments announced a coordinated release of oil from their emergency inventories and 2) it became clear that the coalition forces had total air supremacy. On that day the settle price for the March contract fell \$9.66 while the settle price for the April contract fell \$7.82. The six panels of Chart 3 trace the evolution of expected PDFs on the days surrounding January 17. Prior to the first air strike (as can be seen in the first two panels), the market was still expecting a fairly significant chance of a major oil market disruption (perhaps Iraqi damage to Saudi Arabian oil facilities) that could push prices to the \$40-\$60 per barrel range. On January 17th these PDFs tightened dramatically, and on ensuing days the PDF generated from MLN moved closer and closer to that from SLN. By January 23, there was little difference between the two PDFs, as the market returned to almost a pre-crisis distribution.

III. Current Events: The Return of Iraq

Oil prices fell from June through mid-August of 1993 (the time this paper went to press), after recovering earlier in the year. OPEC production restraint in the first part of the year had buoyed prices, but Kuwait's decision not to participate in the cartel's third-quarter quota allocations resulted in price declines. Most recently, signs of apparent progress in talks between Iraq and the United Nations, coupled with frictions within OPEC on the appropriate response to this progress, have pushed prices back down to 1992 year-end levels.

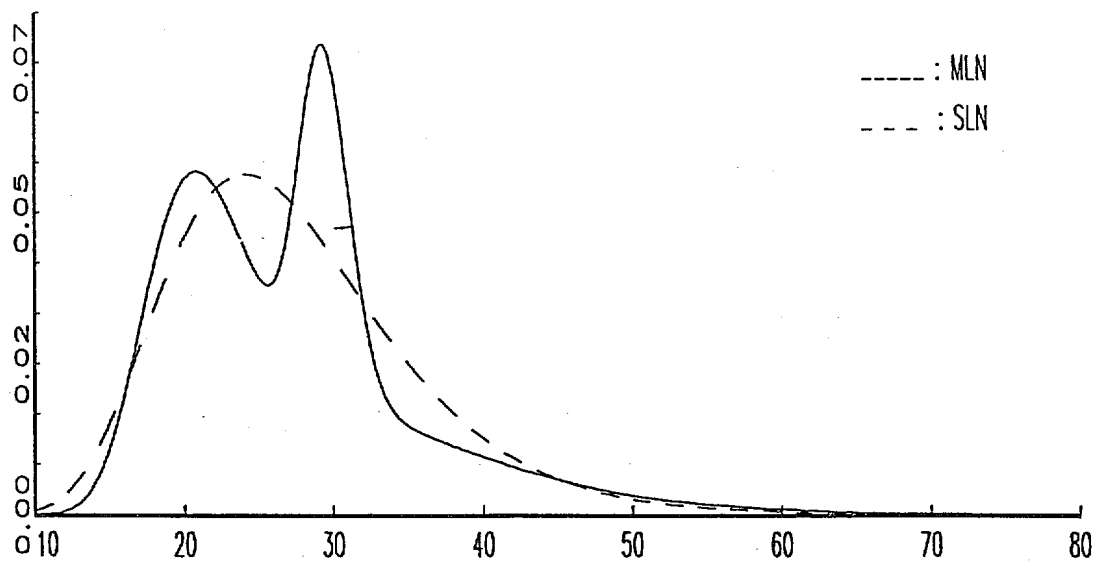
The one-sided nature of Iraq's return to the oil market argues against a standard application à la Overdahl and Matthews to options on crude oil futures. However, the mixture of three lognormals used by MT for the Persian Gulf crisis seems excessive. Therefore, a mixture of two lognormals was used to recover the MLN PDFs drawn in Chart 4. Ex ante it seemed reasonable that market participants might envision a bimodal distribution for oil prices during this period, with a significant chance that Iraq could begin exporting as well as a significant chance of a continued embargo.

As can be seen in the top panels of the chart, on June 30, 1993, before reports of progress in talks with Iraq, expected oil prices at the expiration of both the November 1993 contract (the left panel) and the March 1994 contract (the right panel) were distributed in a fairly smooth bell shape typical of a lognormal distribution. However, by mid July, when indications of progress in the talks were greatest, market perception had changed. Oil prices had softened, and as shown in the middle panels, for both delivery months (but especially March 1994) the market now entertained the possibility of sharply lower prices. Comparing the two middle panels, market participants felt that the chance of an Iraqi sale was increasing with the passage of time, as seen by the much larger left-hand tail for the March 1994 contract than for the November 1993 contract. By March of 1994, the market felt there was a seven percent chance that oil might be selling for less than 15 dollars per barrel.

Data from August 9th were used for the two bottom panels. As prospects for an Iraqi sale have diminished so has the mass in the left-hand tail, with prices once again following a bell-shaped distribution -- albeit with a lower mean price than at the end of June. As of August 9th, the market only saw a 3 percent chance of oil selling for less than 15 dollars per barrel.

Chart 2

JAN Contract on 22-Oct Futures = 28.95



JAN Contract on 25-Oct Futures = 32.19

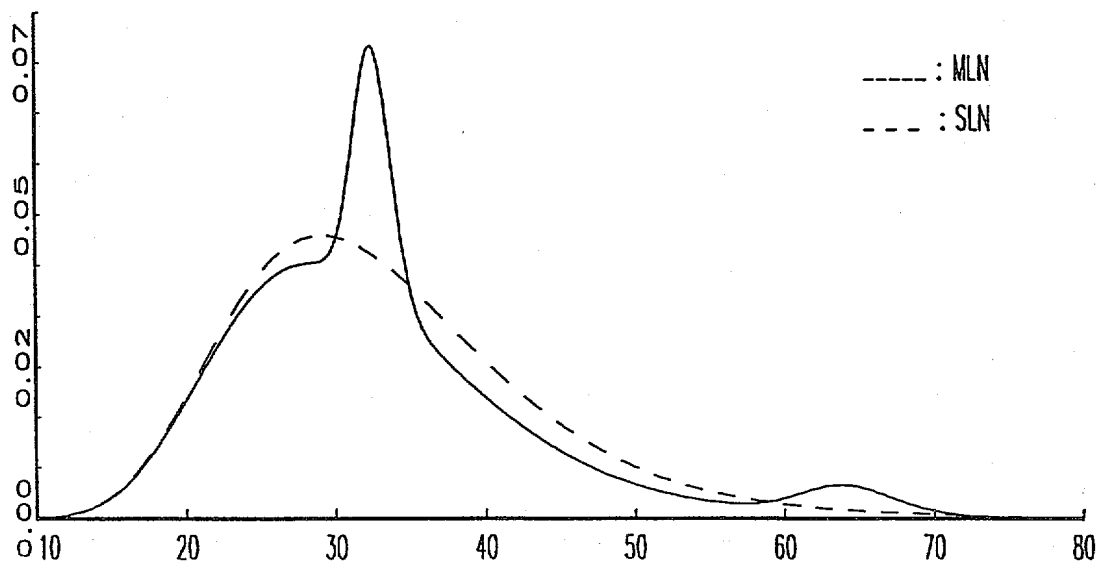
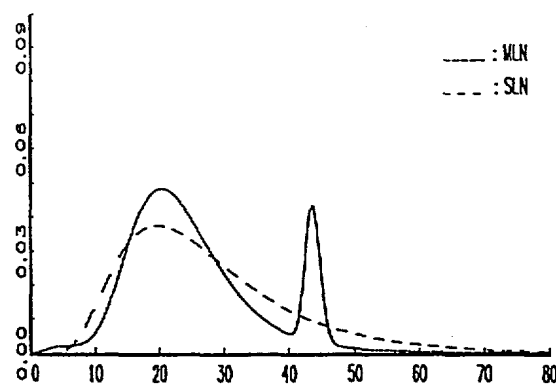
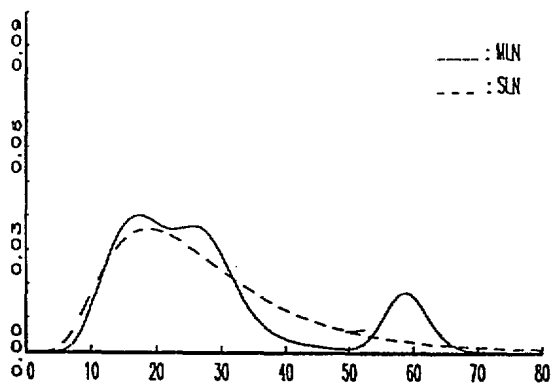


Chart 3

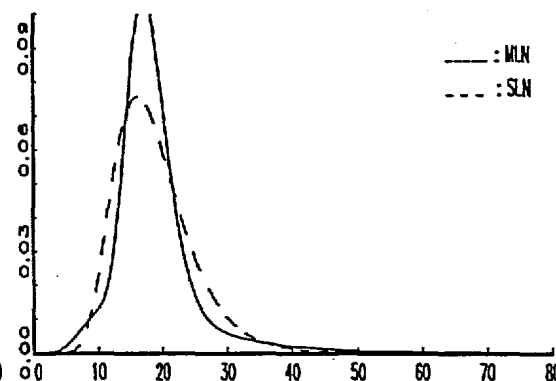
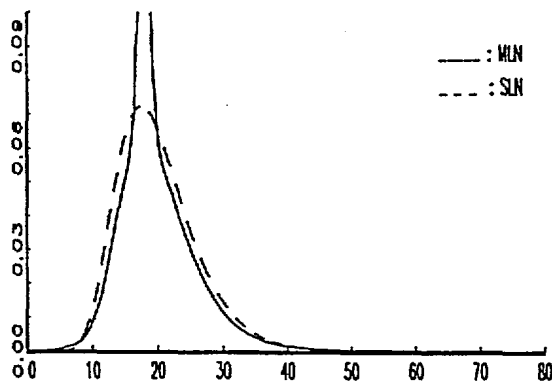
APR Contract on 14-Jan Futures = 26.74

APR Contract on 16-Jan Futures = 27.95



APR Contract on 17-Jan Futures = 20.13

APR Contract on 18-Jan Futures = 18.78



APR Contract on 21-Jan Futures = 19.77

APR Contract on 23-Jan Futures = 21.02

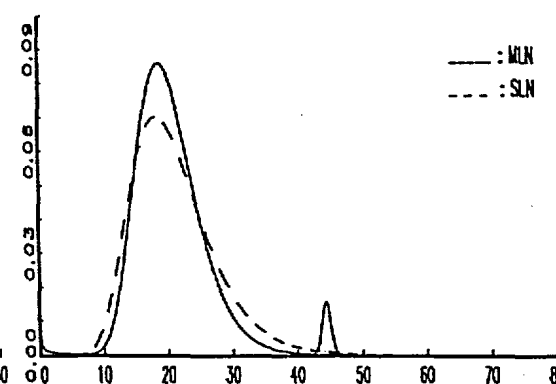
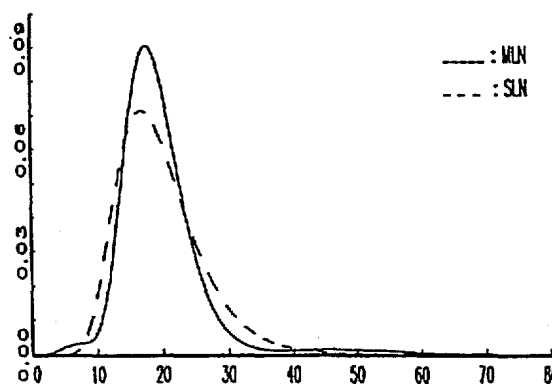
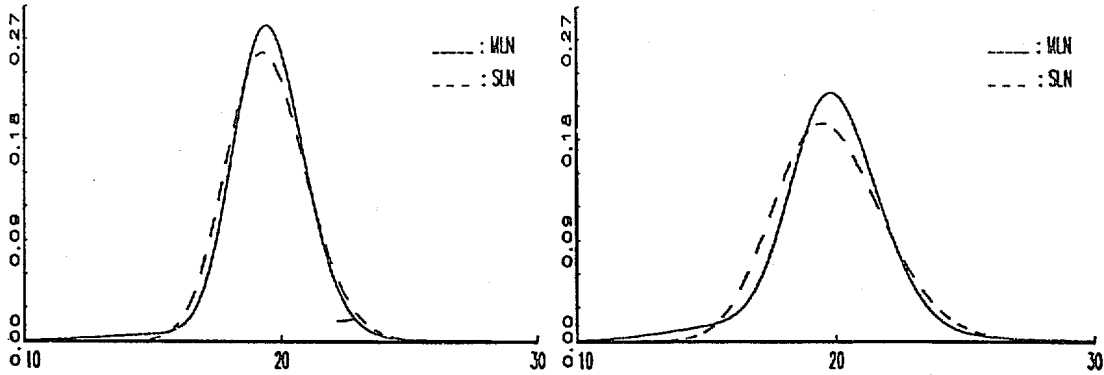


Chart 4

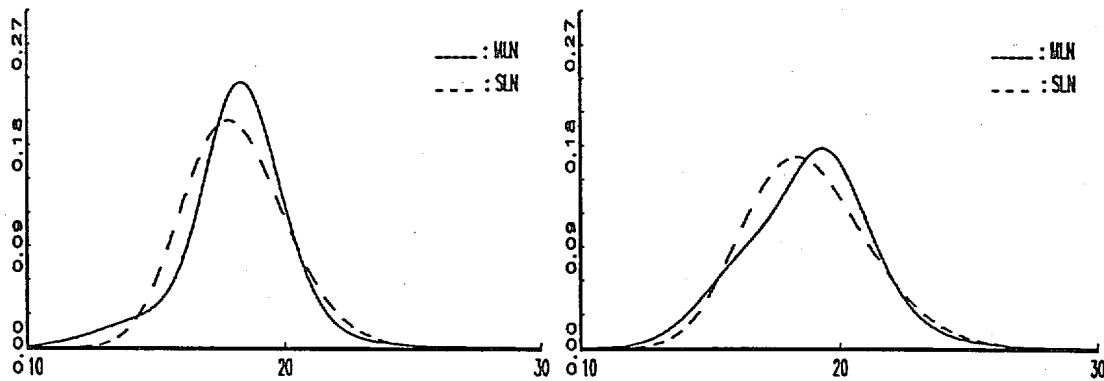
NOV Contract on 30-Jun Futures = 19.46

MAR Contract on 30-Jun Futures = 19.80



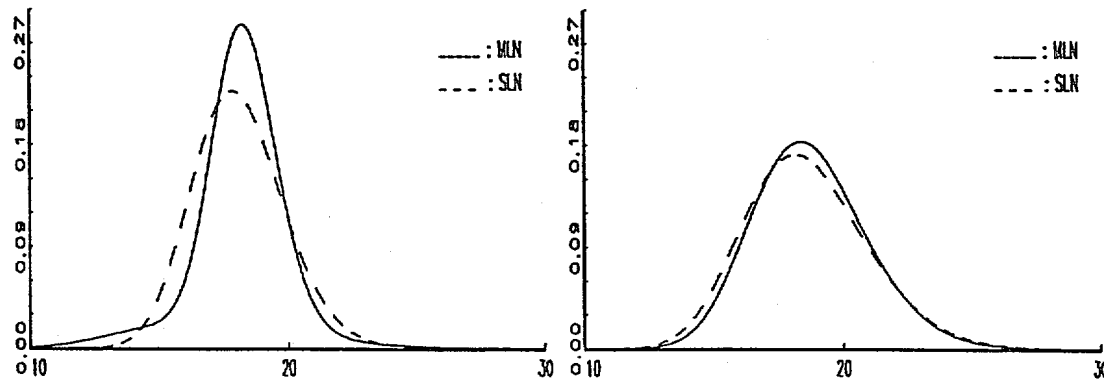
NOV Contract on 14-Jul Futures = 18.13

MAR Contract on 14-Jul Futures = 18.79



NOV Contract on 09-Aug Futures = 18.17

MAR Contract on 09-Aug Futures = 18.73



IV. Conclusion

This paper briefly describes the MT method for extracting the market's probability assessment of future events using American option prices, in particular the PDF for the option's underlying asset. The method had many applications as there is a broad spectrum of indices, commodities, and financial instruments on which option are traded, including interest rates and exchange rates. The major limitation of the method is that it is only applicable to those markets where the underlying asset or index martingales. In the application to the oil market, examination of particular days confirmed the large shift in market expectations that occurred when significant news reached the oil market.

References

- Barone-Adesi, Giovanni and Robert Whaley, "Efficient Analytic Approximation of American Option Values", Journal of Finance 42(2):301-320 (1987).
- Bates, David S., "The Crash of '87: Was It Expected? The Evidence From Option Markets", Journal of Finance 46(3):1009-1044 (July 1990).
- Black, Fisher and Myron Scholes, "The Pricing of Options and Corporate Liabilities", Journal of Political Economy 81:637-659 (1973).
- Deaves, Richard and Itzhak Krinsky, "Risk Premiums and Efficiency in the Market for Crude Oil Futures", The Energy Journal 13(2):93-118 (1992).
- Dominguez, Kathryn, "The Volatility and Efficiency of Crude Oil Futures Contracts," in Oil and Money: Coping With Price Risk Through Financial Markets, Harvard International Energy Studies 5:48-87 (1989).
- Kumar, Manmohan, "The Forecasting Accuracy of Crude Oil Futures Prices", IMF Staff Papers 39(2) (June 1992).
- Melick, William and Charles Thomas, "Recovering an Asset's Implied PDF From Option Prices: An Application to Crude Oil During the Gulf Crisis", mimeo, Federal Reserve Board, (1993).
- Overdahl, James A. and H. Lee Matthews, "The Use of NYMEX Options to Forecast Crude Oil Prices", The Energy Journal 9(4):135-147 (1988).

Does the Commodity Research Bureau Futures Index Predict Inflation?

David Torgerson, Economic Research Service

Abstract

Is the Commodity Research Bureau futures index (CRB) related to inflation? Employing a procedure analogous to that done by Brown and Yucel (1993) I used pairwise Engle-Granger cointegration to search for relationships between the CRB and the Producer Price Index (PPI), the Producer Price Index for food (PPIfood), Consumer Price Index (CPI), and Consumer Price Index for food (CPIfood), respectively. First, all these variables, as is typical for macroeconomic variables, are nonstationary under the usual transformations (levels, logs, logged differences). Second, under the maintained hypothesis of cointegration between the CRB and the measures of inflation except for the PPIfood. The results put in a regression framework suggest that a 1 percent increase in the CRB is associated with an immediate .1 percent rise in the PPIfood and a one-year later increase of .05 percent. I note that this result is consistent with Babula (1993). While this link appeared to be reliable, the size would not suggest that the CRB alone would do a good job as a potential forecaster of PPI food inflation as a futures price is considered as a good predictor of the cash price of a commodity.

PROBLEM The Commodity Research Board and Future Food or General Inflation?

NO

The Commodity Research Board futures index (CRB) has been suggested as a useful indicator of future inflation, either general or food. Some analysts suggested that the recent rise in the CRB, which took place after the extent of the midwest flood became known, would imply sharply higher inflation either general or food. What is the relationship of these variables and the current and lagged values of the CRB?

Common inflation indices are the producer price index (PPI), the consumer price index (CPI), and their respective subindices for food PPIfood and CPIfood, respectively.

I tested all four variables for stationarity using single-variable unit root tests—they all failed. They all failed to be cointegrated, jointly cointegrated in levels, in first differences and logs. All but the PPIfood failed to be cointegrated in logged differences (percent change in the levels) and the logged differences of PPI-food and the CRB were cointegrated at the 10 per cent significance level and at a lag of 12 months the error terms of both variables are significantly correlated. A regression of the percent change PPI-food on the percent change CRB contemporaneously and lagged 12 periods indicated that a 1 percent permanent change in the CRB is associated with an immediate .15 percent change in the PPI-food and .05 percent change 12 months ahead. The time trend in this relationship was essentially zero so the regression was justifiable.

The bottom line is that the CRB does not show great promise as an indicator of food or general inflation. Since only the PPIfood is cointegrated at the 10 percent confidence level and the direct and lagged effects add up "explaining" only 20 percent of the change we are not talking about a fine point such as whether a futures price provides an unbiased estimator of a future cash price and an implied elasticity of .9 is estimated.

If analytically one wants to forecast inflation food or general one needs a richer model than is provided by use of CRB and its lags.

Data

The CRB is a daily index of selected futures prices of twenty commodities has roughly 20 observations per month. The various measures of inflation one could consider such as CPI, PPI, PPI-food and CPI-food which are monthly. Of course, one uses a monthly average for the CRB—which may lose some of the information coming from the data. The data are monthly from January 1974 and May 1993.

Methodology and Results

The unit root testing of the individual variables is routine and can be found in the manuals for most econometric packages. For the time-series aficionado the two-variate case of Engle-Granger (1987) approach has given way to the maximum likelihood several variable methods of Johansen and Juselius (1990). We will sketch how we used the extended Dickey-Fuller test to derive the results outlined above.

We note that in testing the relationship between CRB and the inflation indices we are bending over backwards to make it come out. The only "successful" error correction equation (ECC) equation is based on equilibrium being:

$$(1) \quad PCPPIfood_t - a*PCCRB_t - b*TREND_t = e_t, \text{ where}$$

$PCPPIfood_t$ is the logged difference of the $PPIfood_t$,

$PCCRB_t$ is the logged difference of CRB_t ,

and $TREND_t$ is the time trend,

and e_t is white noise, Normally distributed with mean zero.

All other formulations were statistically inconsistent with the analogous equilibrium condition. Even the above relationship with a constant term is statistically inconsistent with cointegration between the percent changes in CRB and $PPIfood$.

In (1) if we set e_t equal to zero, then the long-run relationship between variables is evident. In doing the cointegration test we allowed lags of up to 24 months. The estimate values of a and b are tested using the residuals or errors from equilibrium. In allowing a lag of up to 24 months one allows for error terms potentially influencing future error terms up to two years later. The calculated t-statistic associated with the estimated b is calculated using the Dickey-Fuller tau statistic since the distribution differs from the ordinary t . See Hall, et al (1990) for general discussion of this technique. The estimated b from this procedure was .22 with significant autocorrelation of the error term at 12 months. As it turned out, the estimated a was not appreciably different from zero so the time trend could be ignored. At this stage one should in principle go to a more sophisticated form of cointegration.

However, all is fair in love, war, and forecasting so given the functional form I re-estimated the relationship in a regression based on:

$$(2) \quad PCPPIfood_t = c*PCCRB_t + d*PCCRB_{t-12} + e_t,$$

where variables are defined as above except that $t-12$ refers to a 12 month lag on the percent change CRB . It turns out that the Ordinary Least Squares (OLS) estimated c was .148 and the estimated d was .046--a total slightly less than the above estimated .22.

Conclusion and Interpretation

As I would agree the above OLS results are not completely rigorous but we note that are broadly consistent with results generated using a Vector Autoregressive (VAR) on similar data. I point out that everything done greatly biases the results toward finding a large significant relationship of the CRB to $PPIfood$ and still this is less than an overwhelming response. In the future we will attempt to test the above relationships using the maximum likelihood methods of Johansen and Juselius (1990).

References

1. Babula, R. "Dynamic Economic Relationships Among Grain-Based U.S. Food Prices," Journal of International Food & Agribusiness Marketing. Vol 5, No. 1, pp. 35-52. 1993.
2. Engle, R. and C. Granger. "Co-integration and Error Correction: Representation, Estimation and Testing," Econometrica. Vol.55, pp. 251-76. 1987.
3. Hall, R. E., R.E. Hall and D. M. Lilien. Micro-TSP User's Manual; Version 7.0. Quantitative Micro Software, Irvine, California, 1991.
4. Johansen, S., and K. Juselius. "Maximum Likelihood Estimation and Inference on Cointegration-with Application to the Demand for Money," Oxford Bulletin of Economics and Statistics.

METAISSUES

MODERATOR

Howard Fullerton, Bureau of Labor Statistics

PRESENTATION 1

George Wesley, Peg Young, Office of Inspector General,
Department of Veterans Affairs

PRESENTATION 2

Karen Hamrick, Economic Research Service and Fred Mills,
Federal Management Partners, Inc.

Metaissues

Forecasting As A Tool For Oversight

George Wesley and Peg Young, Office of Inspector General, Department of Veterans Affairs

As the federal government reinvents itself into the form of anticipatory government, the Office of Inspector General (OIG) moves, in concert, towards anticipatory review. The shift towards proactive review and inspection has opened, for the OIG, numerous avenues for forecasting. The difficulty lies in the fact that there exists no "standard" set of techniques, forecasting or otherwise, for inspections. Rather, each particular instance of inspection requires the development of a technique to solve that particular problem. The purpose of this presentation is to discuss the forecasting applications that have been used, or have been proposed for use, for the purposes of oversight within the Department of Veterans Affairs.

Introduction

The Office of Inspector General (OIG) was created " (1) to conduct and supervise audits and investigations relating to programs and operations ... ; (2) to provide leadership and coordination and recommend policies for activities designed (A) to promote economy, efficiency, and effectiveness in the administration of, and (B) to prevent and detect fraud and abuse in, such programs and operations; and (3) to provide a means for keeping the head of the establishment and the Congress fully and currently informed about problems and deficiencies relating to the administration of such programs and operations and the necessity for and progress of corrective actions." ¹ As the direction of the OIG has shifted from reactive to proactive review, the definition of task for the Office of Inspector General has been expanded to include inspection, which is defined as "a process, other than an audit or an investigation, that is aimed at evaluating, reviewing, studying, and/or analyzing the programs and activities of a Department or Agency for the purposes of providing information to managers for decision making, for making recommendations for improvements to programs, policies or procedures, and for administrative action. The objectives of inspections include providing a source of factual and analytical information, monitoring compliance, measuring performance, assessing the efficiency and effectiveness of operations, and/or conducting inquiries into allegations of fraud, waste, abuse and mismanagement." ²

Within the Department of Veterans Affairs' OIG, three distinct trends and influences (the first largely internal, and other two more general throughout the federal government) have led to the appropriateness of the employment of forecasting methods in the oversight role of the VA's OIG. The first trend is unique to the Department of Veterans Affairs (VA). In 1988, Congress, through the passage of "The Veterans' Benefits and Services Act of 1988," called for the VA's Office of Inspector General to establish oversight of the VA's clinical quality assurance programs. The rationale was that, in order to fulfill the OIG's statutory responsibility to oversee, monitor, and evaluate the Agency's programs, the OIG would require individuals with clinical expertise and analytic capability beyond the realm of classical audits and criminal investigation. As such, the Quality Assurance Review Division, now termed the Office of Healthcare Inspections, was established in 1989 in the VA's OIG. The second trend was the need for flexible and anticipatory analysis, i.e., forecasting techniques, to aid the OIG in keeping the VA's Secretary and the Congress currently informed as to the economy, efficiency, and effectiveness of VA's programs, particularly of its health care programs. The third trend has been the tremendous interest, in general, in health care reform and health care policy throughout the country.

In contrast to the field of audit, the inspection methodologies vary from case to case. Since there exists no standard set of inspection techniques, necessity requires the development of procedures on a case-by-case basis. For those instances that require projections into the future, or estimates of what should have occurred based upon previous patterns of behavior, techniques based upon forecasting procedures are deemed appropriate. This

¹Inspector General Act of 1978

²QUALITY STANDARDS FOR INSPECTIONS, President's Council on Integrity and Efficiency, March 1993

presentation illustrates forecasting techniques that either have been used, or are planned for use, within the Office of Healthcare Inspections in the VA's OIG.

Forecasting Techniques

The evaluation of the complex and broad issues involved in health care inspections requires a wide variety of analytical tools. Numerous forecasting techniques have potential in inspections; this report illustrates how these forecasting techniques can aid in the function of oversight. In order to comprehend the techniques to be discussed, a diagram has been provided that illustrates, overall, the types of forecasting techniques generally employed (see Figure 1). Note that this chart is not meant to be comprehensive; rather, it is a tool for describing the general similarities and differences in the techniques.

Forecasting can be discussed in terms as either qualitative or quantitative techniques. The qualitative methods are designed to deal with soft, or non-numeric, data; the quantitative require hard numerical data gathered over several time periods. Obviously, the better the data, the more accurate the technique; but in those instances when no data exist, the qualitative techniques play a valuable role. Some techniques are quite common - expert opinion is used daily by upper management. Brainstorming allows several individuals to create ideas regarding the future; Delphi and panel consensus permit structured solicitation of group opinions. Scenario development is used to develop alternate plans, based upon potential occurrences in the future. Contextual mapping provides a pattern for technology transfer through its analysis of characteristics, and monitoring is a procedure for gathering information from journals, newspapers and other press material regarding changes in status quo.

Quantitative techniques can be divided into two categories - time series and modeling. Time series techniques utilize patterns of the data set under study to predict the future of that time series; modeling relates the behavior of the data set under study to other, independent, variables. The two fields overlap in multivariate intervention analysis and transfer functions.

For time series techniques, the driving force is the assumption that past behavior predicts future behavior; patterns from the past data in the time series are employed to estimate future patterns in the data. Under the additional assumption that the more recent data points are the more "informative" data, exponential smoothing takes advantage of the related error terms over time to make future predictions. Moving averages also work the emphasis of more current data by averaging recent data points to make forecasts. Both techniques are particularly accurate for short-term forecasting. Trend projections, be they linear, growth, logarithmic, etc., perform long-term projections of the long term trends. If seasonality variation occurs in the data, there exist time series techniques designed to handle such data. And, for the more complex designs, the ARIMA model is designed to handle almost any type of forecast characteristic (constant, trend, seasonal or cyclical, with moving average or autoregressive components).

Modeling provides techniques that relate the variable of interest, termed the dependent variable, with other, independent, variables in order to explain that behavior of the dependent variable. Multivariate regression, in its various forms, is the primary technique employed to combine independent with dependent variables; econometric modeling is a variation of multiple regression. Input/output analysis allows the creation of a matrix that analyzes how the various inputs and outputs in products and services relate to one another; this technique is primarily employed on a national level. Cross-impact also performs a type of sensitivity analysis of multiple variables to determine interrelated behaviors, whereas precursor analysis uses the behavior of a comparable variable to forecast the activity of the variable of interest. Intervention and transfer analysis falls into an overlap area of both time series and modeling, by using time series models to relate the dependent to independent variables.

Not included in the diagram, but worthy of note in the arena of federal forecasting, are the methods labeled as normative techniques, which first establish objectives to occur in a future period of time and then determine the path(s) necessary to achieve those objectives. The normative techniques make use of all the previously mentioned methods to some degree. The emphasis, however, is not where the future will take us, but rather how do we get to where we want to be. Variations of "backcasting" are employed to determine the viability of the goals and the potential paths to those goals.

Health Care Applications in Forecasting

The techniques mentioned in the previous section each have their potential in health care analysis from an oversight, as well as Continuous Quality Improvement (CQI), point of view (Kritchevsky & Simmons, 1991). This

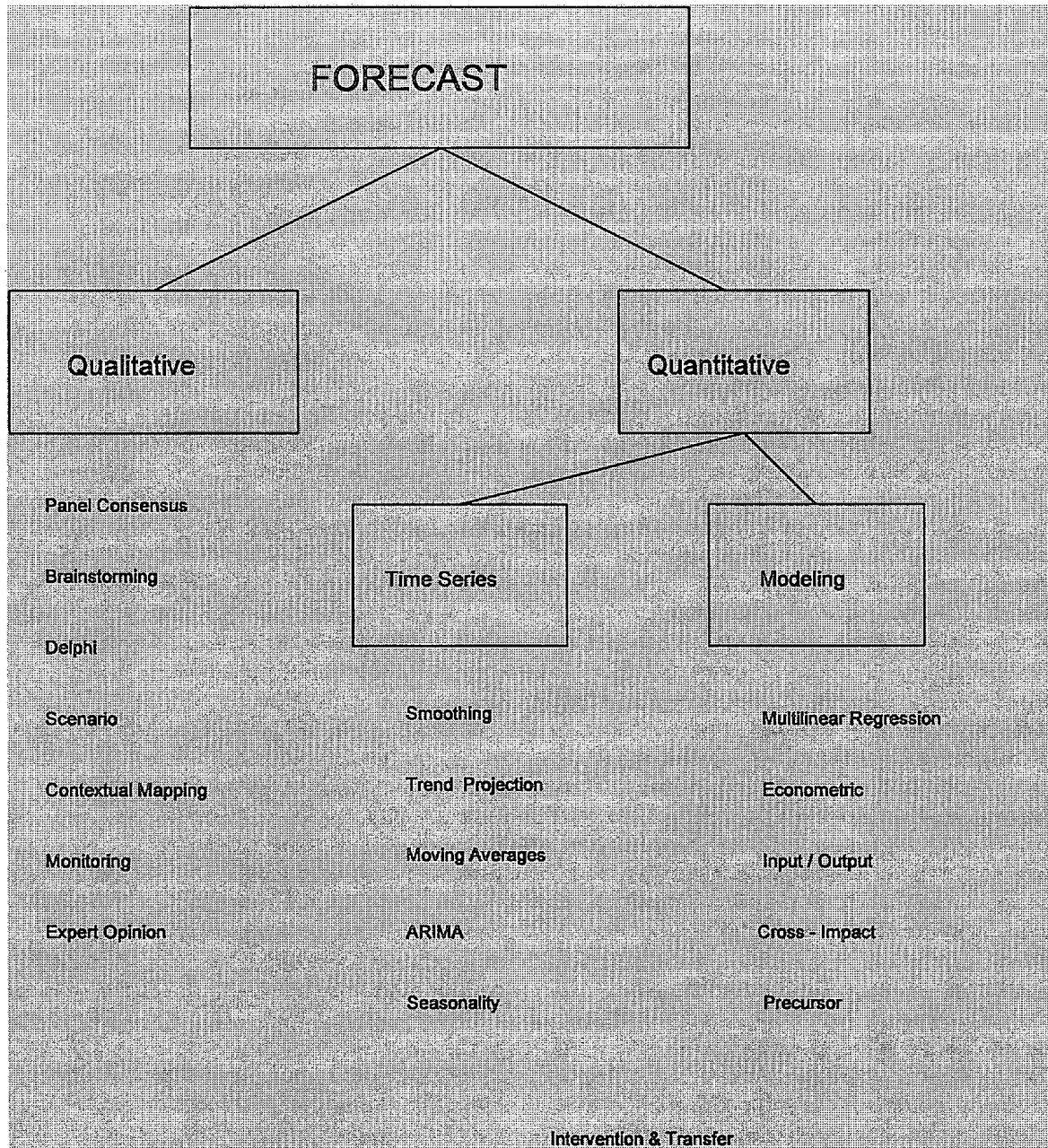


Figure 1. Diagram of forecasting techniques

section of the paper describes, in general, how some of the techniques have been, or are planned to be, used in OHI for health care review.

With respect to the employment of qualitative techniques, scenario analysis plays a key role in determining how the Clinton national health care plan will impact the Veterans Affairs medical centers (VAMCs). Scenario analysis allows us to ask the questions, "If the plan provides all in the US with health care insurance, what is the impact on who will use the VAMCs?" Alternate futures can be hypothesized to permit speculation in a structured format. Another qualitative technique, expert opinion, is also used to elicit opinions where little to no data exist; another term for this is expert judgment. With expert opinion, question may be asked such as, "Given your knowledge of the workings of Congress, what is your opinion as to the likelihood of permitting non-veterans into the VAMCs?"

In the quantitative techniques, one of the simplest techniques is also one of the best techniques for providing forecasts - exponential smoothing. As will be illustrated later, the techniques show great promise in the monitoring of mortality rates, as well as other quality assurance measures. Linear and non-linear trend regressions also have great potential in oversight; such capabilities are already employed in statistical process control (SPC) within health care facilities (VanderVeen, 1992). Data can be reviewed to determine simply if a trend does exist, the most basic form of trend analysis. Growth curves, a specialized form of nonlinear curves, can be used to estimate special populations, such as the growth of the AIDS / HIV infection population, in order to determine if the VAMCs are prepared to handle such an increase in that population. A spin-off from the quantitative time series techniques is the genre of quality control or quality assurance (QA) methods. Some of these techniques are designed to monitor behaviors over time - such as mortality rates (Kritchevsky & Simmons, 1991) or changes in DRG's, or diagnosis-related groups (Cohen et al., 1987).

Also in quantitative forecasting, various modeling possibilities exist to aid in reviewing health care proposals. Modeling is presently being employed to forecast estimates of physician need within the VAMCs, based upon the changes in the demographics of the veteran population. Modeling has also been used to compare costs of the VAMCs to that of their affiliated university hospitals, again permitting a forecast of future requirements in funding. Modeling has also been employed by the Health Care Financing Administration (HCFA) to provide projected mortality rates among hospitals treating Medicare beneficiaries (Krakauer & Bailey, 1991). Intervention analysis has potential in numerous quality assurance analyses, in the attempt to determine which actions or external impacts have altered the pattern of QA measures.

Normative techniques are also employed in health care oversight. A study of health progress in the United States (McGinnis et al., 1992) details several normative goals in health care, along with the measures to determine the degree of achievement towards these goals. The OHI could employ comparable methods to monitor the actions of the VA towards the proposed restructuring of the health care towards the national program being established by the Clinton administration.

Mortality Rate Analysis

As mentioned in previous section, the technique to be studied in detail is the use of exponential smoothing, in this example on mortality rate analysis. Deaths per discharge are reviewed within a medical center to determine if unexpected deviations have occurred in the time series. The difficulty lies in the term "unexpected," when the data themselves are highly variable. When the data exhibit such high degrees of variability from time period to time period, then the task of identifying "unusual" deviations becomes onerous.

But such a technique does exist, as used in production and operations management. Tracking signals are used on smoothed data to signal deviations greater than would be expected from random fluctuations alone. The single exponential smoothing model

$$s_t = \alpha \cdot x_t + (1-\alpha) \cdot s_{t-1} \quad , \text{ where}$$

x_t is the time series under review (mortality rates, in this case) and s_t is the smoothed value at time t , and α is the smoothing constant with a value somewhere between 0 and 1 (Brown, 1963; Gardner, 1985; Gardner & Dannenbring, 1980). This model is appropriate for those instances in which the analyst does not expect to observe any long-term trends upward or downward in the data over time. The smoothed data can then be used to calculate the tracking signals over time to illustrate significant shifts in the data (Mandell & Bretschneider, 1984).

The smoothing model provides "smoothed" mortality rates against which the observed mortality rates can be compared to indicate significant movement not attributable to random error (Batty, 1969; Gardner, 1985a; Gardner, 1985b; Gardner & Dannenbring, 1980; McKenzie, 1978). Extreme deviations based upon variance from the smoothed data could be attributed to shifts in the mortality rate, which can take the form of one or more distinct forms of interventions within a time series. First, suppose an incident of severe food poisoning occurs at a facility; this change in mortality rates would be in the form of a spike (or pulse) in the time series. The spike or pulse intervention, as shown in Figure 2, indicates a short term intervention that causes a shift in the time series over a period of time, and then returns to the previous state. Another change in death rates for a medical ward could be a policy change in sending the more severely ill patients to a different ward; this change in death rates would take the form of a step function, in which the level of the data shifts. The step intervention, as shown in Figure 3, also occurs in a short period of time, but the level of the time series remains at the new level, rather than return to the previous state. The third form of an intervention could occur due to a gradual degradation in quality of care; such a change would alter the slope of the underlying trend. The slope intervention as shown in Figure 4, occurs over a longer period of time as a long-term gradual change in the level of the time series; i.e., the slope changes. All of these interventions would be characterized by some deviation of the observed mortality rate from those forecast by the smoothing model. It is hypothesized in this study that significant deviations between the model and the observed data can be interpreted as indications of interventions in the underlying equilibrium (Trigg, 1964; Trigg & Leach, 1967; Mandell & Bretschneider, 1984).

To create a mechanism of identifying significant changes in the time series, the variance, σ_e^2 , of the forecast error, $e_t = (x_t - \hat{x}_t)$, needs to be taken. For simple exponential smoothing, the standard deviation for a single period forecast error is

$$\begin{aligned} \hat{\sigma}_e(t) &= ((2-\alpha)/2)^{1/2} \cdot \sigma_e(t) \\ &= ((2-\alpha)/2)^{1/2} \cdot (1.25 \cdot \hat{MAD}_t), \text{ where} \\ \hat{MAD}_t &= \alpha e_t + (1-\alpha) \cdot \hat{MAD}_{t-1} \end{aligned}$$

is the estimate of the Mean Absolute Deviation (Montgomery, Johnson & Gardiner, 1990). Dividing each observed forecast error by the above standard deviation, $S.E. \equiv e_t / \hat{\sigma}_e(t)$, provides an indicator of the relative magnitude of the observed error. This term shall be referred to as the standardized error term (S.E.). If used as a type of confidence band, extreme deviations as noted by the indicator could be associated with any of the interventions mentioned, but primarily with a pulse or step intervention.

A tracking signal can also be employed to detect significant changes from the time series pattern. Several types of tracking signals are presently employed to detect underlying shifts in the time series data, such as CUSUM and smoothed-error forecast monitoring schemes (McClain, 1988). The smoothed tracking signal technique will be employed here to allow consistency in smoothing calculations with the previously designed signal.

Using the previous calculation of the estimate of the MAD, the smoothed error tracking signal can be calculated as

$$\begin{aligned} T.S. &= |Q_t / \hat{MAD}_t| \text{ where} \\ Q_t &= \alpha \cdot e_t + (1-\alpha) \cdot Q_{t-1} \end{aligned}$$

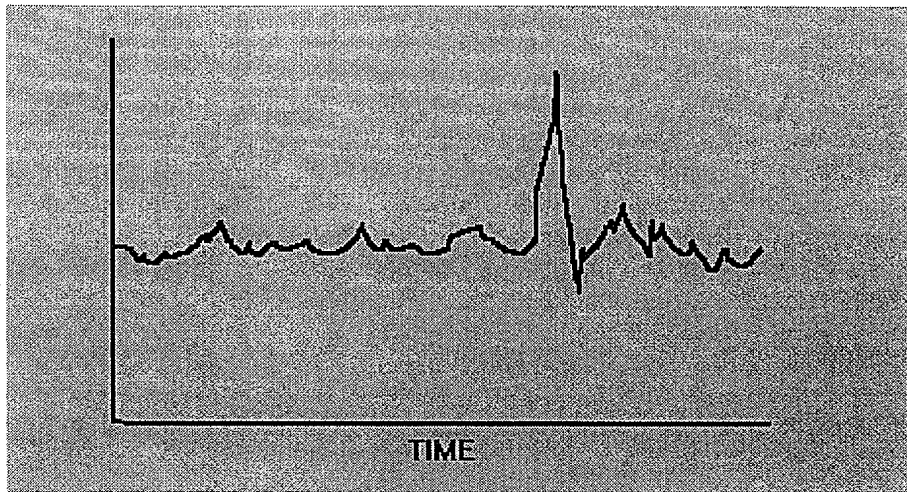


Figure 2. Spike or Pulse Intervention

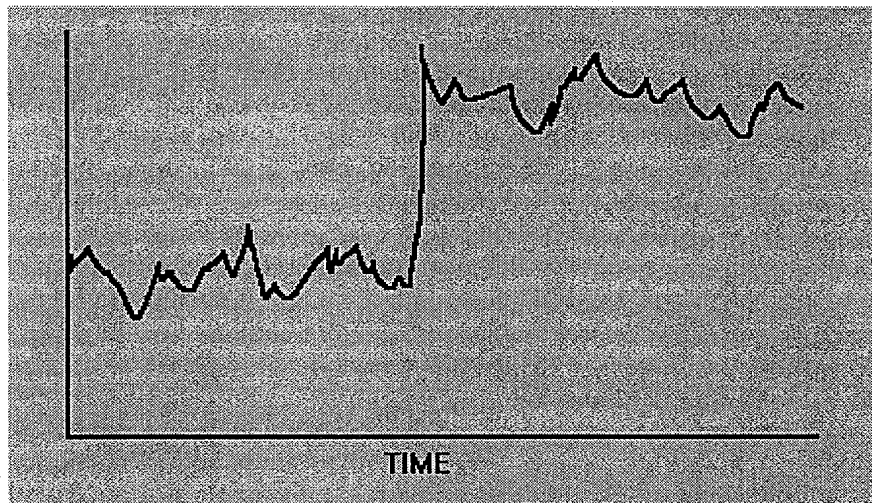


Figure 3. Step Intervention

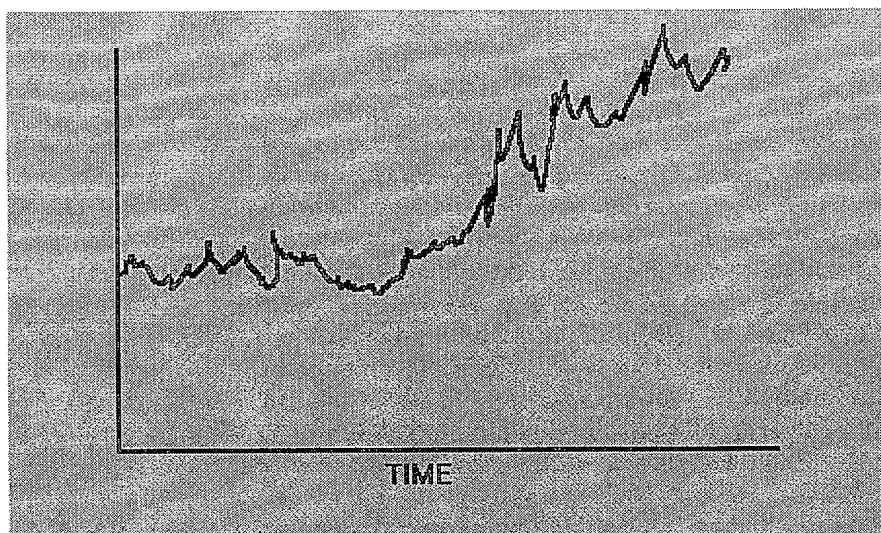


Figure 4. Slope Intervention

If we set

$$\sigma_Q^2 = (\alpha / (1 + \beta)^2) \cdot \sigma_e^2, \text{ then} \\ |Q_t / \text{MAD}_t| \left(\frac{1.25 \cdot \sqrt{\alpha}}{1 + \beta} \right) \cdot \kappa_1 = \kappa_3, \text{ where}$$

κ_1 is a standardized score indicating the degree of confidence desired (usually a value between 2 and 5). When the tracking signal exceeds κ_3 , which is usually between 0.2 to 0.5, this signal would be interpreted as indicating a movement of the observed values away from the smoothed values more than would be expected by random movement. Whereas the standardized error (S.E.) responds to abrupt changes in the time series, the smoothed error tracking signal (T.S.) has the capability of identifying longer-term changes in the data, such as changes in the slope.

The two tracking signals were applied to mortality rates at a medical center. The question being asked is: At what point in time did the intervention, if any, take place? Once the point in time is pinpointed, then the intervention that caused the change in the time series pattern can be possibly ascertained. The empirical analysis to establish the point of intervention consists of (1) the selection of the smoothing parameter for the mortality rate used to forecast the next month's rate, (2) the selection of the parameters for calculating the MAD and the tracking signal, and (3) the conduction of the actual test. The determination of the significant levels of the observed error and the tracking signal was set to that previously determined in a working paper which developed this procedure for the detection of interventions in stock price time series (DeSaix and Young, 1989).

In order to estimate an alpha value (α) for the smoothing process, an ARIMA(0,1,1) was run on the initial segment of the time series; the latter set of data points were those that were being scrutinized for potential changes in the data. The fitting of an ARIMA(0,1,1), which is the Box-Jenkins version of simple exponential smoothing, resulted in an alpha value that is the best fitting estimate of the smoothing parameter. Using the estimated alpha level, the two time signals were calculated - a standardized error (S.E.) and a smoothed tracking signal (T.S.). The standardized error indicates significant activity by some intervention when it becomes larger than $|2.0|$; the smoothed tracking signal indicates significance when it becomes larger than .35 (DeSaix & Young, 1989).

The results of the two signals are portrayed in Figures 5 and 6. The mapping of the standardized error (S.E.) in Figure 5 indicates a spike or step intervention in period 44. The horizontal arrow drawn at the value on the S.E. axes at +2.0 indicates the level up to which variability could be attributed to random error; the time series exceeds that level at period 44, which signals a potential step or pulse intervention. No S.E. value in the graph falls below the -2.0 S.E. value.

The mapping of the tracking signal (T.S.) in Figure 6 also has drawn a horizontal arrow at the value on the T.S. axis of .35, which is the maximum level of the tracking signal that could be attributed to random error. The figure indicates a movement above that maximum level beginning at period 36, as well as activity changes in period 44, which could be interpreted as an intervention that caused a change in slope.

The two tracking signals indicate two time periods of intervention. Whereas the standardized error indicates a pulse or step activity that initiates in period 44, the smoothed tracking signal highlights a slope change initiating at period 36 as well as period 44. What these two statistical measures signal are two potential time periods of change - a change in the slope of the death rate starting around period 36 and an additional intervention occurring near period 44. By being able to pinpoint the periods in time for the occurrence of interventions, an analysis into activity in the ward at those particular points in time could then determine the reasons for the interventions.

Results and Conclusions

Using signals such as the ones developed in this paper allows the time series data to 'inform' the analyst as to the time periods of potential interventions. This type of statistical quality control permits faster response to problems by driving the analyst in finding the interventions that caused the changes in the time series behavior, prior even to the knowledge of the intervention taking place. Time series analysis, as shown above, asks the question "Has any significant activity occurred?"

In general, forecasting techniques offer a new assortment of analytical tools to perform inspections within the Office of Inspector General. As each unique health care issue evolves, a forecasting technique could be devised to perform the analysis of that issue. A repertoire of analytical tools can now be developed to perform inspections through forecasting.

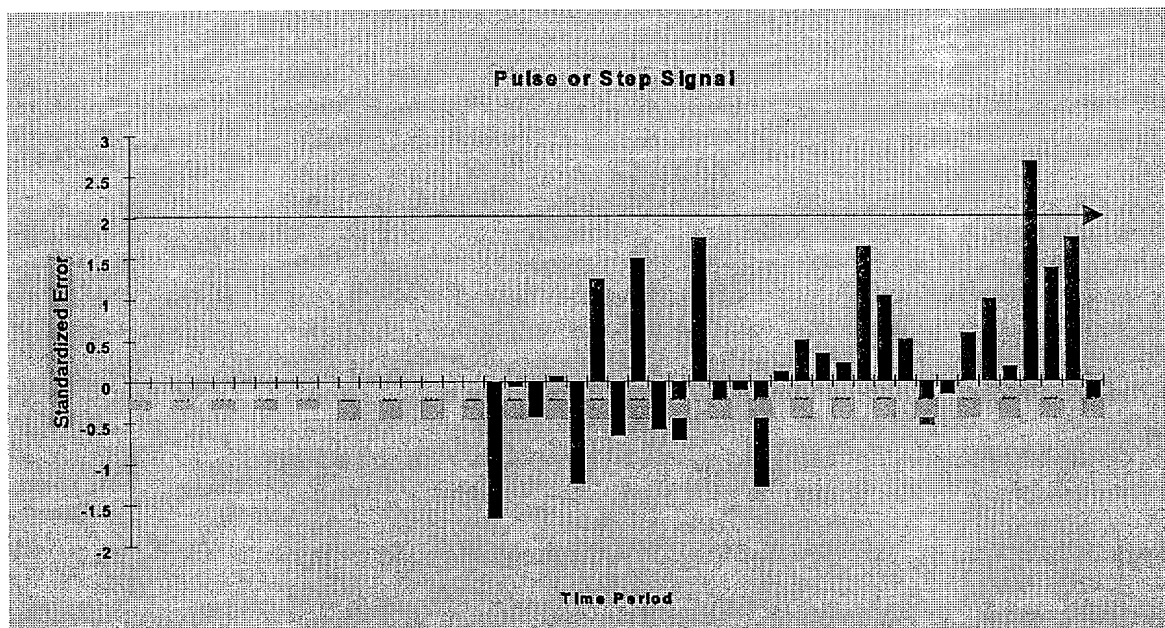


Figure 5. Standardized Error Tracking Signal

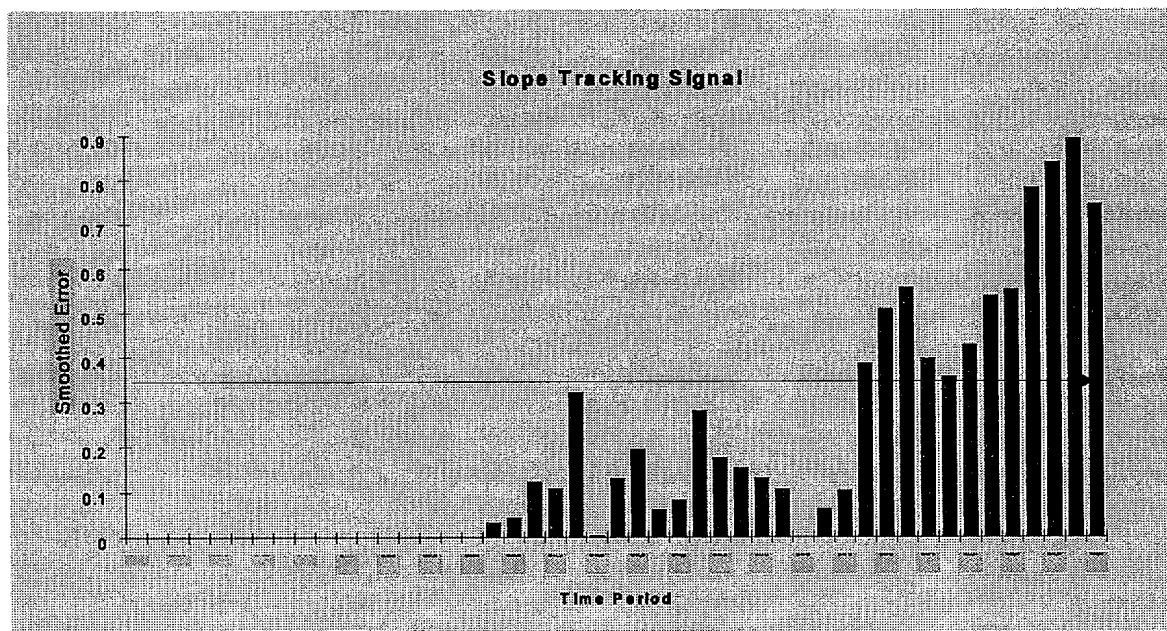


Figure 6. Smoothed Error Tracking Signal

References

- Batty M. *Monitoring an Exponential Smoothing Forecasting System*. Operational Research Quarterly 1969;26:319-325.
- Bretschneider S. *Estimating Forecast Variance with Exponential Smoothing*. International Journal of Forecasting 1986;2:349-355.
- Brown RG. Smoothing, Forecasting and Prediction of Discrete Time Series. Englewood Cliffs, NJ: Prentice-Hall, Inc.; 1963.
- Cohen BB, Pokras R, Meads MS, Krushat WM. *How Will Diagnosis- Related Groups Affect Epidemiological Research?* American Journal of Epidemiology 1987;126(1):1-9.
- DeSaix JC, Young P. *Interventions in Stock Price Time Series*. Working paper. 1989 June.
- Gardner ES, Jr. *CUSUM vs. Smoothed-Error Forecast Monitoring Schemes: Some Simulation Results*. Journal of the Operational Research Society 1985;36:43-47.
- Gardner ES, Jr. *Exponential Smoothing: The State of the Art*. Journal of Forecasting 1985;4:1-28.
- Gardner ES, Jr., Dannenbring DG. *Forecasting with Exponential Smoothing: Some Guidelines for Model Selection*. Decision Sciences 1980;11:370-383.
- Krakauer H, Bailey RC. *Epidemiologic Oversight of the Medical Care Provided to Medicare Beneficiaries*. Statistics in Medicine 1991;10(4):521-540.
- Kritchevsky SB, Simmons BP. *Continuous Quality Improvement: Concepts and Applications for Physician Care*. JAMA 1991; 266(13):1817-1823.
- Mandell MB, Bretschneider SJ. *Using Exponential Smoothing to Specify Intervention Models for Interrupted Time Series*. Evaluation Review 1984;8:663-691.
- McClain JO. *Dominant Tracking Signals*. International Journal of Forecasting 1988;4:563-572.
- McGinnis JM, Richmond JB, Brandt EN, Windom RE, Mason JO. *Health Progress in the United States*. JAMA 1992;268(18):2545-2552.
- McKenzie E. *The Monitoring of Exponentially Weighted Forecasts*. Journal of the Operational Research Society 1978; 29:449-458.
- Montgomery DC, Johnson LA, Gardiner JS. Forecasting and Time Series Analysis, 2nd. Edition. New York: McGraw-Hill Book Company; 1990.
- Trigg DW. *Monitoring a Forecasting System*. Operational Research Quarterly 1964;15:271-274.
- Trigg DW, Leach AG. *Exponential Smoothing with an Adaptive Response Rate*. Operational Research Quarterly 1967;18:53-59.
- VanderVeen. Loretta M. *Statistical Process Control: A Practical Application for Hospitals*. Journal for Healthcare Quality 1992;14(2):20-29.

Federal Forecasting: Occupation or Analytical Tool?

Karen S. Hamrick, USDA Economic Research Service and Fred Mills, Federal Management Partners, Inc.

Abstract

Should forecasting be recognized as an occupation in the Federal Government? Forecasting has developed into its own area of research in the last 10 years, yet the Federal personnel system does not recognize it as either an occupation or as a specialized discipline with unique requirements and qualification standards. Currently economists, statisticians, and other area specialists are hired by agencies as forecasters, with each agency applying its own definition of forecasting in establishing qualification, selection, and training criteria. With the increased emphasis on foresight in government is the potential for expanded roles and responsibilities for forecasters and will require a more rigorous, systematic approach. This paper outlines a method for formalizing forecasting in the Federal government.

Forecasting is more than a collection of techniques, it is a multi-step process. A broad spectrum of skills are needed for the process of forecasting. Setting standards for Federal forecasters would help insure that forecasters are well-versed in the many aspects of forecasting and could lead to better forecasts and improved credibility in Federal forecasts and Federal forecasters.

The Federal government provides a variety of forecasts to the public including those on agricultural markets, labor markets, the national economy, energy, and weather. These forecasts are produced by Federal employees from a variety of disciplines, including economics, statistics, demography, actuarial science, and meteorology. These forecasts are used by individuals, state and local governments, educational institutions, and private companies. For many of these forecasts, it would not be in any one individual's interest or ability to forecast. Federally-provided forecasts facilitate planning for both individuals and organizations.

Federally-produced forecasts are also used for policy-making and planning purposes, such as planning Federal programs and budgets. In an anticipatory government, the need is even greater for more accurate forecasts. Indeed, Wright, et al. stated "Planning for the future at the strategic or any other level is bound up with forecasting."¹ Good forecasts are needed for successful policymaking or program implementation. Klein asserted, "Careful forecasts, as accurate as possible, are central to the successful implementation of policy. There are fundamental reasons why policy makers cannot 'play by ear,' adjusting policy quickly to each unexpected deviation in economic outcomes."² In addition, it is crucial that forecasting be integrated into the policymaking or planning process for success.³

Forecasting as an Area of Study

Forecasting has emerged as an area of study over the last 20 years with a unique body of knowledge. There are now professional organizations of forecasters, whose members come from a variety of fields. In addition to the Federal Forecasters Conference, now in its sixth year, there is the International Institute of Forecasters (13 years), and the International Association of Business Forecasting (12 years).⁴ Four academic journals are devoted

¹Wright, et. al. (1986), p. 149.

²Klein (1984), p. 1.

³From Wright, et al. (1986): "Our aim is to show the danger of separating the production of a forecast from the use of the forecast in decision making and to show how these two aspects can be integrated in practice." (p. 139) And to complete the thought cited in the text above: "Planning for the future at the strategic or any other level is bound up with forecasting. The conclusion of this paper indicate that this bond needs to be more tightly woven than is often the case." (p. 149)

⁴Futurist predictions are not included here as past of forecasting. Many futurists belong to the World Future Society, which publishes The Futurist and the Futures Research Quarterly.

to forecasting, the International Journal of Forecasting, the Journal of Forecasting, Technological Forecasting and Social Change, and the Journal of Business Forecasting. In addition, courses specifically on forecasting now exist at universities, usually in the economics department, the decision sciences program, or the business or engineering schools.

Be that as it may, forecasting has not yet reached the status of a discipline. There are still no degrees in forecasting. The business forecasting community has considered a certification program,⁵ but that effort is in its infancy, and certification is controversial within the community of all forecasters. Many argue that forecasting should never be a separate field, for a forecaster must have substantive knowledge in order to make meaningful forecasts. Jenkins stated, "In trying to understand the relationships between the variables to be forecast and the policy and environmental variables, it is important that the forecaster understands the relationship between his system and its environment--and embodies this understanding in the conceptual model [of the system being studied]."⁶

Although forecasting is not a discipline, it clearly has become an interdisciplinary area of study. A consensus has emerged on what forecasting is and what forecasters need to know. This consensus defines forecasting as a process, not as a forecast product. See figure 1 for a representation of this process.⁷ We have used this flow-chart representation of the process to define the critical elements of forecasting in table 1.⁸ A forecaster would not be expected to know all the items listed under each element, but would need to have an awareness of them.

The Importance of Evaluation

Forecasting research was originally focused on existing techniques and the development of new techniques. Young commented:

In the earlier years of forecasting research, journal articles were inundated with 'competitions' designed to find the ultimate forecasting technique. The classic example is the M-competition, an abbreviation for the Makridakis competition,⁹ which provided over 1000 time series to a panel of forecasting experts. Seven experts in each of 24 methods essentially to determine which technique performed the best overall. What these competitions often failed to take into consideration were the characteristics of the data set, as well as the needs of the forecaster. The search was for the 'best' forecast procedure overall--not the best technique for a given set of time series of data and for a given forecasting need.¹⁰

⁵See Brown, et al. (1992) and Levenbach (1993).

⁶Jenkins (1982), p.6.

⁷This chart is based on the Box-Jenkins Algorithm [Box and Jenkins (1970), p. 19], and the Farnum and Stanton model selection flow chart [Farnum and Stanton (1989), p. 39].

⁸The authors used many references in compiling this table. Included among them are Armstrong (1985), Farnum and Stanton (1989), Granger (1989), Kennedy (1992), the International Journal of Forecasting, the Journal of Forecasting, the Proceedings of the Federal Forecasters Conferences (1990, 1991, and 1992), and discussions with Peg Young and Herman Stekler.

⁹Makridakis, A. et al. "The Accuracy of Extrapolation (Time Series) Methods: Results of a Forecasting Competition," Journal of Forecasting, 1982, Vol. 1, pp. 111-153.

¹⁰Young (1991), p. 35.

Figure 1--The process of forecasting

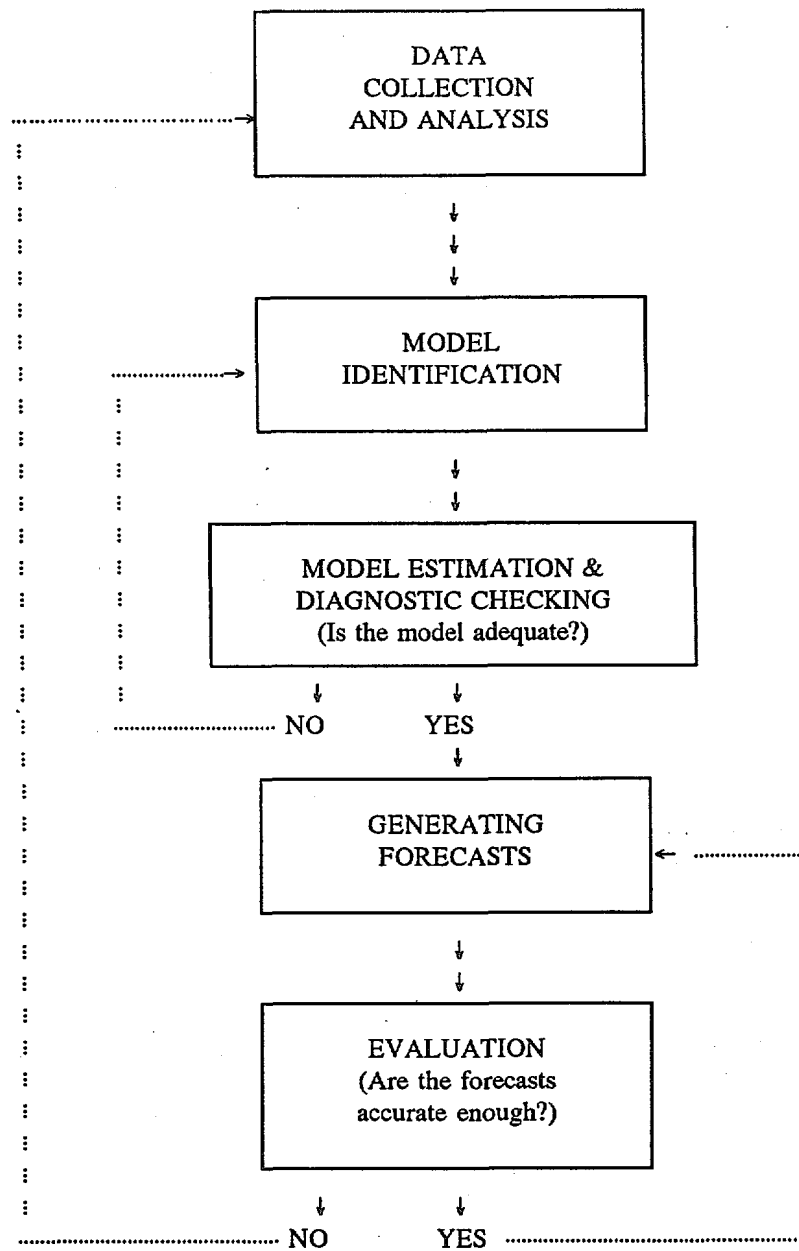


Table 1--The critical elements of forecasting

DATA COLLECTION AND ANALYSIS	MODEL IDENTIFICATION	MODEL ESTIMATION & DIAGNOSTIC TESTING	GENERATING FORECASTS	EVALUATION
<ul style="list-style-type: none"> - survey methods - understanding the psychology of the survey process - knowledge of data sources - using data/data analysis, including: <ul style="list-style-type: none"> - trend/no trend - seasonal adjustment - unit roots, stationarity 	<ul style="list-style-type: none"> - various techniques, including: <ul style="list-style-type: none"> - time series analysis/extrapolation methods - causal/econometric models - leading indicators - Delphi method - combining forecasts - input-output analysis - population forecasting - actuarial methods - segmentation methods - bootstrapping/expert systems - judgmental approaches - technology forecasting - meteorological models 	<ul style="list-style-type: none"> - choosing an appropriate technique - idea of parsimony - awareness of use of judgement in forecasting process: <ul style="list-style-type: none"> - assumptions - choosing model - choosing data - determining what can and cannot be forecast - cost/benefit considerations - diagnostic testing 	<ul style="list-style-type: none"> - point forecasts versus interval forecasts - scenario/sensitivity analysis - confidence interval around the forecast - using forecasts - presenting and communicating forecasts 	<ul style="list-style-type: none"> - variety of forecast accuracy measures - evaluating a single forecast versus evaluating a group of forecasts - choosing appropriate benchmark for comparison of forecast accuracy - accepting/not accepting results - using evaluation to choose a forecasting technique and to improve future forecasts

Later, the forecasting community realized that there is no ultimate technique, and the "the data drives the forecast, not the forecasting technique."¹¹ With that understanding the focus shifted to the importance of evaluating forecasts. In 1989 Bretschneider and Gorr commented,

Over the last ten years, forecasting has made great strides in developing its own research methods and a unique body of knowledge. During this time forecasting research has become less axiomatic and deductive and more empirical. Much of the emphasis in research has moved away from the development of new techniques towards the evaluation of existing techniques.¹²

Jenkins stated, that in order for forecasting to be effective in an organization, "forecasting should be seen as a learning process, so that when the monitored forecast errors are larger than expectation, remedial action can be taken to improve matters in the future."¹³ The General Accounting Office (GAO) engages in the most pointed forecast evaluations of all. Their approach is as follows:

Seven specific evaluation questions are used [by GAO] in assessing forecast accuracy. The methodology emphasizes comparing the forecasts to the actual subsequent event. The evaluation questions include the following:

1. What methodology is used for forecasting the event?
2. Who uses the forecasts?
3. How can forecast accuracy be measured?
4. How accurate are the forecasts?
5. Are the errors we identified "reasonable?"
6. What are the implications of forecast error on private sector as well as on the government policy, program, and budget decisions?
7. How can forecasts be improved?¹⁴

Clearly the forecasting community has adopted the view that forecasting activity is a process, and a crucial part of that process is evaluation. The recognition of evaluation can be considered a benchmark indicating that forecasting has attained status as a separate and distinct area of study and specialization. Evaluation requires that a field "close the circle," as in figure 1, and become self-contained. With evaluation, the field defines its own standards and criteria which make it separate and distinct.

Recognition in the Federal Government

If we accept that forecasting has emerged as a discrete area of study and specialization over the last 20 years, with its own body of knowledge and techniques, a logical question for Federal forecasters and managers to ask is: To what extent does the Federal personnel system recognize and reflect this new specialization? Specially, do the criteria and standards applied in evaluating and selecting candidates for Federal jobs which involve forecasting, as well as those used to establish the grade and pay levels for these positions accurately describe and assign value to the specialized knowledge and skills required for forecasting?

To help answer this question, a sample set of 22 Federal occupations was identified which reflects those job series whose members were thought most likely to be engaged in forecasting activities. (See table 2.) The sample set includes Economist and Statistician, the two occupations with the largest representation among participants in the Federal Forecasters Conference, as well as a number of other obvious candidates such as:

¹¹Young (1991), p. 35.

¹²Bretschneider and Gorr (1989), p. 305.

¹³Jenkins (1982)2p. 17.

¹⁴Solenberger (1991), p. 25.

Table 2--List of 22 Federal Occupations Most Likely to Involve Forecasting

OCCUPATION	SERIES	REFERENCES TO FORECASTING IN OPM DOCUMENTATION	
		Qualification Standards	Classification Standards (year classification standard was written)
Actuary	1510	professional actuarial society training/ examination requirements and standards listed as qualifying at various grade levels	focus on programmatic applications; reference to actuarial science, but no real description of methods, techniques (1967)
Appraiser	1171	none	mentions multiple regression as technique employed by appraisers (1972)
Budget Analyst	560	none	reference to long-range cost estimation, industrial fund budget forecasting, multi-year (3-5) projection; no specifics (1981)
Community Planner	020	none	none (1973)
Economist	110	none	brief reference to forecasting in context of econometrics, identification of methodology as sub- specialty; no discussion of methods, techniques (1963)
Financial Analyst	1160	none	short-term/long-term financial plans, studies; no specifics (1966)
Financial Manager	505	none	none (1969)
Industrial Engineer	896	none	none (1969)
Industrial Specialist	1150	none	none (1961)
Intelligence Analyst	132	none	none (1960)
Management/Program Analyst	343	none	general reference to quantitative methods; no specifics (1990)

Table 2--List of 22 Federal Occupations Most Likely to Involve Forecasting (con't)

OCCUPATION	SERIES	REFERENCES TO FORECASTING IN OPM DOCUMENTATION	
		Qualification Standards	Classification Standards (year classification standard was written)
Mathematical Statistician	1529	none	none (1961)
Mathematician	1520	none	brief reference to forecasting as sample activity at GS-9 level (1965)
Meteorologist	1340	minimum of six semester hours in weather analysis/forecasting methods required	extensive reference to forecasting as a key job function, area of specialization; no discussion of methods, techniques (1972)
Operations Research Analyst	1515	none	discusses modelling in general terms, no specifics; notes that operations research does <u>not</u> have its own body of principles and laws (1967)
Policy Analyst (functional title)	various (e.g., 101)	forecasting listed as sample activity for one area; optional assessment of qualifications in survey, simulation, and regression methodology	reference to quantitative methods, simulation, modelling; forecasting as sample activity in one area, no specifics (1981)
Psychologist	180	none	description of research specialty refers to statistical/mathematical techniques; no specifics (1968)
Sociologist	184	none	list Demography as one of five specialty areas; regional growth forecasts, population projections noted as sample products, no specifics (1988)
Statistician	1530	none	none (1961)
Trade Specialist	1140	none	none (1963)
Transportation Industry Analyst	2110	none	none (1969)
Quality Assurance Specialist	1910	none	none (1983)

Meteorologist and Operations Research Analyst, many of whose practitioners perform forecasting work as part of their ongoing responsibilities. To determine if forecasting is recognized as a part of these occupations, the personnel qualifications standards (used for hiring) and position classification standards (used for setting the grade of a position) issued by the Office of Personnel Management (OPM) covering each were analyzed and reviewed for references or descriptions of forecasting functions, techniques, and methodology.

In general, the OPM documentation examined does not recognize forecasting either as a specialty area within individual occupations, or even as an analytical tool applied in accomplishing the various responsibilities associated with the sample occupations. With the exception of two occupations (Actuary and Meteorologist), the personal qualifications standards examined do not define even a minimum level of preparation or competence in forecasting as a prerequisite to hiring. The position classification standards--used by personnel specialists and managers to assign jobs to General Schedule (GS) grade levels--include some limited references to forecasting, but in no case do they provide any specific discussion of forecasting techniques, methods, or the overall forecasting process, i.e., those elements and criteria defined previously in this paper (figure 1 and table 1).

A partial explanation for this absence of recognition can be inferred from the issue dates of the classification standards reviewed. Most are of 1960's vintage, with the standards for the Economist and Statistician occupations, for example, issued in 1963 and 1961 respectively, predating the emergence of forecasting as a distinct specialty area. In addition, the personal qualifications standards, while generally much newer (three to five years old), are for the most part generic in nature and cover broad groups of occupations, e.g., all "professional" positions, with minimal occupation-specific guidance and criteria.

The net effect of this absence of recognition is that in many cases, Federal managers are evaluating and selecting candidates for jobs involving forecasting without any guidelines or criteria for ensuring that their new hires possess the requisite knowledge and skills to perform successfully. On the position side, jobs are established and grade levels defined absent any consideration of the forecasting work involved. In cases where forecasting represents a substantial portion or even a majority of a given position's duties, the position may be incorrectly graded. These problems are even more pronounced in cases where managers or selecting officials are not themselves forecasters.

Approaches to Formalization

Given this state of affairs, should efforts be made to create a new Federal occupation for forecasting, with a formally defined and officially sanctioned set of personal qualification and job grading standards applicable government-wide? Table 2 includes several examples of existing Federal occupations, such as Actuary and Operations Research Analyst, which are based on highly specialized, quantitative disciplines and whose practitioners--not unlike Federal forecasters--apply their expertise to a variety of policy, programmatic, and management issues.

Table 3 lists the four general criteria applied by OPM in determining whether or not a new Federal occupation should be established: significance of the population to be covered throughout the government; existence of a commonly recognized body of specialized knowledge; whether or not the proposed occupation is clearly separate and distinct from other, existing Federal occupations; and the extent to which the agencies with the largest numbers of covered employees agree that a new occupation is necessary and appropriate. Population significance reflects not only the potential size of the new occupation but also its distribution across agencies and departments, the organizational and grade levels at which covered positions are found, and the general visibility or importance of the group. Consensus among affected agencies, while an informal consideration, is also nevertheless a critically important element in establishing a new occupation.

Application of these criteria to Federal forecasting in its present state suggests several barriers to separate occupational status for forecasters. Most problematic at this point appears to be the separate and distinct issue--i.e., can forecasting logically be detached from the professional and subject matter knowledge associated with the (current) occupations of its practitioners--economists, statisticians, demographers, etc.? The likely answer, as already noted in this paper, is that it cannot; to paraphrase Jenkins, the forecasters must have an indepth understanding of the context and environment in which his or her skills are to be applied. Unlike the Actuary, who need not be a health policy specialist to project the impact on mortality rates of changes in access to primary medical care, or the Operations Research Analyst, who may require only a basic understanding of military science to

construct an effective battlefield simulation, the forecaster must be thoroughly grounded in the particular discipline associated with the forecasts he or she is expected to generate.

Table 3--Establishing a New Federal Occupation

OPM Criteria - New Occupations:

• Population

How many positions/people will be covered by the new occupation government-wide, and does this constitute a "significant" population?

• Common Body of Knowledge

Is there a commonly recognized body of (specialized) knowledge and/or skills required for all practitioners of this occupation across this government?

• Separate and Distinct

Is this body of knowledge, and its various applications and outputs, clearly separate and distinct from other Federal occupations or groups?

• Consensus

Do the agencies most directly affected, i.e., those with the largest number of covered employees/positions, concur in the establishment of a new occupation?

OPM Documentation - New Occupations:

• Classification Standard

Occupational information, i.e., functions, terms, environment; factor level criteria for assigning covered positions to appropriate GS grades.

• Qualification Standard

Description of education, experience, and/or certification and training criteria; entry level test/examination requirements.

And while some practitioners may disagree with this position, the lack of consensus reflected by such disagreement only further weakens the case for establishing forecasting as a separate occupation. Still, forecasting is more than just an analytical tool or statistical technique.

Separate occupational status is not the only approach to addressing the problems resulting from forecasting's current lack of recognition within the Federal personnel system. One possible alternative would be to pursue development of a functional classification guide covering forecasting activities as practiced by members of a variety of Federal occupations and series. Rather than focusing on a single discipline, such guides provide criteria for evaluating specialized functions across multiple occupational series and job titles. Functional guides describe the work performed in essentially generic form, but in sufficient detail to allow for thorough evaluation and appropriate recognition of the activity in determining grade levels and qualification requirements for covered positions.

A good example is the Policy Analyst Grade Evaluation Guide. This document, whose current issuance dates from 1981 (see table 2), provides criteria for evaluating professional and scientific positions which are primarily concerned with the analysis of public policy issues and the provision of expert, objective advice and guidance regarding these issues to policy makers and senior officials. The occupational titles of positions which may be covered by this guide include such diverse series as Social Science Analyst, Physical Scientist, Education Program Specialist, or Environmental Engineer. The common denominator, however, is that the work assigned to positions covered by this guide primarily involves the application of policy analysis skills and techniques, combined with subject matter expertise in a particular discipline or field, to produce a specialized end product or service which adds value to the policy-making process. The grade level evaluation of the analysis function, but retention of the subject matter/discipline title, recognizes the inseparability of these two elements in the practice of this craft within the Federal government.

There appear to be strong parallels in this example to the current state of Federal forecasting. Practiced by members of a variety of occupations and disciplines, forecasting has emerged in recent years as a separate functional area which can now be defined on the basis of a commonly recognized process and knowledge base (see table 1). Many Economist, Statistician, and other positions throughout the Federal government are primarily concerned with forecasting, and the importance of the forecast product can only be expected to increase as efforts to "reinvent" government and move towards a more proactive, "anticipatory" approach proliferate and gather momentum.¹⁵ The role and value of forecasting in the policy-making process, as well as the importance of subject matter expertise to the practice of forecasting, are also well established.

Benefits from Formalization

Formalizing forecasting in the Federal government through development of a functional classification guide could produce a number of benefits. First and foremost, managers would be provided with an up-to-date description of forecasting work and criteria for evaluating candidates for forecasting positions. Over time, the availability of this information should produce higher quality selections, enhancing the credibility of the overall workforce as well as the forecasts it produces. Second, formalization would help insure that positions involving forecasting are accurately classified, and in turn, individuals who occupy those positions are appropriately compensated. In some cases, there may even be classification changes, i.e. grade increases, for particular positions once forecasting work is fully considered in setting grade levels.

Third, formalization may also draw attention to forecasting among agencies interested in establishing a forecast unit or enhancing their "anticipatory" capabilities, helping to expand the field and extending its presence and visibility throughout the government. Fourth, establishing a functional guide applicable to a variety of existing occupations, rather than creating an additional, new Federal occupation, would also avoid adding to the complexity of the current job classification system, which already includes more than 450 separate job series and titles. Rather,

¹⁵Such as that described in Osborne and Gaebler (1993).

a functional guide would provide a valuable tool to assist agencies in managing their current resources more effectively.

Finally, formalizing forecasting in the Federal government and recognizing its role and value is important simply in terms of acknowledging that the nature and complexity of the work performed by Federal employees, and in turn the level of expertise and sophistication required to perform that work, continues to grow and evolve. Recognition of forecasting on this basis, while intangible in its immediate impact, could nevertheless provide an important long-term benefit by communicating to Federal forecasters that their contributions and efforts are an integral and vital part of the process of governance.

Bibliography

Armstrong, J. Scott. Long-Range Forecasting: From Crystal Ball to Computer, second edition. John Wiley and Sons, New York, 1985.

Box, George E.P. and Gwilym M. Jenkins. Time Series Analysis: Forecasting and Control. Holden-Day, San Francisco, 1970.

Bretschneider, Stuart I. and Wilpen L. Gorr. "Editorial: Forecasting as a Science," International Journal of Forecasting. 1989, Vol. 5, No. 3, pp. 305-306.

_____. "Forecasting Practices in the Federal Government," Federal Forecasters Conference/1991: Papers and Proceedings. September 12, 1991, pp. 17-24.

Brown, Robert G., Jeanne M. Doll, Hans Levenbach, Len Tashman, and Lilian Shiao-yen Wu. "Business Forecasting: A Certifiable Profession? A Focus for Bridging the Gap Between Practice and Research," The Forum (The Joint Newsletter of the International Association of Business Forecasting and the International Institute of Forecasters). Fall 1992, Vol. 5, No. 3, p. 5.

Farnum, Nicholas R. and LaVerne W. Stanton. Quantitative Forecasting Methods. PWS-KENT Publishing Company, Boston, 1989.

Fildes, Robert. "Editorial: Influencing Forecasting Practice," International Journal of Forecasting. June 1992, Vol. 8, No. 1, pp. 1-2.

Granger, C.W.J. Forecasting in Business and Economics, second edition. Academic Press, Inc., Boston, 1989.

Hamrick, Karen S., editor. Federal Forecasters Conference 1990: Proceedings. Staff Report No. AGES 9109, U.S. Department of Agriculture, Economic Research Service, March 1991.

Holden, K., D.A. Peel, and J.L. Thompson. Economic Forecasting: An Introduction. Cambridge University Press, New York, 1990.

Jenkins, Gwilym M. "Some Practical Aspects of Forecasting in Organizations," Journal of Forecasting. January-March 1982, Vol. 1, Issue 1, pp. 3-21.

Kennedy, Peter. A Guide to Econometrics, third edition. The MIT Press, Cambridge, Massachusetts, 1992.

Klein, L. R. "The Importance of the Forecast," Journal of Forecasting. January-March 1984, Vol. 3, No. 1, pp. 1-9.

Levenbach, Hans. "Business Forecasters See Benefits to Certification," The Forum (The Joint Newsletter of the International Association of Business Forecasting and the International Institute of Forecasters). Summer 1993, Vol. 6, No. 2, p. 3.

Lienesch, Thomas, editor. Federal Forecasters Conference-1992: Papers and Proceedings. September 17, 1992.

FFC/93 Papers & Proceedings

Makridakis, Spyros. "Editorial: Forecasting in the 21st Century," International Journal of Forecasting. 1991, Vol. 7, No. 2, pp. 123-126.

Monaco, Ralph M. "Developing an Effective Forecasting Program: An Economic Approach," Federal Forecasters Conference/1991: Papers and Proceedings. September 12, 1991, pp. 74-77.

Osborne, David and Ted Gaebler. Reinventing Government. Plume (Penguin Group), New York, 1993.

Saunders, Norman C. and Howard N Fullerton, Jr., editors. Federal Forecasters Conference/1991: Papers and Proceedings. September 12, 1991.

Solenberger, David. R. "U.S. Department of Agriculture Forecasts: Personal Comments on GAO'S Forecast Evaluation Methodology," Federal Forecasters Conference/1991: Papers and Proceedings. September 12, 1991, pp. 25-28.

United States General Accounting Office. "Status of Open Recommendations." Washington, DC, 1993.

United States Office of Personnel Management. "Handbook of Occupational Groups and Series." Washington, DC, October 1990.

_____. "Position Classification Standards for Positions Under the General Schedule Classification System." Washington, DC, 1990.

_____. "Qualification Standards for Positions Under the General Schedule." Washington, DC, 1990.

Wright, D.J., G. Capon, R. Page, J. Quiroga, A.A. Taseen, and F.F. Tomasini. "Evaluation of Forecasting Methods for Decision Support," International Journal of Forecasting. 1986, Vol. 2, No. 2 (1986), pp. 139-152.

Young, Peg. "Developments in Forecasting in the Past Decade," in Federal Forecasters Conference - 1992 Papers and Proceedings, Thomas Lienesch, editor. September 17, 1992, pp. 35-36.