Market Mechanisms and Incentives: Applications to Environmental Policy

A Workshop sponsored by U.S. Environmental Protection Agency's National Center for Environmental Economics (NCEE) and National Center for Environmental Research (NCER)

Resources for the Future
1616 P Street, NW, Washington, DC 20036
(202) 328-5000

Wednesday, April 29, 2009

Disclaimer

U.S. Environmental Protection Agency (EPA) National Center for Environmental Economics (NCEE) and National Center for Environmental Research (NCER)

Market Mechanisms and Incentives: Applications to Environmental Policy

Resources for the Future
1616 P Street, NW, Washington, DC 20036
(202) 328-5000

Wednesday, April 29, 2009

8:30 a.m. – 9:15 a.m. **Registration**

9:15 a.m. – 9:30 a.m. **Introductory Remarks:** *Julie Hewitt, Chief, Economic and Environmental Assessment Branch, Office of Water*

9:30 a.m. – 11:20 a.m. **Session I: Fuel Economy and Gasoline Prices**
Session Moderator: *Cynthia Morgan, EPA, National Center for Environmental Economics*

9:30 a.m. – 10:00 a.m. Imperfect Competition, Consumer Behavior, and the Provision of Fuel Efficiency in Vehicles
Carolyn Fischer, Resources for the Future

10:00 a.m. – 10:30 a.m. New Vehicle Characteristics and the Cost of the Corporate Average Fuel Economy Standard
Thomas Klier, Federal Reserve Bank of Chicago, and Joshua Linn, University of Illinois at Chicago

10:30 a.m. – 10:40 a.m. Discussant: *Gloria Helfand, University of Michigan, and EPA, Office of Transportation and Air Quality*

10:40 a.m. – 10:50 a.m. Discussant: *Winston Harrington, Resources for the Future*

10:50 a.m. – 11:20 a.m. Questions and Discussion

11:20 a.m. – 12:30 p.m. **Lunch** (On Your Own)

12:30 p.m. – 1:30 p.m. Panel Discussion: Role of Market Mechanisms and Incentives to Climate Change
Moderator: *Dick Morgenstern, Resources for the Future*
Panelist: *Joe Aldy, Special Assistant to the President for Energy and the Environment*
David McIntosh, Senior Counsel in the Office of Congressional and Intergovernmental Relations
Brian Murray, Duke University

1:30 p.m. – 1:40 p.m. **Break**

1:40 p.m. – 3:30 p.m. **Session II: Applications of Environmental Trading Programs**
Session Moderator: *Will Wheeler, EPA, National Center for Environmental Economics*

1:40 p.m. – 2:00 p.m. An Experimental Analysis of Compliance in Dynamic Emissions Markets: Theory and Experimental Design
John Stranlund, University of Massachusetts - Amherst, James Murphy, University of Alaska – Anchorage, and John Spraggon, University of Massachusetts - Amherst

	2:00 p.m. – 2:30 p.m.	Can Markets for Development Rights Improve Land Use and Environmental Outcomes? *Virginia McConnell, Elena Safirova, Margaret Walls, and Nick Magliocca, Resources for the Future*
	2:30 p.m. – 2:35 p.m.	Discussant: *Heather Klemick, EPA, National Center for Environmental Economics*
	2:35 p.m. – 3:05 p.m.	Preliminary Findings and Observations on Ohio's Great Miami River Water Quality Credit Trading Program *Richard Woodward, Texas A&M University*
	3:05 p.m. – 3:10 p.m.	Discussant: *Hale Thurston, EPA, National Risk Management Research Laboratory*
	3:10 p.m. – 3:30 p.m.	Questions and Discussion
3:30 p.m. – 3:40 p.m.	**Break**	
3:40 p.m. – 5:30 p.m.	**Session III: Winners and Losers in Cap and Trade** Session Moderator: *Charles Griffiths, EPA, National Center for Environmental Economics*	
	3:40 p.m. – 4:10 p.m.	Paving the Way for Climate Policy: Compensation for Electricity Consumers and Producers Under a CO_2 Cap and Trade Policy *Karen Palmer, Dallas Burtraw, and Anthony Paul, Resources for the Future*
	4:10 p.m. – 4:40 p.m.	When Does Cap-and-Trade Increase Regulated Firms' Profits? *Dave Evans, EPA, National Center for Environmental Economics; Ian Lange, University of Stirling; and Joshua Linn, University of Illinois at Chicago*
	4:40 p.m. – 4:50 p.m.	Discussant: *Ann Wolverton, EPA, National Center for Environmental Economics*
	4:50 p.m. – 5:00 p.m.	Discussant: *Terry Dinan, Congressional Budget Office*
	5:00 p.m. – 5:30 p.m.	Questions and Discussion
5:30 p.m.	**Adjournment**	

Market Mechanisms and Incentives: Applications to Environmental Policy

A Workshop sponsored by U.S. Environmental Protection Agency's National Center for Environmental Economics (NCEE) and National Center for Environmental Research (NCER)

Resources for the Future
1616 P Street, NW, Washington, DC 20036
(202) 328-5000

Wednesday, April 29, 2009

Disclaimer

U.S. Environmental Protection Agency (EPA) National Center for Environmental Economics (NCEE) and
National Center for Environmental Research (NCER)

Market Mechanisms and Incentives: Applications to Environmental Policy

Resources for the Future
1616 P Street, NW, Washington, DC 20036
(202) 328-5000

Wednesday, April 29, 2009

8:30 a.m. – 9:15 a.m. **Registration**

9:15 a.m. – 9:30 a.m. **Introductory Remarks:** *Julie Hewitt, Chief, Economic and Environmental Assessment Branch, Office of Water*

9:30 a.m. – 11:20 a.m. **Session I: Fuel Economy and Gasoline Prices**
Session Moderator: *Cynthia Morgan, EPA, National Center for Environmental Economics*

9:30 a.m. – 10:00 a.m. Imperfect Competition, Consumer Behavior, and the Provision of Fuel Efficiency in Vehicles
Carolyn Fischer, Resources for the Future

10:00 a.m. – 10:30 a.m. New Vehicle Characteristics and the Cost of the Corporate Average Fuel Economy Standard
Thomas Klier, Federal Reserve Bank of Chicago, and Joshua Linn, University of Illinois at Chicago

10:30 a.m. – 10:40 a.m. Discussant: *Gloria Helfand, University of Michigan, and EPA, Office of Transportation and Air Quality*

10:40 a.m. – 10:50 a.m. Discussant: *Winston Harrington, Resources for the Future*

10:50 a.m. – 11:20 a.m. Questions and Discussion

11:20 a.m. – 12:30 p.m. **Lunch** (On Your Own)

12:30 p.m. – 1:30 p.m. Panel Discussion: Role of Market Mechanisms and Incentives to Climate Change
Moderator: *Dick Morgenstern, Resources for the Future*
Panelist: *Joe Aldy, Special Assistant to the President for Energy and the Environment*
David McIntosh, Senior Counsel in the Office of Congressional and Intergovernmental Relations
Brian Murray, Duke University

1:30 p.m. – 1:40 p.m. **Break**

1:40 p.m. – 3:30 p.m. **Session II: Applications of Environmental Trading Programs**
Session Moderator: *Will Wheeler, EPA, National Center for Environmental Economics*

1:40 p.m. – 2:00 p.m. An Experimental Analysis of Compliance in Dynamic Emissions Markets: Theory and Experimental Design
John Stranlund, University of Massachusetts - Amherst, James Murphy, University of Alaska – Anchorage, and John Spraggon, University of Massachusetts - Amherst

	2:00 p.m. – 2:30 p.m.	Can Markets for Development Rights Improve Land Use and Environmental Outcomes? *Virginia McConnell, Elena Safirova, Margaret Walls, and Nick Magliocca, Resources for the Future*
	2:30 p.m. – 2:35 p.m.	Discussant: *Heather Klemick, EPA, National Center for Environmental Economics*
	2:35 p.m. – 3:05 p.m.	Preliminary Findings and Observations on Ohio's Great Miami River Water Quality Credit Trading Program *Richard Woodward, Texas A&M University*
	3:05 p.m. – 3:10 p.m.	Discussant: *Hale Thurston, EPA, National Risk Management Research Laboratory*
	3:10 p.m. – 3:30 p.m.	Questions and Discussion
3:30 p.m. – 3:40 p.m.	**Break**	
3:40 p.m. – 5:30 p.m.	**Session III: Winners and Losers in Cap and Trade** Session Moderator: *Charles Griffiths, EPA, National Center for Environmental Economics*	
	3:40 p.m. – 4:10 p.m.	Paving the Way for Climate Policy: Compensation for Electricity Consumers and Producers Under a CO_2 Cap and Trade Policy *Karen Palmer, Dallas Burtraw, and Anthony Paul, Resources for the Future*
	4:10 p.m. – 4:40 p.m.	When Does Cap-and-Trade Increase Regulated Firms' Profits? *Dave Evans, EPA, National Center for Environmental Economics; Ian Lange, University of Stirling; and Joshua Linn, University of Illinois at Chicago*
	4:40 p.m. – 4:50 p.m.	Discussant: *Ann Wolverton, EPA, National Center for Environmental Economics*
	4:50 p.m. – 5:00 p.m.	Discussant: *Terry Dinan, Congressional Budget Office*
	5:00 p.m. – 5:30 p.m.	Questions and Discussion
5:30 p.m.	**Adjournment**	

Imperfect Competition, Consumer Behavior, and the Provision of Fuel Efficiency in Light-Duty Vehicles

Carolyn Fischer

Abstract

We explore the role of market power on the cost-effectiveness of policies to address fuel consumption. Market power gives manufacturers an incentive to under- (over-) provide fuel economy in classes whose consumers, on average, value it less (more) than in others. Adding a second market failure in consumer valuation of fuel economy, a policy tradeoff emerges. Minimum standards can address distortions from price discrimination but do not provide broad-based incentives for improving fuel economy like average standards. Increasing fuel prices raises demand for fuel economy but exacerbates undervaluation and incentives for price discrimination. A combination policy may be preferred.

Contents

Imperfect Competition, Consumer Behavior, and the Provision of Fuel Efficiency in Light-Duty Vehicles

Carolyn Fischer*

Introduction

The regulation of fuel economy is one of the primary tools for controlling the emissions of greenhouse gases and other pollutants from passenger vehicles in the U.S., as well as for addressing energy security. Heightened attention to these issues has prompted a broader debate over reforming Corporate Average Fuel Economy (CAFE) standards, the current program that requires automobile manufacturers to meet standards for the sales-weighted average fuel economy of their passenger vehicle fleets. Potential reforms include not only strengthening standards, but also allowing fuel economy credits to be tradable, and adjusting standards according to vehicle characteristics like size.[1]

This study addresses an issue that has been overlooked in previous studies of CAFE standards and alternatives: that imperfect competition can affect manufacturer incentives to deploy fuel-saving technologies. While it is well known that market power affects price markups, the distributional effects of regulation, and even the fleet mix, its effects on the choice of fuel economy have been ignored. We explore the impact of this particular brand of market failure on the cost-effectiveness of tradable fuel economy standards and other market-based mechanisms to address automotive fuel consumption.

In particular, we investigate the role of market power among automobile manufacturers and of heterogeneity among consumers in their preferences for fuel economy. In this situation, manufacturers have an incentive to choose fuel economy to differentiate their product line,

---

* Fischer is a Senior Fellow at Resources for the Future, 1616 P Street, NW, Washington, D.C. 20036; e-mail: fischer@rff.org. Support by the US Environmental Protection Agency is gratefully acknowledged.

[1] The latter modification is being implemented for light trucks. Fischer and Portney (2004) discuss the case for making CAFE credits tradable.

segment consumers, and thus obtain higher prices for their fleet of vehicles. Meanwhile, CAFE standards impose certain constraints on these choices, by requiring manufacturers to meet an average rate of fuel consumption. An important question for evaluating reforms to CAFE standards is how do they interact with incentives for price-discrimination that may distort the provision of fuel economy in passenger vehicles?

Common sense dictates that consumers of different car classes are likely to have different preferences for fuel economy, in part because those preferences help determine the class they choose. For example, people more concerned about fuel economy—whether because they drive more, understand the costs better, or care about the environment—would be less likely to choose a large car. They may also be more likely to forego purchasing a car.

Empirical studies support this claim. Goldberg (1995), in her estimation of vehicle demand, finds that while consumers of large and small cars are similarly sensitive to prices, consumer demand for small cars is much more elastic with respect to fuel costs than is demand for large cars (Table 1). Luxury car demand is less sensitive to prices and basically insensitive to fuel costs.

Table 1: Results from Goldberg (1995) Log-Likelihood Demand Estimation

<table>
<tr><th>Model Choice:
Variable</th><th>Small Cars</th><th>Big Cars</th><th>Luxury and Sports Cars</th></tr>
<tr><td>price if purchased model before and income <= $75000</td><td>-4.747 (0.862)</td><td>-4.4 (0.602)</td><td rowspan="2">-1.223 (0.174)</td></tr>
<tr><td>price if purchased model before and income > $75000</td><td>-4.501 (0.356)</td><td>-3.745 (0.332)</td></tr>
<tr><td>price if first-time buyer and income <= $75000</td><td>-2.927 (0.328)</td><td>-3.076 (0.649)</td><td rowspan="2">-0.517 (0.220)</td></tr>
<tr><td>price if first-time buyer and income > $75000</td><td>-2.755 (1.277)</td><td>-2.171 (0.396)</td></tr>
<tr><td>Fuel costs</td><td>-7.143 (0.740)</td><td>-1.381 (0.744)</td><td>0.231 (0.931)</td></tr>
</table>

Similarly, Berry et al. (1995) find that the elasticity of demand with respect to miles per dollar “declines almost monotonically” with the car’s miles per dollar rating. They also conclude that luxury vehicle buyers are unconcerned with fuel economy, while purchasers of high-mileage cars are quite sensitive to it.

Using survey methodology, Kurani and Turrentine (2004) dispute the notion that consumers follow the rational economic framework for computing fuel consumption costs and weighing fuel economy tradeoffs. Still, if one accepts the idea that consumers behave as if they are seeking a certain payback period, "then averages such as the 'three-year' figure that Greene (2002) provides by example are of little interest. Almost every study conducted of consumer payback periods related to energy conservation shows a wide variety of (generally implied) discount rates. This suggests the existence of a market that can be segmented according to how long people are willing to be paid back."[2]

At the same time, there is certainly empirical support for the presence of market power in the automobile industry: the largest four firms account for 75.5% of the value of shipments in the automobile market and 95.7% of the light duty and utility vehicle market, and the Herfindahl Hirschman Index (HHI) for light vehicles overall is 2600 (2002 Economic Census), where above 1800 is the Justice Departments definition of a "highly concentrated" industry.

The accompanying table gives the market shares of light duty vehicle sales, according to NHTSA, for model year 2004, and those shares are significant for the top five. Furthermore, empirical evidence of significant brand loyalty (Train and Winston, *International Economic Review*, forthcoming) may also serve to reinforce the idea that auto manufacturers will recognize demand interactions across models within their fleet (in other words, that the fuel economy of one model is likely to affect demand for other models in the fleet as well). Thus, the conditions are ripe for market power to play a role in determining vehicle quality, including fuel economy.

MY 2004 Market Shares

Manufacturer:	*Market Share*
General Motors	26%
Daimler Chrysler	20%
Ford	18%
Toyota	13%
Honda	8%
Other	16%

However, modelers of automobile markets and their regulation have largely ignored the effects of consumer heterogeneity on the strategy of vehicle manufacturers for providing fuel

2 Kurani and Turrentine (2004), p. III.

economy. A variety of assumptions have enabled researchers to avoid this question. Many studies that allow for imperfect competition among vehicle manufacturers focus on responses in fleet composition, assuming that fuel economy and marginal production costs for each vehicle model are exogenously determined (Jacobsen 2008; Bento et al. 200?; Berry et al. 1995; Goldberg 1995, 1998; Kwoka 1983; Petrin 2002). While this assumption is useful for modeling short-run responses to policy or gas price changes, these studies cannot incorporate the longer-run response of changing the fuel consumption characteristics of the vehicle.

Other modelers have allowed manufacturers to choose fuel economy, but they have avoided the strategic problem by assuming perfect competition or by aggregating the market (e.g., Fischer et al. 2005, Kleit 2004, Greene et al. 2005). Similarly, Rubin et al. (2006) abstract from imperfect competition in the product market, while they do evaluate the impact of market power in the market for tradable fuel economy credits. Austin and Dinan (2005) allow imperfectly competitive firms to choose both price and fuel consumption rates for their vehicle models; however, they simplify the problem by assuming that consumers respond to average fuel costs in the same way as they respond to price changes. By this assumption, any fuel economy change then changes the fully-loaded vehicle price (ownership and operating costs) the same amount for all consumers, in which case manipulating fuel economy is no more effective at segmenting consumers than changing the retail price. However, that individual consumers would base their decisions on average consumer behavior is a strong assumption.

Given the degree of concentration among auto manufacturers and the wide range of consumer traits, none of these assumptions is satisfying. We show that when we incorporate consumer heterogeneity into a model of Bertrand price and quality competition, the results are very similar to those in the classic price-discrimination framework (e.g. Fischer 2005, Plourde and Bardis 1999). In this situation, fuel economy will tend to be over-provided in classes whose consumers value it more than others, and underprovided in classes whose consumers value it less than in others. In this manner, fuel economy represents a way to solidify market segmentation; by offering less fuel economy to consumers of large cars, for example, they can charge higher prices to small-car consumers, without worrying they will switch classes. Similarly, they can charge higher prices for large cars when they are charging more for highly efficient small cars than the large-car buyers are willing to pay. As a result, imperfect competition in the product market creates a market failure in the provision of fuel economy. Overlooking this market failure leaves out an important motivation for fuel economy regulation and will bias estimates of policy cost-effectiveness. On the other hand, as Fischer (2005) shows, average fuel economy regulation is not necessarily the best response to the distortions caused by price discrimination. We will thus consider modifications to tradable CAFE standards that can improve welfare.

This study extends important theoretical underpinnings for improving models of fuel economy policy and for conducting future empirical estimates of consumer and market behavior. These issues are critical for understanding the cost-effectiveness of policies like CAFE and whether they can enhance welfare as well as fuel economy. We complement the analytical work with numerical simulations to evaluate the potential magnitude of the problem. The goal is to inform policymakers about the extent to which fuel economy policy needs to keep an eye on market power issues, and the corresponding sensitivity analysis will also help identify key parameters for further empirical research.

Model

Theory of Producer Behavior

Consider a representative firm in our automobile manufacturing sector. For each vehicle class, the manufacturer chooses a retail price P_i, and a fuel consumption rate ϕ_i. We specify a model with Bertrand competition and product differentiation that can easily be extended to any number of manufacturers. A given manufacturer will care about how its choices will affect its entire product line, taking the choices made by other manufacturers as given.

The costs of manufacturing a vehicle of class *i* are $C_i(\phi_i)$, a function that is decreasing and convex in fuel consumption ($\partial C_i / \partial \phi_i < 0$ and $\partial^2 C_i / \partial \phi_i^2 > 0$). Consumer demand for class *i* is a function of the vector of prices and fuel consumption rates for all vehicles ($q_i(\mathbf{P}, \boldsymbol{\varphi})$). Demand in class *i* is decreasing in its own price and fuel consumption rate, and weakly increasing in those of other classes. Profits *V* for the representative manufacturer are the retail price less production costs, multiplied by the output of each model class:

$$V(\mathbf{P},\boldsymbol{\varphi}) = \sum_i (P_i - C_i(\phi_i))q_i(\mathbf{P},\boldsymbol{\varphi}) \tag{1}$$

Price. Maximizing profits with respect to the price of each vehicle class *i* leads to the following first-order condition:

$$\frac{\partial V(\mathbf{P},\boldsymbol{\varphi})}{\partial P_i} = q_i + \sum_j \pi_j \frac{\partial q_j}{\partial P_i} = 0 \tag{2}$$

where $\pi_j = P_j - C_j(\phi_j)$ is the own marginal profit (or total markup) for vehicle type *j*. Let $\eta_{ji} = \frac{\partial q_j}{\partial P_i}\frac{P_i}{q_j}$ be the cross-price elasticity of demand for vehicle class *j* with respect to a change in the price of *i*. Then we can rewrite the pricing condition as

$$P_i = \sum_j \pi_j (-\eta_{ji}) \frac{q_j}{q_i} \tag{3}$$

Rearranging, we can express the price as the sum of the vehicle's own costs, with a markup according to its own-price elasticity, and the cross-price responses, weighted by the marginal profits of the other vehicles in the manufacturer's fleet:

$$P_i = C_i(\phi_i)\frac{\eta_{ii}}{\eta_{ii}+1} + \sum_{j\neq i} \pi_j \frac{-\eta_{ji}}{\left(\eta_{ii}+1\right)}\frac{q_j}{q_i} \tag{4}$$

From (4) we see that a change in one model's costs, all else equal, causes a proportional increase in the price, with that ratio depending on the own-price elasticity of demand: $\Delta P_i = \frac{\eta_{ii}}{\eta_{ii}+1}\Delta C_i(\phi_i)$. Note that this result implies that more than 100% of the marginal cost increases are passed through to consumers. Equilibrium price changes, however, will reflect both ost changes and the demand interactions for all the vehicle classes. Thus, the effective pass-through rates for different model classes could be more or less than 100% in equilibrium. In the perfectly competitive case, as $\eta_{ii} \to -\infty$ the firm becomes a price taker and we get a 100% pass-through of cost changes into retail prices. However, most empirical studies have found positive markups, validating models of oligopolistic competition.[3]

Fuel Consumption Rate. Next, we consider the incentives with respect to fuel economy. The first-order conditions are

$$\frac{\partial V(\mathbf{P},\boldsymbol{\varphi})}{\partial \phi_i} = -\frac{\partial C_i(\phi_i)}{\partial \phi_i} q_i + \sum_j \pi_j \frac{\partial q_j}{\partial \phi_i} = 0 \tag{5}$$

Using the first-order condition with respect to the retail price, this equation simplifies to

$$-\frac{\partial C_i(\phi_i)}{\partial \phi_i} = \frac{\sum_j \pi_j \frac{\partial q_j}{\partial \phi_i}}{\sum_j \pi_j \frac{\partial q_j}{\partial P_i}} \tag{6}$$

Let $g\bar{\rho}_i$ be the average willingness to pay for decreases in the fuel consumption rate among consumers of car class *i* (the fuel price *g* multiplied by a factor reflecting annual VMT, discounting, and preferences). Efficiency, at least in allocating fuel economy, would require that $-\frac{\partial C_i(\phi_i)}{\partial \phi_i} = g\bar{\rho}_i$, meaning the per-vehicle cost increase equals that average willingness to pay for lower fuel consumption.

With Bertrand pricing, this condition holds if $\sum_{j\neq i} \pi_j \frac{\partial q_j}{\partial \phi_i} = g\bar{\rho}_i \sum_{j\neq i} \pi_j \frac{\partial q_j}{\partial P_i}$. For example, a sufficient situation would be $\frac{\partial q_j}{\partial \phi_i} = g\bar{\rho} \frac{\partial q_j}{\partial P_i}$ for all *j*, that is, if consumers in all classes respond to a fuel consumption change in class *i* in proportion to the way they respond to a price change in that class, with that proportion being the average willingness to pay among all consumers. This

---

[3] Bresnahan (1981) and Feenstra and Levinsohn (1995) found markups in the range of 4%-25% for individual models. The NRC study assumed a 40% markup for cost increases, shared across parts and auto manufacturers and retailers. This and subsequent studies use published dealer markups and the estimated ratio of dealer and manufacturer markups from Bresnahan and Reiss (1986). Bento et al. (2008) find markups in the range of 14-46.

situation occurs in Austin and Dinan (2005), since consumers in all classes are assumed to have on average the same sensitivity to fuel consumption rates ($\bar{\rho}_i = \bar{\rho}$ for all *i*), although there could still be different utilization and internalization rates within classes. The other obvious situation is if $\pi_j = 0, \forall j \neq i$, as with perfect competition.

The proportionality assumption (the first condition above) has attractive properties for modelers of CAFE policy. Note that if consumers respond to fuel costs in the same way as price changes, the pricing strategy does not directly affect fuel economy choice in the maximization problem (by the Envelope Theorem). In other words, imperfect competition does not create an incentive to over- or underprovide fuel economy. Rather, firms wish to provide all the fuel efficiency demanded, in order to maximize the rents from the price markups.

However, as we have discussed, it seems more reasonable to believe that consumers of different car classes have different preferences for fuel economy, since those preferences help determine the class they choose. Suppose consumers do respond "rationally" to fuel costs in the same way as prices, but they differ in their valuation of the fuel consumption rate. For example, suppose the cost to the average consumer of vehicle type *j* for driving vehicle *i* would then be $g\bar{\rho}_j\phi_i$. In this situation, $\frac{\partial q_j}{\partial \phi_i} = g\bar{\rho}_j \frac{\partial q_j}{\partial P_i}$. Furthermore, $\frac{\partial q_j}{\partial P_i} = \frac{\eta_{ji} q_j}{P_i}$. Substituting, we get

$$-\frac{\partial C_i(\phi_i)}{\partial \phi_i} = g\left(\bar{\rho}_i + \frac{\sum_j (\bar{\rho}_i - \bar{\rho}_j)\pi_j \eta_{ji} q_j}{-\sum_j \pi_j \eta_{ji} q_j}\right) \tag{7}$$

(Recall that $\partial q_j / \partial P_i > 0$ for $j \neq i$ and that from (2) the denominator is positive if $q_i > 0$, meaning simply that the own-price effect dominates the cross-price effects.) In other words, fuel economy will tend to be over-provided in classes whose consumers, on average, value it more than in others, and underprovided in classes whose consumers value it less than in others. In this manner, fuel economy represents a way to solidify market segmentation; by offering less fuel economy to consumers of large cars, for example, they can charge higher prices to small-car consumers, without worrying they will switch classes. Similarly, they can charge higher prices for large cars when they are charging more for highly efficient small cars than the large-car buyers are willing to pay.

This same result can in theory be extended to any vehicle quality. Quality competition can occur over several characteristics, not just one, as long as the valuation of each characteristic varies across product classes. For our purposes, however, we assume that other features are held constant.

Fuel Economy Regulation and Producer Behavior

In this section, we consider how different kinds of policy interventions affect the distortions that may arise out of price discrimination incentives. We find that most either do little or exacerbate them, with the potential exception of minimum fuel economy standards.

Higher Gasoline Prices

One policy for improving fuel economy is increasing gasoline prices through taxation or other means. The effects of higher gasoline prices on producer incentives are evident in Equation (8): they raise the average consumer willingness to pay for fuel economy, and they also proportionately magnify the strategic incentives for distorting fuel economy provision to facilitate price discrimination.

CAFE Standards

The CAFE standards require that each manufacturer's fleet must meet or surpass a harmonic average for fuel economy, measured in miles per gallon, for all the vehicles of that type. We consider a stylized version of the domestic new vehicle market, in which we initially abstract from the differentiation between cars and light trucks. In this first case, we consider the uniform CAFE standard, as is currently applied to passenger cars. (In essence, this assumption is equivalent to zero cross-price elasticity between cars and trucks, which is obviously strong.) However, in the second case, we consider size-based standards, as are being implemented in the light truck category, or could also reflect the different standards for cars and trucks.

The uniform CAFE standard is equivalent to mandating that the average fuel consumption rate for the fleet be below the corresponding standards, expressed as $\bar{\phi}$. That is, if

q_i is the sales of vehicles in class *i*, CAFE standards mandate that for each fleet of autos, $\Sigma_i \phi_i q_i \leq \overline{\phi}\, \Sigma_i q_i$.[4] The manufacturer then maximizes profits, subject to the prevailing fuel economy constraint, defined as an average fuel consumption rate (or a harmonic average of MPG). The Lagrangian is

$$L = \Pi(\mathbf{P},\boldsymbol{\varphi}) - \lambda \sum_i (\phi_i - \overline{\phi}) q_i(\mathbf{P},\boldsymbol{\varphi}) \tag{8}$$

Maximizing profits with respect to the price of each vehicle class *i* leads to a similar first-order condition as in (3), but the full marginal profit for vehicle type *j* includes the shadow value of the extent to which its fuel consumption rate is above or below the standard. (Furthermore, it is possible that marginal profits excluding the shadow value can now be negative.)

$$P_i = \sum_j \left(\pi_j - \lambda(\phi_j - \overline{\phi}) \right) (-\eta_{ji}) \frac{q_j}{q_i} \tag{9}$$

Let $\hat{\pi}_j = P_j - C_j(\phi_j) - \lambda(\phi_j - \overline{\phi})$. Rearranging, as in Equation (4), we can express the price as the sum of the vehicles own costs, including the implicit net tax or subsidy from the fuel consumption standard, with a markup according to its own-price elasticity, and the cross-price responses, weighted by the marginal profits of the other vehicles in the manufacturer's fleet:

$$P_i = \left(C_i(\phi_i) + \lambda(\phi_i - \overline{\phi}) \right) \frac{\eta_{ii}}{(\eta_{ii}+1)} + \sum_{j \neq i} \hat{\pi}_i \frac{-\eta_{ji}}{(\eta_{ii}+1)} \frac{q_j}{q_i} \tag{10}$$

Here we see again that the markup ratio depends on the own-price elasticity of demand, but the basis for cost changes also includes the implicit net tax/subsidy.

The first-order conditions with respect to fuel economy are

---

[4] Although paying a fine is an alternative, we assume that all firms choose to meet the standard, as has been the case. Manufacturers must pay a penalty of $55 per vehicle for every 1 mpg that their fleet average falls below the relevant standard. Vehicles weighing more than 8,500 pounds (such as the Hummer H2 and Ford Excursion) are exempt from CAFE.

$$\frac{\partial L(\mathbf{P},\boldsymbol{\varphi})}{\partial \phi_i} = \left(-\frac{\partial C_i(\phi_i)}{\partial \phi_i} - \lambda \right) q_i + \sum_j \hat{\pi}_j \frac{\partial q_j}{\partial \phi_i} = 0 \qquad (11)$$

Using the first-order condition with respect to the retail price, this equation simplifies to

$$-\frac{\partial C_i(\phi_i)}{\partial \phi_i} = \frac{\sum_j \hat{\pi}_j \frac{\partial q_j}{\partial \phi_i}}{\sum_j \hat{\pi}_j \frac{\partial q_j}{\partial P_i}} + \lambda \qquad (12)$$

Thus, the CAFE constraint shifts up the marginal benefit from decreasing the fuel consumption rate by the same amount for all vehicles in the regulatory category, without directly changing the strategic incentives for price discrimination. However, it does have indirect effects on these strategic incentives.

Assuming again that $\frac{\partial q_j}{\partial \phi_i} = g\bar{\rho}_j \frac{\partial q_j}{\partial P_i}$ and substituting, we see that CAFE standards change the effective marginal profits and thereby the relative weights on the induced demand changes for other vehicles in the fleet:

$$-\frac{\partial C_i(\phi_i)}{\partial \phi_i} = g\left(\bar{\rho}_i + \frac{\sum_j (\bar{\rho}_i - \bar{\rho}_j)\left(\pi_j - \lambda(\phi_j - \bar{\phi})\right)\eta_{ji} q_j}{-\sum_j \left(\pi_j - \lambda(\phi_j - \bar{\phi})\right)\eta_{ji} q_j} \right) + \lambda \qquad (13)$$

In the absence of CAFE, vehicles with higher-than-average consumer willingness to pay for fuel economy have lower-than-average fuel consumption rates, and vice-versa. With CAFE, marginal profits are relatively higher for vehicles with lower-than-average fuel consumption rates. This creates countervailing effects for some vehicle types. In the numerator, larger differences in willingness to pay are correlated with larger differences in fuel economy component effective marginal profits, which tends to magnify the strategic effects. On the other hand, for fuel efficient cars, larger marginal profits also raise the denominator, which is dominated by the own-price effects, thereby dampening this term. For fuel inefficient cars, however, the reduction in marginal effective profits in the denominator magnify the strategic incentives to underprovide fuel economy in larger vehicles.

Size-Based Standards

With size-based standards, CAFE standards are modified such that each manufacturer's fleet must meet or surpass a harmonic average for fuel economy that depends on the size distribution of its fleet. This method is currently being implemented for light trucks and may be extended to cars. Formally, if q_i is the sales of vehicles in class i, size-based standards mandate that for each fleet of vehicles, $\Sigma_i \phi_i q_i \le \Sigma_i \overline{\phi}_i q_i$. Sized-based standards can improve overall cost-effectiveness over uniform standards if manufacturers are sufficiently heterogeneous and cannot trade credits, as the reduction targets can be better tailored to costs (Elmer and Fischer, 2009). With that rationale in mind, it is useful to consider the case in which these standards better approximate desired fuel economy than the uniform standard for all classes ($|\phi_j - \overline{\phi}_j| < |\phi_j - \overline{\phi}|, \forall j$). The new Lagrangian for the manufacturer is

$$L = \Pi(\mathbf{P}, \boldsymbol{\varphi}) - \lambda \sum_i (\phi_i - \overline{\phi}_i) q_i(\mathbf{P}, \boldsymbol{\varphi}) \tag{14}$$

The (rearranged) first-order conditions are modified from those of the uniform standards to reflect the different allocations of fuel economy credits:

$$P_i = \left(C_i(\phi_i) + \lambda(\phi_i - \overline{\phi}_j)\right)\frac{\eta_{ii}}{(\eta_{ii}+1)} + \sum_{j\neq i} \tilde{\pi}_i \frac{-\eta_{ji}}{(\eta_{ii}+1)} \frac{q_j}{q_i} \tag{15}$$

where $\tilde{\pi}_j = P_j - C_j(\phi_j) - \lambda(\phi_j - \overline{\phi}_j)$. The main difference from the uniform standards is that the deviation in fuel consumption rates from the standard, and thereby the influence of the standard on marginal costs and profits, is mitigated. Of course, the shadow value of fuel economy is also affected by the change in the stringency of the effective standard; for manufacturers specializing more in larger vehicles, this shadow value tends to fall, while for manufacturers of smaller vehicles, the standard tends to become more stringent.

Similarly, in the choice of fuel economy, the first-order conditions are similar to (12), but modified by the change in the distribution of marginal profits. We do assume here that changing size is not an available means for improving fuel economy. See Elmer and Fischer (2009) for the influence of market power on the distortionary effects of weight-based standards. With the same assumptions and substitutions as before, the first-order condition for fuel economy can be written as

$$-\frac{\partial C_i(\phi_i)}{\partial \phi_i} = g\left(\bar{\rho}_i + \frac{\sum_j (\bar{\rho}_i - \bar{\rho}_j)\left(\pi_j - \lambda(\phi_j - \bar{\phi}_j)\right)\eta_{ji} q_j}{-\sum_j \left(\pi_j - \lambda(\phi_j - \bar{\phi}_j)\right)\eta_{ji} q_j}\right) + \lambda \qquad (16)$$

In general, size-based standards tend to reduce the change in fully loaded marginal profits relative to uniform standards. This is especially true for large-vehicle manufacturers, who also see the shadow value of fuel economy fall, relative to uniform standards. In the extreme case in which size-based standards accurately reflect the equilibrium fuel economy by class for the manufacturer (so $\bar{\phi}_j = \phi_j$), this condition reduces to that in the absence of regulation, just shifted up by the shadow value of fuel economy. In other words, size-based standards have little effects on the strategic incentives to distort fuel economy provision across vehicle types.

Minimum Standards

An alternative standard to CAFE standards would be minimum fuel economy standards. Such standards are used in China, for example, which imposed fuel consumption limits on light-duty passenger cars based on the weights of the vehicles, beginning in 2005.[5]

Under this fuel economy constraint, the Lagrangian for the manufacturer is

$$L = \Pi(\mathbf{P}, \boldsymbol{\varphi}) - \sum_i \lambda_i (\phi_i - \bar{\phi}_i) q_i \qquad (17)$$

such that $\lambda_i(\phi_i - \bar{\phi}_i) = 0$ for all *i*. (By multiplying the constraint by the quantity of vehicle sales, we are effectively scaling the shadow value, for consistency with the previous analysis.)

Maximizing profits with respect to the price of each vehicle class *i* leads to the same first-order condition as in Equations (3) and (4), with the constraint not directly affecting marginal

---

[5] "Limits of Fuel Consumption for Passenger Cars" was jointly issued by the State Administration of Quality Supervision, Inspection and Quarantine and the Standardization Administration in 2004.

profits: $\pi_j = P_j - C_j(\phi_j)$, since $\lambda_i(\phi_i - \bar{\phi}_i) = 0$. In other words, this regulation does not create an incentive for fleet-mix shifting via pricing (other than by changes in actual costs).

The first-order conditions with respect to fuel economy look identical to those in Equation (11); the difference here is that the shadow value varies by each class, as opposed to just across cars and trucks. From (12), then, we see that the minimum fuel economy standard shifts up the marginal benefit from decreasing the fuel consumption rate by different amounts for all classes, but only when it is binding.

$$-\frac{\partial C_i(\phi_i)}{\partial \phi_i} = g\left(\bar{\rho}_i + \frac{\sum_j (\bar{\rho}_i - \bar{\rho}_j)\left(\pi_j\right)\eta_{ji}q_j}{-\sum_j \left(\pi_j\right)\eta_{ji}q_j}\right) + \lambda_i \tag{18}$$

In this way, minimum standards can potentially counteract strategic incentives to underprovide fuel economy if they are binding for those market segments. However, they cannot directly address overprovision.

A Simple Application

Much of the intuition can be illustrated by considering a manufacturer with two types of cars: large, relatively fuel inefficient cars (*L*) and small, relatively fuel efficient cars (*S*). Let the *q*'s represent fleet shares, such that $q_L = 1 - q_S$. We will express markups *m* as a share of the price, so $\pi_i = m_i P_i$, and let $B = P_L / P_S$ be the ratio of large car prices to small car prices. The following simplifications also allow us to represent the willingness to pay for fuel economy and fuel consumption rates as a function of the averages and differences:

$$\rho_S = \bar{\rho} + \Delta_\rho, \quad \rho_L = \left(\bar{\rho}(1 - q_S) - q_S\Delta_\rho\right) / q_L$$
$$\phi_S = \bar{\phi} - \Delta_\phi, \quad \phi_L = \left(\bar{\phi}(1 - q_S) - q_S\Delta_\phi\right) / q_L$$

Let us focus on the strategic incentives to manipulate fuel economy, or the fuel economy premium ("FE Premium"). Our measure will be difference between the marginal reduction benefits (MRB) to the manufacturer for providing fuel economy—the right hand side of the first-order conditions for ϕ_i—from the average consumer willingness to pay. In the absence of regulation ("NR"), this simplifies to

$$MRB_L^{NR} - g\bar{\rho}_L = -\frac{g m_S \eta_{SL} q_S \Delta_\rho}{\chi_L},$$
$$MRB_S^{NR} - g\bar{\rho}_S = \frac{g B m_L \eta_{LS} \Delta_\rho}{\chi_S}$$

where $\chi_L = \left(-\eta_{LL} B m_L (1-q_S) - m_S \eta_{SL} q_S\right)(1-q_S)$ and $\chi_S = \left(-\eta_{SS} m_S q_S - B m_L \eta_{LS}(1-q_S)\right)$.

The other main policy of interest is the uniform CAFE standard ("U"), and how it might differ from incentives without regulation. Here,

$$MRB_L^{U} - MRB_L^{NR} = \lambda - g\frac{\lambda}{P_S}\left(\frac{q_S \eta_{SL} \Delta_\rho \Delta_\phi}{\chi_L}\right)$$
$$MRB_S^{U} - MRB_L^{NR} = \lambda - g\frac{\lambda}{P_S}\left(\frac{q_S \eta_{LS} \Delta_\rho \Delta_\phi}{(1-q_S)\chi_S}\right)$$

Thus, uniform CAFE standards in part raise the marginal benefits to reductions by a uniform amount for each type, but they also have secondary effects that lower the marginal benefits to reductions.

As we observed from the previous theory section, size-based standards tend to mitigate these secondary effects.

$$MRB_L^{SBS} - MRB_L^{NR} = \lambda - g\frac{\lambda}{P_S}\left(\frac{q_S \eta_{SL} \Delta_\rho (\Delta_\phi - \Delta_{\bar{\phi}})}{\chi_L}\right)$$
$$MRB_S^{SBS} - MRB_L^{NR} = \lambda - g\frac{\lambda}{P_S}\left(\frac{q_S \eta_{LS} \Delta_\rho (\Delta_\phi - \Delta_{\bar{\phi}})}{(1-q_S)\chi_S}\right)$$

where $\Delta_{\bar{\phi}} = \bar{\phi} - \bar{\phi}_S$ is the difference between the uniform standard and the size-based standard for small cars.

And for minimum standards, as we know from the previous section, $MRB_L^{M} - MRB_L^{NR} = \lambda_L$ and $MRB_S^{M} - MRB_S^{NR} = \lambda_S$, although the constraint on small cars may not be binding.

To parameterize this simplified model, we draw on existing data and estimates in the literature. From the Wards 2006 model year, we find that average (sales-weighted) prices of small and large cars are \$22,562 and \$29,422, respectively, with small cars representing 49% of the national automobile fleet. We draw on the recent study by Bento et al. (2006) to calibrate marginal profits. With their markups by manufacturers, we calculate average sales-weighted markups for compact cars of roughly 22%, while markups for mid- and full-sized cars average 25%.

Next, we assume modest, symmetric cross-price elasticities of demand between small and large cars of 0.1, which falls within the range found by Kleit (2002) and Jacobsen (2008) (and will be a target of sensitivity analysis). Then, solving from the price equation (4) for both small and large cars, we use these markups and other parameters to calibrate the own-price elasticities of demand. In other words,

$$\eta_{LL} = -\frac{B(1-q_S)+\eta_{LS} m_S q_S}{B m_S (1-q_S)}, \quad \eta_{SS} = -\frac{B\eta_{LS} m_L (1-q_S)+q_S}{m_S q_S}$$

We find that the own-price elasticities consistent with our other parameter assumptions are $\eta_{LL} = -4.6, \eta_{SS} = -4.7$. Finally, we use a gasoline price of \$2.70 per gallon.

With these parameters, we find that

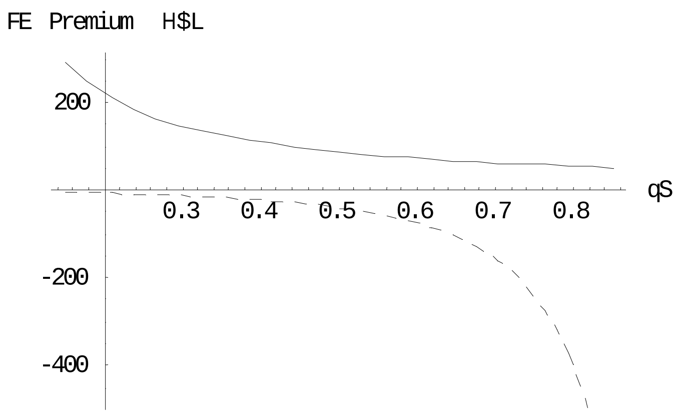
$$MRB_L^{NR} - g\bar{\rho}_L = -0.075\Delta_\rho, \quad MRB_S^{NR} - g\bar{\rho}_S = 0.180\Delta_\rho$$

In other words, to the extent that small-car consumers are willing to pay more than the average for increased fuel economy, the MRB to the manufacturer increases by an additional 18% of that amount for small cars. Meanwhile, it decreases the MRB for large cars by 8% of that extra small-car consumer willingness to pay. Interestingly, for this representative manufacturer, the distortion for overprovision of fuel economy in small cars is more than twice as large as the underprovision of fuel economy in large cars.

This distribution does depend in good part on the share of small and large cars in the manufacturer's fleet. On average, small and large cars are fairly evenly represented in the new vehicle market; however, some manufacturers sell much higher proportions of one or the other.

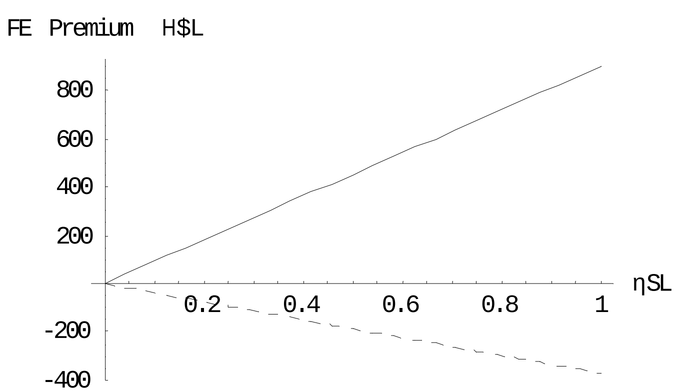
Figure 1 reveals that as the market share of small cars goes up, the fuel economy premium tapers down for small cars (solid line), while it gets increasingly negative for large cars (dashed line).[6] Meanwhile, for producers concentrating on large cars, the underprovision incentive is fairly low, although the incentive to overprovide fuel economy in small cars gets quite large.

Figure 1: Fleet share of small versus large cars and the fuel economy premium



Another important factor is the assumed cross-price elasticity. The distortions to the marginal reduction benefits increase in proportion to the cross-price elasticity across vehicles. The greater price sensitivity evidently makes quality differentiation more important.

Figure 2: Sensitivity of the fuel economy premium to cross-price elasticities of small and large cars



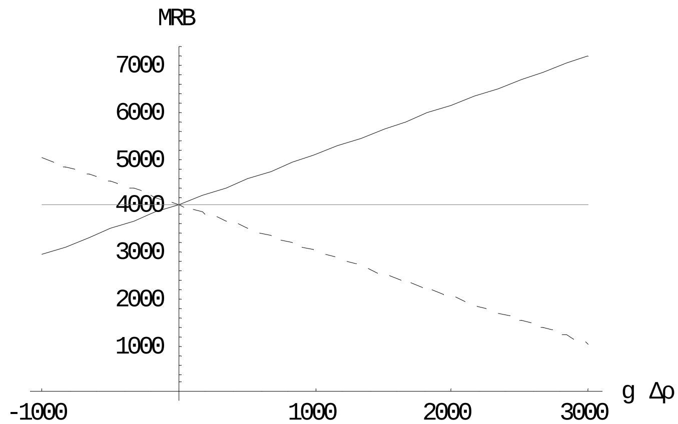
[6] The figures assume Δ_{ρ} =500.

On the other hand, the distortions get smaller as the own-price elasticities get larger; this result is evident from the previous equations, since the χ 's (the denominators) increase with the own-price elasticities.

The additional distortion from CAFE standards appears to be relatively small. Assuming a rather substantial shadow value of λ =\$2000, and a $\Delta_\phi = .3$ from the baseline data, we find that $MRB_L^U - g\bar{\rho}_L = \lambda - 0.084\Delta_\rho$, $MRB_S^{NR} - g\bar{\rho}_S = \lambda + 0.166\Delta_\rho$. Thus, CAFE does mitigate some of the fuel economy premium for small cars but exacerbates the distortion for large cars. Of course, this takes the shadow value of fuel economy as given and does not account for the influence of price discrimination on that value. Since the fuel economy premium falls in both cases, given any fleet standard, the shadow value would have to rise, compared to the absence of such a distortion.

Finally, it is worth considering the effect of Δ_ρ on the overall MRB for each car type. Recall that most models for evaluating CAFE standards and other policies for fuel economy assume that all consumers have the same willingness to pay for fuel economy. Using the NAS study assumptions of annual travel (15,600 miles in the first year, declining by 4.5% annually), vehicle lifetime (14 years), discount rate (5%), and onroad shortfall (15%), this translates into a willingness to pay for a farsighted consumer of \$1491 per \$1 of gasoline price (or about \$4000 at our assumed price of \$2.70). By considering that consumers may sort by type and on average have different preferences, MRB will deviate from the average both by the direct effect on willingness to pay and by the additional effect on the fuel economy premium. The results are depicted in Figure 3.

Figure 3: Influence of disparity in willingness to pay for fuel economy on marginal reduction benefits



Failure to capture this kind of consumer heterogeneity can lead to significant errors in predicting the distribution of effort in complying with CAFE, as well as the calculation and distribution of the benefits.

Policy Discussion

We find that market power gives manufacturers a strategic incentive to over-provide fuel economy in classes whose consumers, on average, value it more than in others, and underprovide it in classes whose consumers value it less than in others. In this manner, manufacturers can better segment their markets and charge higher prices, with less worry that consumers will switch classes.

If one combines this kind of imperfect competition with a second market failure in consumer willingness to pay for fuel economy, a tradeoff in policy prescriptions emerges. Minimum fuel economy standards may better deal with distortions from price discrimination, but they do not provide broad-based incentives for improving fuel economy like average standards. Furthermore, increasing fuel prices can exacerbate both the incentives for price discrimination and the undervaluation of fuel economy. Therefore, a combination policy of both average and minimum standards may be preferred.

Extensions of this research will investigate the effects of market power on fuel economy choice in a more complicated model with multiple manufacturers and vehicle classes. Given the full interdependency of pricing and fuel economy decisions across models, solving such an equilibrium is much more challenging. However, it offers the opportunity to also gauge the welfare implications of market power distortions and alternative policy interventions.

Another interesting question for future research involves imperfect competition in credit markets, in addition to that in product markets. Obviously, if market power is an issue in the latter, and the same firms are active in the credit markets, then market power is likely to be an issue there as well. The influence is not always clear, since in the credit market, both monopoly and monopsony power may be exercised. Rubin et al. (2006) show that market power in the credit markets can mean that some of the gains from trade are left on the table. This concern is relevant for trading across manufacturers, but not for all of the policy alternatives (e.g., trading across a manufacturer's car and light truck fleets or switching to feebates). More important, though, is the question of whether it would be relevant for fuel economy decisionmaking.

In general, the results in this paper indicate the importance of additional empirical research on the demand for fuel economy, with greater attention paid to how that might vary by vehicle classes. Indeed, although we have focused on a potential market failure in the supply of fuel economy, a market failure in the demand for fuel economy is the most powerful justification for regulation.[7] Still, few empirical studies consider such complexities in consumer response to fuel costs.

[7] See Fischer et al. (2007).

References

Austin, D. and T. Dinan. 2005. Clearing the air: The costs and consequences of higher CAFE standards and increased gasoline taxes, *Journal of Environmental Economics and Management* 50 (2005) 562–582.

Bento, A., L. H. Goulder, E. Henry, M. R. Jacobsen, and R. H. von Haefen. 2005. Distributional and Efficiency Impacts of Gasoline Taxes: An Econometrically Based Multi-market Study, *American Economic Review* 95(2): 282–287.

Bento, A., L. H. Goulder, M. R. Jacobsen, and R. H. von Haefen. 2006. Distributional and Efficiency Impacts of Increased U.S. Gasoline Taxes. Working Paper.

Berry, S., J. Levinsohn & A. Pakes. 1995. Automobile Prices in Market Equilibrium. *Econometrica,* 63, 4, 841–890.

Bresnahan, T.1981. Departures from Marginal-Cost Pricing in the American Automobile Industry: Estimates for 1977–1978. *Journal of Econometrics,* 17, 201–227.

Bresnahan, T.F. and P. Reiss. 1986. Departures from marginal-cost pricing in the american automobile industry: estimates for 1977–1978, *Journal of Econometrics* 11 (1985) 201–227.

Feenstra, R and J. Levinsohn. 1995. Estimating Markups and Market Conduct with Multi-Dimensional Product Attributes, *Review of Economic Studies,* 62, 19–52.

Elmer, K.F. and C. Fischer. 2009. Designing Vehicle CO2 Emission Standards: Uniform vs. Weight-Based Standards. Working Paper. Washington, DC: Resources for the Future

Fischer, C., I. Parry, and W. Harrington. 2005. “Economic Impacts of Tightening the Corporate Average Fuel Economy (CAFE) Standards,” Report prepared for Environmental Protection Agency and National Highway Traffic Safety Administration.

Fischer, C., W. Harrington, and I.W.H. Parry (2007) “Do Market Failures Justify Tightening Corporate Average Fuel Economy (CAFE) Standards?” The Energy Journal 28 (4): 1-30.

Fischer, C. and P. R. Portney. 2004. "Tradable CAFE Credits," *New Approaches on Energy and the Environment: Policy Advice for the President*, R. D. Morgenstern and P. R. Portney, Ed. Washington, DC: RFF Press.

Goldberg, Pinelopi Koujianou, 1998. "The Effects of the Corporate Average Fuel Efficiency Standards in the US." *Journal of Industrial Economics* 46:1-33.

Goldberg, Pinelopi Koujianou, 1995. "Product Differentiation and Oligopoly in International Markets: The Case of the US Automobile Industry." *Econometrica* 63: 891-951.

Greene, David L., 1996. *Transportation and Energy*. The Eno Transportation Foundation, Inc., Washington, DC.

Greene, D.L. 2002. "Effectiveness and Impact of Corporate Average Fuel Economy (CAFE) Standards," Report of the Committee, National Research Council, National Academy Press, Washington, DC.

Greene, D.L., P.D. Patterson, M. Sing and J. Li. 2005. "Feebates, Rebates and Gas-Guzzler Taxes: A Study of Incentives for Increased Fuel Economy," *Energy Policy* 33 (6): 721-827.

Jacobsen, Mark R. 2008. "Evaluating U.S. Fuel Economy Standards In a Model with Producer and Household Heterogeneity." Manuscript, University of California at San Diego.

Kleit, Andrew N., 2004. "Impacts of Long-Range Increases in the Corporate Average Fuel Economy (CAFE) Standard." *Economic Inquiry* 42: 279-294.

Kurani, Kenneth S. and Thomas Turrentine. 2004. Analysis of Consumer Response to Automobile Regulation and Technological Change in Support of California Climate Change Rulemaking. Final Report, prepared for the California Air Resources Board and California Environmental Protection Agency. University of California, Davis: Institute for Transportation Studies.

Kwoka, J.E., Jr., 1983. "The Limits of Market-Oriented Regulatory Techniques: the Case of Automotive Fuel Economy." *Quarterly Journal of Economics*, November: 695-704.

NRC, 2002. *Effectiveness and Impact of Corporate Average Fuel Economy (CAFE) Standards*. National Research Council, Washington, DC, National Academy Press.

Petrin, A. 2002. Quantifying the Benefits of New Products: The Case of the Minivan, *Journal of Political Economy*, 110:705-729.

Plourde, C. and V. Bardis. 1999. Fuel economy standards in a model of automobile quality, *Energy Economics* 21: 309-319.

Rubin, J., P. Leiby, and D. Greene. 2006. “Tradable Fuel Economy Credits: Competition and Oligopoly,” Working Paper, University of Maine.

New Vehicle Characteristics and the Cost of the Corporate Average Fuel Economy Standard*

Thomas Klier
Federal Reserve Bank of Chicago

Joshua Linn
University of Illinois at Chicago

November 2008

Preliminary Draft – Please Do Not Cite Without Permission

Abstract

Recent legislation has increased the Corporate Average Fuel Economy (CAFE) standard by 40 percent, which represents the first major increase in the standard since its creation in 1975. Previous analysis of the CAFE standard has analyzed the short run effects (1-2 years), in which vehicle characteristics are held fixed, or the long run effects (10 years or more), when firms can adopt new power train technology. This paper focuses on the medium run, when firms can choose characteristics such as weight and power, and have a limited ability to adopt technology. We first document the historical importance of the medium run and then estimate consumers' willingness-to-pay for fuel efficiency, power and weight. We employ a novel empirical strategy that accounts for the characteristics' endogeneity, which has not been addressed in the literature, by using variation in the set of engine models used in vehicle models. The results imply that an increase in power has a similar effect on vehicle sales to a proportional increase in fuel efficiency. We then simulate the medium run effects of an increase in the CAFE standard. The policy reduces producer and consumer welfare and causes substantial transfers across firms, but the effects are significantly smaller than found in previous studies.

* We thank conference participants at the NBER Summer Institute and seminar participants at Resources for the Future for helpful comments. Taft Foster, Vincent Liu and Christina Miller provided excellent research assistance. We are grateful to Eric Fedewa from CSM providing data on engine specifications. Authors' email addresses: tklier@frbchi.org and jlinn@uic.edu.

1 INTRODUCTION

The Corporate Average Fuel Economy (CAFE) standard is the minimum fuel efficiency that manufacturers of new vehicles must attain in the U.S. market. After a lengthy period of public debate, the Energy Independence and Security Act of 2007 increased the CAFE standard for new vehicles by about 40 percent, to be effective by the year 2020. The legislation represents the first significant increase in the standard since it was first created in 1975, and followed a period of vigorous public debate. The law's proponents argued that it would reduce carbon dioxide emissions and oil imports without undermining the automobile industry. Opponents claimed that the costs to vehicle manufacturers and consumers would not justify the benefits, and that other policies would be more effective at reducing emissions and oil imports.

Coinciding with the recent policy debate, a sizeable literature has analyzed the costs to consumers and producers of using the CAFE standard to reduce gasoline consumption. These studies simulate the effect of an increase in the standard on market equilibrium and can be classified into two categories. Some, including Goldberg (1998), have used a short run model, pertaining to one or two years after a change in the standard, in which vehicle characteristics and technology are held constant. Firms respond to an increase in the CAFE standard by adjusting vehicle prices, i.e., by changing the "sales mix." Other studies, such as Austin and Dinan (2005), use a long run model, which pertains to 10 years or more after a change in the standard, to estimate costs. In this model, firms choose vehicle prices and power train (engine and transmission) technology.

Yet casual observation of the new vehicles market suggests that the preceding analysis is overly simplified. Firms typically select vehicle prices every year and make major changes to power train technology every ten years. But every four or five years, firms can redesign vehicles by changing their characteristics, such as interior cabin features. Of particular relevance to the CAFE standard is the fact that firms can increase the fuel efficiency of a vehicle by reducing weight and power or by making minor changes to the engine technology. For example, removing components or using lighter materials can reduce the vehicle's weight. Firms can also modify the engine to reduce the number of cylinders that power the vehicle at low speeds (by contrast, the long run analysis includes major changes to the power train, such as adopting hybrid technology). Relatively minor changes are made routinely in the new vehicles market, and are expected to occur in response to the new CAFE regulation. For example, in the spring of 2008 Honda introduced the 2009 version of the Acura TSX model, which has less power and greater fuel efficiency than the previous version. The Vice President of corporate planning for Honda announced at the time of the introduction that "We feel comfortable there's plenty of horsepower already and wanted to focus on improving fuel efficiency and emissions. For us generally, you'll see more of that," (Ohnsman, 2008). Similarly, GM has announced, "Never mind the fuel cells, plug-ins or diesels. To achieve quick improvements in fuel efficiency, General Motors is adopting an off-the-shelf technology: small engines with turbochargers," (Kranz, 2008). There is thus a medium run response to the CAFE standard that is distinct from short run price changes and long run technology adoption.

The CAFE literature has concluded that the regulation is far more costly than using the gasoline tax to reduce gasoline consumption. However, because the previous analysis does not incorporate the medium run, total discounted costs may be significantly overstated. To the extent that reductions in weight and power or modifications to the power train are less costly than adjusting the sales mix, actual costs a few years after a change in the standard could be much lower than the short run analysis suggests. Medium run changes in characteristics may also reduce the need to equip vehicle models with expensive advanced engine technologies in the long run, implying that the long run estimates may also be too high. Finally, the short run/long run distinction may overstate the length of time before significant improvements in fuel economy can be realized. But it is an empirical question whether the medium run is quantitatively important.

We first document the importance of changes in weight and power following the imposition of the initial CAFE standard in 1978. Changes in the sales mix reduced fuel efficiency by a small amount and for only a few years after the standard was imposed. Reductions in weight and power explain much of the increase in fuel efficiency in the late 1970s and early 1980s, after which technology adoption becomes increasingly important. These patterns suggest that the medium run response to CAFE lasts about five years.[1]

These results motivate the main analysis, in which we simulate the short and medium run effects of the CAFE standard on market equilibrium. The difference between the short and medium run is that in the short run all vehicle characteristics are fixed, while in the medium run firms choose vehicle prices and characteristics but cannot change the power train technology. As such, this paper is the first to characterize the medium run effects of the regulation. But the analysis of the medium run poses a major empirical challenge, which is to consistently estimate consumers' willingness-to-pay for characteristics while taking account of their endogeneity. The large literature on consumer demand in the new vehicles market has ignored this issue. For example, Berry, Levinsohn and Pakes (1995) construct a set of instrumental variables that is

[1] A number of studies in the 1980s analyzed the changes in weight, power and fuel efficiency after CAFE was adopted. Similarly to this study, Greene (1987 and 1991) concludes that short run changes in the sales mix explain a small share of the increase in fuel efficiency and that technology explains about half of the increase in fuel efficiency. Greene and Liu (1988) calculate the change in consumer surplus after CAFE was adopted using changes in these characteristics and willingness-to-pay estimates from other studies However, the earlier studies do not perform the analysis at the engine level, as this paper does, and they pertain to a shorter time period.

valid only if characteristics observed by the econometrician are uncorrelated with unobserved characteristics, which seems unlikely to be the case; e.g., a larger vehicle may have worse handling.

Several recent studies of other industries have confronted this empirical challenge (e.g., Ishii, 2005), but the new vehicles market poses the additional difficulty that unobserved characteristics are also endogenous and are potentially correlated with observed characteristics. In this case, estimation requires an identifying assumption on the joint distribution of the observed and unobserved variables. For example, Sweeting (2007) assumes that changes in unobserved characteristics of radio stations occur after the firm has chosen the observed characteristics.[2] We use an instrumental variables strategy that is similar to Hausman *et al*. (1994) and exploits a particular feature of the new vehicles market: firms often sell vehicle models in different vehicle classes with the same engine. For example, the Ford F-Series (a pickup truck) and the Ford Excursion (a sports utility vehicle) have the same engine. We instrument for a vehicle's endogenous characteristics using the engine characteristics of vehicles located in different classes that have the same engine. Combined with the estimated demand for fuel efficiency that we report in Klier and Linn (2008), the results imply that consumers are willing to pay roughly the same amount for a proportional increase in power as for fuel efficiency.

We use the empirical estimates to simulate the medium run cost of the CAFE standard. Similarly to the short run analysis, an increase in the CAFE standard causes large transfers across firms and would particularly harm U.S. firms in the medium run. However, the medium run costs are about one-half of the short run costs, which implies that the cost of the CAFE standard, in dollars per gallon of gasoline saved, is much smaller than the short run analysis suggests.

[2] In Sweeting (2007), unobserved station quality is exogenous, but is potentially correlated with observed characteristics. Sweeting uses the timing assumption to construct a valid set of instruments using lagged variables.

Furthermore, the long run analysis does not reveal the substantial improvements in fuel efficiency that can be attained only a few years after a new standard is adopted. On the other hand, the cost of reducing gasoline consumption in the medium run is probably greater using the CAFE standard than the cost of using the gasoline tax.

2 DATA

This paper uses a detailed data set of vehicle and engine characteristics and vehicle sales from 1975-2008. Klier and Linn (2008) describe the vehicle characteristics and sales data in more detail. Vehicle sales are from the weekly publication Ward's Automotive Reports for the 1970s and from Ward's AutoInfoBank in subsequent years. Sales are matched to vehicle characteristics by vehicle model from 1975-2008.[3] The characteristics data are available in print in the annual Ward's Automotive Yearbooks (1975-2008), and include horsepower, curb weight, length, fuel efficiency and retail price. Note that the data do not include fuel efficiency from 1975-1977, as fuel efficiency was not reported prior to the CAFE program. We impute fuel efficiency from the other vehicle characteristics during these years, using the estimated relationship among characteristics for 1978-1979.

The data coverage for cars is far more extensive than for light trucks. The sample includes all car models produced in the U.S. during the 1970s, but does not have any light trucks in the 1970s. Consequently, the historical analysis in this paper focuses on cars, which account for most of the vehicle market during the late 1970s and early 1980s. According to the U.S. EPA

[3] The match is not straightforward because the two data sets are reported at different levels of aggregation. Vehicle characteristics data are reported at the "trim level" to recognize differences in the manufacturer suggested retail price (MSRP); for example, the data distinguish the 2- and 4-door versions of the Honda Accord sedan. We aggregate the characteristics data to match the model-based sales data, and calculate four statistical moments for the distribution of the vehicle characteristics by model line (minimum, maximum, mean and median).

(2007), the share of light trucks in the new vehicles market was between 20 and 30 percent between the years 1975 and 1988.

We have obtained data on detailed engine specifications for the years 2000-2008 from CSM, a Michigan-based consulting firm for the automobile sector. The engine data distinguish two levels of aggregation. An engine program refers to a distinct engine technology, and a platform is a collection of related programs. For example, the Volkswagen Passat and Audi A4 are sold with the same engine program. The Volkswagen Jetta has a different engine program from the Passat and the Audi, but both engine programs belong to the same platform. Firms may produce different versions of the same engine program that vary by power and size. Note that engines in the same program have the same number of cylinders, but the number of cylinders may vary across engines in a platform.

For each vehicle model, we construct a list of engine programs that are sold with that model. For a given vehicle, there are three sources of variation over time in the engine technologies that are sold with it. First, the engine may be redesigned, in which case the program identifier changes. Second, firms may discontinue selling a vehicle model with a particular engine, as Honda recently did with the hybrid Accord. Third, a firm can introduce a new version of the vehicle model that is sold with an engine that had previously been sold only with other vehicle models. We have matched engine and vehicle model characteristics for 2000-2008, which limits the estimation of consumer demand for vehicle characteristics to those years; future work will extend the sample to 1995-2008, and possibly further.

3 FUEL EFFICIENCY REGULATION AND ENGINE TECHNOLOGY

3.1 THE CAFE STANDARD

Following the 1973 oil crisis, Congress passed the Energy Policy and Conservation Act in 1975 in order to reduce oil imports.[4] The Act established the CAFE program and required automobile manufacturers to increase the average fuel efficiency of passenger and non-passenger vehicles sold in the United States. There are separate standards for cars and light trucks, which have varied slightly over time; for model-year 2007, the standards are 27.5 miles per gallon (MPG) for cars and 22.2 MPG for light trucks. Firms may also earn credits for over-compliance that can be used in future years. The standards are administered by the U.S. Department of Transportation (DOT) on the basis of the U.S. Environmental Protection Agency's test procedure for measuring fuel efficiency.

The recently passed Energy Independence and Security Act of 2007 requires DOT to raise fuel-efficiency standards, starting with model year 2011, until they achieve a combined average fuel efficiency of at least 35 mpg for model year 2020. The CAFE standard continues to be extremely controversial, as the 2007 law has been called "a victory for America" (Senator Carper, D-Del, Stoffer 2007), as well as "unnecessary at best and damaging at worst," (Wall Street Journal op-ed, Ingrassia, 2008). Note that firms are evaluated for compliance with the new standard using a different formula that is based on a vehicle's "footprint" (the product of length and width).

3.2 CAFE AND MARKET OUTLOOK

As Section 4 shows in more detail, when the original CAFE standard was introduced, automobile manufacturers rather quickly reduced horsepower and weight in order to raise fuel efficiency.

[4] This section draws extensively from National Research Council (2008).

Engine technologies improved over time, which allowed firms to improve a vehicle's performance while continuing to meet the CAFE standard.

Many industry analysts believe that because many of the "easy" improvements to engine technology were made in response to the initial CAFE standard, the future increase in the standard may be much more costly to producers and consumers. While new power train systems, such as those relying on hybrid electric and diesel technologies, have begun to penetrate the U.S. market, the vast majority of vehicles are powered by conventional gasoline-powered spark-ignition engines. While essentially every vehicle manufacturer is advertising its alternative power train research, as of 2007, sales of hybrid vehicles represent about 2 percent of total sales of cars and light trucks.[5] Thus, once again, the performance characteristics of the existing gasoline engine technology, as well as the related transmission technologies, are the focus of attention.

3.3 THE MEDIUM RUN

We define the medium run as the period of time in which engine technology is constant, but firms can adjust weight, power and fuel efficiency. In the new vehicle market, the short, medium and long run arise from the timing of firms' major decisions. Firms typically choose vehicle prices each year, although firms can also offer price incentives during the year. Large changes in vehicle characteristics typically occur every 4-5 years during major model redesigns. Engine technologies change more slowly, as engines are redesigned roughly every 10 years. Thus, following an unexpected increase in the CAFE standard, firms may adjust prices in the short run; weight, power and fuel efficiency in the medium run; and power train technology in the long run.

[5] In that context it is interesting to note that the hybrids available in the market today represent one of two types: mild hybrids (micro-hybrids or integrated starter-generator hybrids) and parallel hybrids. The Toyota Prius and the GM two-mode hybrid fall into the latter category (National Research Council 2008).

More specifically, in the medium run a firm can modify a vehicle in two ways. First, the firm may improve fuel efficiency by reducing weight or power. Using lighter weight components or replacing a six-cylinder engine with a four-cylinder engine would increase fuel efficiency. Note that the former change would likely increase production costs while the latter change might decrease costs; Section 6 returns to this issue.

The second type of modification is that the firm can adopt a limited set of fuel efficiency-improving technologies, which do not require the firm to redesign the engine or transmission. Engines are intentionally designed with this flexibility to allow firms to respond to demand shocks without completely redesigning the power train. Table 1 provides examples of medium and long run changes to the engine or transmission, taken from NHTSA (2008). Relative to the long run changes, the medium run changes are simple to implement and generally cost less, but result in lower fuel efficiency gains.

4 Response to the Initial CAFE Standard

This section documents changes in fuel efficiency, weight and power in the late 1970s and early 1980s. Much of the increase in fuel efficiency during the 5-10 years following the imposition of the initial standard was due to changes in weight and power. This result motivates the use of a medium run model to simulate the effect of CAFE, which is done in sections 5 and 6.

Figure 1 provides summary information on changes in characteristics in the new vehicles market over time. The figure shows the CAFE standard and changes in weight, power and fuel efficiency for all cars sold in the U.S. from 1975-2007, using data reported in U.S. EPA (2007). Average fuel efficiency increased dramatically in the late 1970s and early 1980s as the standard was phased in. During the same period, power and weight decreased and then increased.

The increase in fuel efficiency in Figure 1 could be due to short run changes in the sales mix; medium run changes in power, weight or technology; or the long run adoption of power train technology. This section decomposes the total increase in fuel efficiency into these three effects. The analysis in this section focuses on cars sold by U.S. automobile manufacturers (Chrysler, Ford and GM) for two reasons. First, as Jacobsen (2008) notes, there have been three categories of firms: firms that consistently exceed the standard by a large amount (e.g., Honda and Toyota); firms that are constrained by the standard and typically meet it (e.g., Ford); and firms that consistently pay a fine for not meeting the standard. U.S. firms account for the vast majority of sales from the constrained category, so the response of U.S. firms to the CAFE standard is of particular interest. The second reason for focusing on U.S. cars is that the light truck data are incomplete, and do not allow for a complete analysis for trucks in the 1970s and 1980s.

For comparison with Figure 1, Figure 2 reports fuel efficiency, weight and power of cars sold by U.S. firms. The figure shows that changes in the characteristics of U.S. firms' cars were similar to the overall market, which reflects the dominance of U.S. firms during this time period. Between 1975 and 1978, which was the first year the CAFE standard was in effect, fuel efficiency increased by about 2 MPG. Gasoline prices were fairly stable during this time period, suggesting that the increase was in anticipation of the standard. It should be recalled, however, that fuel efficiency from 1975-1977 is imputed, and this result should be treated with caution. From 1978 until the early 1980s, fuel efficiency increased by an additional 4 MPG, during which time the U.S. automakers remained above the standard. From the mid 1980s until the end of the sample period, average fuel efficiency was slightly higher than the standard.

At the same time as fuel efficiency was increasing, weight and power were decreasing. Both power and weight decreased by about 25 percent between 1975 and 1982, after which they increased steadily. In summary, the increase in fuel efficiency following the imposition of the CAFE standard coincided with a large decrease in power and weight. Subsequently, weight and power increased while fuel efficiency did not change.

The remainder of this section assesses the magnitudes of the short, medium and long run responses to CAFE. We first separate the short run from the medium and long run. We abstract from entry and exit decisions and analyze a balanced panel of vehicle models that have positive sales each year from 1975-1984, which Figure 2 shows to be the main period in which fuel efficiency increased.[6] The first data series in Figure 3 is the sales-weighted fuel efficiency of the vehicle models in the sample, which follows a very similar pattern to Figure 2. Two counterfactual series are constructed for this figure, which separate the short run changes in average fuel efficiency from the medium and long run. The first series is the sales-weighted average fuel efficiency, which is calculated using the actual sales of the vehicle models in each year and the fuel efficiency in 1975; this series illustrates the effect of changes in the sales mix, as an increase in the sales of vehicle models that initially have high fuel efficiency would cause the sales-weighted average fuel efficiency to increase. The second series plots average fuel efficiency using the sales weights in 1975 and the actual fuel efficiency of the vehicle model each year, which includes medium and long run changes in fuel efficiency.[7] The short run series shows that changes in the sales mix increased average fuel efficiency by about 0.5 MPG between 1978 and 1981. The other counterfactual series is very close to the average MPG, however, implying that within-model changes in fuel efficiency explain nearly all of the overall change.

[6] The models account for about 45 percent of the sales included in the sample in Figure 2.

[7] Note that the change in sales-weighted average fuel efficiency equals the sum of the effect of the change in sales mix, plus the effect of within-model changes in MPG, plus a cross-term: $\Delta\overline{M}_t = \sum_j \Delta s_{jt} M_{j0} + \sum_j s_{j0} \Delta M_{jt} + \sum_j \Delta s_{jt} \Delta M_{jt}$. Figure 2 reports changes in MPG due to changes in the sales weights and within-model changes in fuel efficiency; i.e., the final term is omitted. In practice, the omitted term explains less than 10 percent of the overall change in all years, and is not shown for clarity.

Thus, within the first 10 years of the introduction of the CAFE standard, firms largely complied by increasing fuel efficiency rather than adjusting the sales mix.

Within-model changes in fuel efficiency in Figure 3 could be due to medium or long run changes in vehicle characteristics and technology. Recall that firms can increase fuel efficiency while holding constant weight and power in both the medium and long run. Unfortunately, detailed engine technology data are not available, and it is not possible to separate medium and long run changes to power trains. However, we can estimate the effect of weight and power on fuel efficiency, which provides a lower bound to the full medium run response.

We first estimate the within-engine technology tradeoff between fuel efficiency, weight and power. We use data from 2000-2008 to estimate the following equation:

$$\ln M_{jet} = \delta_0 + \delta_1 \ln H_{jet} + \delta_2 \ln W_{jt} + \eta_e + \varepsilon_{et} \qquad (1)$$

The dependant variable is the log of the fuel efficiency of vehicle j with engine e in year t and the first two variables are the logs of power and weight. Equation (1) includes engine fixed effects, and the coefficients on power and weight are the within-engine elasticity of fuel efficiency with respect to power and weight; by definition, such changes correspond to the medium run.

Table 2 reports the results of estimating equation (1). The two columns include engine program and engine platform fixed effects (recall that multiple engine programs belong to the same platform). The reported coefficients are the within-program and -platform effects of power and weight on fuel efficiency. The two specifications should be considered to be lower and upper bounds of the medium run effect of weight and power on fuel efficiency. The within-program elasticity of fuel efficiency with respect to power is -0.07 and for weight is -0.33; the estimate for power is larger in column 2 with platform fixed effects. On the other hand, the effect of weight on fuel efficiency is the same, which is as expected because weight varies at the vehicle level and not the engine level.

Overall, Table 2 suggests that firms can increase fuel efficiency by decreasing power and weight. Assuming the elasticities have not changed over time, we can use the estimated parameters in equation (1) to obtain a lower bound of the medium run response to CAFE. In particular, we use the actual weight and power each year from 1975-2007 for the sample in Figure 2, combined with the estimates in column 1 of Table 2, to predict the fuel efficiency of each vehicle. The predicted series captures medium run changes in weight and power, but does not include medium run technology adoption. The difference between the actual and predicted series can be interpreted as the effect on fuel efficiency of medium and long run technology adoption. Figure 4 shows the actual and predicted fuel efficiency from 1975-2007. The figure demonstrates that decreases in power and weight explain about one-third of the increase in fuel efficiency in the late 1970s and early 1980s.[8] Given that this is probably a lower bound, we conclude that the medium run response to the CAFE standard has been historically important.

5 ESTIMATING WILLINGNESS-TO-PAY FOR ENGINE POWER AND WEIGHT

This section specifies and estimates the parameters of the market for new vehicles, and the following section reports simulations of an increase in the standard.

5.1 THE NEW VEHICLES MARKET

We model the market for new vehicles, particularly focusing on firms' choices of vehicle characteristics. The model is static and in each period firms select vehicle prices and

[8] Similarly, Greene (1987) concludes that about half of the increase in fuel efficiency between 1978 and 1985 was due to technology.

characteristics for the vehicles they sell. Consumer demand for each vehicle model depends on its price and characteristics, and each period there is a market clearing vector of prices, quantities and characteristics.

Consumer demand follows a standard nesting structure. We define seven classes based on the vehicle classification system in the Wards database (McManus, 2005). Consumers first decide whether to purchase a vehicle, and then select a class, and finally, a vehicle model. Following Berry (1994), the market share of each vehicle model can be expressed as:

$$\ln s_{jt} - \ln s_{0t} = \alpha p_{jt} + \beta_D D_{jt} + \beta_H HW_{jt} + \beta_W W_{jt} + \xi_{jt} + \sigma \ln s_{jt|c} \tag{2}$$

The left hand side of equation (2) is the difference between the log market share of vehicle model *j* and the log market share of the outside good, which is a used vehicle; the denominators in the market shares include new and used vehicles. The first variable on the right hand side is the price of the vehicle model, p_{jt}, and the coefficient α is the marginal utility of income. The next three independent variables are expected fuel costs, D_{jt}, the ratio of power to weight, HW_{jt}, and weight, W_{jt}. Similarly to Klier and Linn (2008), we define the variable D_{jt} as dollars-per-mile, which is equal to the price of gasoline divided by the vehicle's fuel efficiency. The variable is proportional to expected fuel costs if the price of gasoline follows a random walk over the life of the vehicle. Note that the price of gasoline is taken to be exogenous, but the firm can change the expected fuel costs of a vehicle by changing its fuel efficiency. Power-to-weight is a proxy for acceleration, and weight may capture nonlinear effects of acceleration as well as serve as a proxy for safety. This specification allows power-to-weight and weight to enter the utility function separately, while many other studies omit weight, e.g., Petrin (2002).

The next term in equation (2), ξ_{jt}, is the average utility derived from the vehicle's unobserved characteristics. The final term in equation (2) is the log share of the vehicle's sales in the total sales of the vehicle class, c, where σ is the within-class correlation of market shares.

The supply side of the model is static, following Berry, Levinsohn and Pakes (1995) (henceforth, BLP). A set of multi-product firms competes in a Bertrand-Nash manner. Each firm is subject to the CAFE standard, that the harmonic mean of its car and truck fleets must exceed particular thresholds. If the firm does not satisfy the constraint it would have to pay a fine, but we assume that in equilibrium the constraint is satisfied exactly; this assumption is not important for the empirical analysis and is relaxed in the simulations.

To compare with the medium run model, we first specify the firm's optimization problem in a standard short run model. Vehicle characteristics are exogenous and the firm chooses the vector of prices of its set of vehicles J_f:

$$\max_{\{p_t\}_{j\in J_f}} \sum_{j\in J_f} (p_{jt} - c(X_{jt}))q_{jt}(p_{jt}, X_{jt}, \xi_{jt}) \qquad \text{(SR)}$$

$$\text{s.t.} \sum_{j\in J_f} q_{jt}(p_{jt}, X_{jt}, \xi_{jt})/C_{jt} \geq \sum_{j\in J_f} q_{jt}(p_{jt}, X_{jt}, \xi_{jt})/M_{jt} \,,$$

where X_{jt} is a vector of (exogenous) characteristics: fuel efficiency, weight and power; and $c(X_{jt})$ is the marginal cost of the vehicle, which depends on the characteristics. The parameter C_{jt} is the CAFE standard that applies to vehicle model j in year t.

We now specify the medium run optimization problem, in which firms choose prices and characteristics each period:

$$\max_{\{p_t, X_{jt}, \xi_{jt}, T_{jt}\}_{j\in J_f}} \sum_{j\in J_f} (p_{jt} - c(X_{jt}))q_{jt}(p_{jt}, X_{jt}, \xi_{jt}) \qquad \text{(MR)}$$

s.t. $\sum_{j \in J_f} q_{jt}(p_{jt}, X_{jt}, \xi_{jt}) / C_{jt} \geq \sum_{j \in J_f} q_{jt}(p_{jt}, X_{jt}, \xi_{jt}) / M_{jt}$ (a)

$$\ln M_{jt} = \delta_0 + \delta_1 \ln H_{jt} + \delta_2 \ln W_{jt} + T_{jt} \quad \text{(b)}$$

$$\ln c_{jt} = \gamma_0 + \gamma_1 \ln H_{jt} + \gamma_2 \ln W_{jt} + \gamma_3 \ln T_{jt} \quad \text{(c)}$$

Equation (b) specifies that the fuel efficiency of vehicle model *j* depends on the engine's horsepower, the vehicle's weight and the level of the engine technology. The engine technology is continuous and is scaled so that a unit increase raises log fuel efficiency by one.[9] The marginal cost of the vehicle model is given by equation (c), and depends on the power of the engine, the weight of the vehicle and the engine technology. Note that improving engine technology raises fuel efficiency and therefore demand for the vehicle, but also raises costs; this tradeoff is governed by the coefficient on dollars-per-mile in equation (2) and the cost elasticity in (c). Analogous tradeoffs exist for increasing weight and power. In equilibrium, firms choose the profit-maximizing vectors of prices and vehicle characteristics and consumers choose vehicles based on the prices and characteristics.

The equilibrium depends on supply and demand parameters, but also on the CAFE standard. Similarly to past research, we are interested in the effect of the CAFE standard on the market equilibrium. To answer this question, it is necessary to estimate the parameters in equation (2). Estimating the demand for fuel efficiency, β_D, is straightforward, using the same approach as Klier and Linn (2008). Specifically, we use within model-year variation in gasoline prices and sales to estimate β_D, which controls for unobserved vehicle model-specific parameters, ξ_{jt}. Identification arises from within model-year variation in fuel costs, but it is not possible to use

[9] Equation (b) is similar to equation (1) above, but the subscripts are different. Equation (1) is estimated using observations at the engine-vehicle model level. Sales data are only available by vehicle model and year, however, and the analysis in this section is aggregated to that level.

this approach to estimate the coefficients in equation (2) for the variables that do not vary within the model-year, α , β_H , β_W ,and σ . Therefore, we use the estimate of β_D to obtain equation (2'):

$$\ln s_{jt} - \ln s_{0t} - \hat{\beta}_D D_{jt} = \alpha p_{jt} + \beta_H HW_{jt} + \beta_W W_{jt} + \xi_{jt} + \sigma \ln s_{j|c} \qquad (2')$$

The transformation reduces the number of parameters needed to be estimated.

Estimating equation (2') is far more challenging than in a short run setting. Firms choose the characteristics of each vehicle, taking as given the characteristics of the vehicles sold by other firms in the market. From the first order conditions for (MR), the observed characteristics are correlated with the unobserved characteristics of the same vehicle model, and with both observed and unobserved characteristics of other vehicles. For example, if Honda increases the power of one of its Acura car models, Toyota may increase the power of the Lexus car models that are substitutes for the Acura.

Because of this correlation, estimating equation (2') by Ordinary Least Squares (OLS) would yield biased estimates of all coefficients. The endogeneity of vehicle characteristics implies that three standard approaches would also yield biased estimates. First, including vehicle fixed effects would only address the problem if one assumes that unobserved characteristics do not change over time (i.e., $\xi_{jt} = \xi_j$). In that case, the parameters would be identified by within-model changes in prices, power and weight. This assumption is not appropriate because there are many unobserved characteristics, such as interior cabin space, that firms can change as readily as power and weight.

The second approach would be to follow many previous studies of automobile demand, such as BLP, and use moments of vehicle characteristics of other vehicles in the same class or other vehicles sold by the same firm to instrument for the price and within-class market share. The instruments are valid if characteristics are exogenous, in which case the instruments would be correlated with vehicle prices (via first order conditions in model SR), but would not be correlated with the unobserved characteristics. Such an argument cannot be made in the medium run analysis, however, in which characteristics are endogenous. A similar argument can be made for the third approach, performing a hedonic analysis (e.g., McManus, 2005).

5.2 ESTIMATION STRATEGY

We use an estimation strategy that is similar in spirit to Hausman *et al*. (1994), in that we take advantage of common cost shocks across subsets of the market. The difference is that we use characteristics of other vehicle models to instrument for characteristics and prices, rather than instrumenting solely for prices, and we exploit the technological relationships across vehicle models sold by the same firm.

Many vehicle models in different classes contain the same engines. This practice is common for SUVs and pickup trucks, but is not confined to those classes; Section 5.3 documents the prevalence of this behavior across the entire market. As a result, when vehicles in different classes have the same engines, they have very similar engine characteristics. For example, the Ford F-Series, a pickup truck, has the same engine as the Ford Excursion, an SUV, and both vehicles have very similar fuel efficiency and power.

Consider two vehicle models, j and j', which have engines e and e' that belong to the same engine platform. The vehicles are in different vehicle classes and the profit-maximizing power of vehicle j depends on the cost of increasing power for the particular engine platform, and similarly for vehicle j'. Therefore, the power of vehicle j will be a function of the power of vehicle j', plus a constant:

$$H_{jec} = f(H_{j'e'c'}) + \eta_c \tag{3}$$

The power of the two vehicles is correlated because they have the same engine. The class intercepts, η_c, are arbitrary, potentially nonlinear, functions of the characteristics of other vehicles in the same class, as well as non-engine characteristics of the same vehicle. The intercepts allow for class-specific demand and supply shocks, so that the power of the two vehicles will differ because of variation across classes in consumer preferences and the characteristics of the other vehicles in the respective classes.

The instrumental variables (IV) strategy is based on equation (3), in which we instrument for a vehicle's price, power-to-weight, weight and within-class market share. The instruments are the means of eight engine characteristics of vehicle models that are located in other classes, but which have the same engine platform.[10] The IV strategy yields unbiased estimates of the demand for power and weight if the error term in equation (3) is uncorrelated across classes for vehicles that have the same engine.[11] Note that this assumption is considerably weaker than the standard assumption that observed and unobserved characteristics are uncorrelated.[12]

Although this approach relaxes the assumption that vehicle characteristics are exogenous, there are several potential sources of bias. First, there may be unobserved brand-specific fixed effects or trends, which would cause η_c to be correlated across classes. To address this concern, the specification includes brand-year interactions; for example, the approach would be robust if

[10] The instruments are listed in Appendix Table 1 and include fuel efficiency, power, weight, power-to-weight, torque, the number of valves, the number of cylinders and displacement. The instruments are calculated as the mean deviation from the class mean to account for the class intercepts in equation (3). The results are similar if means rather than mean deviations are used to construct the instruments. We prefer to construct the instruments using engine platforms rather than engine programs because the sample size is much larger and the instruments for a particular vehicle are constructed from a wider range of other vehicles, which probably reduces bias. Note that the results are sensitive to this distinction, however, as the demand for power is small and not statistically significant using program-based instruments.

[11] We assume that demand is uncorrelated across vehicle classes. Strictly speaking, this is not the case in the nested logit framework, but cross-class demand elasticities are second order in magnitude.

[12] Estimating equation (2') is preferable to equation (2) because the same set of instruments is available for both equations, but (2') has one less endogenous variable. An additional advantage is that power, weight and fuel efficiency are highly correlated with one another, making it difficult to obtain robust estimates of the coefficients on dollars-per-mile, power and weight if all variables are included in the IV estimation.

all Honda models share common unobserved characteristics. Second, the estimates would be biased if there were unobserved engine characteristics. However, we believe that the included variables in equation (2') capture the main features that consumers use to differentiate engines, as the results are robust to adding other engine characteristics, such as the number of cylinders or the engine's torque. Finally, the decision to use a particular engine in a vehicle model may be endogenous. The identifying assumption is that the correlation of characteristics across vehicle models is driven by the common engine technology, but this may not be valid if unobserved vehicle characteristics are also correlated across models with the same engine. We can partially address this issue by using lagged engine characteristics as instruments, which takes advantage of the fact that engines are redesigned at longer time scales than the rest of the vehicle. Consequently, the correlation between the instruments and endogenous variables is more likely to be driven by a common engine technology, rather than common unobserved characteristics. The results are not sensitive to using lagged values to construct the instruments (see section 6.3 and Table 7 for robustness checks).

5.3 Variation in Engines and First Stage Results

Before reporting the results of estimating equation (2'), we summarize the engine variation across vehicle models and discuss the first stage estimates for equation (2'). Each row in Table 3 includes a different vehicle class. Column 1 shows the number of vehicle models in 2008 and column 2 shows the number of vehicle models in the sample for 2008. The sample only includes vehicles that have an engine found in a vehicle from a different vehicle class, i.e., for which the instruments can be constructed. Only about two-thirds of the vehicles are in the sample, but columns 3 and 4 show that the sample includes 87 percent of total sales. Furthermore, except for small cars, the sample includes nearly all of the sales for each class. It is important to note that it would be possible to increase the sample size by defining narrower vehicle classes. There is a tradeoff between sample size and bias, however, because with narrower classes it is more likely that demand shocks are correlated across classes, invalidating the IV approach.

Table 4 reports summary statistics for the dependent variable and four endogenous right-hand-side variables in equation (2’). For the final estimation sample, the two columns show the means and standard deviations of the variables. Price is reported in thousands of dollars, power-to-weight is measured in horsepower per pound and weight is in tons.

Appendix Table 1 reports the first stage estimates. The dependent variables are the four endogenous variables from Table 4. All specifications include brand-year interactions and the reported engine-based instruments. The instruments are jointly strong predictors of the endogenous variables.

5.4 The Demand for Power and Weight

Table 5 reports the estimates of the demand for power and weight from equation (2’). The dependent variable is the log of the vehicle model’s market share and the independent variables are the price of the vehicle, power-to-weight, weight, the within-class market share and a set of brand-year interactions.

Column 1 reports the OLS estimates of (2’) for comparison with the IV estimates. The coefficient on the price of the vehicle is statistically significant but is small in magnitude, as the average own-price elasticity of demand is -0.16. The coefficient on power-to-weight is negative and is not significant. The price coefficient is likely biased towards zero because the price should be positively correlated with unobserved variables, but the direction of the bias for the characteristics is ambiguous because they may be positively or negatively correlated with unobserved characteristics.

Previous studies, such as BLP, use observed vehicle characteristics to instrument for the vehicle's price. As noted above, this approach is only valid if the instruments are uncorrelated with the unobserved characteristics. Column 2 of Table 5 reports a specification that follows the previous literature and uses other characteristics as instruments, in particular, the sum of the characteristics of other vehicles in the same class and the sum of characteristics of other vehicles sold by the same firm. The coefficient on the vehicle's price is larger in magnitude than the OLS estimate, and implies an average elasticity of demand of -2.02, which is somewhat smaller than previous studies. The coefficient on power-to-weight is close to zero, however.

Column 3 reports the baseline specification using the engine-based instruments. The estimated coefficient on the vehicle's price is larger than the other estimates and the average elasticity of demand is -2.6. The coefficient on power-to-weight is much larger and is statistically significant. The estimate implies that a one percent increase in power raises willingness-to-pay for the average vehicle by about the same as a one percent increase in fuel efficiency. Because of the steep technological tradeoff between power and fuel efficiency (see Table 2), this result is consistent with Figures 2 and 4, which show that as engine technology improved, firms have increased power and weight while keeping fuel efficiency constant.

5.5 EFFECT OF CHANGES IN CHARACTERISTICS ON WILLINGNESS-TO-PAY FOR U.S. CARS

If the demand for weight and power is sufficiently large relative to the demand for fuel efficiency, the decrease in weight and power in the late 1970s and 1980s for U.S. cars would have reduced willingness-to-pay for these vehicles. Figure 5 plots the change in willingness-to-pay for the average car sold by U.S. firms from 1975-2007, using the characteristics in Figure 2, the estimates from column 3 of Table 5, and holding the price of gasoline fixed. The figure shows that willingness-to-pay decreased soon after CAFE was implemented, but increased steadily beginning around 1980.[13] Note that the willingness-to-pay calculations are properly interpreted as the effect of the CAFE standard on willingness-to-pay only if all characteristics and prices would have remained constant in the absence of the policy. Thus, Figure 5 does not allow for an inference about the causal effect of CAFE, but is useful for summarizing the relative demand for fuel efficiency, power and weight.

6 Simulation Results and Interpretation

This section uses the empirical estimates from Section 5 to compare the short and medium run costs of the CAFE standard. We simulate the equilibrium under a 2 MPG increase in the CAFE standard for all vehicles.

6.1 Short Run Effects of an Increase in the CAFE Standard

In the simulation model firms maximize profits subject to the CAFE standard. For comparison with the previous literature and with the medium run analysis, we first simulate the short run effects of the CAFE standard. The model is summarized in Section 5.1. Firms choose a vector of prices to maximize profits subject to the CAFE standard. Firms are separated into three categories: unconstrained firms that exceed the standard, constrained firms that meet the standard, and firms that pay the fine for not meeting the standard. Firms are assigned to the three categories based on past behavior. Honda, Toyota and several smaller Asian firms have

[13] Greene and Liu (1988) perform a similar analysis and reach the same conclusion using estimates of willingness-to-pay for characteristics from other studies performed in the 1970s and 1980s.

consistently exceeded the standard by a wide margin and are unconstrained; Chrysler, Ford and GM and a few other firms have generally been close to the standard and are constrained; and all other firms have been well below the standard. The constrained firms solve problem (SR), while the other firms do not have a constraint; unconstrained firms that do not satisfy the constraint pay a fine. In performing the simulations, we assume that firms do not change categories as a result of the increase in the standard.

Table 6 shows the estimated effects of a 2 MPG increase in the CAFE standard. The columns report the changes in consumer surplus, total profits, profits of U.S. firms, market share of U.S. firms, overall fuel efficiency, horsepower and weight. Consumer surplus declines by about $19 billion because of the changes in vehicle prices under the increased standard. Total profits decrease by about $17 billion. Columns 3-5 show that the increase in the standard causes a transfer in profits from U.S firms to Honda and Toyota, which can be explained as follows. In response to the higher CAFE standard, U.S. firms must change their sales mix in order to increase average fuel efficiency. The resulting price changes cause consumers to substitute to competing vehicle models, which increases the profits of firms that are not constrained by the new standard. The table shows that the increase in the CAFE standard raises average fuel efficiency by less than 2 MPG because many firms are not constrained and do not increase fuel efficiency. Finally, power and weight decrease because constrained firms adjust prices so that consumers purchase more fuel efficient vehicles, which tend to be less powerful and lighter.

6.2 Medium Run Effects (preliminary)

The second row of Table 6 reports the results of simulating a 2 MPG increase using the medium run model from Section 5.1, (MR). All firms choose prices and vehicle characteristics to maximize profits. Firms are classified among the same three categories as before.

The medium run simulation model includes two important differences from the short run model. First, each vehicle's fuel efficiency is endogenous and depends on weight, power and technology. The simulation uses the elasticities of fuel efficiency with respect to power and weight that were estimated in Section 4.

The second difference of the medium run model is that marginal costs are now endogenous. Because firms do not change characteristics in the short run, marginal costs are not affected by the CAFE standard in the short run.[14] However, marginal costs play an important role in the medium run analysis. For example, if marginal costs increase significantly when firms reduce weight, firms would be unlikely to do so. We assume a CES cost function, where the elasticity of costs to power is estimated using proprietary engine cost data. Similarly to Austin and Dinan (2005), the elasticities of costs to weight and engine technology are estimated using data on the costs and efficacy of engine and weight reduction technologies from NHTSA (2008).[15] It is important to note that in the medium run analysis, only a limited set of engine technologies can be adopted. Therefore, the elasticity of costs to engine technology is greater in the medium run than in the long run (the short run elasticity is infinite).

The second row of Table 6 reports summary statistics from a preliminary simulation of the medium run effects of the standard. The differences between the short and medium run

[14] We assume throughout that there are no economies of scale, so that marginal costs only depend on vehicle characteristics.

[15] The constant terms in the cost and technology equations are estimated using the initial fuel efficiency and marginal cost of each model (i.e., before the increase in the standard). The final fuel efficiency and marginal cost are calculated using the deviations from the initial values of power, weight and technology.

simulations underscore the importance of accounting for the endogeneity of vehicle characteristics. The overall changes in producer and consumer surplus are roughly half as large in the medium run as in the short run. This result is consistent with Jacobsen (2008), who finds that the long run cost is roughly one-third of the short run cost, so that the medium run costs lie between the two extremes. Section 4 suggests that short run changes in the sales mix are important for at most one or two years, while medium run changes in vehicle characteristics are important for roughly 5 years. Thus, previous studies significantly overstate the annual cost of the CAFE standard for horizons of about 2-5 years.

Many previous studies compare the cost of reducing gasoline consumption using the gasoline tax with the cost of using the CAFE standard. Although the medium run costs of the CAFE standard are much lower than the short run costs, the magnitudes do not overturn the conclusions of other studies that the gasoline tax is much less costly than the CAFE standard. Jacobsen (2008) finds that the short run cost of the gasoline tax is roughly one-sixth the cost of the CAFE standard. Therefore, even in the medium run, CAFE is more expensive than the gasoline tax.

6.3 Robustness and Limitations

Table 7 reports a number of robustness checks for equation (2’). Columns 1-4 assess the importance of including brand-year interactions, add vehicle class-year interactions and address potential serial correlation. The coefficient on power-to-weight is considerably smaller if class-year interactions are added to equation (2’). Columns 5 and 6 address functional form assumptions by including power and weight separately and adding other engine characteristics on the right-hand-side; the results are similar in both cases. Column 7 shows that the estimated coefficient on power-to-weight is smaller if additional instruments are included. The estimate is not affected using lagged instruments (columns 8 and 9), which addresses the potentially endogenous choice of which engines are paired with which vehicles (see Section 5.3). Overall, the results are somewhat sensitive to the alternative specifications, although the estimate on power-to-weight is positive in all specifications and is statistically significant in most. We use the specification in Table 5 for the simulations because of the relatively large estimate on power-to-weight. The fact that the large estimate is used implies that the decrease in costs between the short and medium run may be at least as large as reported in Table 6.

We believe that the sensitivity of estimated willingness-to-pay to alternative specifications has not been emphasized enough in the previous literature, where the standard practice is to report one or two specifications. Furthermore, Appendix Table 2 shows that the BLP specification is at least as sensitive as the engine-based specification.

A few limitations of the analysis should be noted. The model used to perform the simulations uses the original structure of the CAFE standard, which was based on the harmonic mean of a firm's fuel efficiency for cars and light trucks. Future work will incorporate the new version of the standard, which is based on a vehicle's footprint. More difficult to address is the assumption in the simulations that unobserved characteristics do not change in response to the increase in the standard.

Finally, the policy scenario discussed above considers the medium run effect of the CAFE standard, in which there is no entry (exit is modeled in the simulation, however). Explicitly allowing for the entry of vehicle models is a potential direction for future research.

7 CONCLUSION

The upcoming increase in the CAFE standard will significantly affect the new vehicles market. This paper analyzes the medium run effect of the standard, which we define as the response when engine technology is held constant but firms can change vehicle characteristics. This paper first shows that in response to the initial standard, firms significantly reduced the power and weight of vehicles sold in the late 1970s and early 1980s in order to increase fuel efficiency, but technological progress caused power to recover in the long run.

We then estimate consumers' demand for power and weight in order to analyze the medium run effects of the CAFE standard. Estimating demand is complicated by the fact that firms select vehicle characteristics endogenously, which previous empirical work has not addressed. We propose an instrumental variables strategy that controls for endogenous and time-varying unobserved characteristics. The estimates suggest that consumers value an increase in power roughly the same as a proportional increase in fuel efficiency. We use a static model of the new vehicles market to simulate the effect of an increase in the standard. The policy causes considerable transfers from constrained firms (U.S. firms, for the most part) to other firms. The medium run costs are substantially lower than the short run costs, however. Given the small role of changes in the sales mix documented in Section 4, this result implies that the short run analysis substantially overestimates the cost of the regulation. Furthermore, the results suggest that firms can attain larger improvements in fuel efficiency in a shorter amount of time than is suggested by a long run analysis. That is, both the short and long run analysis likely overstate the total discounted cost of the CAFE regulation by a significant margin. However, the magnitudes reported in this paper still do not suggest that the CAFE standard compares favorably to a gasoline tax in terms of the cost of reducing gasoline consumption.

8 REFERENCES

1 Atkinson, S. (1981). "Rising Gasoline Prices and Federal Automotive Efficiency Standards: Their Impact on Consumer Choice." Research Study #23, American Petroleum Institute, Washington D.C.

2 Austin, David and Terry Dinan (2005). "Clearing the Air: The Costs and Consequences of Higher CAFE Standards and Increases in Gasoline Taxes." *Journal of Environmental Economics and Management*: vol. 50, 562-582.

3 Bento, Antonio M., Lawrence H. Goulder, Mark R. Jacobsen and Roger H. von Haefen (2006), "Distributional and Efficiency Impacts of Increased U.S. Gasoline Taxes."

4 Berry, Steven (1994). "Estimating Discrete Choice Models of Product Differentiation." RAND Journal of Economics, vol. 25: 242-262.

5 Berry, Steven, James Levinsohn and Ariel Pakes (1995). "Automobile Prices in Market Equilibrium," *Econometrica*: vol. 63(4), 841-890.

6 Goldberg, Penelope Koujianou (1998), "The Effects of the Corporate Average Fuel Efficiency Standards in the U.S.," *The Journal of Industrial Economics*: vol. 46, n1, 1-33.

7 Greene, David L. (1987). "Advances in Automobile Technology and the Market for Fuel Efficiency, 1978-1985." *Transportation Research Record*, vol.1155, 18-27.

8 Greene, David L. (1987). "Short-Run Pricing Strategies to Increase Corporate Average Fuel Economy". *Economic Inquiry*, vol. 29, 101-114.

9 Greene, David L. and Jin-Tan Liu (1988). "Automotive Fuel Economy Improvements and Consumers' Surplus." *Transportation Research A*, vol. 22-A, 203-218.

10 Hausman, Jerry, Gregory Leonard and J. Douglas Zona (1994). "Competitive Analysis with Differentiated Products." *Annales D'Economique et de Statistique*, v34, 159-180.

11 Ingrassia, Paul. 2008. "Detroit's (Long) Quest for Fuel Efficiency," The Wall Street Journal, February 19, p. A19.

12 Ishii, Joy (2005). "Compatibility, Competition and Investment in Network Industries: ATM Networks in the Banking Industry."

13 Jacobsen, Mark (2008). "Evaluating Fuel efficiency Standards in a Model with Producer and Household Heterogeneity."

14 Kleit, Andrew N. (2004). "Impacts of Long-Range Increases in the Fuel Economy (CAFE) Standard." *Economic Inquiry*, vol. 42, 279-294.

15 Klier, Thomas and Joshua Linn (2008). "The Price of Gasoline and the Demand for Fuel Efficiency: Evidence from Monthly New Vehicles Sales Data."

16 Kranz, Rick (2008). "GM cars to get smaller engines with turbos," *Automotive News*, http://www.autonews.com/apps/pbcs.dll/article?AID=/20080128/ANA06/801280366 (accessed January 28 2008).

17 McManus, Walter S. (2005). "The Effects of Higher Gasoline Prices on U.S. Vehicle Sales, Prices and Variable Profit by Segment and Manufacturer Group, 2001 and 2004."

18 National Research Council - Committee on Assessment of Technologies for Improving Light-Duty Vehicle Fuel efficiency, 2008. Interim Technology Assessment. Letter to national Highway Traffic Safety Administration. February 14

19 National Highway Traffic Safety Administration, 2008. Preliminary Regulatory Impact Analysis for the Corporate Average Fuel Economy for MY 2011-2015 Passenger Cars and Light Trucks.
20 Ohnsman, Alan. 2008. "Honda bucks industry with TSX favoring fuel efficiency over power," Bloomberg. February 15
21 U.S. EPA, 2007. Light-Duty Automotive Technology and Fuel efficiency Trends: 1975 through 2007. EPA 420-R-07-008
22 Petrin, Amil (2002). "Quantifying the Benefits of New Products: The Case of the Minivan." *Journal of Political Economy*: vol. 110, 705-729.
23 Sweeting, Andrew (2007). "Dynamic Product Repositioning in Differentiated Product Markets: The Case of Format Switching in the Commercial Radio Industry." NBER Working Paper #13522.
24 Ward's Automotive Yearbook, 1980-2003, Ward's Communications.
25 Ward's AutoInfoBank, Ward's Automotive Group.

Table 1

Examples of Medium and Long Run Engine and Transmission Changes

Medium Run			Long Run		
Technology	Cost ($)	Percent Increase in MPG	Technology	Cost ($)	Percent Increase in MPG
Low Friction Lubricants	3	0.5	Turbocharge/ Downsize	120	5-7.5
Variable Valve Timing	59-209	1-3	Continuously Variable Trans	139	3.5
5-speed Automatic Transmission	76-167	0.5-2.5	Automatic Manual Transmission	141	4.5-7.5
Cylinder Deactivation	203	4.5-6	PHEV	6750	28

Source: NHTSA (2008). All figures represent estimates for a mid-size car.

Table 2

Tradeoff Between Fuel Efficiency, Weight and Power for Cars

	Dependent Variable: Log Fuel Efficiency	
	(1)	(2)
Log Horsepower	-0.06 (0.03)	-0.15 (0.03)
Log Weight	-0.33 (0.07)	-0.33 (0.09)
R^2	0.90	0.84
Number of Observations	1989	1989
Fixed Effects	Engine Program	Engine Platform

Notes: Standard errors in parentheses, clustered by engine. Observations are by engine and year for 2000-2007. All specifications are estimated by Ordinary Least Squares. The dependent variable is the log of the fuel efficiency of the corresponding vehicle model. All columns include the log of the engine's power and the log of the vehicle model's weight. Column 1 includes engine program dummies and column 2 includes engine platform dummies.

Table 3

Sample Coverage by Vehicle Class, 2008

	(1)	(2)	(3)	(4)
Vehicle Class	Number of Vehicle Models	Number of Vehicle Models with Instruments	Fraction Sales	Fraction Sales with Instruments
Small Cars	36	15	0.16	0.10
Mid-Size Cars	38	22	0.20	0.19
Large, Luxury and Specialty Cars	68	46	0.12	0.10
Small SUVs	56	40	0.18	0.16
Large SUVs	43	34	0.11	0.11
Vans	15	10	0.07	0.06
Pickup Trucks	21	18	0.16	0.16
Total	277	185	1.00	0.87

Notes: Vehicles are assigned to the vehicle classes, which are defined in the Wards database. The number of vehicle models is the number of unique models in each class in the 2008 model-year. The number of vehicle models with instruments is the number of models for which there is another model that belongs to a different class and has the same engine. Fraction sales is the share of sales of vehicle models in the class in total sales in the 2008 model-year. Fraction sales with instruments is the fraction of sales in total sales for the vehicle models with instruments.

Table 4

Summary Statistics

Variable Name	Mean	Standard Deviation
Log Market Share	-4.717	1.490
Vehicle Price	33.192	18.002
Power-to-Weight	0.059	0.014
Weight	1.911	0.421
Log Within-Class Market Share	-4.076	1.445

Notes: The table reports the mean and standard deviation of log market share, vehicle price (thousands of dollars), power-to-weight (horsepower per pound), weight (tons) and the log of the within-class market share.

Table 5

Willingness-to-Pay for Power and Weight

	Dependent Variable: Log Market Share		
	(1)	(2)	(3)
Vehicle Price	-0.004 (0.001)	-0.026 (0.007)	-0.050 (0.017)
Power-to-Weight	4.656 (0.977)	1.544 (4.752)	32.785 (10.686)
Weight	0.603 (0.030)	0.895 (0.132)	1.350 (0.295)
Log Within-Class Share	0.924 (0.010)	0.420 (0.070)	0.628 (0.120)
R^2	0.96	0.83	0.88
N	1804	1804	1804
Estimation Model	OLS	IV, BLP Instruments	IV, Engine Instruments

Notes: The table reports the results from estimating equation (2'). Standard errors are in parentheses, robust to heteroskedasticity. The dependent variable is the difference between the log share of sales of the vehicle model in total sales, and the log share of sales of used vehicles in total sales, where total sales include used and new vehicles. The independent variables are the price of the vehicle, in thousands of dollars; power-to-weight, in horsepower divided by weight, in pounds; weight, in tons; the log of the within class share of sales; and a full set of brand-year interactions. Column 1 is estimated by Ordinary Least Squares and columns 2 and 3 are estimated by Instrumental Variables. Column 2 instruments for vehicle price using the sum of characteristics of vehicle models in the same category produced by other firms and the sum of characteristics of other models produced by the firm. Column 3 uses as instruments the independent variables in the Appendix Table.

Table 6

Effects of a 2 MPG Increase in the CAFE Standard

	Change in Cons Surplus (Billion $)	Change in Total Profits (Billion $)	Change in U.S. Firms' Profits (Billion $)	Change in Profits for Honda/Toyota (Billion $)	Percent Change in U.S. Market Share	Change in Fuel Efficiency (MPG)	Change in Horsepower	Change in Weight (Pounds)
Short Run	-19.37	-17.46	-25.43	7.68	-8.82	1.33	-11.36	-184.46
Medium Run	-8.16	-8.18	-8.26	2.14	-3.46	1.42	-24.11	-421.19

Notes: The table reports the effect of a 2 MPG increase in the CAFE standard on consumer surplus total profits, profits of U.S. firms, profits of Honda and Toyota (all in billions of 2007 dollars), the percent change in market share of U.S. firms, and the change in fuel efficiency (MPG), the change in horsepower and the change in weight (pounds). The two rows report the results of different simulations. In the first row, weight, power and fuel efficiency of each vehicle model are held constant, while in the second row these characteristics are chosen by the firm. See text for details on the simulations.

Table 7

Alternative Specifications

	Dependent Variable: Log Market Share								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Vehicle Price	-0.051 (0.017)	0.001 (0.005)	-0.050 (0.030)	-0.004 (0.012)	-0.058 (0.021)	-0.050 (0.017)	-0.028 (0.008)	-0.034 (0.023)	-0.081 (0.047)
Power-to-Weight	33.100 (11.369)	6.646 (2.622)	32.785 (19.452)	21.003 (10.898)		23.990 (9.190)	20.943 (6.785)	39.913 (22.044)	66.969 (36.952)
Power					0.010 (0.004)				
Weight	1.377 (0.299)	0.483 (0.103)	1.350 (0.536)	0.214 (0.238)	0.485 (0.248)	0.026 (0.541)	1.020 (0.129)	1.104 (0.307)	1.726 (1.888)
Log Within-Class Share	0.620 (0.125)	0.968 (0.029)	0.628 (0.223)	0.421 (0.119)	0.591 (0.137)	0.819 (0.076)	0.781 (0.060)	0.718 (0.204)	0.367 (0.366)
Lag Dep Var				0.565 (0.102)					
R^2	0.86	1.00	0.88	0.91	0.87	0.88	0.94	0.90	0.69
N	1804	1804	1804	1496	1804	1804	1804	1089	1151
Spec	Year and Brand Dummies	Add Class-Year Interactions	Cluster by Model	Add Lag Dep Var	Separate Power, Weight	Add Torque and Disp	Other Engine Instr	3-yr Lagged Instr	Lagged 3-yr Mean Instr

Notes: The table reports the specifications indicated in the bottom row, using column 3 of Table 5 as the baseline. Standard errors are robust to heteroskedasticity, except in column 3 where standard errors are clustered by vehicle model. Column 1 includes brand and year dummies instead of brand-year interactions. Column 2 adds vehicle class-year interactions, and does not demean the instruments. Column 4 includes the lag of the dependent variable. Column 5 includes weight and power separately. Column 6 adds torque and displacement (not reported). Column 7 uses additional instruments for vehicle price, log within-class market share and length, which are constructed similarly to the other instruments. Column 8 uses the 3-year lags of the instruments from the corresponding engine platform, and column 9 uses the means of the instruments from 2, 3 and 4 years earlier.

Appendix Table 1

First Stage Estimates

	Dependent Variable: Vehicle Price (Thousand $)	Power-to-Weight (Horsepower/Pound)	Weight (Tons)	Log Within-Class Share
Fuel Efficiency	-0.168	-0.236	-0.415	-0.637
	(0.082)	(0.104)	(0.315)	(1.339)
Power	-0.107	-0.088	-0.043	1.655
	(0.034)	(0.039)	(0.058)	(0.338)
Weight	2.596	12.046	12.437	-1.288
	(6.197)	(5.515)	(13.955)	(58.066)
Power-to-Weight	-0.041	-0.151	-0.014	1.533
	(0.040)	(0.066)	(0.139)	(0.614)
Torque	0.054	-0.045	0.327	-0.025
	(0.031)	(0.021)	(0.065)	(0.298)
Number of Valves	0.945	1.167	-1.024	-10.968
	(0.126)	(0.154)	(0.390)	(1.455)
Number of Cylinders	0.840	-3.253	4.330	-17.415
	(0.915)	(1.081)	(3.501)	(12.606)
Displacement	0.006	0.009	0.009	-0.061
	(0.002)	(0.003)	(0.005)	(0.028)
R^2	0.66	0.38	0.56	0.39
N	1804	1804	1804	1804

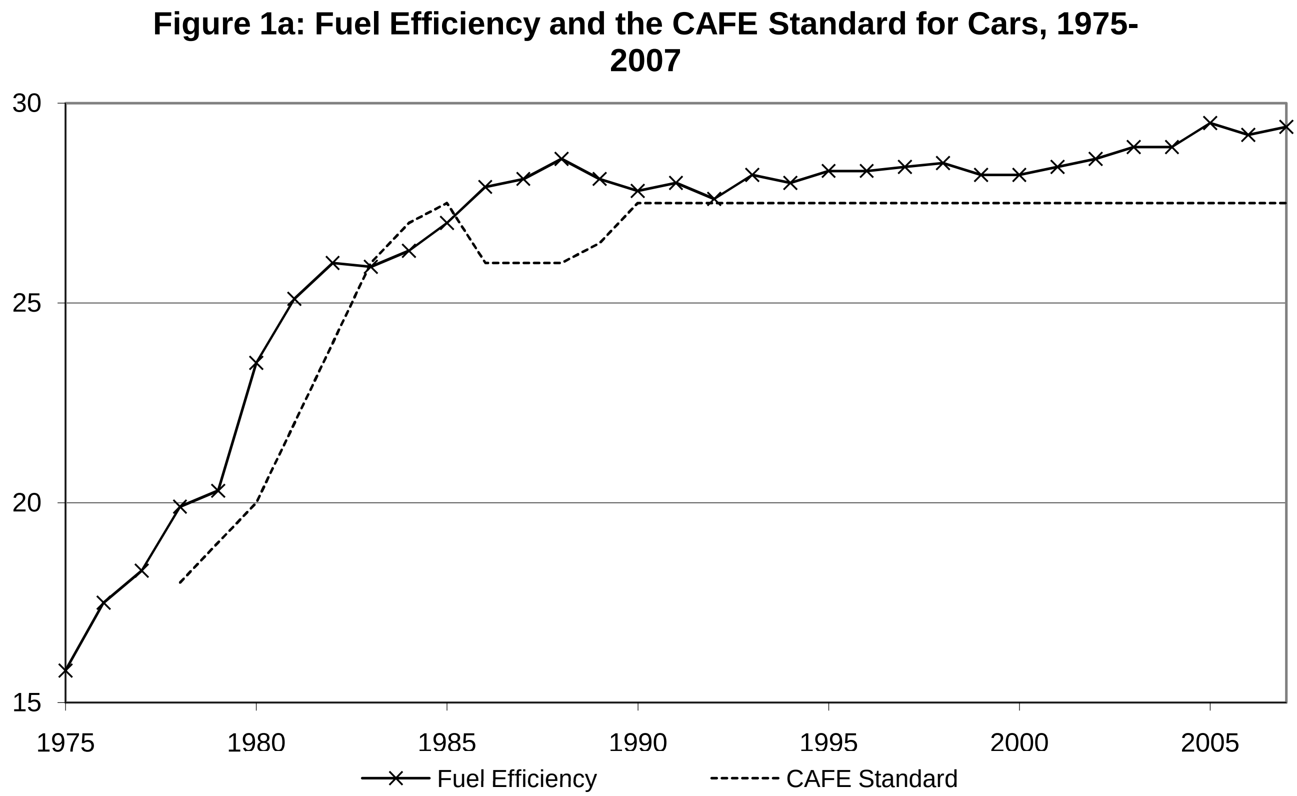
Notes: Instruments for vehicle price, power-to-weight, weight, and within-class market share are constructed from the matched engine model-vehicle model data set. The instruments are the mean of within-class deviations of vehicles belonging to other classes that have the same engine. The sample includes all models for which the instruments can be calculated, and spans 2000-2008. The table reports coefficient estimates with standard errors in parentheses. All regressions include brand-year interactions. Standard errors are robust to heteroskedasticity. For readability, the power-to-weight instrument is divided by 1000, coefficients in column 2 are multiplied by 1000, and the coefficients in columns 3 and 4 are multiplied by 100.

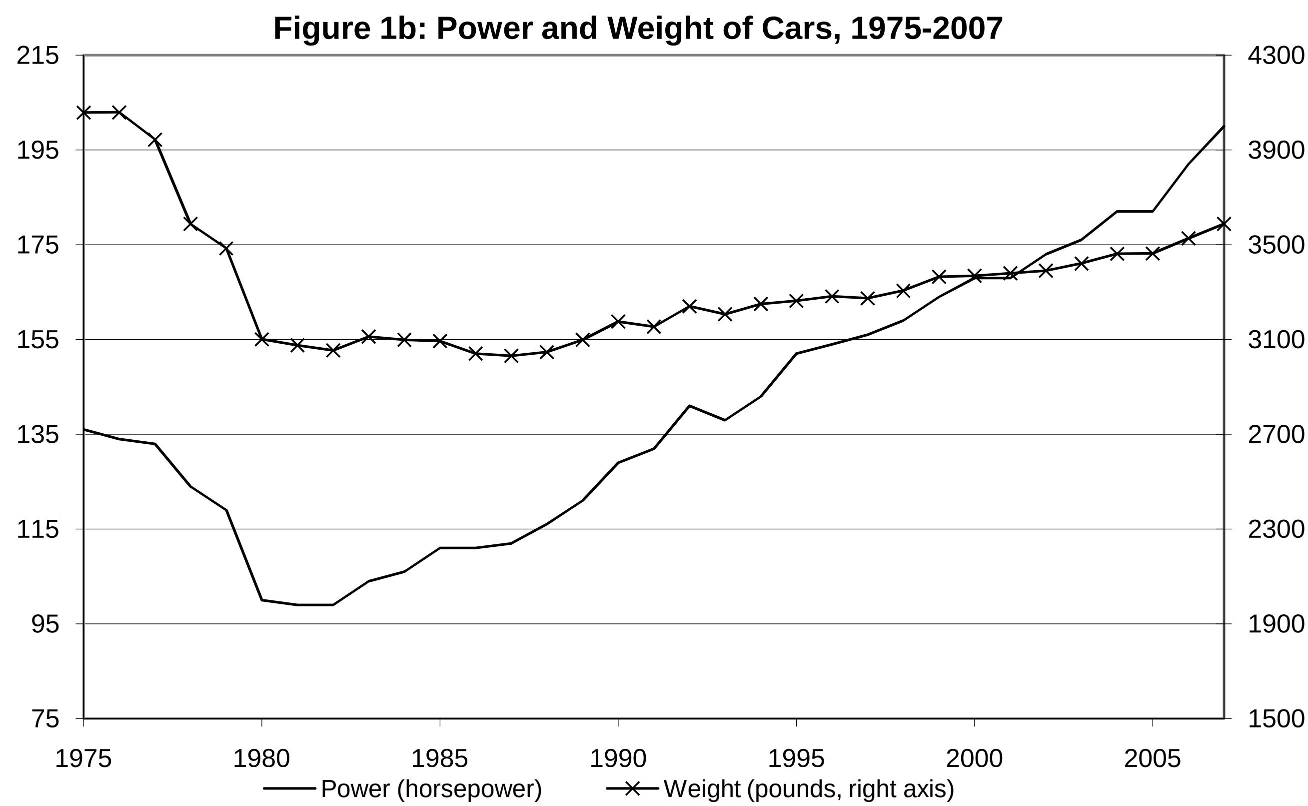
Appendix Table 2

Alternative Specifications With BLP Instruments

	Dependent Variable: Log Market Share						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Vehicle Price	-0.070 (0.014)	-0.021 (0.009)	-0.026 (0.013)	0.004 (0.006)	-0.123 (0.018)	-0.026 (0.007)	-0.010 (0.005)
Power-to-Weight	29.738 (9.071)	-5.779 (5.990)	1.544 (8.482)	-5.705 (3.499)		-4.483 (4.400)	-1.552 (2.893)
Power					0.019 (0.003)		
Weight	1.710 (0.262)	1.159 (0.186)	0.895 (0.245)	0.123 (0.110)	0.681 (0.123)	0.582 (0.147)	0.805 (0.080)
Log Within-Class Share	0.430 (0.074)	0.356 (0.092)	0.420 (0.115)	0.181 (0.063)	0.346 (0.097)	0.419 (0.070)	0.675 (0.042)
Lag Dep Var				0.694 (0.061)			
R^2	0.76	0.81	0.83	0.89	0.59	0.83	0.95
N	1804	1804	1804	1496	1804	1804	1804
Specification	Year and Brand Dummies	Add Class Dummies	Cluster by Model	Add Lag Dep Var	Separate Power, Weight	Add Torque and Disp	Add Car/Truck Nest

Notes: The table reports the specifications indicated in the bottom row. All specifications are the same as the corresponding columns in Table 7, except that the BLP instruments from column 2 of Table 5 are used, rather than the engine-based instruments





Notes: Figures are constructed using data reported in U.S. EPA (2007).

Figure 2a: Fuel Efficiency, Weight and Displacement for Cars of U.S. Manufacturers, 1975-2007

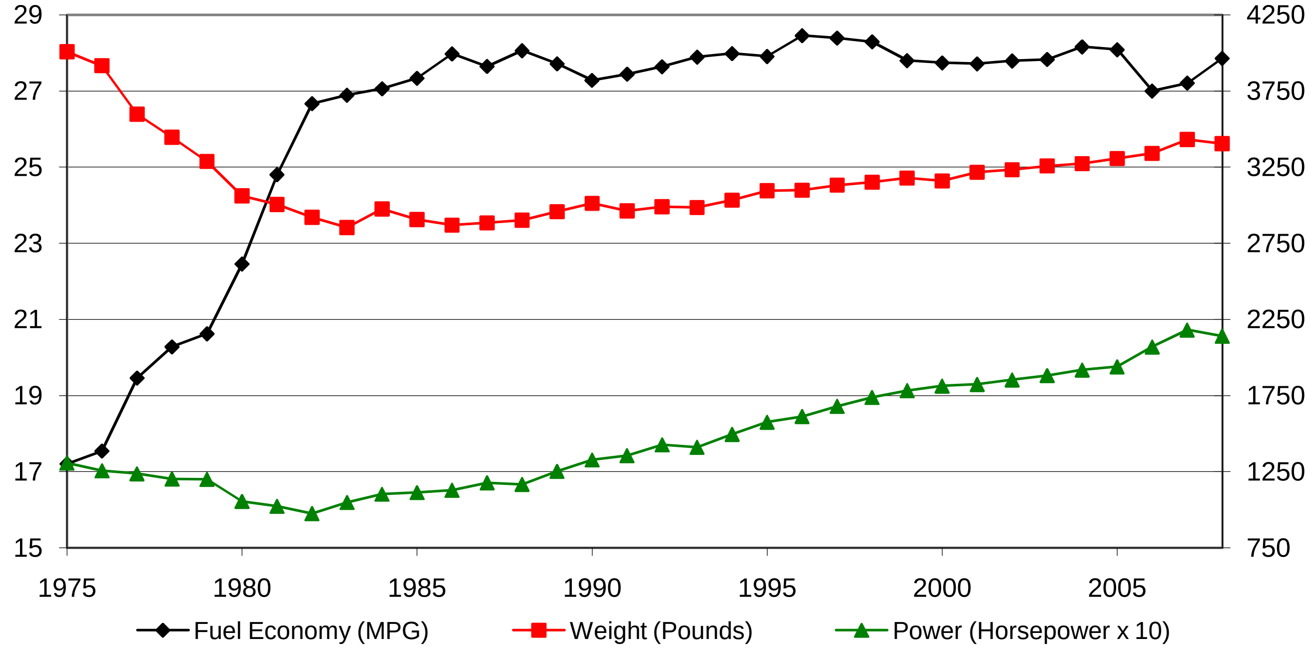
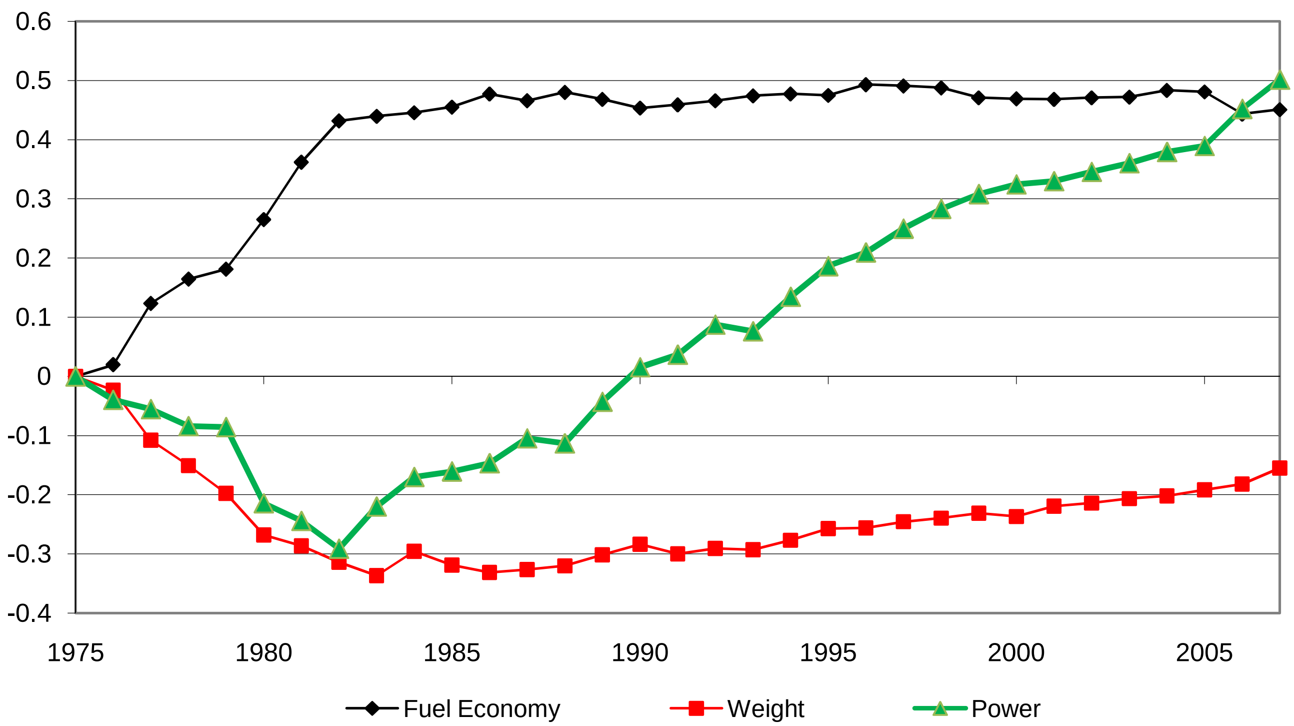
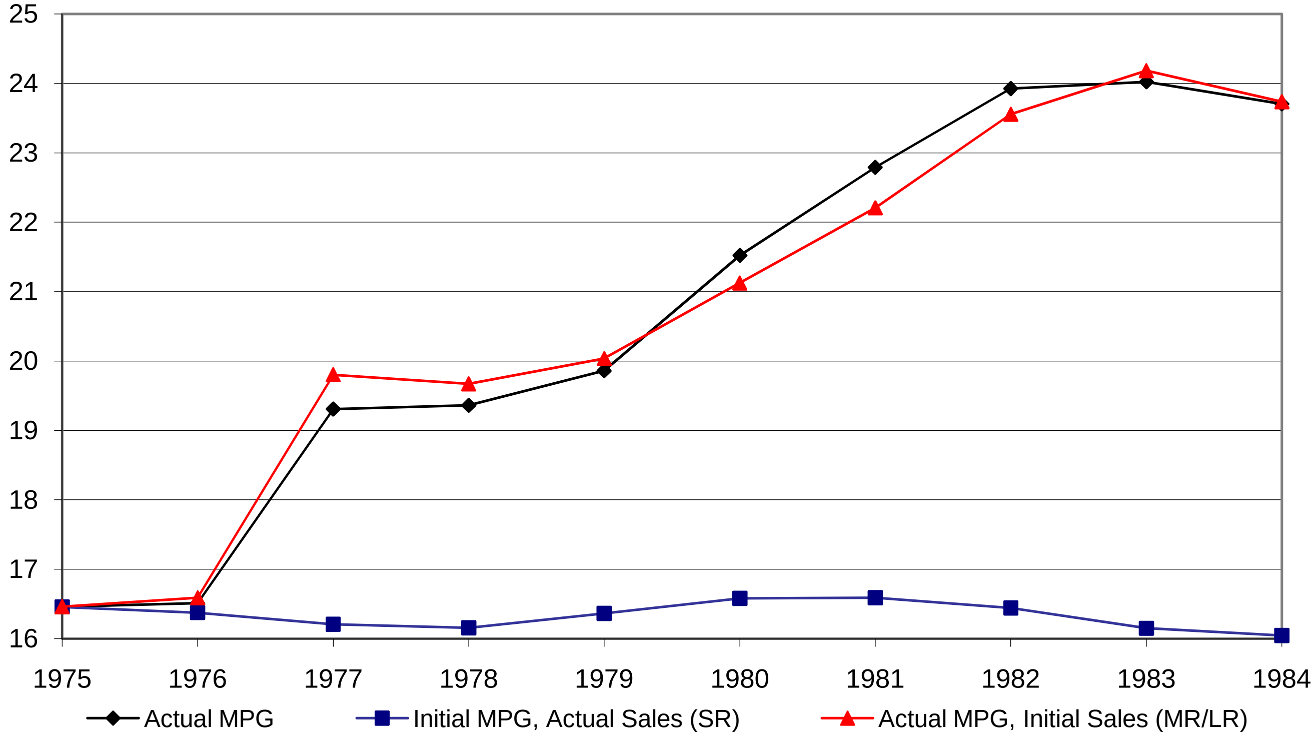


Figure 2b: Change in Fuel Efficiency, Weight and Power, 1975-2008



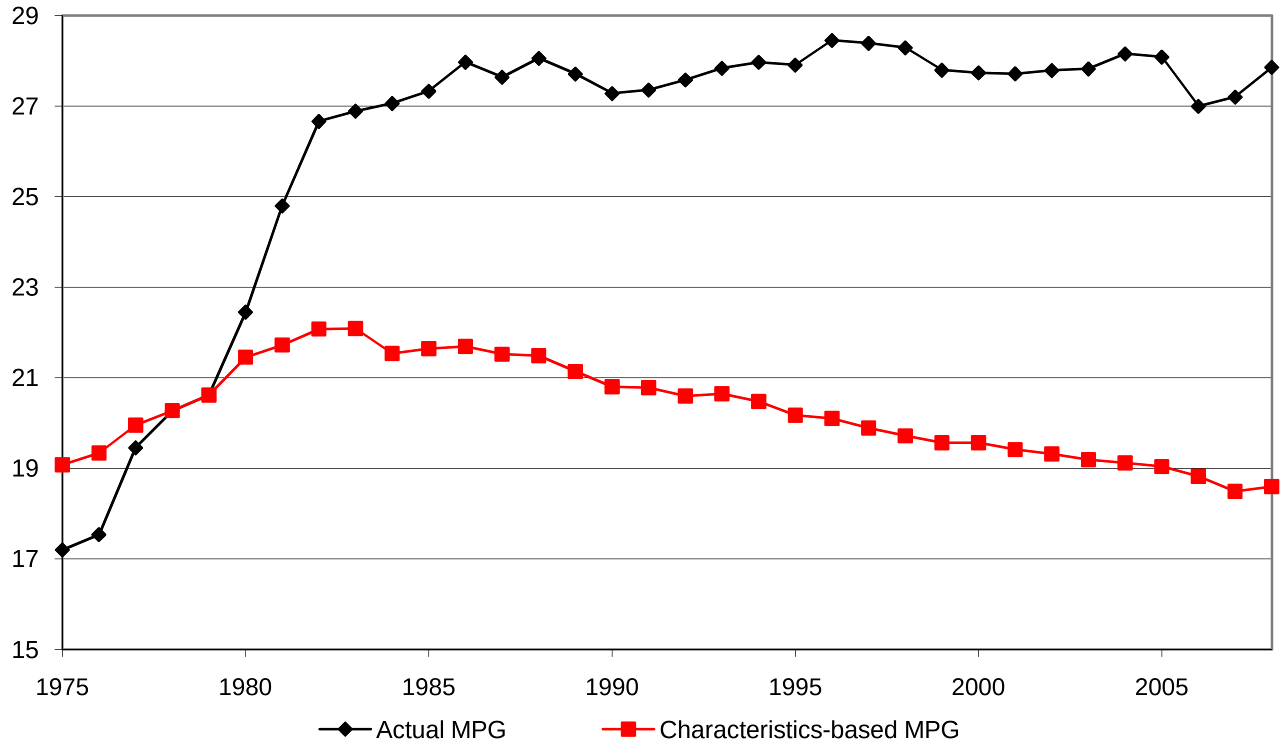
Notes: Figure 2a reports the sales-weighted mean fuel economy (in MPG), weight (in pounds) and horsepower (multiplied by 10) of all cars sold by U.S. companies for each year. Figure 2b reports the percent change in each variable, relative to 1975.

Figure 3: The Effect of Changes in Sales and Fuel Efficiency, Balanced Panel of U.S. Cars, 1975-1984



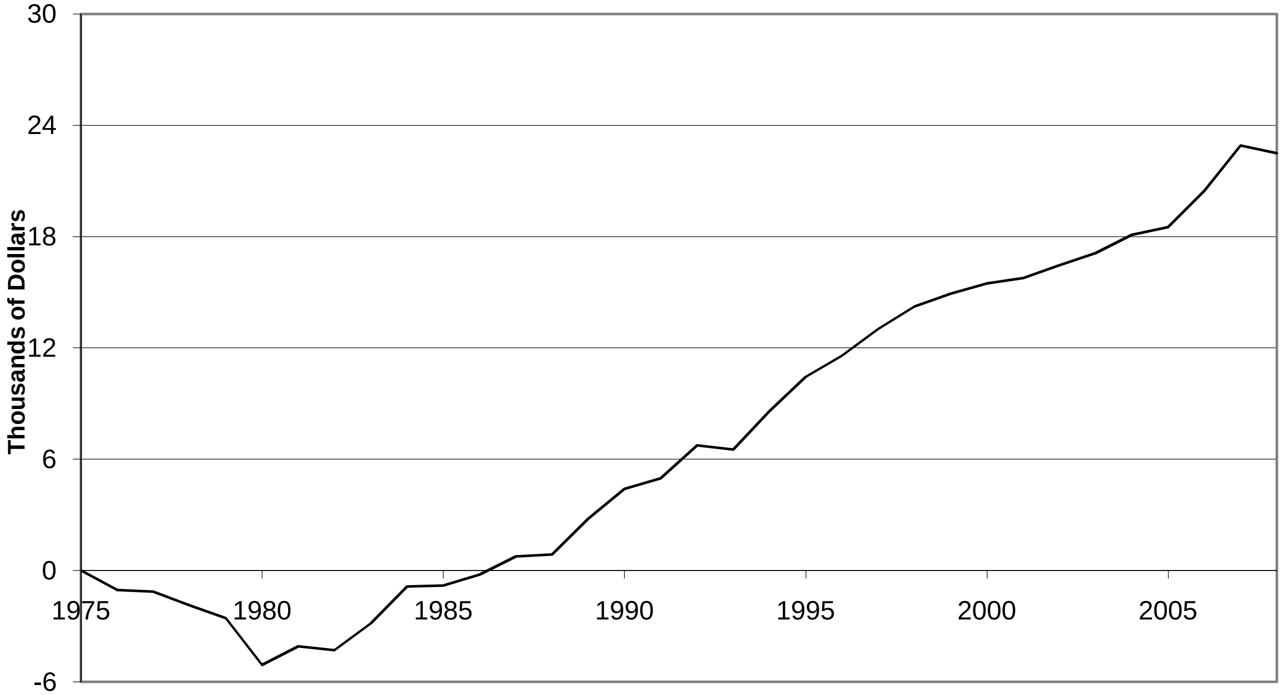
Notes: Actual MPG is the sales-weighted mean MPG of all cars sold by U.S. firms that have positive sales for each year, 1975-1984. The initial MPG series is the sum of the actual MPG in 1975 and the inner product of the change in sales weights and the 1975 MPG of each vehicle model. The actual MPG series is the sum of the actual MPG in 1975 and the inner product of the change in MPG of each vehicle model with the 1975 sales weight. See text for details.

Figure 4: Effect of Power and Weight on Fuel Efficiency for U.S. Manufacturers, 1975-2008



Notes: The actual MPG series is the same series as reported in Figure 2. The change in predicted MPG is calculated using equation (1), the estimated coefficients reported in column 1 of Table 2 and the change in sales-weighted power and weight from Figure 2. The characteristics-based MPG is equal to the sum of the actual MPG in 1978 and the change in predicted MPG.

Figure 5: Change in Willingness-to-Pay Due to Changing Vehicle Characteristics for U.S. Firms, 1975-2008



Notes: The figure plots the change in willingess-to-pay for U.S. cars, using 1975 as the baseline year. Change in willingness-to-pay is calculated using the change in sales-weighted power and weight from Figure 2 and the estimates from column 3 of Table 5.

Is the Auto Market Getting Fuel Economy Right?
Comments on Fischer, Klier & Linn

Gloria E. Helfand

Office of Transportation & Air Quality

U.S. Environmental Protection Agency

What I'll quickly take from Fischer

- A gas tax increases fuel economy but increases distortions
- A minimum fuel economy standard raises an interesting alternative to a CAFE standard
- A "footprint" CAFE standard may lead to less distortion than a "uniform" CAFE standard.

A Few (Random) Observations

- Why isn't price a function of fuel economy too, instead of asserting a willingness to pay for fuel economy independently?
- The policies don't actually aim for a target fuel consumption or other environmental goal to be on the same footing.
- Is there a story here related to actual behavior?
  - Hummers could have better fuel economy
  - Priuses have more than they should

What I'll quickly take from K&L

- Further evidence that automakers put effort into power, not fuel economy
- Fuel economy technology is likely to be cheaper than changing market shares to achieve fuel economy standards
- Comparison of short- and medium-runs
  - Allowing time for technology changes reduces costs substantially – approximately 50%.
- Effects on domestic automakers
  - Since they're the only ones constrained by CAFE, they're worse off from tightening fuel economy

A Few (Random) Observations

- Why is fuel economy a function of horsepower & weight?
  - It implies that it's determined by the exogenous choices of HP & weight, not a choice by itself.
- What's the value of fuel economy used in the fuel cost calculation?
  - Do people get the value "right"?
  - It is a little odd to see a key parameter vanish via subtraction from the estimated equation
- Data sets from different sets of years
  - Fuel economy-HP-weight equation uses 2000-2008 data
  - Counterfactuals use 1975-1984 data
  - Unclear what data used for estimation
- Simulations assume fuel economy averaging over all manufacturers, not within manufacturers
- Is the dependent variable both positive and negative?
  - Where does the share of used cars come from?

Policy Perspective: What do consumers want?

- Do consumers buy the cost-minimizing amount of fuel economy?
  - Do consumers view fuel economy as only a component of the cost of driving?
  - Is fuel economy a vehicle attribute not subject to the same constrained behavior?
- How well can we explain what vehicles people buy?
  - Are consumer choice models good predictive tools?

Do automakers provide the amount of fuel economy that consumers want?

- Fischer suggests that automakers may be operating strategically
- K&L suggest that
  - automakers have invested in power instead of fuel economy
  - a 1% increase in power raises WTP the same as a 1% increase in fuel efficiency
- Are automakers getting it right?

Taxes vs. Standards

- EPA can't do much (anything?) on fuel taxes
- EPA doesn't set fuel economy standards
  - NHTSA does that.
- It's worth pointing out the opportunity costs
  - But Congress needs to hear about them at least as much as EPA

Recommendations for Future Work

- Consumer, producer tradeoffs between fuel economy and other vehicle attributes
  - How much is fuel economy worth to consumers?
  - How much does it cost automakers?
- Efficiency of markets for fuel economy
  - Do consumers buy the cost-minimizing amount?
  - Do automakers offer the choices that consumers most want?
  - If not, why not?
- The footprint CAFE standard
  - Size may become an important endogenous parameter

Discussion of the Klier-Linn and Fischer papers

In the last couple of years we have seen a remarkable amount of work being completed on new car fuel economy and policies to improve it. In addition to the papers presented at this workshop, we have see work by Jacobson, Gulati and coauthors, Train and Winston, Alcott and Muelleger, Bento and Goulder, among others. This work has made use of new and previously unavailable (and still often proprietary) datasets and new methodological approaches. The two papers presented at this workshop are excellent examples of this new body of work. They are very innovative and they reach interesting policy conclusions. They also have another important aspect in common: Each might be summarized by, "And now, a few kind words about CAFE." Other than that, they are about as different as two papers about CAFE can be.

Klier and Linn

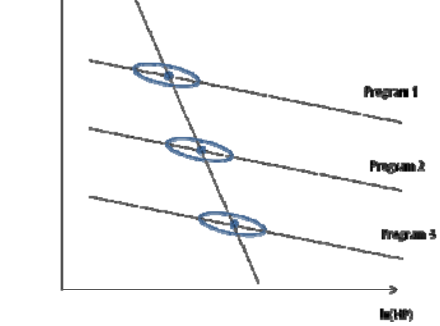
This paper makes at least two significant contributions, I think. First, the authors consider, in greater detail than has been previously the case, the technologies available to improve vehicle fuel economy. To do so, they must have employed an army of research assistants to plow through the old issues of Ward's Automotive Yearbooks, and because of data limitations they are able to analyze cars, but their analysis of these data is interesting and provocative. As a result, they are able to introduce a new time horizon into the discussion of vehicle fuel-use policies. In addition to the short term, when vehicle characteristics must be taken as given, and the long term, when new technologies can increase the fuel economy of vehicles without sacrificing other characteristics, they define a "medium term," in which manufacturers can improve fuel economy by changing other vehicle characteristics, in particular by shedding weight or reducing power. Their second contribution is a novel estimation strategy to correct for the problem endogeneity in estimating consumer WTP for various vehicle characteristics. They make use of their detailed engine database to construct new instruments that are highly correlated with those observed characteristics yet uncorrelated with the unobserved characteristics. But rather than go on about how much I like this paper, I think it will be more useful if I raise a couple of questions.

First, in their data engines are classified in a hierarchy. At the top level there are a number of engine platforms, and within each platform are several programs. But the distinction between program and platform—what defines this nesting structure—is never given. We know it has something to do with the level of technology, but it's not clear what. In particular, what is the relationship, if any, between the hierarchical level and the lead time required by manufacturers to make engine changes? We are told that platform changes are typically not made more than once a decade. Could they be made more often if necessary? Ten years is also time for improvements in technology to be introduced, so would it be correct to say that platform changes overlap between the medium and long term? The authors don't really say; they just offer a static example. (i.e. the Passat and Audi A4 have the same engine program, and the Jetta has a different program, but all three belong to the same engine platform). I'm not sure how important the distinction is for the paper; the authors don't really make much of it right away, they just talk about engines. However, the distinction pops up again in the discussion of the elasticity of substitution between fuel economy and horsepower. They compare models with fixed effects for platform and program and find very different elasticities: -0.06 for fixed program effects and -0.15 for fixed platform effects. They refer to these as "the upper and lower bounds on the medium-term effect of weight and power on fuel efficiency." elasticity of substitution. But isn't this a standard fixed-effects issue, as illustrated in the figure below? When you estimate with platform fixed effects you are failing to observe a lot of heterogeneity across programs, it seems to me. Rather than a contrast between upper and lower bounds, isn't this just a contrast between two estimating strategies, one of which is more appropriate than the other?

Let me make two other small points:

First, throughout the paper the authors use the term "fuel efficiency" to refer to the outcome of interest, vehicle fuel use per mile. This is what other authors have called "vehicle fuel economy." I believe engineers use "fuel efficiency" to refer to something else: the amount of useful energy you're able to get out of the process, relative to the amount you put in. If you get more energy out of the fuel in the tank, you have a choice on how to use it: drive further on the same amount of fuel, increase power, towing capacity, better climate control, etc. The authors say on p. 12 that the "main period" during which fuel efficiency increased was 1975-1984. I think an engineer might say it differently. Definitely, fuel economy improved, with about half that improvement occurring before 1980. But fuel efficiency increased very little in that period before 1980; there just wasn't time. And in the period after 1985, fuel efficiency improved tremendously, but fuel economy improved very little. In any case, the distinction between fuel efficiency, a technical concept, and fuel economy, a techno-economic concept, is a useful one.

And finally, the authors show pretty convincingly, I think, in the medium term the costs of improving CAFE are lower than the short-term estimates. However, the new cost estimates apply to other policies for improving fuel economy as well, so it's not clear that the relative cost effectiveness of CAFE has improved.



Fischer

Where the Klier-Linn paper is mostly empirical, Carolyn's paper is mostly theoretical, with a policy simulation at the end to examine magnitudes. And where Klier and Linn look closely at CAFE costs over time and develop a novel estimation strategy, the main subject in the Fischer paper is market structure. The world auto industry consists of a small number of very large firms, with the predictable effects of oligopoly on prices. As far as I know, Carolyn is the first to point out that oligopolistic distortions can also affect the demand for fuel economy, or indeed any other vehicle attribute. This distortion works to cause manufacturers to oversupply fuel economy for those vehicles appealing to buyers who are likely to value it most, and to undersupply it in vehicle and to undersupply it in vehicles whose customers value it least. The effect is to make fuel taxes, as well as other quasi-market-based instruments involving fleet averaging, less attractive than they might be if market distortions were not considered.

This analysis makes perfect sense to me, and about all I can do at this point is provide a little perspective. I wasn't sure how large the fuel economy distortion is relative to the fundamental oligopolistic distortion in vehicle prices and whether the former affects the latter. Perhaps the fuel economy distortion in simpler models has been picked up and attributed to price distortion. One would think there is a limit to the total amount of distortion in prices, and it must be distributed among all the discriminatory elements available to the manufacturer.

I'd also like to put in a good word for fuel taxes, which Carolyn finds that gasoline taxes exacerbate "both the incentives for price discrimination and the undervaluation of fuel economy." Raising fuel prices will certainly increase consumers' actual valuation of fuel economy, but what I guess she means is that it will increase the proper valuation of it even more. In any case we should remember that CAFE standards only provide fuel conservation incentives in the new vehicle purchase decision. Fuel taxes, by contrast, provide conservation incentives in use: they encourage consumers to drive less; they provide greater incentives for proper vehicle maintenance; they provide incentives for faster fleet turnover (whereas raising CAFE standards probably retards fleet turnover ; and above all, raising fuel prices applies these incentives of every vehicle in the fleet.

Let me conclude by commending the authors of both papers for some really creative, interesting and useful work.

Market Mechanisms and Incentives: Applications to Environmental Policy

A Workshop sponsored by U.S. Environmental Protection Agency's National Center for Environmental Economics (NCEE) and National Center for Environmental Research (NCER)

Resources for the Future
1616 P Street, NW, Washington, DC 20036
(202) 328-5000

Wednesday, April 29, 2009

Disclaimer

U.S. Environmental Protection Agency (EPA) National Center for Environmental Economics (NCEE) and
National Center for Environmental Research (NCER)

Market Mechanisms and Incentives: Applications to Environmental Policy

Resources for the Future
1616 P Street, NW, Washington, DC 20036
(202) 328-5000

Wednesday, April 29, 2009

8:30 a.m. – 9:15 a.m. **Registration**

9:15 a.m. – 9:30 a.m. **Introductory Remarks:** *Julie Hewitt, Chief, Economic and Environmental Assessment Branch, Office of Water*

9:30 a.m. – 11:20 a.m. **Session I: Fuel Economy and Gasoline Prices**
Session Moderator: *Cynthia Morgan, EPA, National Center for Environmental Economics*

9:30 a.m. – 10:00 a.m. Imperfect Competition, Consumer Behavior, and the Provision of Fuel Efficiency in Vehicles
Carolyn Fischer, Resources for the Future

10:00 a.m. – 10:30 a.m. New Vehicle Characteristics and the Cost of the Corporate Average Fuel Economy Standard
Thomas Klier, Federal Reserve Bank of Chicago, and Joshua Linn, University of Illinois at Chicago

10:30 a.m. – 10:40 a.m. Discussant: *Gloria Helfand, University of Michigan, and EPA, Office of Transportation and Air Quality*

10:40 a.m. – 10:50 a.m. Discussant: *Winston Harrington, Resources for the Future*

10:50 a.m. – 11:20 a.m. Questions and Discussion

11:20 a.m. – 12:30 p.m. **Lunch** (On Your Own)

12:30 p.m. – 1:30 p.m. Panel Discussion: Role of Market Mechanisms and Incentives to Climate Change
Moderator: *Dick Morgenstern, Resources for the Future*
Panelist: *Joe Aldy, Special Assistant to the President for Energy and the Environment*
David McIntosh, Senior Counsel in the Office of Congressional and Intergovernmental Relations
Brian Murray, Duke University

1:30 p.m. – 1:40 p.m. **Break**

1:40 p.m. – 3:30 p.m. **Session II: Applications of Environmental Trading Programs**
Session Moderator: *Will Wheeler, EPA, National Center for Environmental Economics*

1:40 p.m. – 2:00 p.m. An Experimental Analysis of Compliance in Dynamic Emissions Markets: Theory and Experimental Design
John Stranlund, University of Massachusetts - Amherst, James Murphy, University of Alaska – Anchorage, and John Spraggon, University of Massachusetts - Amherst

	2:00 p.m. – 2:30 p.m.	Can Markets for Development Rights Improve Land Use and Environmental Outcomes? *Virginia McConnell, Elena Safirova, Margaret Walls, and Nick Magliocca, Resources for the Future*
	2:30 p.m. – 2:35 p.m.	Discussant: *Heather Klemick, EPA, National Center for Environmental Economics*
	2:35 p.m. – 3:05 p.m.	Preliminary Findings and Observations on Ohio's Great Miami River Water Quality Credit Trading Program *Richard Woodward, Texas A&M University*
	3:05 p.m. – 3:10 p.m.	Discussant: *Hale Thurston, EPA, National Risk Management Research Laboratory*
	3:10 p.m. – 3:30 p.m.	Questions and Discussion
3:30 p.m. – 3:40 p.m.	**Break**	
3:40 p.m. – 5:30 p.m.	**Session III: Winners and Losers in Cap and Trade** Session Moderator: *Charles Griffiths, EPA, National Center for Environmental Economics*	
	3:40 p.m. – 4:10 p.m.	Paving the Way for Climate Policy: Compensation for Electricity Consumers and Producers Under a CO_2 Cap and Trade Policy *Karen Palmer, Dallas Burtraw, and Anthony Paul, Resources for the Future*
	4:10 p.m. – 4:40 p.m.	When Does Cap-and-Trade Increase Regulated Firms' Profits? *Dave Evans, EPA, National Center for Environmental Economics; Ian Lange, University of Stirling; and Joshua Linn, University of Illinois at Chicago*
	4:40 p.m. – 4:50 p.m.	Discussant: *Ann Wolverton, EPA, National Center for Environmental Economics*
	4:50 p.m. – 5:00 p.m.	Discussant: *Terry Dinan, Congressional Budget Office*
	5:00 p.m. – 5:30 p.m.	Questions and Discussion
5:30 p.m.	**Adjournment**	

May, 2009

An Experimental Analysis of Compliance in Dynamic Emissions Markets: Some Preliminary Results

John K. Stranlund*
Department of Resource Economics
University of Massachusetts-Amherst

James J. Murphy
Department of Economics
University of Alaska-Anchorage

John M. Spraggon
Department of Resource Economics
University of Massachusetts-Amherst

This paper contains preliminary results from a work in progress. Please do not quote or cite this paper without permission from the authors.

Abstract: Whether pollution sources should be allowed to bank transferable emissions permits and what restrictions should be placed on this activity are fundamental design choices for market-based pollution control. Recent theoretical work examines compliance incentives and enforcement strategies for trading programs with banking provisions. This paper reports preliminary results from economic experiments to test hypotheses from this theoretical work. The experiments are designed to address questions about how to construct enforcement strategies to motivate truthful emission reporting and permit compliance in a dynamic trading environment. We also investigate the consequences of inadequate enforcement on compliance and banking behavior.

Keywords: Compliance, Enforcement, Emissions trading, Laboratory experiments, Permit markets, Permit banking

JEL Codes: C91, L51, Q58

Acknowledgements: The research is funded by the U.S. EPA – Science to Achieve Results (STAR) Program. Daniel Peterson and Ben Laine provided outstanding research assistance, Glenn Caffery and Emily Caffery programmed the software for this project, and Karen Pelto provided valuable administrative support.

---

* Correspondence to John K. Stranlund, Department of Resource Economics, 214 Stockbridge Hall, University of Massachusetts-Amherst, Amherst, MA 01003, USA. Phone: (413)545-6328, Fax: (413)545-5853, E-mail: stranlund@resecon.umass.edu.

An Experimental Analysis of Compliance in Dynamic Emissions Markets: Some Preliminary Results

1. Introduction

Whether pollution sources should be allowed to bank transferable emissions permits, and what restrictions should be placed on this activity are fundamental design choices for emission trading policies. We report on the initial stages of efforts to design and conduct economic experiments to test hypotheses related to enforcement and compliance in emission trading programs with banking provisions. Several existing trading programs allow some form of permit banking.[1] Perhaps the most well known is the SO_2 Allowance Trading program, which allows unrestricted banking of permits, but not borrowing. Pollution sources have made good use of this feature. For example, for the 2006 compliance year banked allowances accounted for nearly 40% of available allowances. In addition, there have been only a few incidences of noncompliance since the inception of the program. By the measure of maintaining compliance, the enforcement apparatus of the SO_2 program has been very successful (US EPA 2007).

Policy analysts usually point to two features of the SO_2 program to explain its success in maintaining near-perfect compliance. First, if a source fails to hold sufficient allowances to cover its emissions in a compliance period it is automatically assessed a financial penalty at a rate which has usually been many times higher than going allowance prices.[2] Second, all facilities in the program are required to install a continuous emissions monitoring system or an approved alternative. These systems are fully automated, thereby minimizing the opportunities for submitting false emissions data. Two additional program features have received less attention, but are likely to be at least partially responsible for the high rate of compliance. Misreporting of emissions is a separate violation under the SO_2 program that is distinct from the failure to hold sufficient permits. The Clean Air Act authorizes the use of criminal sanctions for false or misleading reporting. Each source must identify a single individual who bears the responsibility of submitting truthful reports, and who faces criminal liability for misreporting (Tietenberg

[1] Including the EPA's SO_2 Allowance Trading program, its NO_x Budget Trading program, and the EU's Emissions Trading Scheme. Many proposals for carbon trading schemes include banking provisions (e.g., Stavins 2008).
[2] The penalty for excess emissions was set at $2,000 per ton in 1990, and is adjusted for inflation every year. In the 2005 compliance year the penalty was $3,152 (US EPA 2007). The penalty has always been many times higher than going allowance prices, except for a brief period near the end of 2005 when the allowance price spiked to about $1,600 (US EPA 2006).

2006). Finally, in addition to facing a high financial penalty, a firm's excess emissions in one period are offset by a one-to-one deduction from its allocation of permits in the next period.

Recent theoretical work by Stranlund *et al.* (2005) provides results about the relative contributions of these enforcement characteristics to maintaining compliance in trading programs with banking.[3] This work is motivated by emission trading programs that include permit banking provisions when regulators cannot rely on the perfect (or near-perfect) emissions monitoring provided by continuous emissions monitoring systems. Stranlund *et al.* highlight the importance of detecting and punishing under-reported emissions in these trading programs. Self-reporting of emissions is a necessary enforcement component when permits can be banked and monitoring is imperfect. If a firm is not monitored in a particular period its emissions report is the only piece of information available to a regulator to determine how many permits the firm is using for current compliance purposes and how many are carried into the future. Moreover, misreporting and the failure to hold sufficient permits must be distinct violations. This is so because a firm that holds enough permits to cover its emissions in a period may be motivated to under-report its emissions to increase the size of its permit bank.[4]

In addition, Stranlund *et al.* suggest that a high unit penalty for permit violations is not warranted. When permit borrowing is not allowed, a permit violation penalty has only limited deterrence value; in particular, increasing this penalty does not reduce the amount of monitoring necessary to maintain compliance. In contrast, a penalty for under-reported emissions allows regulators to maintain compliance with imperfect monitoring, and setting this penalty as high as is practicable conserves monitoring costs. This suggests that it is possible to maintain compliance in a dynamic emissions trading program with imperfect monitoring and low permit violation penalties, but doing so requires focusing on punishing reporting violations rather than on punishing permit violations.

---

[3] The theoretical literature on compliance and enforcement in emissions trading is extensive (including contributions by Keeler 1991, Malik 1990, 1992, and 2002, vanEgteren and Weber 1996, Stranlund and Dhanda 1999, Stranlund and Chavez 2000, Chavez and Stranlund 2003, Stranlund 2007). However, only Innes (2003) and Stranlund *et al.* (2005) allow for noncompliance in models with bankable permits. Innes argues that giving sources the ability to bank and borrow permits eliminates the need to impose costly sanctions to maintain compliance in these programs. He does not, however, examine the design of monitoring and punishment strategies that is the focus of Stranlund *et al.*

[4] Requiring self-reporting and making misreporting a distinct violation is different from most of the literature on self-reporting in law enforcement. Malik (1993), Kaplow and Shavell (1994), Livernois and McKenna (1999), and Innes (1999, 2000, and 2001) all assume that self-reporting is a voluntary activity that can be encouraged by offering a lower penalty for self-reported violations. In fact, self-reporting is not necessary to achieve compliance in most of the models in this literature.

While there is a substantial body of economic theory about compliance and enforcement in emissions trading programs, and readily available information about how existing emissions trading programs are enforced, to our knowledge there are no empirical analyses of the determinants of compliance decisions in emissions trading programs with real world data. To fill this empirical gap, some authors have turned to data generated by economic experiments. Cason and Gangadharan (2006) conducted emissions trading experiments with stochastic emissions, bankable permits, and noncompliant subjects. They identified interesting interactions among random emissions shocks, permit banking, and compliance, but they did not attempt to draw conclusions about the appropriate design of enforcement strategies. Murphy and Stranlund (2006 and 2007) examined compliance behavior under several combinations of monitoring and penalties, but their design was based on a static model of emissions trading that did not allow for permit banking.[5]

In this paper we report preliminary results from experiments designed to address questions about how to construct enforcement strategies for emission trading programs that feature imperfect emissions monitoring and bankable emissions permits. In particularly we test the hypothesis that it is possible to motivate truthful emissions reporting and permit compliance with imperfect monitoring and permit violation penalties that are lower than predicted permit prices. Our preliminary results suggest qualified support for this hypothesis. Moreover, it appears that the main consequence of weak enforcement is to increase aggregate emissions through reporting violations rather than permit violations.

2. A Sketch of the Theory of Compliance in Dynamic Emissions Trading Programs

In this section we provide a brief sketch of the theory of compliance in a dynamic emission trading program developed by Stranlund et al. (2005). While they allowed for the possibility that firms could borrow against future permit allocations, we do not allow borrowing in our initial experimental design. Therefore, in this section we focus on compliance incentives in programs that allow permit banking but not borrowing.

Consider a risk-neutral firm in a dynamic emissions trading program that lasts T periods. Let x_t be the number of emissions permits the firm holds at the beginning of period t. Each

---

[5] Muller and Mestelman (1998) review a number of emission trading experiments that include banking provisions. None of them deal with the problem of noncompliance.

permit allows the release of one unit of emissions. During t the firm chooses how many permits l_t to purchase ($l_t > 0$) or sell ($l_t < 0$). Permits trade in period t at a competitive price p_t. A system is in place to track emissions permits so that at any point in time the regulator has perfect information about the number of permits held by each firm. During period t the firm also chooses its emissions e_t. The firm has an abatement cost function, $c(e_t)$, which is strictly decreasing and convex and does not vary over the life of the program.

The firm's emissions are unknown to the regulator unless it conducts an audit of the firm. Because a trading program with bankable permits but imperfect emissions monitoring must include a self-reporting provision, the firm is required to submit a report, r_t, of its emissions in t. The firm can commit two types of violations. A *reporting violation* occurs in period t if the firm under-reports it emissions; that is, $e_t > r_t$. A *permit violation* occurs when the firm does not hold enough permits to cover its emissions; that is, $e_t > (x_t + l_t)$. Because borrowing against future permit allocations is not allowed, a permit violation can occur in any period.

The firm's emissions report is also its report of its compliance status and whether it is banking permits. If $r_t > (x_t + l_t)$, then the firm is reporting a permit violation. If $r_t \leq (x_t + l_t)$, then the firm is reporting that it is permit compliant and banking permits if $r_t < (x_t + l_t)$. Of course, we must distinguish reported permit violations, permit compliance, and permit banking from their actual values. If actual emissions exceed permit holdings, $e_t > (x_t + l_t)$, then there is an actual permit violation. If $e_t < (x_t + l_t)$ the firm has excess permits to bank, and if $e_t \leq (x_t + l_t)$, then the firm is permit compliant.

Monitoring for compliance by authorities is potentially imperfect in the sense that the probability that the authority is able to make a determination of a firm's compliance status is $\pi_t \in [0,1]$. Permit violations in period t (whether they are revealed in a firm's emissions report or discovered by the authorities) are penalized at ϕ_t per unit. This penalty corresponds to the permit violation penalty in the SO_2 program. Reporting violations that are discovered through an audit are penalized at γ_t per unit. Both ϕ_t and γ_t are constants known by all parties.

Bringing the enforcement features together, the expected penalty for a firm that is violating its permits and under-reporting its emissions is

$$f(e_t, l_t, r_t) = \phi_t(r_t - (x_t + l_t)) + \pi_t \{\gamma_t(e_t - r_t) + \phi_t[(e_t - (x_t + l_t)) - (r_t - (x_t + l_t))]\}$$

$$= \phi_t(r_t - (x_t + l_t)) + \pi_t(\gamma_t + \phi_t)(e_t - r_t). \quad [1]$$

To understand how $f(e_t, l_t, r_t)$ is constructed, note that a firm that reports a part of its permit violation faces an automatic penalty of $\phi_t(r_t - (x_t + l_t))$. If the firm is audited so that its reporting violation is discovered (this occurs with probability π_t) the penalty for this violation, $\gamma_t(e_t - r_t)$, is assessed. Of course, if a firm does not hold enough permits to cover its emissions and also under-reports its emissions, it has not reported the full extent of its permit violation. If this is discovered the firm is liable for its unreported permit violation, $\phi_t(e_t - r_t)$.

Combining the elements defined thus far yields the firm's single-period expected costs,

$$v(e_t, l_t, r_t, x_t) = c(e_t) + p_t l_t + \phi_t(r_t - (x_t + l_t)) + \pi_t(\gamma_t + \phi_t)(e_t - r_t). \quad [2]$$

Let us now characterize the evolution of the firm's stock of permits. Let the firm's expectation in period *t* of the number of permits it starts the next period with be $E_t(x_{t+1})$. The subscript on the expectation operator indicates that the expectation is from the perspective of period *t*. From this perspective, x_{t+1} is potentially a random variable because of incomplete monitoring and the possibility of under-reporting of emissions in *t*. $E_t(x_{t+1})$ is also determined by rules about the rate at which permits trade across time, and possible offset penalties for permit violations. In our experiments we focus on the simple case of programs that allow banking permits across time on a one-to-one basis, and permit violations in a period are offset by a one-to-one reduction in next period's endowment of permits. Assuming that the firm receives a predetermined endowment of permits in $t + 1$ of $\overline{l}_{t+1}$,

$$E_t(x_{t+1}) = \overline{l}_{t+1} + [\pi_t(x_t + l_t - e_t) + (1 - \pi_t)(x_t + l_t - r_t)]. \quad [3]$$

Note that if an audit is conducted in *t* the firm's actual permit shortfall, $(x_t + l_t) - e_t < 0$, or bank, $(x_t + l_t) - e_t > 0$, is carried into the next period. If an audit is not conducted the firm's reported permit shortfall, $(x_t + l_t) - r_t < 0$, or reported bank, $(x_t + l_t) - r_t > 0$, is carried forward.

We can now specify a firm's decision problem in its entirety. Its objective is to choose a time path of emissions, permit transactions, and emissions reports to minimize its discounted sum of expected costs, subject to [3], and non-negativity constraints for emissions, reported emissions, and permit holdings in every time period. In the final period the firm will never find it advantageous to hold excess permits or report that it holds excess permits, because excess permits at the end of T have no value. Therefore, we impose the constraint that $e_T - (x_T + l_T) \geq 0$. Formally, the firm's problem is to choose $\{e_t, l_t, r_t\}$, $t = 0, \ldots, T$, to solve:

$$\min E\left[\sum_{t=0}^{T} \beta^t \left(c(e_t) + p_t l_t + \phi_t (r_t - (x_t + l_t)) + \pi_t (\gamma_t + \phi_t)(e_t - r_t)\right)\right]$$

$$s.t. \quad x_{t+1} = \begin{cases} \overline{l}_{t+1} + (x_t + l_t - e_t) \text{ with probability } \pi_t \\ \overline{l}_{t+1} + (x_t + l_t - r_t) \text{ with probability } 1-\pi_t, \quad t = 0,1,...,T-1, \end{cases}$$

$$e_t \geq 0,\ r_t \geq 0,\ x_t + l_t \geq 0,\ t = 0, 1, \ldots, T,$$

$$e_T - (x_T + l_T) \geq 0,$$

$$x_0 = \overline{l}_0. \qquad [4]$$

In the objective function β is the discount factor, which is assumed to be constant over the life of the program.

The objective function and constraints specified in [4] define a discrete-time stochastic dynamic programming problem. The uncertainty in our problem stems from incomplete monitoring: there are no other stochastic elements in the problem. In particular, we assume that each firm can accurately forecast equilibrium permit prices over the life of the policy. Define $J_t(x_t)$ as minimum expected discounted costs from period t on through the last period, given that the firm has x_t permits at the beginning of t. Stranlund *et al.* show that the stochastic dynamic programming equation associated with [4] in periods $t = 0,...,T-1$ is

$$J_t(x_t) = \min_{e_t, l_t, r_t} c(e_t) + p_t l_t + \phi_t (r_t - (x_t + l_t)) + \pi_t (\gamma_t + \phi_t)(e_t - r_t) + \beta E_t[J_{t+1}(x_{t+1})]$$

$$s.t. \quad E_t[J_{t+1}(x_{t+1})] = -p_{t+1}[\pi_t (x_t + l_t - e_t) + (1-\pi_t)(x_t + l_t - r_t)] + \tilde{C}_{t+1}$$

$$e_t \geq 0,\ r_t \geq 0,\ x_t + l_t \geq 0,$$

for some constant $\tilde{C}_{t+1}$. The dynamic programming equation balances the effects of decisions e_t, l_t, and r_t on period t expected costs against the effects of these decisions through the state equation [3] on minimum discounted expected costs from $t + 1$ on through the last period. In the last period, there are no future costs to consider so the dynamic programming equation is

$$J_T(x_T) = \min_{e_T, l_T, r_T} \quad c(e_T) + p_T l_T + \phi_T(r_T - (x_T + l_T)) + \pi_T(\gamma_T + \phi_T)(e_T - r_T)$$

$$s.t. \quad e_T \geq 0, r_T \geq 0, x_T + l_T \geq 0, \text{ and } e_T - (x_T + l_T) \geq 0.$$

This model provides the compliance incentives of the firm. Let us first deal with the firm's reporting incentives. A firm that is violating its permits truthfully reports its emissions in any $t = 0,...,T-1$ if $\pi_t(\gamma_t + \phi_t) \geq (1-\pi_t)\beta p_{t+1} + \phi_t$. To interpret this condition, note that there are two reasons a firm may choose to under-report its emissions. One is to cover up a permit violation while the other is to carry additional permits into the next period. Thus, $\pi_t(\gamma_t + \phi_t)$, is the expected marginal penalty for a reporting violation and the undisclosed part of a permit violation. The expected marginal benefit of under-reporting emissions is $(1-\pi_t)\beta p_{t+1} + \phi_t$, in which $(1-\pi_t)\beta p_{t+1}$ is the expected discounted marginal benefit of carrying additional permits into the next period because emission are under-reported, and ϕ_t is the certain unit penalty for the part of the permit violation that the firm avoids by under-reporting its emissions.

On the other hand, if the firm is permit compliant, and perhaps has a positive permit bank, its only incentive to under-report its emissions is to increase the size of its permit bank. Thus, a permit compliant firm provides a truthful emission report in any $t = 0,...,T-1$ if and only if $\pi_t\gamma_t \geq (1-\pi_t)\beta p_{t+1}$.

Given some γ_t, ϕ_t and βp_{t+1}, it is straightforward to demonstrate that the monitoring required to induce truthful emissions reporting is higher when the firm is permit compliant than when it is violating its permits. However, a regulator does not know if a firm is complying with its permits unless it audits the firm. Therefore, it cannot choose a different monitoring strategy for permit compliant firms than for firms that are violating their permits. This suggests that inducing truthful reporting by all firms requires $\pi_t(\gamma_t + \phi_t) \geq (1-\pi_t)\beta p_{t+1} + \phi_t$, which further suggests a monitoring probability of

$$\pi_t \geq (\phi_t + \beta p_{t+1})/(\gamma_t + \phi_t + \beta p_{t+1}) \qquad [5]$$

in all periods $t = 0,...,T-1$. In the last period, a firm is motivated to under-report its emissions only to cover up a permit violation. Thus, under-reporting in the last period is deterred by monitoring so that

$$\pi_T \geq \phi_T/(\gamma_T + \phi_T). \qquad [6]$$

Now let us turn to a firm's decision to hold permits, given that it has the proper incentive to truthfully report its emissions. In $t = 0,...,T-1$, the firm holds enough permits to cover its emissions (and perhaps banks permits) if

$$\phi_t \geq p_t - \beta p_{t+1}. \qquad [7]$$

In the last period the firm is permit compliant if

$$\phi_T \geq p_T. \qquad [8]$$

Under the assumption that permit borrowing is not allowed, previous work has shown that intertemporal equilibrium in a permit market under certainty requires that real permit prices be non-increasing across time periods, and that firms will bank permits only when real permit prices are expected to remain constant (Cronshaw and Kruse 1996; Rubin 1996, Kling and Rubin 1997).[6] In all periods but the last, [7] indicates that the permit violation penalty serves no

[6] Whether a firm is permit compliant or not, $p_t < \beta p_{t+1}$ implies that all firms would demand an unbounded number of permits in *t*, because they are more valuable in the future. Since this cannot be true in equilibrium, we must have $p_t \geq \beta p_{t+1}$. If $p_t > \beta p_{t+1}$ firms will not hold excess permits because they are less valuable in the future. Thus, firms bank permits only if $p_t = \beta p_{t+1}$. When $p_t > \beta p_{t+1}$ the only reason to hold permits is for compliance purposes.

deterrence role when real permit prices are expected to remain constant across periods. When real permit prices are falling and firms are not banking permits, the permit violation penalty needs to only make up the difference between real prices across time periods. Finally, in the last period the permit violation penalty only needs to cover the equilibrium permit price.

Examining the role of the permit violation penalty in determining the monitoring requirements [5] and [6] suggests that increasing this penalty cannot conserve monitoring costs. In fact, [5] and [6] are both increasing in ϕ_t, $t = 0,...,T$. In contrast, increasing the reporting violation penalty γ_t, $t = 0,...,T$, reduces the monitoring required to induce truthful emissions reporting. From these results Stranlund *et al.* (2005) conclude that enforcement of an emissions trading program with a banking provision and imperfect monitoring should focus primarily on inducing truthful emissions reporting and punishing misreporting with as high a penalty as is practicable to conserve monitoring effort. A testable hypothesis that emerges from Stranlund *et al.* is that using an enforcement strategy that satisfies [5] through [8] should achieve compliance in an emissions trading program with bankable permits despite imperfect monitoring and a permit violation penalty that is lower than predicted permit prices. Our experiments were designed primarily to test this hypothesis.

3. Experimental Design

Our experimental design extends the static model of permit market enforcement used in Murphy and Stranlund (2006 and 2007) to the case in which permits can be banked into the future. Subjects were given a predetermined endowment of production permits at the beginning of each market period. The permit endowment for all periods was known at the beginning of the session. During each market period subjects simultaneously chose to produce units of a fictitious good and trade in a continuous double auction for permits that conveyed the right to produce. The experiments were framed as a production decision in which permits were a license to produce, rather than an emissions decision, to avoid introducing potential biases due to individual attitudes about the environment or emissions trading.[7]

Subjects received a benefit from their choice of production according to the "Earnings from Production" schedules shown below in Table 1. Each experiment consisted of two subjects of each type A through D for a total of eight subjects in each group. These marginal benefit

[7] The experiment instructions are available upon request.

schedules are similar to those used by Cason and Gangadharan (2006). From these data the aggregate marginal benefit function in a period is presented in Figure 1. Individual and aggregate marginal benefit functions were stationary across time periods and stages in an experiment.

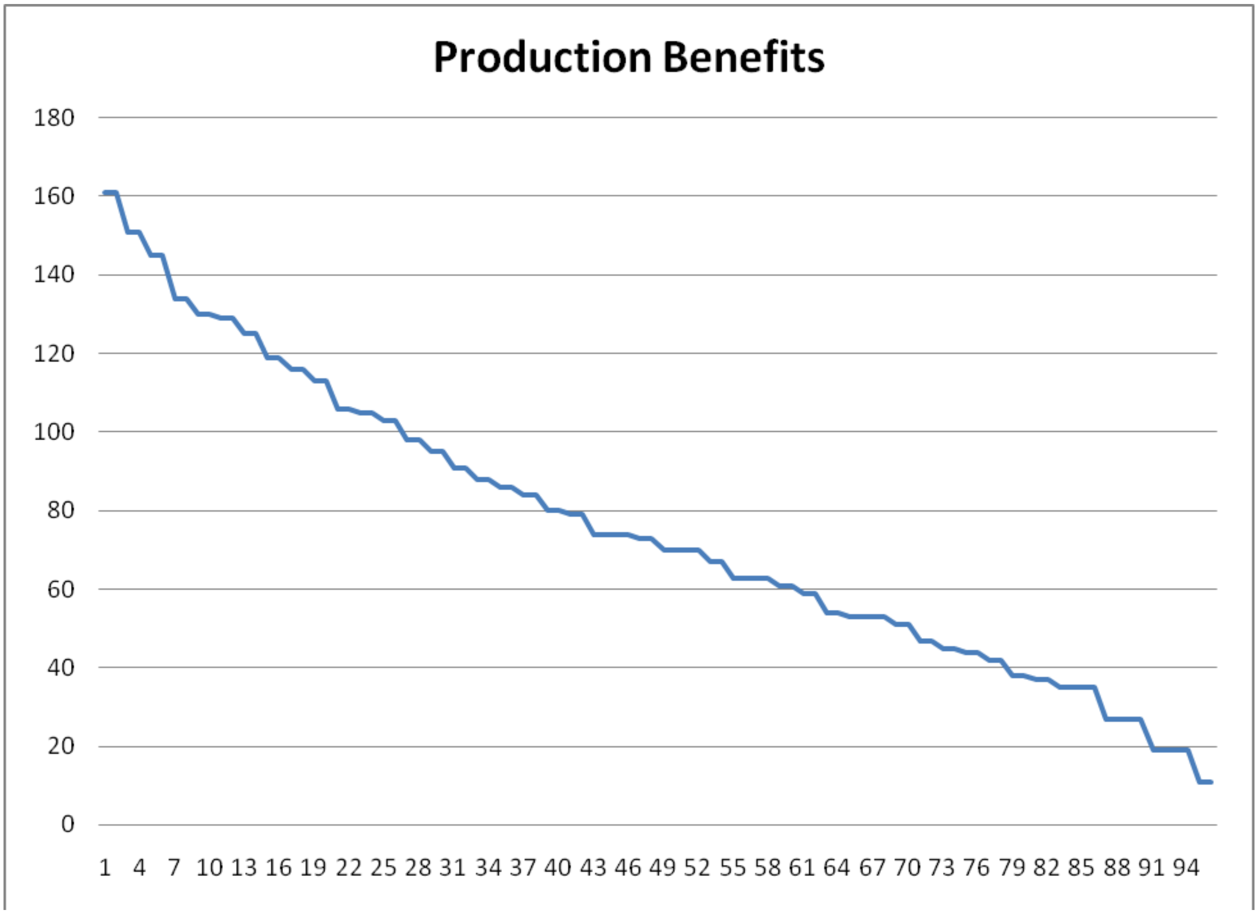
Table 1: Earnings from Production

	Subject Type			
Production	**A**	**B**	**C**	**D**
1	161	151	129	125
2	145	134	113	105
3	130	119	98	88
4	116	106	84	74
5	103	95	73	63
6	91	86	63	54
7	80	79	53	47
8	70	74	44	42
9	61	70	35	38
10	53	67	27	35
11	45	59	19	27
12	37	51	11	19

At the end of each market period each subject was required to submit a report of his or her production for the period. After all reports were submitted, subjects' production choices were audited with a known probability. Permit shortfalls that were either self-reported or uncovered by an audit were punished with a constant unit financial penalty. In addition, a subject's permit shortfall in any period but the last was offset by a one-for-one reduction in the subject's endowment of permits in the next period. If a subject's reported or discovered permit shortfall ever exceeded his endowment in the next period, the subject was declared bankrupt and was not allowed to participate in the remainder of the stage. Finally, if an audit revealed that a subject had under-reported his or her production, he or she was assessed an underreporting penalty that is different from the permit shortfall penalty.

With inter-temporal experiments such as these, learning in earlier periods can affect outcomes in the remainder of a session. To minimize learning effects, sessions were organized into three separate stages each of which contained six market periods, similar to the design used by Anderson and Sutinen (2006) in their study of trading in fishery quota markets. This provided subjects with an opportunity to gain experience in a short six-period setting, adapt, and try again.

Figure 1: Aggregate Marginal Benefit of Production.



Since marginal benefits were stationary for all six periods of a session and production was non-stochastic, we induced permit banking by reducing the aggregate supply of permits for the last three trading periods of each stage. In particular, we distributed a total of 68 production permits for each of the first three periods of a stage, and a total of 16 for each of the last three periods. The perfect foresight and perfect compliance equilibrium consists of 42 units of total production and an equilibrium permit price of about $79 in each period.

Treatments

Our initial experimental design consists of three treatments. In all treatments, individuals were allowed to bank permits and were motivated to do so because of the reduction in the aggregate supply of permits in the middle of each stage. Individual permit endowments were allocated in the following way. In the first three periods, five permits were allocated to each of subject types A and B and twelve permits were allocated to each of subject types C and D. In the last three periods, subject types A and B received 1 permit and subject types C and D received three permits. These permit endowments were the same for each treatment. The three treatments differ according to enforcement aspects.

Full Compliance: This treatment was designed to induce full compliance using an enforcement strategy derived from equations [5] through [8] and the distribution of production permits described above. We decided to over-enforce somewhat so we assumed a monitoring probability of $\pi_t = 0.7$, $t = 1, 2, ..., 6,$ and assessed permit and reporting violation penalties that are about \$20 higher than they need to be to satisfy [5] through [8]. Since our perfect foresight and perfect compliance equilibrium produces the permit price $p_t = \$79$ for each $t = 1, 2, ..., 6,$ [7] and [8] suggest that the permit violation penalty can be set to zero for periods 1 through 5 and \$79 for period six. Therefore, we chose $\phi_t = \$20$ for t = 1, …, 5 and $\phi_6 = \$100$.

Next we chose the reporting penalty. For $t = 1, 2, ..., 5,$ [5] implies that γ_t needs to be set so that $\gamma_t \geq (1 - \pi_t)(\phi_t + p_{t+1}) / \pi_t$ to induce truthful reporting. With $\pi_t = 0.7$, $\phi_t = \$20$, and $p_{t+1} = \$79,$ $\gamma_t \geq \$42.43$. Setting this penalty about \$20 higher led us to choose $\gamma_t = \$60$, $t = 1, 2, ..., 5.$ We set $\gamma_6 = \$60$ in the last period as well, because from [6], $\gamma_6 = \$60$ and $\phi_6 = \$100$ implies that a monitoring probability of at least 0.625 should be sufficient to induce truthful reporting in the last period. Therefore, $\gamma_6 = \$60$ with $\pi_6 = 0.7$.

This treatment was designed to test the hypothesis that compliance can be maintained in dynamic trading programs with imperfect monitoring and low permit violation penalties. Note that we set the permit violation penalty at about ¼ of the predicted permit price for all periods but the last. Note as well that the theoretical model allows us to set the reporting violation penalty below the predicted permit price.

Weak Enforcement: This treatment is the same as the Forced Compliance treatment except that the monitoring probability was reduced by half to $\pi_t = 0.35,\ t = 1,...,6.$ In this treatment we expected significant noncompliance. Weak enforcement reduces the aggregate demand for permits in every period so we expected that permit prices would be significantly lower than in the Forced Compliance and Full Compliance treatments. However, we still expected individuals to bank permits to smooth the effects of the decrease in the aggregate supply of permits.

Forced Compliance: In this treatment, we did not allow subjects to violate their permits or to submit false production reports. This treatment is a baseline treatment against which we will judge the effects of allowing noncompliance.

Subject Recruitment

Participants were recruited from the student population at the University of Massachusetts, Amherst. Subjects were told that to be eligible they had to participate in four two-hour sessions (two days a week for two consecutive weeks). Subjects were paid $5 for agreeing to participate and showing up on time for the first session, and were then given an opportunity to earn additional money in each experiment.

Experiment procedures

Table 2 summarizes the key aspects of the experiments. The experiments were conducted in a computer lab at the University of Massachusetts-Amherst using software designed in Visual Basic specifically for this research. To familiarize subjects with the experiments the first of the four sessions were for training purposes. During the training session, subjects first reviewed the instructions which included interactive questions to ensure that they understood the instructions before proceeding. Subjects then participated in a two-stage experiment. Stage 1 of the trainer followed the same rules as the Forced Compliance treatment (but with different parameters), and Stage 2 followed the rules of the Full Compliance treatment (again with different parameters). The data from the training sessions were not included in the analysis. Prior to the start of the real data sessions, subjects read a summary of the instructions. To mitigate possible order and experience effects, we varied the order of the treatments as shown in Table 3. A total of 72 subjects participated in nine real data sessions, with eight subjects per group.

Table 2. Experiment Summary

- Subjects
  - All subjects participated in a 2-hour training session prior to participating in real data sessions.
  - 72 University of Massachusetts-Amherst students participated in real data sessions.
  - Paid $5 for participating, plus experiment earnings (mean about $23 per session).
- Number and Type of Subjects
  - 8 subjects, 2 of each of four types described in Table 1
- Sessions
  - 3 Stages of 6 four-minute periods during which subjects produced a fictitious good and traded production permits for the right to produce the good.
- Production
  - Production generates "Earnings from Production".
  - Production allowed only during first three minutes of the period.
  - Each unit produced sequentially; production takes 10 seconds per unit.
  - Each subject could produce a maximum of 12 units.
- Permit Market
  - Permit market open for entire four-minute period.
  - Permits traded in a continuous double auction.
  - Permits could be banked for future periods.
- Reporting (for the Full Compliance and Weak Enforcement treatments)
  - After the market closed subjects were asked to report their production.
  - If they reported more production than they had permits they were charged a permit shortfall penalty for each unit they produced without a permit. Also, reported excess emissions were subtracted from next period's endowment.
- Auditing (for the Full Compliance and Weak Enforcement treatments)
  - Within a session, each individual faced the same probability of being audited.
  - Random audits occurred after production and market trading period was over.
  - Permit shortfall penalty applied if an audit revealed production exceeded permit holdings. The permit shortfall penalty is a constant penalty per unit of permit violation. In addition, uncovered excess emissions were subtracted from next period's endowment.
  - Reporting penalty applied if an audit revealed production report less than actual production. The reporting penalty is a constant penalty per unit of underreported production.
  - Subjects could become bankrupt if they required more permits to cover current production than they would be allocated in the next period. Bankrupt subjects were excluded from the remainder of the stage.

Table 3. Sequence of Treatments

Cohort	Session 1	Session 2	Session 3	Session 4
1	Trainer	Weak Enforcement	Forced Compliance	Full Compliance
2	Trainer	Forced Compliance	Full Compliance	Weak Enforcement
3	Trainer	Full Compliance	Weak Enforcement	Forced Compliance

Because individual risk attitudes may have played an important role in the outcomes of our experiments, we conducted a Holt-Laury risk test (Holt and Laury 2002) with all subjects after the training session was completed. The analysis of individual compliance behavior, including how results from the Holt-Laury test are correlated with individual choices, will be conducted at a later date.

Each experiment consisted of three identical six-period stages. At the start of each period, the eight subjects were each given an initial allocation of permits. Each permit conveyed the right to produce one unit of output. During the experiment, subjects earned experimental dollars (E$) that were converted to US dollars at a pre-announced exchange rate.

Subjects produced each unit of the good sequentially by clicking on a button that initiated the production process. Production of a single unit took 10 seconds. After production of the unit was completed the “Earnings from Production” were immediately added to the individual’s cash balance. Subjects were able to “plan” future production within the period by indicating the total number of units to produce. Once production of a unit was completed, if there were any “planned” units, the 10-second production process for the next began automatically. Subjects could increase or decrease their “planned” production, but units that were “in progress” or “completed” were committed and could not be changed. That is, subjects could alter planning decisions about units not yet produced, but they could not reverse production of a good after the 10-second production process begins.

During a market period, and concurrent with the production decision, subjects could alter their permit holdings by trading in a continuous double auction (CDA). In the CDA, individuals could submit bids to buy or asks to sell a single permit (provided that they had a permit available to sell). The highest bid and lowest ask price were displayed on the subjects’ computer screens. A trade occurred whenever a buyer accepted the current ask or a seller accepted the current bid.

After each trade, the current bid and ask were cleared and the market opened for a new set of bids and asks. The trading price history was displayed on the subjects' screens.

Each period lasted a total of four minutes. The permit market was open for the entire period, but production had to be completed in three minutes. The three-minute production time was more than sufficient for a subject to produce up to his or her capacity constraint. We provided the additional minute of permit trading after production was completed to give subjects a final opportunity to adjust their permit holdings. The computer screen displayed the time remaining for both the production and the permit markets.

Since it was possible for individuals to lose money either through permit trading or penalties, we implemented two bankruptcy rules. If an individual's cash balance ever fell below negative E$800, he or she was declared bankrupt and no longer allowed to participate. A subject was also bankrupt if the permit offset penalties for reported and uncovered permit violations exceeded his or her permit allocation in the next period. We also instituted a price ceiling of E$400 above which offers to trade permits were not allowed. This ceiling set well above the highest possible marginal benefit from production so there was little chance that it would ever be reached.

To help subjects understand how to optimally allocate their permits across time periods, in all treatments we explained in the instructions that the best way to use their permits was to spread them out evenly over time. In addition to this we also provided an 'Even use' calculation that indicated how an individual would "even out" the number of permits he or she owned plus all their future allocations. This is consistent with the planner used by Godby et al. (1997). Providing this information to the subjects did not reveal their optimal paths of production and permit demands. It suggested to them that these should be spread evenly through time, but it did not tell them what their aggregate production and permit demands should be. In general these are functions of their Earnings from Production, the enforcement parameters, and the behavior of the others in their session.

As soon as a market period ended in the Full Compliance and Weak Enforcement treatments, each subject was required to submit a report of its production for the period. Subjects could not over-report their production. Subjects were told the probability of random audits prior to each market period. Recall that the Full Compliance and Weak Enforcement treatments are the same except the probability of an audit after each period was 0.7 in the Full Compliance treatment and 0.35 in the Weak Enforcement treatment. After audits were conducted for a period, all penalties were assessed. At the end of each period, subjects were given their production earnings, receipts or expenditures from permit trades, and any penalties for the period, along with a running tally of their cumulative cash balance.

4. Preliminary results

In this section, we present some preliminary results from the experiments. At this date, the results have not been subjected to rigorous statistical tests; therefore, they should not be taken as conclusive. Moreover, the results presented are highly aggregated; we will examine individual compliance and banking behavior in the future.

We begin by noting that subjects made good use of the banking provision in all treatments. Figures 2a, 2b, and 2c contain average production and permit banks across the nine groups in stage 3 of each treatment. We present only the third stage results to conserve space and because the subjects' understanding of each treatment was at its highest in the third stage. The blue bars in the figures represent the supply of permits in each period, 68 in the first three periods and 16 in the last three. The green line indicates average production and the yellow bars represent the average size of the permit bank at the start of periods 2 through 6. Recall that the perfect foresight/perfect compliance equilibrium results in total production of 42 units in each period. This path of production and the distribution of permits reveal that the equilibrium path of banked permits at the start of periods 2 through 6 is 26, 52, 78, 52, and 26.

Although these results may mask important differences across the nine groups in each treatment, it is clear that on average the groups smoothed production across the six periods by building up a permit bank in the first three periods when the supply of permits was high, and then drawing the bank down in the last three periods when the permit supply was low. Note that it is hard to detect differences in average group production and permit banks between the Full and Forced Compliance treatments. This is to be expected because subjects were predicted to be fully compliant in the Full Compliance treatment. Note also that the mean production and bank values tended to be close to the perfect foresight/perfect compliance equilibrium values. Production tended to be higher in the Weak Enforcement treatment. This is a consequence of significant noncompliance in this treatment, which we will turn to momentarily.

Figure 2a: Third stage production and permit bank, Full Compliance

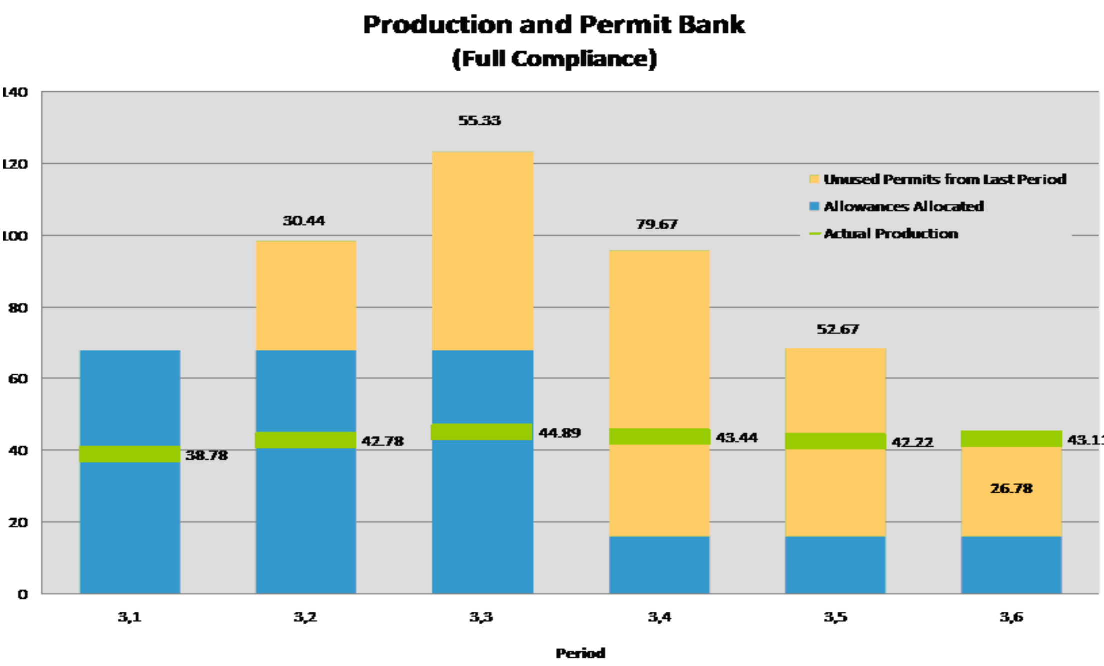


Figure 2b: Third stage production and permit bank, Forced Compliance

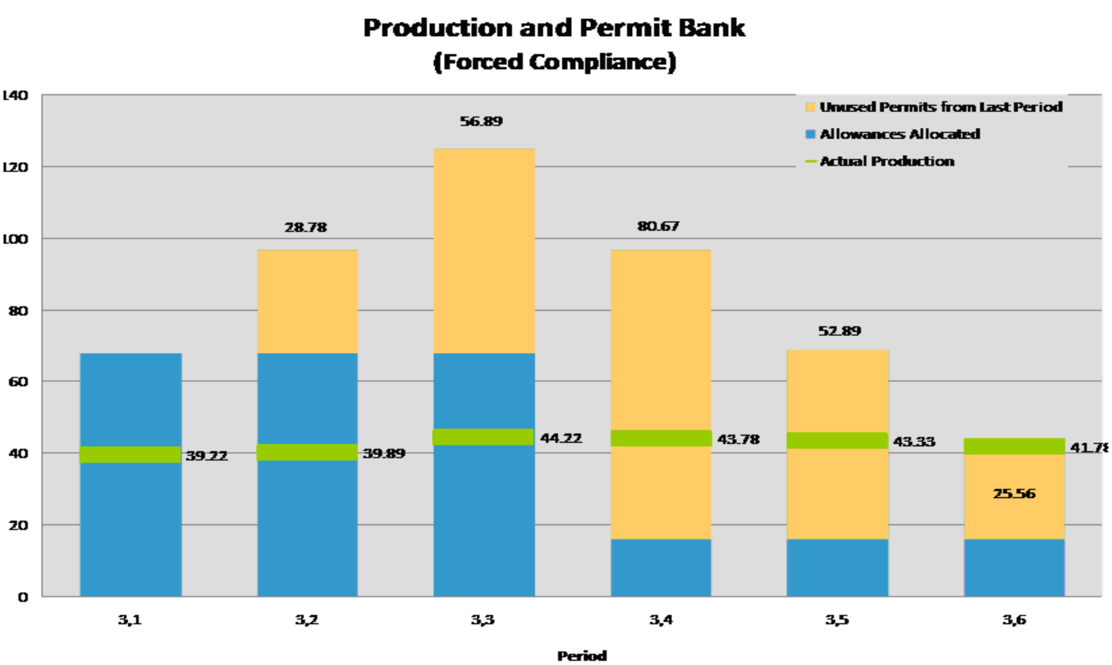
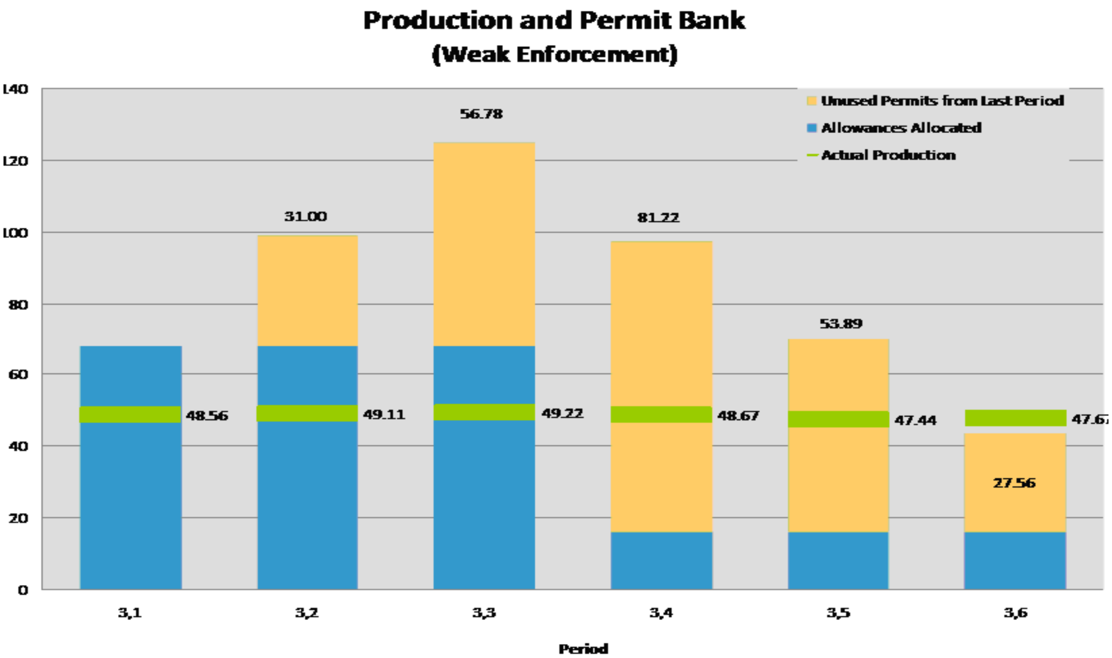


Figure 2c: Third stage production and permit bank, Weak Enforcement



Now let us turn to the main hypothesis of this study, that subjects are compliant under the Full Compliance treatment despite imperfect monitoring and a low permit violation penalty. Tables 2a and 2b are the frequency distributions of third stage permit violations and reporting violations for the Full Compliance treatment. Although the subjects were not perfectly compliant as theory predicts, the enforcement strategy of this treatment did achieve very high rates of compliance. Note that subjects complied with their permits 96.5% of the time and that two-thirds of the violations were one unit violations. Thus, most of the violations in the third stage of this treatment are more likely to be the result of subjects probing the extent to which they can violate their permits and get away with it, rather than systematic noncompliance choices. There are more reporting violations than permit violations, but again most of these are one unit violations.

It is likely that 3.5% permit violation rate and the 7.2% reporting violation rate are statistically significant. Thus, strictly speaking our hypothesis is rejected. However, it appears that an enforcement strategy for an emissions trading program with permit banking can achieve \high rates of compliance despite imperfect monitoring, low permit violation penalties, and modest reporting violation penalties.

Table 2a: Third stage permit violations, Full Compliance

Permit Violations	Frequency	Percent
0	417	96.53
1	10	2.31
2	4	0.93
9	1	0.23
Total	432	100.00

Table 2b: Third stage reporting violations, Full Compliance

Reporting Violations	Frequency	Percent
0	401	92.82
1	19	4.4
2	7	1.62
4	2	0.46
> 4	3	0.69
Total	432	100.00

As expected there was significant noncompliance in the Weak Enforcement treatment. Tables 3a and 3b are the frequency distributions of third stage permit violations and reporting violations for this treatment. Although the permit compliance rate is lower than in the Full Compliance treatment it is still close to 90%. Moreover, about half of the permit violations were one unit violations. It may seem surprising that there are not more permit violations in this treatment, but this is consistent with the significant amount of banking we observe in Figure 2c. One cannot simultaneously bank permits and violate them. Table 3b reveals that the main effect of weak enforcement is on production reporting. About 32.3% of reporting choices were violations, and half of these were 3 units or greater.

Table 3a: Third stage permit violations, Weak Enforcement

Permit Violations	Frequency	Percent
0	384	88.89
1	24	5.56
2	7	1.62
3	8	1.85
> 3	9	2.08
Total	432	100.00

Table 3b: Third stage reporting violations, Weak Enforcement

Reporting Violations	Frequency	Percent
0	293	67.82
1	35	8.1
2	34	7.87
3	27	6.25
4	17	3.94
5	5	1.16
>5	21	4.86
Total	432	100.00

Moreover, the main effect of reporting violations that are undetected is to allow total production to exceed the number of permits supplied. Table 4 contains total production and undetected reporting violations for the six periods in the third stage of the Full Compliance and Weak Enforcement treatments averaged across groups. Recall that 252 permits were supplied in each stage, 68 permits for the first three rounds and 16 for the last three. The percentages in Table 4 are with respect to the 252 permit cap. The data for the Full Compliance treatment reflect the high rate of compliance we've already observed. However, under the Weak Enforcement treatment total production tended to be about 38.67 units (15.34%) above the supply of permits. This matches the average number of undetected reporting violations. Undetected and unreported permit violations averaged about 14 units in this stage of the Weak Enforcement treatment. Thus, about 36% of undetected reporting violations were successful attempts to cover up permit violations while the remaining 64% were successful attempts to simply increase individual permit banks.

Table 4: Third stage total production and undetected reporting violations

Treatment	Total production	Undetected reporting violations
Full Compliance	255.22 (1.01%)	2.89 (1.15%)
Weak Enforcement	290.67 (15.34%)	38.22 (15.17%)

The higher total production in the Weak Enforcement treatment is reflected in average permit prices. In the third stage of this treatment the average permit price was $70.12. Recall that the predicted price is $79. Average prices in the third stage of the Full Compliance and Forced Compliance treatments were $75.30 and $77.68, respectively.

We also calculated the percentage of realized potential gain over the no-trade, no-banking, and no-violations outcome (autarchy) for each treatment. In the third stage of the Forced Compliance treatment subjects earned about 95% of the potential gain. This suggests that the subjects understood the basic incentives of trading within and across time periods very well; consequently the baseline markets functioned well. The efficiency gain in the third stage of the Full Compliance treatment was significantly lower at 77%. This drop in efficiency is likely due to the fact that the need to make decisions about permit and reporting compliance significantly increased the complexity of the environment. Despite this we think it is important to stress that the subjects were able to capture a large part of potential gains. The efficiency gain in the third stage of the Weak Enforcement treatment was 118%. The subjects in this treatment did very well largely by getting away with reporting violations, which allowed them to increase production beyond the permit cap.

5. Concluding remarks

In this paper we have described a study that uses economic experiments to test hypotheses about compliance behavior and enforcement in emission trading programs with bankable permits. Imperfect monitoring of emissions requires firms to self-report their emissions when permit banking is allowed. Consequently, recent theoretical work highlights the importance of motivating truthful emissions reporting in maintaining compliance in dynamic trading programs with less than perfect monitoring. In fact, very high penalties for permit violations have limited deterrence value in these programs.

Our experiments address two important issues. The first is a test of the theoretical result that an enforcement strategy with imperfect monitoring and a permit violation penalty that is set below predicted permit prices can induce subjects to report truthfully and to remain permit compliant. The second concern is about the effects of weak enforcement on the performance of emission markets with banking. Clearly, it is premature to draw any firm conclusions from our preliminary presentation of the results. However, it appears that we may have qualified support for our main hypothesis. That is, it appears possible to achieve high rates of compliance with imperfect emissions monitoring, a low permit violation penalty, and a modest reporting violation penalty. A weaker enforcement strategy (low monitoring probability in our case) results in aggregate emissions well beyond the permit cap, but this is mainly through reporting violations rather than permit violations. Rigorous analysis of the experimental data will reveal whether these results hold up and will reveal further results about compliance behavior in emissions markets with bankable permits.

References

Anderson Christopher, M. and Jon G. Sutinen. 2006. "The Effect of Initial Lease Periods on Price Discovery in Laboratory Tradable Fishing Allowance Markets." *Journal of Economics Behavior and Organization* 61(2), 164-180.

Cason, Timothy N. and Lata Gangadharan. 2006. "Emissions Variability in Tradable Permit Markets with Imperfect Enforcement and Banking." *Journal of Economic Behavior and Organization* 61(2), 199-216.

Cronshaw, Mark and Jamie Brown Kruse. 1996. "Regulated Firms in Pollution Permit Markets With Banking." *Journal of Regulatory Economics* 9: 179-189.

Godby, Robert W., Stuart, Mestelman, R. Andrew Muller, and J. Douglas Welland. 1997. "Emissions Trading with Shares and Coupons when Control over Discharges is Uncertain." *Journal of Environmental Economics and Management* 32(3), 359-381.

Holt, Charles A. and Susan K. Laury. 2002. "Risk Aversion and Incentive Effects." *American Economic Review* 92(5), 1644-55.

Innes, Robert. 2003. "Stochastic Pollution, Costly Sanctions, and Optimality of Emission Permit Banking." *Journal of Environmental Economics and Management* 45(3): 546-68.

Keeler, Andrew 1991. "Noncompliant Firms in Transferable Discharge Permit Markets: Some Extensions." *Journal of Environmental Economics and Management* 21(2), 180-189.

Kling, Catherine and Jonathan Rubin. 1997. "Bankable Permits for the Control of Environmental Pollution." *Journal of Public Economics* 64: 101-115.

Malik, Arun S. 1990. "Markets for Pollution Control when Firms are Noncompliant." *Journal of Environmental Economics and Management* 18(2), 97-106.

Malik, Arun. S. 1992. "Enforcement Cost and the Choice of Policy Instruments for Controlling Pollution." *Economic Inquiry* 30(4), 714-721.

Muller, R. Andrew and Stuart Mestelman. 1998. "What Have we Learned From Emission Trading Experiments?" *Managerial and Decision Economics* 19, 225-238.

Murphy, James J. and John K. Stranlund. 2006. "Direct and Indirect Effects of Enforcing Emissions Trading Programs: An Experimental Analysis." *Journal of Economic Behavior and Organization* 61(2), 217-233.

Murphy, James J. and John K. Stranlund. 2007. "A Laboratory Investigation of Compliance Behavior under Tradable Emissions Rights: Implications for Targeted Enforcement." *Journal of Environmental Economics and Management* 53(2), 196-212.

Rubin, Jonathan. 1996. "A Model of Intertemporal Emissions Trading, Banking, and Borrowing." *Journal of Environmental Economics and Management* 31: 269-286.

Stavins, Robert N. 2008. "Addressing climate change with a comprehensive US cap-and-trade system." *Oxford Review of Economic Policy* 2008 24(2), 298-321.

Stranlund, John K. and Kathy K. Dhanda. 1999. "Endogenous Monitoring and Enforcement of a Transferable Emissions Permit System." *Journal of Environmental Economics and Management* 38(3), 267-282.

Stranlund, John K. 2007. "The Regulatory Choice of Noncompliance in Emissions Trading Programs." *Environmental and Resource Economics* 38(1), 99-117.

Stranlund, John K., Christopher Costello, and Carlos A. Chavez. 2005. "Enforcing Emissions Trading when Emissions Permits are Bankable." *Journal of Regulatory Economics* 28(2): 181-204.

Stranlund, John K. and Carlos Chavez. 2000. "Effective Enforcement of a Transferable Emissions Permit System with a Self-Reporting Requirement." *Journal of Regulatory Economics* 18(2): 113-131.

Stranlund, John K. and Kanwalroop K. Dhanda. 1999. "Endogenous Monitoring and Enforcement of a Transferable Emissions Permit System." *Journal of Environmental Economics and Management* 38(3): 267-282.

Tietenberg, Tom H., 2006. *Emissions Trading*: *Principles and Practice, 2nd Edition*. Resources for the Future. Washington DC.

U.S. Environmental Protection Agency. 2006. *Acid Rain Program 2005 Compliance Report.* Washington DC: U.S. EPA Acid Rain Program.

U.S. Environmental Protection Agency. 2007. *Acid Rain Program 2006 Compliance Report.* Washington DC: U.S. EPA Acid Rain Program.

vanEgteren, Henry and Marian Weber. 1996. "Marketable Permits, Market Power, and Cheating." *Journal of Environmental Economics and Management* 30: 161-173.

Can Markets for Development Rights Improve Land Use and Environmental Outcomes?

Virginia McConnell, Elena Safirova, Margaret Walls, and Nicholas Magliocca
Resources for the Future

Prepared for
Workshop on Market Mechanisms and Incentives:
Applications to Environmental Policy

Environmental Protection Agency
NCEE and NCER

Resources for the Future
April 29, 2009

Can Markets for Development Rights Improve Land Use and Environmental Outcomes?

I. Introduction

Undeveloped open space can provide public benefits to communities by maintaining wildlife habitat, protecting sensitive ecological resources, preserving prime farmland or historically significant areas, protecting scenic views, and maintaining buffers for key aquatic resources. These social values are not likely to be fully captured by private land owners in the decisions they make, leading open space to be under-provided by private land markets. To provide these public goods, some communities are relying on publicly funded programs such as Purchase of Development Rights (PDRs) to preserve land. In PDR programs, the local government purchases and retires the development rights, placing a conservation easement or other restriction on the property. Other communities are instituting large lot or agricultural zoning restrictions. Still others have required clustering of development, essentially requiring open space be set aside.

Another approach that has received a good deal of attention lately is to use a more private market-based approach known as Transfer of Development Rights (TDRs). TDRs aim to preserve land from development in some areas by transferring that development to another location. Developers purchase rights from landowners, whose land is then placed under easement; the developers are permitted to use the rights to develop alternative properties more densely than would be allowed under baseline zoning regulations. Although a TDR market must be set up by government, the transactions in development rights are made by individual land owners, and do not require the kind of government funding that many other programs require, such as PDRs and other conservation easement purchase programs.

Thus far, studies of TDR markets have relied on case studies of real-world programs, some theoretical economic models, and a very few econometric studies where data is available. But to fully assess land use outcomes and identify the features of the programs that determine those outcomes, a detailed simulation model of the land market is required. This model needs to capture both the spatial heterogeneity and the dynamics of the land and housing markets. Spatial heterogeneity is key as policies such as TDRs work to protect land of particular types in particular locations, while allowing development in other locations. It is essential for a land use model to capture that geographic detail. Incorporating market dynamics is also key. When a TDR or other program is introduced, the time path of future development is likely to be altered. To capture these changes, the model must represent the economic decisions of all of the parties involved: landowners, consumers, and developers.

In this paper, we develop a land use model that includes these features. We build an agent-based simulation model that, on one hand, is based on economic theory and, on the other hand, features the degree of heterogeneity that is adequate for our purposes. This framework captures agent optimizing behavior, the functioning of markets, spatial heterogeneity, and a variety of other key features of real-world land and housing markets. We show how the model can be used to represent the dynamics of land use decisions over time. We describe the individual agents and their optimizing behavior, the market interactions among the agents, and the time path and spatial distribution of development and preservation outcomes that result both with and without a TDR program. We show how the model can be used to compare TDRs to a zoning policy and a PDR program. We also discuss alternative TDR program designs. At this is work in progress, the model has not been fully calibrated; only the farmer/landowner module is running and has been used for simulations. In this paper, we describe the framework of the full model but provide additional specific details only about the farmer/landowner component.

In the next section of the paper, we describe how TDRs work and review the literature on TDRs from both economics and planning. We discuss the design features and land market conditions that are key to determining whether TDR programs are successful in meeting their goals of land preservation. In section III, we describe some specific programs; we spend most of our time on two very active programs in Maryland but also briefly describe some other newer programs. This discussion provides the motivation for the simulation model. Section IV provides an overview of the agent-based modeling approach and briefly describes our model.

II. Transfer of Development Rights Markets

The most commonly used tool in the local planner's toolbox is zoning. Zoning laws set broad categories of uses for areas of land in a jurisdiction, as well as the intensity of use on that land. In the case of residential development, this means there are usually limits set on the density of development, most often expressed as the number of dwelling units permitted per acre. To protect open space, local government may set very restrictive density limits. But because it is impossible for the government to have all the information it needs to set the density limits optimally for every land parcel, it usually sets limits for broad categories covering large land areas. Moreover, these density limits are set at one point in time, and are not easy to change because landowners develop expectations about the value of land based on rules governing how it can be developed. For these reasons, broad zoning categories are a blunt instrument for internalizing the externalities associated with development and for protecting land-based environmental resources and open spaces.

TDR programs work in concert with zoning density limits to protect land in some areas while still allowing development in others. Land targeted for protection is the TDR "sending area." Property owners in this area are the suppliers of development rights to the TDR market. Demanders of development rights are usually developers who can use them in areas known as "receiving areas." The government allows more dense development in the receiving areas than is permitted by baseline zoning. TDR programs are voluntary thus individual parcels may be developed or preserved, depending on parcel characteristics and landowner preferences.

The advantages of TDRs for achieving improved land use allocations are numerous. Because they are voluntary, they give landowners more flexibility compared to strict mandates or broad zoning rules. Early studies argued that there can be substantial efficiency advantages to these types of tradable systems over a zoning-only approach (Carpenter and Hefley, 1982, Mills, 1980). By giving individual landowners the flexibility to go over or under the number of building rights allocated under zoning regulations, while maintaining a cap on the overall number of rights, land parcels can be allocated to their most efficient uses depending on their relative values in preservation and development (Mills, 1989, Thorsnes and Simon, 1999).

On the distributional side, TDRs have the potential to compensate landowners in areas targeted for preservation who may be subject to relatively restrictive zoning, thereby avoiding "takings" arguments by those landowners. Another political advantage is that TDR transfers occur through a private market, and therefore no tax dollars are needed for ensuring that land is preserved. This is in contrast to PDRs which can require many millions of dollars for preserving significant areas from development. And TDRs can achieve this preservation, while still accommodating growth in the region. Thus in comparison with PDRs and many other measures, which are growth controls, TDRs are more of a growth management tool.

Whether or not a TDR program in practice achieves the results found in theory depends heavily on the design of the program and the underlying fundamentals of the housing and land markets. TDR programs have been in existence since 1968 and proliferated in the 1970s and 1980s as concerns over suburban "sprawl" and loss of open space on the urban fringe began to mount (Johnston and Madison, 1997). Currently, about 140 TDR programs exist in the U.S. (Pruetz, 2003; Messer, 2007). However, only a handful of these programs have actual working TDR markets with significant numbers of transactions taking place each year. Many other programs lie dormant. Several authors offer suggestions for why programs fail and identify features that lead to success. Pruetz and Standbridge (2009) catalog these studies and highlight the factors that each study identifies as a key determinant of success. In our previous work, we have emphasized the importance of an adequate demand for density in TDR "receiving areas" (Walls and McConnell, 2007; Kopits, McConnell, and Walls, 2008). However, TDR program design is complex and multi-faceted, and a number of factors can be critical for success. The government is essentially creating a market, which is a difficult undertaking. In the next section, we walk through the key issues in program design.

II.1. What makes effective TDR markets so difficult to design and implement?

II.1.1. Designation of sending and receiving areas. Local government must decide which areas of the community are allowed to sell TDRs, i.e., the designated sending areas, and which are allowed to use TDRs to develop more densely, the receiving areas. Sending areas are usually clear: they are the areas with the resources that the community wants to protect. But designation of receiving areas is often more difficult. A narrow designation, in which receiving areas are limited to higher density areas and town centers, may be intended to achieve some "smart growth" objectives but may severely limit the functioning of the TDR market if demand for additional density does not exist. Moreover, existing residents in those areas may object to higher density.

II.1.2. Baseline zoning density limits. One factor that determines whether demand for additional density exists is the baseline zoning in the receiving area. The more restrictive the zoning, the more likely there will be a residual demand for density and thus TDRs. Density limits in the sending area are also important. The more restrictive the limits in TDR sending areas, i.e., the fewer the houses permitted per acre of land, the less lucrative the development option vis-à-vis preservation and the more likely a landowner will be willingto sell her TDRs. TDRs will only work when there is a supply of and demand for development rights and the baseline density limits in both sending and receving areas are critical determinants of supply and demand..

II.1.3. The density bonus. The density bonus defines how many more dwelling units are allowed per acre in the receiving area relative to the baseline number of units allowed. For example, if the baseline zoning allows 1 dwelling unit per acre but 4 units per acre may be built with TDRs, the density bonus is 300%. In areas where building is constrained by the baseline density limits, a higher density bonus will increase demand for TDRs. However, as discussed above, if there is little underlying demand for additional density, it is possible that a higher density bonus will have no effect.

II.1.4. The TDR allocation rate. The TDR allocation rate is the number of TDRs a landowner in the sending area is permitted to sell and is usually based on acreage. .[1] Often, the allocation of TDRs is also based on the zoning density limits in the sending area. In cases where the sending area has been downzoned prior to, or in conjunction with, establishment of the TDR program, the TDR allocation rate is often related to the original density limits and therefore can provide some measure of compensation to landowners. [2]

II.1.5. Housing and land market conditions. A working TDR market is only established if the underlying housing and land market conditions support additional density in receiving areas and preservation of land as farmland or open space. For receiving areas, it is essential to know the additional value of housing densities beyond the baseline zoned densities. For sending areas, the agricultural value of land vis-à-vis the value in development is likely to be crucial. The issue of uncertainty looms large here, given the time dimensions involved and the irreversible nature of the development decision.

---

[1] This is not always the case. For example, the complex program in the Lake Tahoe region of California and Nevada determines the number of TDRs a property owner may sell based on several property characteristics related to stormwater runoff. See Solimar Research Group (2003) for an analysis of this program. Other TDR programs, such as the one in Chesterfield, New Jersey, allocate the number of rights based on soil quality and slope to better reflect the actual number of housing units that could potentially be built on the land; Malibu, California's program works similarly.
[2] Planners often refer to the TDR "transfer ratio," the ratio of the number of TDRs allocated per acre to the allowable density (dwelling units per acre) in a sending area.

II.1.6. Other potential issues with created markets. Created markets, in general, come with several potential problems. One is market power. In the case of TDR markets, there may be market power on the part of either developers or landowners in the market for development rights. If developers have market power, for example, TDR prices may be inefficiently low and fewer TDR transactions take place than is optimal. Similarly, markets may be thin. Especially in the case of TDRs, this can be a problem, as these programs are established at the local level and some communities may have a limited number of buyers and sellers. Another problem centers around transaction costs. If it is costly to get information about prices or to actually transact with other parties, it may be difficult to make efficient trades (Stavins, 1995).

Given these difficulties, it is no wonder that many TDR programs have not been able to create working markets and protect land from development as expected. In the next section of the paper, we examine two TDR programs in detail and then briefly discuss problems in other programs and new approaches being tried.

III. TDR Programs in Practice

III.1. Calvert County, Maryland

The community with the most flexible and consistently active TDR market is Calvert County, Maryland. Calvert County's TDR program has been in existence since 1978 and since the first TDR sale in 1982, it has had an active private market with many buyers and sellers of development rights engaging in trades each year.

Calvert's program is focused on farmland preservation. The county targets for protection the land in its Farm and Forestry District, the prime agricultural land in the county. It allows TDRs to be transferred to town centers and residential zones and also to a zoning district called the Rural Community District (RCD), which covers about 40% of the land area of the county and which has relatively low baseline density limits. The RCD is also a designated sending area. Thus landowners in the RCD have a great deal of flexibility: they can sell TDRs and preserve their land from development or they can sell to a developer who could either build to the baseline density limits or purchase TDRs and build more densely. This treatment of the RCD zone in Calvert is unique among TDR programs. As we explained above, many communities have limited receiving areas to more dense residential areas and no other program that we know of has an overlap in sending and receiving areas.

In an analysis of the demand for TDRs in Calvert County, Kopits, McConnell, and Walls (2008) argue that designation of the RCD as a receiving zone led to the robust demand for TDRs there. They show that most TDRs used over the 1978-2002 period were in the RCD; nearly 50% of RCD subdivisions used TDRs compared with only 9% of subdivisions in the residential and town center zones.

In a separate study, the same authors analyze development in the county over the 1965-2001 period – a period that covers years with and without the TDR program – and conclude that the 1 house/3 acres and 1 house/5 acres density limits in the RCD constrained development in some subdivisions during the period prior to the TDR program (McConnell, Walls, and Kopits, 2007). They estimate that approximately 46% more houses would have been built in those constrained subdivisions, which would have led to 11% more houses overall in the county. The density limits established in the TDR program, 1 house/2.5 acres in the RCD, seemed to satisfy demand for density during the post-1978 period, according to the authors.[3]

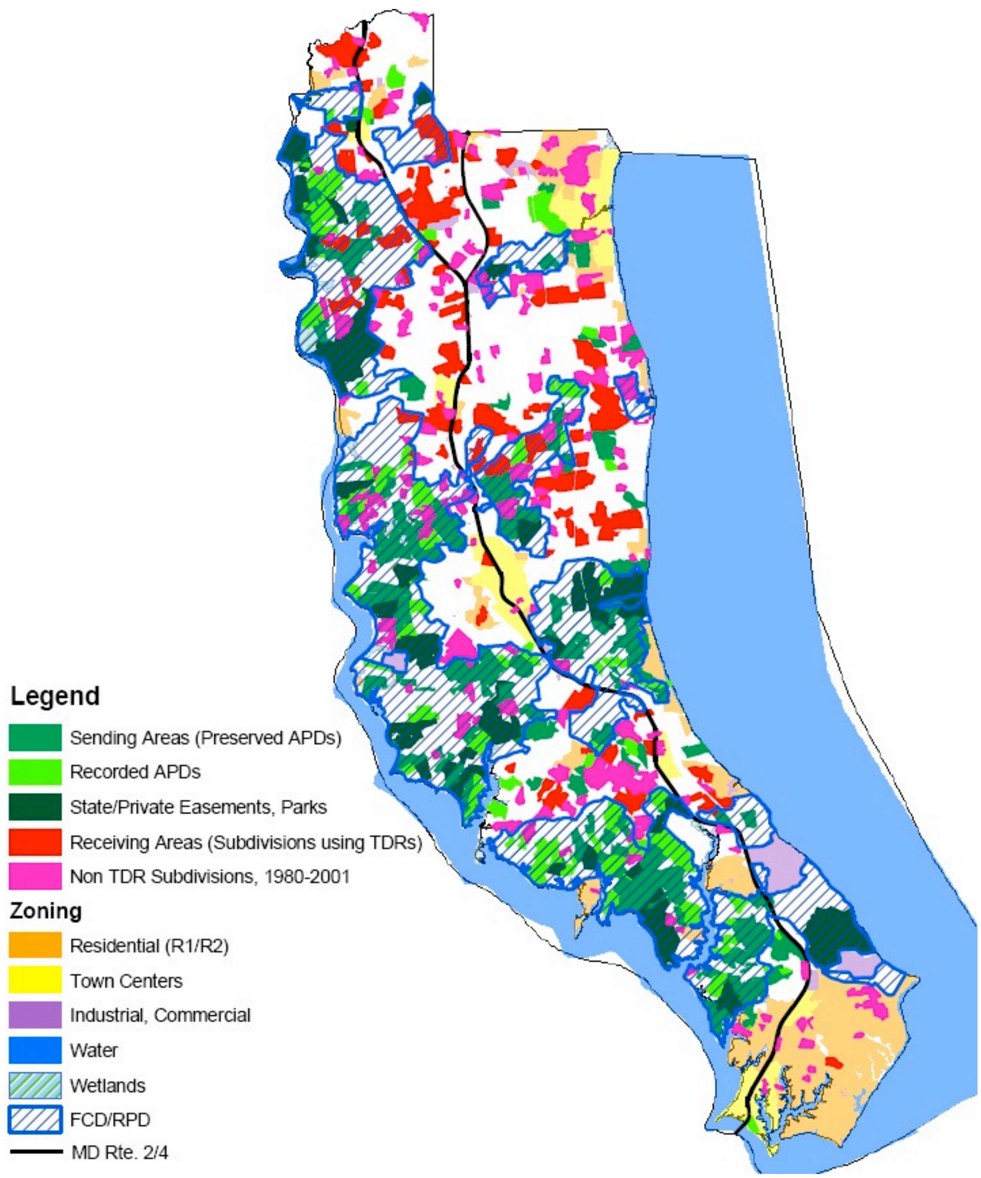
Figure 1 shows the spatial patterns of development and preservation that existed in Calvert County as of 2002. The green areas are protected lands. APDs are Agricultural Preservation Districts, with light green being APDs that are not yet under permanent easement and the medium green being lands under easement;[4] the dark green areas are parks and other protected lands. Pink and red areas are subdivisions; pink are the subdivisions that have been built without TDRs and red are TDR subdivisions. The zoning categories are also shown on the map. Orange and yellow are the residential and town center zones; the cross-hatched blue is the Farm and Forestry District (previously known as the Farm Community and Resource Preservation Districts and thus denoted as FCD/RPD in the figure); and the white area is the RCD.

The map shows that most subdivisions are located in the northern part of the county, closer to the employment centers of Washington and Baltimore. It also shows that most of these subdivisions are in the RCD, the white areas that are rural and can be either sending or receiving areas, as we explained above. The protected lands, shown in green, are primarily in the prime agricultural areas, the blue hatched "FCD/RPD" zones. As of January 2008, the Calvert TDR program had permanently preserved 13,260 acres of land.[5]

While we argued in our earlier work that the Calvert program has had success because of its flexibility and the resultant healthy demand for TDRs, some planners and farmland preservation advocates contend that this flexibility has led to more sprawl than would otherwise have occurred (Dehart and Etgen, 2007, pg. 108). Essentially, they are critical of the fact that there are developments (red areas) going into rural zones (white areas) as shown in Figure 1. In their view, all of the rural zones – both the FCD/RPD and the RCD, (the blue-hatched and white areas of the map) – should have been downzoned

[3] Two county-wide downzonings – one in 1999 and another in 2003 – reduced the RCD density limit to 1 house/10 acres and then 1 house/20 acres, where it currently stands today. However, with TDRs, developers were permitted to build up to 1 house/2.5 acres during the 1978-1998 period; today, they can go to 1 house/4 acres. This current density bonus is 500% -- i.e., developers are permitted to build up to 5 more houses per acre with TDRs than without. See Walls and McConnell (2007) for more details about these downzonings and other features of the program.
[4] A farmer's first step to selling development rights and preserving his land is to have the land designated as an APD. Once it is an APD, development rights can be lifted and sold.
[5] These numbers were provided by Bowen (2008); 5,366 acres have been preserved by the county's PDR program and 25,722 acres in total are under easement from all programs.

Figure 1. Land Use in Calvert County, Maryland



Source: Map created using subdivision, agricultural land, and TDR data from Calvert County Department of Planning and Zoning. See McConnell, Kopits, and Walls (2006).

and protected from development with TDR use permitted only in the town center and residential zones (the yellow and orange areas of the map). This view and the alternative view we have expressed in the publications cited above are clearly at odds. We have argued that an alternative program design in which receiving areas are more limited often dooms TDR programs to failure, because there is often no residual demand for density in those zones. If there is no TDR market created because of insufficient demand, then the goals of preserving farmland in the targeted preservation areas does not occur.

Sorting out these competing views of the Calvert outcome is difficult without knowing the counter-factual – i.e., what would have occurred had there been a different kind of TDR program or no program at all. This is an issue we return to below.

III.2. Montgomery County, Maryland

The alternative program design often held up by planners as a more appropriate model is embodied in the Montgomery County, Maryland, TDR program (Johnston and Madison, 1997). Montgomery County's program, which also targets farmland preservation, began in 1980 with downzoning of a 90,000-acre rural area of the county designated the Rural Density Transfer (RDT) zone. Original limits were 1 house/5 acres; the downzoning restricted development to 1 house/25 acres. To compensate landowners and to avoid a sharp reduction in growth in the county, landowners in the RDT zone were allocated TDRs equivalent to the original 1:5 density limit. These TDRs could be purchased and used by developers in a variety of other areas of the county ranging from very high density urban areas to suburban zones with baseline density limits of 1 house/2 acres.

The Montgomery program has preserved more land, by far, than any other TDR program. As of June 30, 2008, nearly 52,000 acres had been preserved.[6] Most of the TDR transactions took place in the 1980s, a period of high growth in the county. We have argued elsewhere that this growth, due in large part to the county's proximity to Washington, DC, created a strong TDR demand, and the RDT downzoning combined with the generous TDR allocation created a strong supply (McConnell and Walls, 2009). These two key ingredients led to an active TDR market.

Although the farmland preservation and "smart growth" advocates admire many features of the Montgomery program, it is important to first, highlight the lack of flexibility in the program compared with the Calvert program and second, dig deeper to look at the patterns of TDR demand and use in the county.

In contrast to Calvert, where there was no downzoning of the agricultural areas, Montgomery imposed a very restrictive downzoning, leaving landowners in the RDT with essentially only one option: to sell TDRs. Moreover, no other region of the 317,000-acre county was designated as a sending area, only RDT landowners can sell their development rights. And the RDT is a sending area only; not only can it not be a receiving area, the baseline density limits are very restrictive, making development a less attractive option. Essentially, this treatment of the 90,000-acre area is an example of "one-size-fits-all" zoning regulations that we described at the beginning of the paper. While it protects open space, it is likely to do so at high cost because of its lack of flexibility across landowners.

[6] Acreage totals for the TDR program and other farmland preservation programs are available from the county's Department of Economic Development at http://www.montgomerycountymd.gov/content/ded/agservices/pdffiles/pie08.pdf.

By digging deeper into the workings of the program, we have found that despite the county's intent to have many receiving areas designated in high-density zones, most of the TDR use has been in the relatively low-density areas. Figure 2 shows, by zoning category, the maximum number of TDRs allowed by county rules, the maximum allowed by the individual Planning Areas of the county, which set their own, usually more restrictive rules, and the actual number of TDRs used. The figure makes clear that most TDRs have been used in the relatively low-density zone that allows 2 houses/acre, followed by the 1 house/2 acre zones. And even in these areas, the Planning Areas are not allowing as much density as the county would have permitted and developers are using far fewer TDRs than even the Planning Areas designate.

There are two likely factors leading to this outcome: (1) limited demand for density beyond baseline limits, and (2) high transaction costs due to the fact that TDRs do not confer a "by-right" density in receiving areas; in effect, each development must be negotiated individually, including the number of units and the density, through a lengthy review process.

Figure 2. Permitted and Actual TDR Use, by Zoning Category in Montgomery County, Maryland

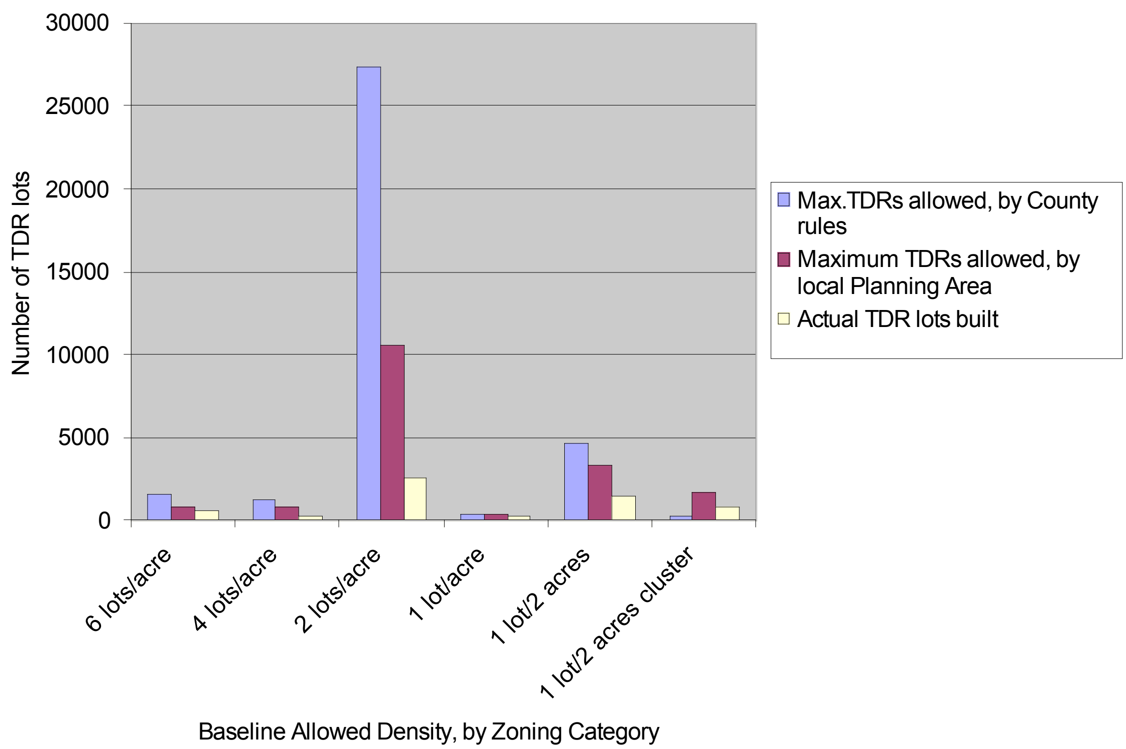
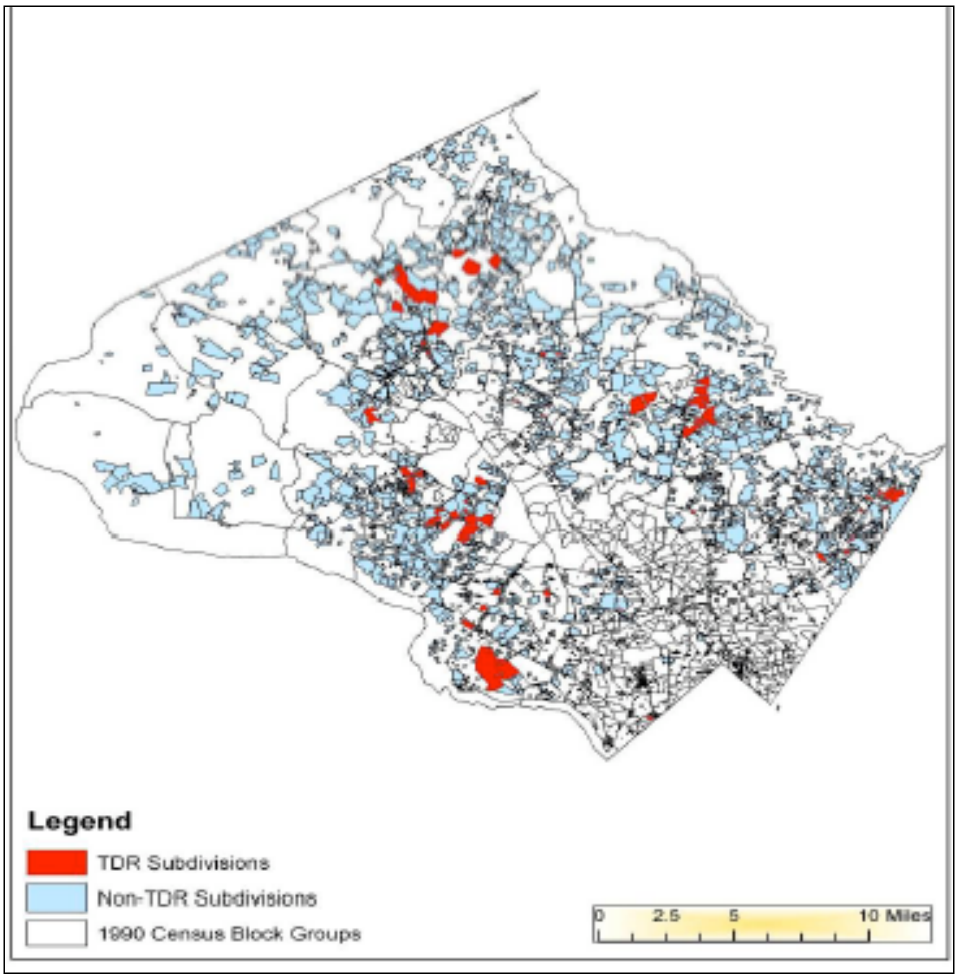


Figure 3 shows a map of the county with Census block groups outlined and all subdivisions built in the county between 1973 and 2004; TDR subdivisions are highlighted in red and non-TDR subdivisions in blue. As we explained in the context of Figure 2 above, each of the Planning Areas can choose where to designate TDR receiving areas and what density to allow with TDRs. This is a partial explanation for the spatial patterns of TDR use that show up on the map; some areas of the county are limiting TDR use more than others. The other explanation is that developers are not using as many TDRs as they could, thus there are many more blue subdivisions than red ones on the map. And finally, it can be seen that TDR subdivisions are not located in the most densely developed area of the county, closer to Washington, DC, which is the lower portion of the map. The area designated for preservation, the RDT, is not highlighted on the map but lies in the north and west portion of the county where it is clear that there is little development.

Figure 3. Development Patterns in Montgomery County, Maryland



Source: Map constructed using subdivisions built (1973-2004) and TDR data (1980-2004) from Montgomery County Parks and Planning Department. See Walls and McConnell (2007) and references therein.

III.3. Other TDR Programs

While Montgomery and Calvert County are among the best-known programs and are by far the most successful in terms of protecting land from development, there are many other programs in existence. By Pruetz's (2003) count, there are approximately 140 communities in the U.S. with some kind of TDR program on the books. However, several programs that we have studied exhibit the problems we discussed above: inactive markets that appear to be a result of very limited demand for TDRs. In some programs, changes in receiving areas, baseline density limits, or other factors that impact the demand side of the market have been implemented and have either dried up demand or increased it, depending on the change. Examples include Queen Anne's and St. Mary's Counties in Maryland and Malibu, California. Newer programs are trying innovative approaches on the demand side. Several communities have discarded the idea of trying to force additional density into established, high-density zones and rather, designated receiving areas in new "greenfield" locations. Chesterfield Township in New Jersey, Collier and Sarasota Counties on the Florida gulf coast, and Larimer County in Colorado are three examples.

Whether these newer programs will have active markets, and what the patterns of development and preservation will look like if they do, is unclear. Moreover, while we have speculated about how land uses have changed in the programs described above, it is difficult to know what the development and preservation outcome would have been had the program not been implemented -- the counter-factual. What would land use look like with either no TDR program or a program with different design features? We would like to be able to analyze the effects of different program parameters and designs to determine which are most important for creating effective markets for land preservation. It is also important to be able to assess trade-offs in program design – such as flexibility in the receiving areas – and land use outcomes.

These questions cannot be addressed without the aid of a spatially explicit model of land use. The model needs to incorporate decision-making on the part of landowners, consumers, and developers. It needs to capture market dynamics, as well as spatial heterogeneity. In the next section, we describe more carefully the features that need to be included and describe the model we are developing. We then describe the simulations we will undertake.

IV. A Land Use Simulation Model

To accomplish our goals – analysis of the spatial and time patterns of land use under alternative policies – we develop a simulation model of land use in a hypothetical community. As with all simulation models, we are facing many trade-offs. The major trade-off, as usual, is between the model's simplicity (that includes analytical simplicity and fast run times) and its ability to adequately represent all (or at least most) features that, based on current knowledge, are important for understanding outcomes of TDR programs. Below we list and briefly explain the features that we believe are critically important for modeling TDR programs in a simulation context.

IV.1. Model Features.

Heterogeneity of agents. The presence of heterogeneous consumers is important for accurate representation of housing demand – in particular the relative demand for housing of different types. Some of the desired dimensions of consumer heterogeneity are income level and preference for lot size and open space. At the same time, farmer/landowners should be modeled as heterogeneous agents with respect to their commitment to farming and initial land endowment.

Heterogeneity of land parcels. Three significant features of land parcels that should be reflected in the model are (i) heterogeneous agricultural productivity, (ii) differing sizes of existing farms, and (iii) spatial location of parcels and therefore, their attractiveness for development.

Heterogeneity of dwelling units. Heterogeneity of dwelling units is critically important for a representation of demand for housing. In particular, in order to effectively represent the interaction between zoning rules, construction of new housing, and the overall housing supply, there should be a quite rich representation in the variety of housing. In particular, individual houses should differ by the house size (in square feet of floor space), lot size, location and possibly other characteristics, such as the presence of nearby open space.

Endogenous prices. In order to evaluate and compare the modeling outcomes, the prices of land, housing units, and TDRs should be determined through market forces. Also, modeling of counterfactuals and alternative policies is meaningful only if the changes can be propagated through the entire system of prices represented in the simulation model.

Spatially explicit modeling. Evaluating spatial outcomes is at the heart of measuring TDR program effectiveness. Thus, it is essential for the modeling framework to distinguish more rural areas from urbanized regions and environmentally sensitive areas, or other target preservation areas, from non-sensitive areas. In addition, spatially explicit modeling is important for correct determination of the location-related price differentials in the housing and land markets.

Dynamic modeling. The process of land development is an intrinsically dynamic process, as is the land preservation decision. Farmer/landowners evaluate their choices each period in the context of expectations for the future and the recognition that both development and preservation are irreversible decisions. In addition, we are interested in the outcomes of land preservation over a reasonably long period of time, on the order of several decades, rather than in a single period.

Explicit representation of housing density, zoning, and other land use regulations. In order to be able to test the effectiveness of land use policies and regulations, they have to be explicitly represented in the model. Moreover, policies such as TDRs work together with zoning regulations, therefore it is critical to represent this real-world setting.

Optimizing behavior of individual agents. The assumption that all individual agents try to achieve the best outcomes for themselves and not just follow some rules of thumb is a standard feature of economic models. Sometimes, due to imperfect information or uncertainty about the future, the individual agents might achieve outcomes that are inferior to the true optima. Nevertheless, it is important to assume that each agent intends to achieve the best possible outcome, conditionally on the limited access to market information.

Ability to evaluate outcomes (welfare analysis). The goal of the TDR programs, first of all, is land preservation. Therefore, it is natural to assume that the amount and the location of preserved land is the desired measure of the program effectiveness. However, in addition to preserving land, a TDR program would alter the construction pattern that would otherwise occur, affect land and housing prices, and ultimately would impact the well-being of all economic agents in the model. For example, one particular design of a TDR program may result in land preservation but a welfare loss, while another program design might preserve as much land with no welfare loss. The welfare calculation adds an important dimension to the program evaluation. Given the complexity of the effects, it is highly desirable to be able to evaluate the outcomes of the program from a broader standpoint. Moreover, it is likely that the presence of a TDR market would have a different impact on different groups of agents (e.g. low income consumers vs. high income consumers). The ability of the modeling framework to produce a welfare analysis is a strong feature that is very helpful in producing quantitative comparisons for policy analysis.

Faced with this long list of the model's features and capabilities that the model should have, we have very limited options. Given the importance of agent heterogeneity and an explicit treatment of space, we cannot employ simple and elegant urban economics models with closed-form solutions. Essentially the only two sensible alternatives are large-scale spatially disaggregated equilibrium models (e.g. Safirova et al. 2006) and agent-based models (ABM). Following the preliminary modeling and analysis, we came to the conclusion that an accurate representation of salient features of TDR markets requires levels of heterogeneity (especially among agents and dwelling types) that cannot be adequately handled in a spatial equilibrium model. In the next sections we describe advantages and drawbacks of the agent-based models and contrast them with those of alternative modeling approaches.

IV.2. Why use ABMs for Land Markets?

Economic models of urban land use and land markets vary in their representations of space, spatial heterogeneity, price expectations, and levels of uncertainty, but they typically fall into the category of either econometric or spatial equilibrium approaches (Irwin, 2009). Econometric models of land use change typically rely on micro-data to estimate models of land conversion and other phenomena and then use the estimated model to simulate outcomes. Examples include Irwin and Bockstael (2002), Irwin et al. (2003), and Towe et al. (2008). These models are based on theoretical economic models in which individuals are assumed to maximize their net expected returns by choosing the most profitable land use (Arnott, 1980; Arnott and Lewis, 1979; Capozza and Helsley, 1990). Intertemporal land use decisions are driven by land owners' forward-looking rent expectations. As outlined by (Irwin, 2009) this method allows for the inclusion of substantial spatial heterogeneity, but data requirements are often restrictive and methodological issues such as spatial dependence, endogenous regressors, spatial instability of parameters, and selection biases can arise. Furthermore, econometric modeling is limited in its ability to represent land use dynamics (Parker et. al., 2003). Econometric parameters ineffectively capture feedbacks that occur over multiple spatial and temporal scales that often drive land use change dynamics (Irwin, 2009).

Spatial equilibrium models are foundational to urban economics (Glaeser, 2008). They are based on the assumption that given a sufficiently long period of time, mobile consumers will locate across a spatially heterogeneous landscape so that utility is equal for all consumers and location disamenities are perfectly offset by prices. The spatial structure of these models is often based in the monocentric city model with travel costs to the central business district accounting for spatial variations (see Bento et al. (2006), for example). Newer versions of these models abandon the monocentricity assumption (Epple and Sieg, 1999; Walsh, 2007). The spatial equilibrium assumption provides analytical tractability for modeling urban land use patterns and captures the capitalization process (Irwin, 2009). However, these models quickly become complex when adding the spatial detail often necessary to analyze land use policies such as TDRs. And the equilibrium requirement does not allow for an explicit representation of land use change dynamics in the model, which makes the integration of economic and ecological models extremely difficult (Irwin et. al., 2007). Furthermore, in reality, cities tend not to be in spatial equilibrium. The importance of out-of-equilibrium market dynamics has gained growing recognition among economists (Arthur, 2006).

Driven by these limitations, the use of ABMs for land markets has grown rapidly (e.g. Arthur, Durlauf, and Lane, 1997; Epstein and Axtell, 1996; Kirman and Vriend, 2001; LeBaron, 2006; Lux, 1998; Tesfatsion, 2006). ABMs offer several advantages over traditional economic models because they can explicitly represent causal drivers of land market dynamics. These advantages are outlined in Parker and Filatova (2008) and we summarize them here:

1) *Heterogeneous goods can be traded by heterogeneous agents.* Unique attributes of land (productivity, neighborhood characteristics, etc.) can be represented explicitly and are selected and traded by heterogeneous agents. For example, farmers selling their land have heterogeneous agricultural returns derived from their land's productivity and operating costs, and they may have different heuristics for decision-making (prediction models and preferences for farming

versus selling). Developers differ in their costs and the housing demands they face based on the zones in which they operate. Consumers differ in preferences and income. Heterogeneity in all of these agents drives their interactions and ABMs allow for that. Transaction prices then emerge dynamically.

2) *Spatial and agent-agent interactions are explicitly represented*. Location-specific land uses and option values affect the use and value of surrounding land through spatial externalities. The dynamics of these uses and values are driven by competition between agents' bid and asking prices and future price expectations. Heterogeneous consumers value living in specific locations differently. ABMs capture these differences by modeling competing consumer bids for a specific location, therefore explicitly modeling the emergence of neighborhood effects. Furthermore, transitional dynamics of that location will be dependent on which consumers value that location more in the future. Thus, feedbacks emerging from consumer competition for spatially unique land uses are preserved locally and dynamically adjusted.

3) *By modeling agent-agent and agent-environment interactions, market non-equilibrium outcomes can be represented.* Land markets are rarely in equilibrium and typically display such complex behaviors as cyclical growth and decline and bubbles (Arthur, 2006). Land markets and associated urban growth patterns tend to be path-dependent and driven in large part by out-of-equilibrium dynamics (Arthur, 2006; Tesfatsion, 2006). ABMs allow these dynamics to emerge rather than exclude them through abstract market adjustments.

ABMs have some drawbacks. Many of the models developed and used to analyze land use dynamics and spatial outcomes, for example, are highly complex and detailed – the UrbanSim model, for example (Waddell, 2002) – making it difficult to understand model results and attribute those results to particular factors. In addition, the models can generate multiple plausible outcomes and developing rules for choosing among them can be difficult and sometimes seem arbitrary. However, land use change and urban growth are spatially complex and path-dependent phenomena; ABM modeling provides the necessary flexibility and richness to address such challenges without the traditional spatial equilibrium and representative agent simplifications needed for analytical tractability.

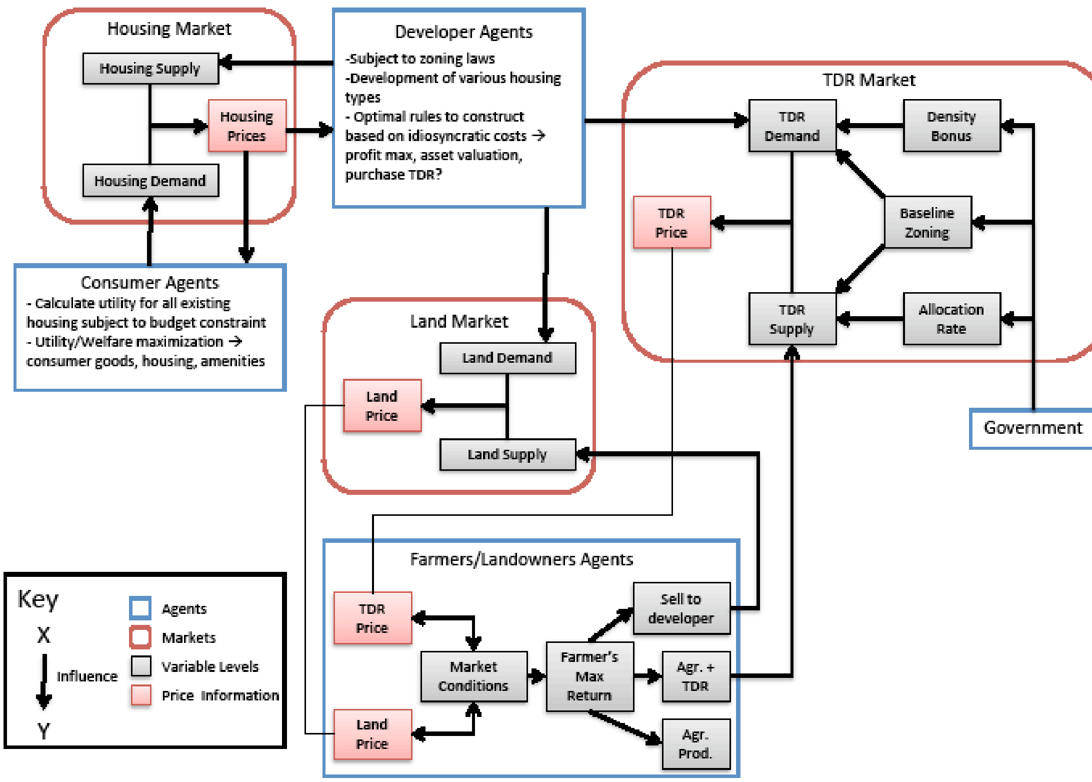
In the next section, we describe our modeling framework and give examples how it can be used to simulate scenarios and policies.

IV.3. Preliminary Description of the Simulation Model

The agent-based simulation model we are currently building will represent three interconnected markets – the land market, the housing market, and the TDR market. While the land market and the housing market are both spatially explicit, the TDR market will encompass the entire modeling area. At present, we have built a first version of the land market and continue to work on the details of modeling the other two markets.

Figure 4 provides a visual description of the model, including the interactions among the markets. In what follows, we briefly describe the agents in the model and their decisions, the role of individual markets, and the interaction between different markets. Then, for illustration purposes, we provide several examples of how scenarios of interest can be modeled in this framework.

Figure 4. Interdependence of Agents and Markets in the Land Use Simulation Model



IV.3.1. Behavior of Agents in the Model

Consumers. Consumer agents in the model are the force that determines the demand for housing. Consumers are endowed with exogenous income and intend to optimize their utility functions that depend on housing characteristics and a non-housing consumption good. The housing characteristics we explicitly incorporate in the model include square footage of the house, lot size, house location, and a general, spatially-differentiated environmental amenity. The housing supply in each period consists of houses, each of which is a bundle of housing characteristics. As a result of bidding for houses, housing is allocated among consumers, and prices of individual transactions are determined endogenously in the model.

Developers. Developers in this model play two roles. On the one hand, they serve as landlords for existing housing and accept or reject consumers' bids. On the other hand, they produce new housing by first purchasing land from farmers and then constructing the new housing. Developers intend to optimize their profits from constructing new housing and select the most profitable types (bundles of characteristics) of housing to satisfy their goal. Their costs of construction are exogenously determined. In making their decisions, developers are faced with zoning density limits that vary spatially. These limits constrain the lot sizes of the houses that the developer can build. If the most profitable lot size is smaller than allowed by the zoning regulations – i.e., density is greater –purchase of TDRs may occur if that is the most profitable option. Thus, the developers engage in transactions with farmers both by purchasing land and TDRs.

Farmer/Landowners. Farmers in the model provide a supply of land for future residential development. In each model period, a farmer can decide to sell his land to the developer, sell TDRs to the developer and continue farming in perpetuity, or not sell land or TDRs but just continue farming until the next period. Farmers are endowed with heterogeneous plots of land that differ from each other by their size, agricultural productivity, and operating costs. Farmers' decisions are described in more detail in the last section of the paper.[7]

The actions of agents who participate in more than one market (developers and farmers) establish the connections between the markets in the model. And consumers, although they ultimately purchase housing in one particular location, bid on houses over the entire region. Those features lead to a tightly integrated market system where each impact in one part of the system is propagated through several channels and affect the rest of the region.

IV.3.2. Characterization of Markets in the Model

Housing Market. The housing market determines the overall level of housing prices in the entire region. If housing supply lags demand, housing prices increase and provide a signal for developers to construct more housing. A similar mechanism is engaged for specific housing characteristics. For example, when there is a relative shortage of houses on small lots, their relative profitability will increase and the developers will be encouraged to build more of such houses.

Land Market. The land market determines the dynamics of land development over time. If a developer escalates his willingness to pay for land due to greater housing demand, the farmers are more likely to sell their land. On the other hand, new housing construction temporarily reduces the pressure on housing prices and reduces the developers' willingness to pay for land. Such mechanisms work in different segments of the land marker and can lead to interactions between the segments. For example, a massive housing construction in one segment will indirectly reduce housing prices and therefore the developers' willingness to pay for land in other markets.

---

[7] At this point, the farmer module has been fully programmed and run under various scenarios, thus more details of this module are available than the others.

TDR Market. The TDR market is closely connected with other markets, and especially the land market. The sharp difference between the TDR market and other markets is that this is the single market for the whole region, since the TDRs as a commodity is not a local good. As such, the presence of the TDR market reinforces the interconnectedness between local land markets, since developers and farmers can buy/sell TDRs anywhere in the region. Note also from Figure 4 the role that government plays. Several zoning and TDR design parameters affect the TDR market.

IV.3.3. Examples of How the Model Can be Used

Although some specific modeling details have yet to crystallize, the connections between the markets in the model allow us to hypothesize how some of the scenarios of interest will be represented in this framework once the full model is complete. Each of the following could be implemented and compared to a baseline run of the model which would include a baseline TDR program of a design we specify. That baseline program will include a large rural sending area targeted to certain areas that have the highest value in preservation, and a receiving area that is also broadly defined to include a fairly large area with the ability to add density to a range of baseline lot sizes. This baseline model would yield a time path for the sale and use of TDRs over time, in addition to the results in the markets for land and housing development over time. Here we describe farmland protection, but TDR programs can apply to protection of other land-based resources.

Elimination of the TDR program. Here we assume that the TDR program was never implemented and no land has been preserved. This is the counter-factual to which the results of any TDR program must be compared. What would have happened if we did not have this program? Elimination of the TDR program will take away the opportunity to preserve land from development, particularly those areas that were designated for preservation under the TDR program. Farmers with the highest agricultural land productivity in these areas would be more likely to enter their land into the pool of potentially developable lands. However, based on their high value in agriculture, those lands are unlikely to be developed early. Instead, the housing construction in urbanized areas will be limited to lower density units since TDRs are unavailable. The overall outcome will depend on housing market conditions and the parameters of the developers' profitability conditions in different regions, but it likely to result in lower density development and less land preserved.

Assessment of the Effects of Changing the Receiving Areas. There is controversy about whether receiving areas should be designated only in high density urbanized areas, and not at all in low-density areas. We discussed this issue above in Section III . If the TDR program were changed to allow TDRs only in areas that already have high baseline densities, markets throughout the urban area could change. The demand for TDRs might fall, because of the reduced options to developers for building with extra density. The model can be used to explore the parameters of housing demand that affect this outcome. In other words, under what conditions would the TDR market continue to function, additional housing be built in these limited areas, and land preserved?

A Decrease in Baseline Zoning Density Limits in Urbanized Areas. A decrease in the density limits in urbanized areas (downzoning) will lead developers to either build housing units on large lots, or to purchase more TDRs to be able to construct higher density housing. Initially, this policy is likely to increase the demand for TDRs and increase land conservation in the sending areas. However, if TDRs cannot be used to reach the original density levels or if TDR costs are too high, there is likely to be pressure to develop in other parts of the region, including more rural areas. In fact, "leapfrog" development is often a criticism of downzoning, urban growth boundaries, and other growth controls (Pendall, 1999). Ultimately, the overall impact of the down-zoning policy in our model will depend on model parameters and we will explore these results. Without computer simulation of a well-parameterized, spatially-explicit, dynamic model, it is impossible to determine the overall regional effects of this change.

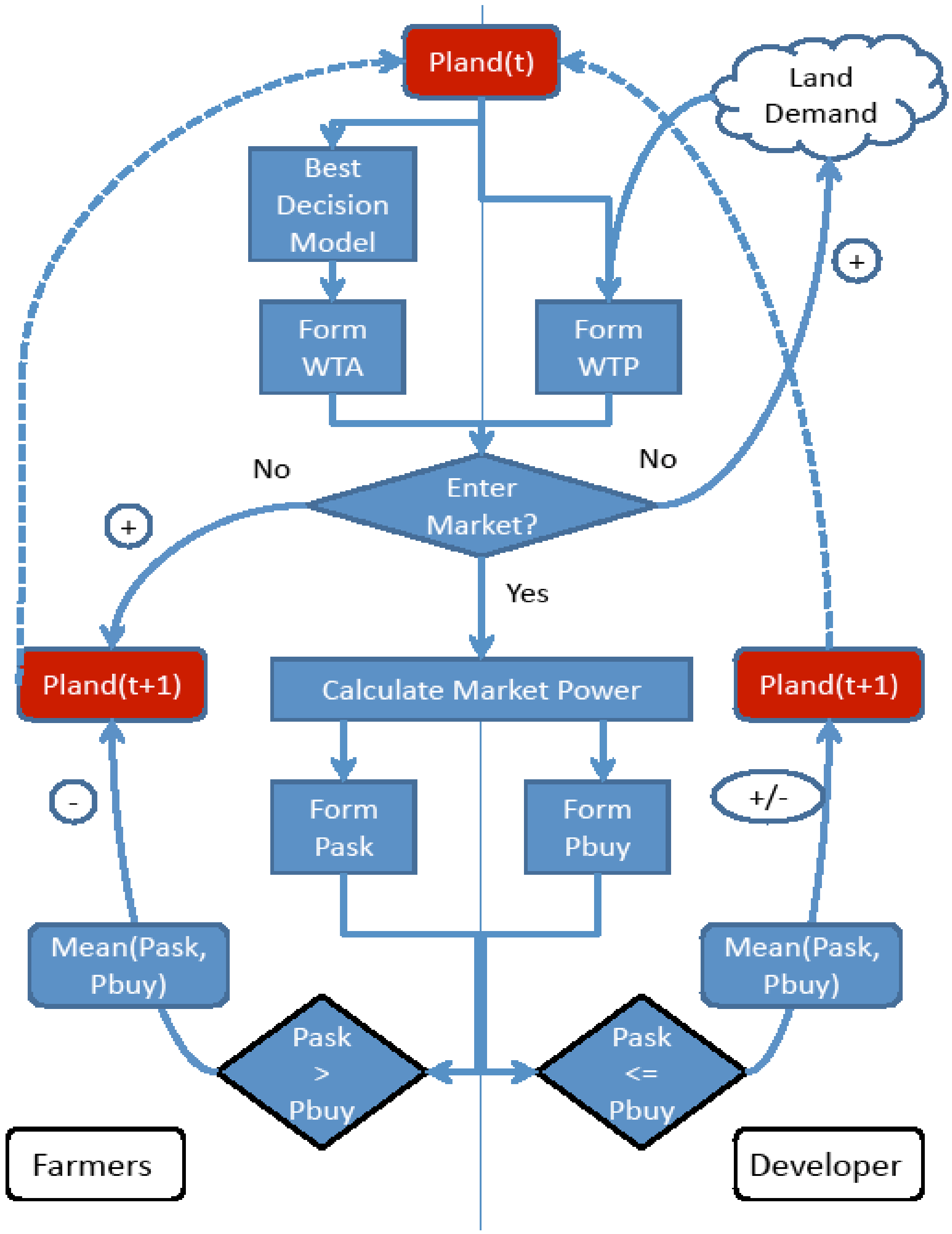
As we stated in the discussion of agent-based models, their development requires a construction of a set of rules that govern the model's propagation from one period to the next. While such sets of rules are not unique and at times may seem arbitrary, the development of a consistent set of rules that produces a realistic development pattern is a complicated task. At present, we have constructed a set of rules that depict the interactions between farmers and developers in the land market. In the next section we describe the workings of the land market in detail.

IV.3.4. Detailed Description of Land Market Interactions

Land market interactions between farmers and developers that endogenously generate land prices are depicted in Figure 5 and are described below. Farmers are each endowed with a randomly generated set of characteristics including acreage, land productivity, operating costs, and farming versus selling preference. Acreage is allocated randomly (within a spatially-delineated zone) then assigned an agricultural productivity that is uncorrelated with acreage. Agricultural return is calculated as the revenues from agricultural output, considering the land's productivity, minus each farmer's idiosyncratic operating cost. A farmer's baseline willingness to accept (WTA) payment to sell his land is determined by calculating the value of his agricultural return per acre in perpetuity. This WTA is adjusted dynamically through price prediction models. Farmers attempt to predict the price of land in a given period using information about past prices. Prediction models allow farmers to follow trends in land prices and secure speculative profits if they exist.

To represent the heterogeneity of farmers, each farmer is randomly assigned twenty prediction models of six different prediction methods. Predictions models are adapted from those described by Arthur (1994) in the "El Farol Café" problem to describe consumers' decisions over spatial goods. Prediction models apply prediction method-specific rules to past price information over a specified time period before the present (i.e. anywhere from 1 to 10 years before the present). Each price prediction model utilizes one of the six different methods for predicting the price of land in each farmer's zone in the next period. Twenty are assigned to ensure diversity in the parameters of prediction models of the same type (i.e. a farmer may have multiple versions of the same type of prediction model, but each version may specify a different time span of past price information over which current predictions are based).

Figure 5. Land Market Interactions between Developers and Farmers



For example, one prediction method looks back into past price information ‘x’ years ago up to the present and extrapolates a trend that is used to predict next period’s price. Another prediction method calculates a mean price over ‘x’ years ago up to the present and uses that mean as a prediction for next period’s price. Others are purely numerical prediction methods, which also look over varied time spans into the past, use cycle predictors, price change analysis, and trend mirroring. In addition, behavioral prediction methods, which use the behavior of other farmers to predict next periods’ price, are included. A farmer in Zone ‘z’ uses the rate of farmers selling their land 1) in their zone and 2) over the entire region to inform their own decision.

All models in the farmer’s ‘projection model bank’ are used to make a prediction for the next time period’s price (t+1). In the next time period, after that period’s price is known, error squared is calculated for each model by squaring the difference between the predicted price and the actual price, and the projection model with least error is used as the prediction of the current period’s price. This same process of prediction and evaluation is used every period so that the most successful prediction model is used every time.

IV.3.5. Determination of Land Prices and Farmers’ Decisions

Based on current and expected conditions of the land market, farmers are making the decision whether to continue farming, sell their land to developers, or continue farming and sell their development right on the TDR market (Figure 5). Farmers are making these decisions based on limited information. For instance, they know their current agricultural return, current and past land prices within their zone, and the number of farmers in their zone, but they do not know the explicit rate of population growth, future land prices, nor the decisions or sale prices of other farmers in their zone. Therefore, farmers use their prediction models to fill-in these knowledge gaps through time and choose the option that maximizes their expected profit.

Given the price of land in period t (which was generated by farmer-developer interactions in period $t-1$), farmers use their most successful prediction to set their WTA in the current period. This represents a price floor below which the farmer will not sell his land, and ensures that he receives at least the agricultural value of his land in perpetuity. Since a farmer’s WTA is dynamically adjusted based on the farmer’s most successful prediction of the current period’s price, it will follow market trends as predictions are utilized and discarded based on their success at predicting changing prices. Simultaneously, a developer will generate his own willingness to pay (WTP) based on the profitability of conversion of purchased land to a particular housing type. The developer’s WTP is calculated for each possible housing type and is based on the maximum positive difference between rent at t for each housing type and its corresponding construction costs.

At the beginning of each time step, the model checks if any farmers’ WTAs are below a developer’s WTP. If so, mutually beneficial transactions may exist and the corresponding farmers enter the land market. At this point, land supply and demand are known to all parties in the current period and each perceives the market power of buyers and sellers according to the procedure proposed by Parker and Filatova (2008):

$$P_{ask} = \text{WTA} \cdot (1 + \varepsilon); \quad P_{bid} = \text{WTP} \cdot (1+\varepsilon)$$

where ε= (Land Demand -Land Supply)/(Land Demand +Land Supply) and thus provides a measure of the relative size of excess demand for land. If farmers supply more land than is demanded by developers, market power lies with the developers and they formulate a bid price (P_{bid}) below their initial WTP. Conversely, if developers demand more land than farmers supply, market power belongs to the farmers and they increase their asking price (P_{ask}) above their initial WTA. If P_{bid} falls below P_{ask}, then the farmer withdraws and is returned to the farmers' pool to await the same decision process next time step. However, if P_{bid} remains above P_{ask}, then a mutually beneficial transaction is made. Through individual transactions, the developer buys all of an individual farmer's land beginning with the lowest asking price. Market land prices are the aggregate result of individual exchanges between farmers who sell their land and the developers who buy it. We consider the equilibrium mean price for each zone in $t+1$ to be the mean of all sale prices in period t, weighted by farm size.

This method for endogenously generating land prices has several important consequences. First, the land market is responding to the behavior of farmers making decisions under incomplete information. Predictions are made based on past and present trends in land prices and substantial uncertainty about future trends. In general, farmers with the highest productivity will have the lowest probability of selling and the highest asking prices all else equal. As a consequence, land prices will start low because farmers with low productivity/low asking prices will sell first and the mean selling price will be low. But as farmers and their land leave the land market, the market price for land will increase -- the result of land supply decreasing and the average asking price of remaining farmers increasing. More farmers will sell as the price increases. Fluctuations in land prices are caused by interactions between farmers and developers within the current market conditions as relationships of market power change through time. Farmers will try to do the best for themselves, but they may end up choosing to sell at a time that is not necessarily optimal.

The developer and consumer decision frameworks are well along, but not all aspects of those two components of the overall model are well enough developed to include detailed explanations here.

V. Concluding Remarks

In this paper, we described a market-based policy for preserving land as open space, Transfer of Development Rights. TDRs have many virtues, not the least of which is the potential to achieve open space protection at least cost. Because of the flexibility given to landowners and developers, TDRs can be more efficient than more command-and-control based policies such as restrictive zoning, urban growth boundaries, clustering requirements, and so forth. However, TDR program success depends on the functioning of created markets and oftentimes these markets fail to work. Many communities have programs on the books that achieve very little. Moreover, even the programs that have worked best have achieved controversial outcomes and have been much debated. We described the Calvert and Montgomery County, Maryland, programs and the issues that have been put forth about those programs. In order to fully address these issues it is best to construct a simulation model of local land markets that captures all of the features of the land, housing, and TDR markets along with agent behavior in those markets and use that model to evaluate alternative policies. That is our objective in this research project.

We described the framework of our model and our plans for assessing alternative TDR programs. Our development of an agent-based model of the land market that fully incorporates economic behavior of agents, market-clearing conditions, and allows for welfare analysis will be unique in the economics literature on land markets. Although the model is not yet fully developed, we provided detail about the farmer/landowner module and how the agent-based approach plays out. In future work, we will show the consumer and developer components, baseline model runs, and analysis of alternative land use policies.

References

Arnott, R.J. 1980. "A simple urban-growth model with durable housing". *Regional Science and Urban Economics*, 10(1): 53-76.

Arnott, R.J., and Lewis F. 1979. "The Transition of Land to Urban Use". *Journal of Political Economy*, 87(11): 161-169.

Arthur, W. B. 2006. "Out-of-equilibrium economics and agent-based modeling". In L. Tesfatsion and K.L. Judd (Eds.). *Handbook of computational economics: Agent-based computational economics* (2, pp. 1551–1564). Amsterdam: Elsevier B.V. (2, pp. 1551–1564). Amsterdam: Elsevier B.V.

Arthur, W. B., Durlauf, S. N., and Lane, D. 1997. *The economy as an evolving complex system II. Santa fe institute studies in the science of complexity (XXVII)*. Addison-Wesley.

Arthur, B.W. (1994). "The El Farol Problem". *Amer. Econ. Assoc. Papers and Proc.*, 84, pg. 406.

Berke, Peter et al. 2003. "Greening Development to Protect Watersheds," *Journal of the American Planning Association* 69(4): 397-413.

Bowen, Greg. 2008. Director of Calvert County Department of Planning and Zoning. Personal communication, August.

Capozza D., and Helsley R. (1990). "The Stochastic City". *Journal of Urban Economics*, 28(2): 187-203.

Carpenter, Bruce E., and Dennis R. Heffley. 1982. Spatial-Equilibrium Analysis of Transferable Development Rights. *Journal of Urban Economics* 12: 238-261.

Dehart, Grant and Robert Etgen. 2007. *Feasibility of Successful TDR Programs for Maryland's Eastern Shore*. Report for the Maryland Center for Agro-ecology, Inc. January.

Epple, Dennis and Holger Sieg. 1999. "Estimating Equilibrium Models of Local Jurisdictions," *Journal of Political Economy* (August).

Epstein, J. M., & Axtell, R. (1996). Growing articial societies: Social science from the bottom up. Washington D.C: Brookings Institution Press and MIT Press.

Glaeser Ed. (2008). *Cities, Agglomeration and Spatial Equilibrium*. Oxford University Press: Oxford.

Irwin, Elena. 2009. "New Directions for Urban Economic Models of Land Use Change: Incorporating Spatial Heterogeneity and Transitional Dynamics". Invited Paper presented at Journal of Regional Science 50th Anniversary Symposium, Federal Reserve Bank of New York, April 23-24.

Irwin, Elena G., Kathleen Bell, and Jacqueline Geoghegan. 2003. "Modeling and managing urban growth at the rural-urban fringe: A parcel-level model of residential land use change," *Agricultural and Resource Economics Review*, 32(1): 83-102.

Irwin, Elena G. and Nancy E. Bockstael. 2002. "Interacting agents, spatial externalities, and the endogenous evolution of residential land use pattern," *Journal of Economic Geography*, 2(1), January: 31-54.

Irwin, Elena G., Ciriyam Jayaprakash and Yong Chen. 2007. "A Dynamic Model of Household Location, Regional Growth and Endogenous Natural Amenities with Cross-Scale Interactions." Paper presented at the "Frontiers of Environmental Economics" conference, Resources for the Future and US Environmental Protection Agency, Washington D.C., February 26-27.

Kirman, A.P., and Vriend, N.J. (2001). "Evolving market structure: An ACE model of price dispersion and loyalty". *Journal of Economic Dynamics and Control*, 25(3–4), 459–502.

Johnston, R.A. and Mary E. Madison. 1997. From Landmarks to Landscapes: A Review of Current Practices in the Transfer of Development Rights. *Journal of the American Planning Association*. Vol. 63.

Kopits, Elizabeth, Virginia McConnell and Margaret Walls. 2008. "Making Markets for Development Rights Work: What Determines Demand?" *Land Economics*, February 84(1):1-16.

LeBaron, B. (2006). "Agent-based computational finance". In L. Tesfatsion and K. L. Judd (Eds.). *Handbook of computational economics: Agent-based computational economics* (2, pp. 1187–1233). Amsterdam: Elsevier.

Lux, T. (1998). "The socio-economic dynamics of speculative markets: interacting agents, chaos, and the fat tails of return distributions". *Journal of Economic Behavior and Organization*, 33, 143–165.

McConnell, Virginia, Elizabeth Kopits, and Margaret Walls. 2006a. Using Markets for Land Preservation: Results of a TDR Program. *Journal of Environmental Planning and Management* 49(5) (September): 631–51

McConnell, Virginia, Margaret Walls, and Elizabeth Kopits. 2006b. Zoning, TDRs, and the Density of Development. *Journal of Urban Economics* 59 (September): 440–457.

Messer, Kent D. 2007. Transferable development rights programs: An economic framework for success. *Journal of Conservation Planning*, vol. 3: 47-56.

Mills, David E. 1989. Is Zoning a Negative Sum Game? *Land Economics* 65(1): 1-12.

Mills, David E. 1980. Transferable Development Rights Markets. *Journal of Urban Economics* 7: 63-74.

Parker, D.C. and Filatova, T. (2008). "A conceptual design for a bilateral agent-based land market with heterogeneous economic agents". *Computers, Environment and Urban Systems*, 32: 454-463.

Parker, D.C., Manson, S.M., Janssen, M.A., Hoffmann, M., and Deadman, P. (2003). "Multi-agent systems for the simulation of land-use and landcover change: A review". *Annals of the Association of American Geographers*, 93(2), 314–337.

Pendall R. 1999. "Do land-use controls cause sprawl?" *Environment and Planning B: Planning and Design* 26(4): 555 – 571.

Pruetz, Rick. 2003. *Beyond Takings and Givings: Saving Natural Areas, Farmland, and Historic Landmarks with Transfer of Development Rights and Density Transfer Charges.* Burbank, CA: Arje Press.

Pruetz, Rick and Noah Standbridge. 2009. "What Makes Transfer of Development Rights Work? Success Factors from Research and Practice," *Journal of the American Planning Association* 75(1): 78-87 (Winter).

Safirova, E., et al. 2006. "Congestion Pricing: Long-Term Economic and Land-Use Effects," Resources for the Future Discussion Paper 06-37, Washington, DC.

Stavins, Robert N. 1995. Transactions Costs and Tradeable Permits. *Journal of Environmental Economics and Management* 29(2): 133-148.

Solimar Research Group. 2003. *Tahoe Basin Marketable Rights Transfer Program Assessment.* Report prepared for the League to Save Lake Tahoe (Ventura, CA: Solimar Research Group) March.

Thorsnes, Paul, and Gerald P.W. Simon. 1999. Letting the Market Preserve Land: The Case for a Market-Driven Transfer of Development Rights Program. *Contemporary Economic Policy* 17(2): 256 – 266.

Towe Charles, Nickerson Cynthia, and Bockstael Nancy. 2008. An empirical examination of the timing of land conversions in the presence of farmland preservation programs. *American Journal of Agricultural Economics* 90(3): 613–626.

Waddell, P. (2002). "UrbanSim: Modeling urban development for land use, transportation, and environmental planning". *Journal of the American Planning Association*, 68(3): 297-314.

Walls, Margaret and Virginia McConnell. 2007. *Transfer of Development Rights in U.S. Communities: Evaluating Program Design, Implementation, and Outcomes*. Resources for the Future Report (Washington, DC: RFF), September.

Walsh, Randall. 2007. ""Endogenous Open Space Amenities in a Locational Equilibrium", *Journal of Urban Economics* 61(2): 319-44 (March).

Can Markets for Development Rights Improve Land Use and Environmental Outcomes?

By Virginia McConnell, Elena Safirova, Margaret Walls, and Nicholas Magliocca
Resources for the Future

Discussant comments – Heather Klemick

The project "Can Markets for Development Rights…" seeks to assess the efficiency and environmental outcomes of tradable development rights (TDR) and other programs to preserve open space using a new simulation model of land use dynamics. The report reviews the TDR approach to land use preservation and describes the key design issues that affect the performance of TDR programs. It also presents two examples of real-world TDR programs based in Maryland's Calvert and Montgomery Counties and discusses features that have led to their high levels of activity. The authors propose to build an agent-based model, which can represent heterogeneous consumer, farmer, and developer behavior and allow for out-of-equilibrium land market dynamics. Once the model is complete, they propose to examine outcomes under a variety of hypothetical policy scenarios, including a no-TDR baseline, TDR, tighter zoning restrictions, and other relevant options.

The project looks to be an exciting contribution to the literature on land use policies. I appreciated the careful justification of the selection of the agent-based model and the discussion of its pros and cons relative to other modeling approaches. The representation of heterogeneous farmer preferences strikes me as a particularly interesting way of adding complexity to behavior while adhering to the assumption that agents are rational profit-maximizers. The relevance of the work could even extend beyond local amenities to inform the development of policies to mitigate land-based greenhouse gas emissions, considering that more and more cities and counties are participating in voluntary carbon markets like the Chicago Climate Exchange.

I have four major comments on the report. Even if the authors cannot feasible address all of these points in the modeling framework as part of the current project, it would at least be helpful to see more discussion of these issues in the final paper or future extensions to address them.

- *Heterogeneity of amenities/externalities*. Although after further discussion with the authors, I understand that this issue will be addressed in future model extensions, I would still have liked to see a more detailed discussion of the heterogeneity and spatial distribution of open space externalities provided by land parcels targeted for preservation in the current paper. The report did not make clear whether amenities provided by open space parcels are assumed to be homogeneous (with distance from homes to the parcels the only source of variation), or whether the model allows open space parcels to provide different levels or types of amenities (whether recreational opportunities, scenic views, storm water runoff management, etc.). These local externalities could be uncorrelated or even negatively correlated with farmland productivity, and could even depend on proximity to other open space parcels. To what extent will these complexities be addressed by the model in the future?

The report also led me to wonder whether there any TDR policies in place that establish trading ratios based on the value of environmental amenities. If TDR programs neglect to do so but these differences are important, what are the implications for the efficiency of outcomes?

- *Aggregate cap on development*. Can the model be used to solve for an "optimal" cap on development in the community? My understanding is that the environmental improvements are being driven by the aggregate cap (implemented through zoning), and the trading regime is a mechanism to achieve the environmental benefits at least cost. Even if it is not possible to use the model to solve for the cap, it would be helpful to see more discussion about how caps could be chosen to maximize welfare.
- *Distributional impacts*. The authors discussed the importance of welfare analysis for evaluating outcomes, and it seems that they will have the capability to evaluate welfare impacts for different sub-groups, not just the community as a whole. I encourage the authors to report on these impacts in the final report, since it could be quite interesting to see which groups (homeowners vs. renters; farmers with different agricultural productivity and open space externalities) benefit or lose out from the different policy options considered.
- *Relevance of the policy scenarios*. The no-TDR baseline and change in sending/receiving areas seem like interesting scenarios to consider, but I wondered whether downzoning in an urban area is the most relevant scenario. Presumably a community interesting in preserving open space would instead target downzoning to rural areas. (Though, based on my discussion with the authors, it seems that this is indeed a common scenario actively considered by several TDR programs.)

Understanding that the report represents a work in progress, I do also have two minor comments:

- Figure 3 is confusing. Even though the sending areas and urban areas are discussed in the text, it would be helpful to see these labeled, and to ultimately include a map more akin to Figure 1.
- I'd recommend tempering any statements claiming that Calvert and Montgomery Counties' TDR programs are the most "successful," recognizing (as the authors do later in the report) that there hasn't been sufficient analysis to assess whether current programs have been welfare-improving relative to counterfactual scenarios with alternative or even no preservation policies.

Preliminary Findings on
Ohio's Great Miami River Water Quality Credit Trading Program*

Richard T. Woodward and David A. Newburn

Department of Agricultural Economics

Texas A&M University

r-woodward@tamu.edu danewburn@ag.tamu.edu

Abstract: Water quality trading (WQT) has been advocated and attempted since Dales first introduced the idea in 1968. Yet this policy approach has achieved relatively little success in creating and sustaining WQT program. The Great Miami River Water Quality Credit Trading Program, run by the Miami Conservancy District (MCD), has shown to be an exception. Through innovative design and effective leadership, the program has come to resemble an actual market in which buyers are paying for pollution reductions and sellers are offering to provide those reductions in response to a price signal. A critical program design element has been to rely on county-level extension agents at Soil and Water Conservation District (SWCD) offices, which acts as intermediaries between farmers and the MCD program managers. Because of their longstanding relationship with farmers, these agents are able to recruit and advise effectively the potential applicants into the program. In this paper we summarize what can be learned from this experience to date and provide preliminary findings about the implications for WQT programs nationally. We draw primarily upon the bids from farmers to the program and interviews with county-level SWCD agents. Two main conclusions can be drawn at this stage in our research. First, bids have varied systematically across counties, influenced more by the agents who are assisting the farmers than variation of the properties of the farms and practices. Second, we find evidence that bids are influenced by strategic behavior at the county level and are also influenced by the opportunity available to farmers to receive conservation payments from the Farm Bill programs. Hence, despite its use of a reverse-auction structure, the MCD program has not pushed farmers to reveal their direct cost of implementing a conservation practice. Overall, the MCD experience demonstrates how effective design can improve the prospects for WQT, but it also demonstrates some of the limitations this approach has in achieving a cost-minimizing allocation of pollution reduction efforts.

* This article was developed under a STAR Research Assistance Agreement No. RD-83367401-0 awarded by the U.S. Environmental Protection Agency. It has not been formally reviewed by the EPA. The views expressed in this document are solely those of the authors and the EPA does not endorse any products or commercial services mentioned in this publication. The authors also receive funding through Texas AgriLife Research with support from the Cooperative State Research, Education & Extension Service,, Hatch Project TEX8604.

Preliminary Findings on

Ohio's Great Miami River Water Quality Credit Trading Program

I. Water Quality Trading in the U.S.

In his seminal article, Dales (1968) used the case of water quality to outline how markets could be more efficient to meet pollution reduction goals. Although his depiction of a cap-and-trade program looks prescient today, his speculation that this approach would be used to address water pollution has been far from realized. While markets for air pollutants trade daily with volumes valued in the millions of dollars, the water quality trading (WQT) markets that exist are small and local and most have few if any actual trades.

The failure of WQT to lead to substantial trading activity is not for lack of potential or institutional effort. If cost heterogeneity were all that is required for a market to prosper, water quality trading would be an instant success. The prototypical trade is between a highly regulated point-source polluter such as a waste water treatment plant and agricultural nonpoint sources. There is usually a wide gap between the marginal abatement costs, as reported by the USEPA (1994), which predicted that trading could save between $700 million and $7.5 billion.

In terms of institutional support, overall EPA has been a strong supporter of WQT. The agency has provided substantial financial and institutional support as have many state-level environmental agencies. There are still important institutional barriers, including the lack of specific authorizing language in the Clean Water Act and some resistance within the agency (USEPA, 2008). Nonetheless, on balance, the EPA has not been a systematic barrier to trading.

Yet despite these economic and institutional forces that seem to favor WQT, in practice the policy instrument have had very limited effect to date. While transferable permit programs for air pollutants and lead in gasoline expanded rapidly in the 1980's and 90's,[1] applications to water pollution problems were rare and typically resulted in no water quality changes or trades of significance. Early efforts such as Wisconsin's Fox River program were largely unsuccessful

(David et al., 1980) and other programs such as Colorado's Lake Dillon program sat unused for over a decade (Woodward, 2003). By the time of their 2004 review of WQT programs, Breetz et al. identified 34 programs nationally that include some authorization for trading between point and nonpoint sources, yet few of these have led to more than a handful of real trades.

There are several programs that might be deemed successful, in which "trading" has facilitated economically productive enterprises or reduced costs (USEPA 2007). The Rahr Malting Company and the Southern Minnesota Beet Sugar Cooperative have both paid farmers to implement nutrient-reducing practices as part of an agreement to allow them to emit nutrients into the Minnesota River. In the Connecticut portion of the Long Island Sound watershed, the "trading" that exists is more like a tax-subsidy scheme, occurring each year after emissions have been reported. Those emitting in excess of their permit must pay, while those emitting less than their permit are paid, both at the same rate. Other programs involve trading among point sources (Tar Pamlico and Neuse River in North Carolina and the Grassland Area Farmers in California) operating under a group permit. However, none of these "success stories" seems to be delivering on the potential that Dales (1968) envisioned. In the Minnesota programs, the demand for pollution offsets comes from a single source, and there is limited diversity in the practices that can reduce pollution. The remaining programs do not include nonpoint sources, and therefore forego the cost savings that WQT has the potential to provide.

What is it about water pollution that makes the development of markets for its control so much more difficult than for air pollution? Numerous reviews and critiques of WQT have been published including Hoag and Hughes-Popp (1997), Jarvie and Solomon (1998), Shabman et al. (2002), and Horan et al. (2004). Together, they identify some important challenges faced by WQT. First, the geographic scope of such a market is defined by a watershed, which tends to be smaller than airsheds that define air pollution market. Second, monitoring and enforcement of trades is more difficult in WQT, particularly when a substantial part of loading problem is attributable to nonpoint sources such as agricultural practices. These factors combine to create

---

[1] Stavins (2008) includes a concise summary of these early efforts.

thin markets with high transaction costs per trade, substantially limiting the trading that might occur in such programs.

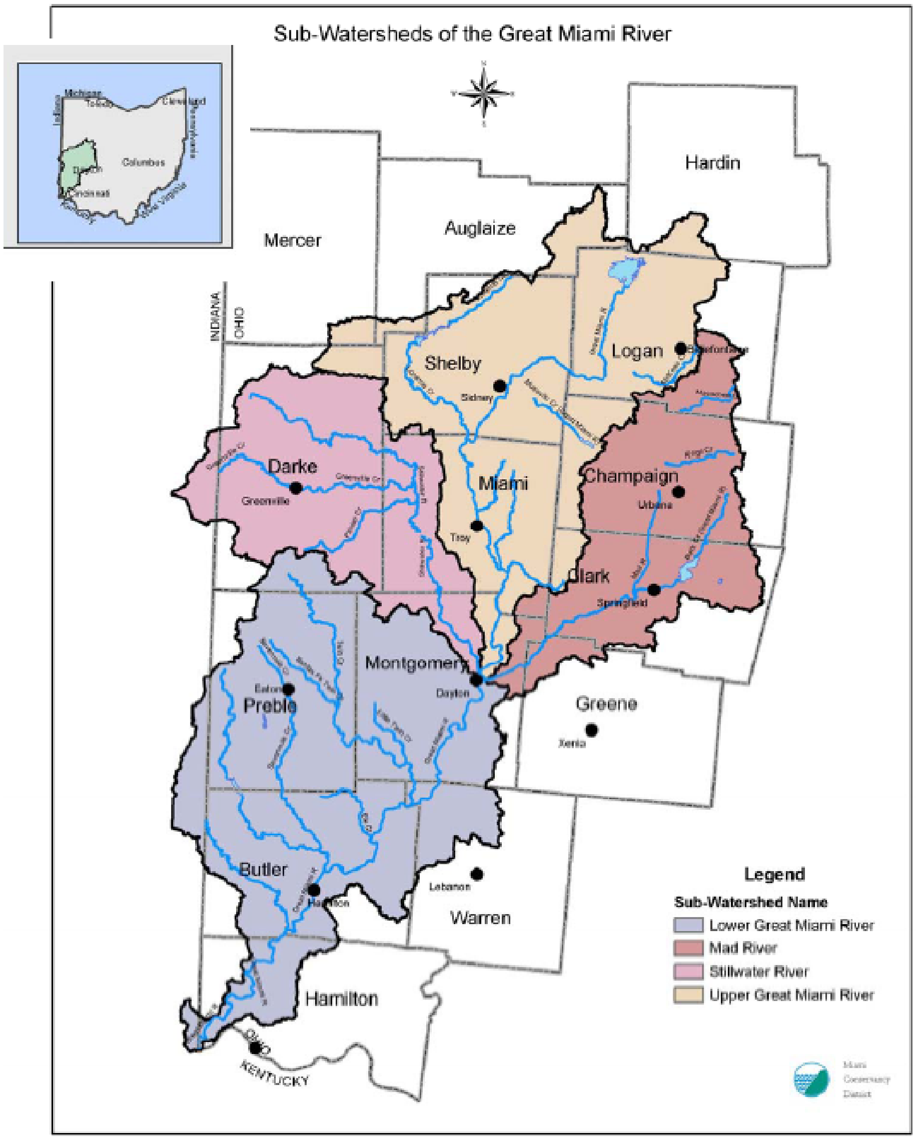
WQT programs in the U.S. have been very limited in scope and have not looked much like a true market in which a variety of buyers and sellers are trading nutrient credits. The program that is the focus of our research project is an exception. In Ohio's Great Miami River Water Quality Credit Trading Program, run by the Miami Conservancy District (MCD), point sources have paid over $1 million to purchase offsets of their emissions. Agricultural landowners have participated in five rounds of bidding program in which 51 projects have been funded from 171 bids submitted. While the program lacks a central exchange like the SO_2 program, it definitely resembles something akin to what economists would call a market.

In this paper we summarize what we have learned so far about the MCD program, highlighting the program's features that we believe are responsible for its success and some of the implications for WQT in general. A main observation is that the program's choice to use county-level agents as intermediaries to provide the conduit through which bids are submitted was an institutional innovation that has led to much higher level of bidding by farmers. However, this feature has also distorted bidding in a way that appears to have created a wedge between the farmers' reservation price and the bids that are actually submitted. That is, some agents have learned the screening criteria set by the MCD program managers, and they have the incentive to help their farmers obtain the highest bid possible. We also discuss how the program's design and policy setting raise interesting issues about the feasibility of obtaining bids that will lead to a cost-minimization in a WQT program.

II. Background and history of the MCD Program

The Great Miami River Watershed is located in southwestern Ohio and covers an area of nearly 4,000 square miles. Though dominated by agricultural lands (83% of all land), the watershed also includes major metropolitan areas including Dayton and part of Cincinnati. The watershed includes over 1,000 stream miles, of which 19.8% are rated as partially attaining aquatic life standards and 21.4% do not meet the standards (MCD, 2005).

Figure 1: Map of the Great Miami River Watershed



Source: Miami Conservancy District. 2005

For WQT programs, a Total Maximum Daily Load (TMDL) report serves as roughly analogous to the cap in standard cap and trade program. A complete TMDL report defines the maximum allowable levels of specific pollutants and allocates pollution reduction responsibilities to different groups of polluters. The Great Miami watershed is composed of four subwatersheds and TMDLs will eventually be developed for each of these. A TMDL was completed in 2004 for the Stillwater, and the Mad River TMDL is currently being prepared. The upper Great Miami watershed was assessed in 2008 and the process of developing a TMDL is ongoing and the

TMDL for the lower Great Miami watershed is scheduled to start in 2010 (Personal communication, Beth Risley). When implemented, TMDLs usually place a substantial and binding burden on the point sources in the watershed. Kieser and Associates (2004) reports 334 point source polluters in the watershed with most of the load coming from large waste water treatment facilities. The prospect that the TMDLs would impose restrictions on these sources, combined with an estimate that the cost of nutrient reductions by point sources might be more than thirty times more expensive than from nonpoint sources (Kieser & Associates, 2004), served as the impetus for demand for the trading program.

In the Great Miami watershed, the management of water quality issues is primarily the responsibility of the Miami Conservancy District (MCD). The MCD was established under the Conservancy Act of Ohio in 1915 with the core responsibility to provide flood protection for communities within the watershed. As political subdivisions of the state of Ohio, conservancy districts can form at the initiative of local land managers or communities to solve water management problems, usually flooding. In addition to flood protection, other approved purposes of conservancy districts include conserving and developing water supplies, treating wastewater and providing recreational opportunities. The MCD has accepted a lead role in developing a strategy for the reducing nutrient limitations in order to comply with the TMDLs.

Faced with TMDLs to be completed throughout its watershed, the MCD began exploring the possible use of WQT as a means to cost effectively address the water quality concerns of the district. The MCD's interest in WQT coincided with the development of guidance from US EPA on the issue (2003) and the initiation of a process to develop rules of trading with Ohio EPA. Kieser & Associates (2004) carried out a preliminary economic analysis of the opportunities for trading and that report found evidence that WQT could lead to a substantial reduction in the cost of meeting nutrient load restrictions.

The MCD plays a central role in the WQT market. Its responsibilities include acting as a broker for trades, supervising the collection of WQ data, issuing requests for nonpoint WQ improving projects, maintaining data on credits, and managing an insurance pool of credits (MCD 2005). The program became fully operable in 2006. Under the rules authorizing the program, point sources can purchase credits from MCD in order to offset nutrient load reductions that will be required in their National Pollutant Discharge Elimination System (NPDES) permits. MCD in turn purchases credits from nonpoint sources.

Though often described as acting as a broker (USEPA 2008), MCD's role is more like a clearinghouse as described by Woodward et al. (2002). MCD plays a central role, purchasing all credits from nonpoint sources and being the sole seller of the credits to the nonpoint sources. There are no direct contractual agreements between the point sources and the nonpoint sources. The MCD program was initiated using a $1 million USDA NRCS grant, but wastewater treatment plants have since made payments of $1.2 million to the program to obtain credits for reductions.

As mentioned above, the TMDLs in the watershed serve as the drivers for participation by point sources. However, the salience of these limits has been diminished by the slow pace of TMDL development. In addition, Ohio EPA admitted in April 2008 that it likely overstated the phosphorus loads in the 2004 TMDL for the Stillwater, the only one that has been completed (Sutherly, 2008). Without a firm requirement that point sources be required to reduce their loads, it is unlikely that demand for credits will persist. As of the end of 2008, the MCD had essentially exhausted the funds it had available for credits (personal communication, Dusty Hall, MCD), and new funds are unlikely to arise until point sources face a credible and impending threat of regulations on their loads.

III. Innovative design features of the Great Miami Program

As we have indicated above, the MCD program stands apart from the vast majority of the WQT programs that have been attempted to date. Seven wastewater treatment plants have made payments into the program and have received credits in return. More impressive is that 171 bids have been submitted by nonpoint sources offering to implement nutrient-reducing practices on their land and these practices have come from a wide range of sources, with bids coming from 10 of the 15 counties within the watershed. If WQT is to become a central part of the nation's approach to address water quality problems, such widespread participation will be necessary.

There are several features of the MCD program that have, in our opinion, been critical to its success. First, one cannot understate the importance of an influential and inspired leadership.

The manager of the MCD program has been an effective leader for the program and organized over 100 meetings with various stakeholders during the process of formulating the program. This effort led to widespread support for the trading program from the stakeholder, representing farmer groups, municipalities, regulatory agencies, and others. While it may not be possible to replicate this inspired leadership, the lesson of taking the time to develop widespread support for the program should not be ignored.

The second feature of the program that we want to highlight is MCD's intermediary role as a clearinghouse in the program. This role means that point sources do not need to negotiate independently with farmers, nor do they sign contracts directly with the farmers. Instead, the point sources in the program deal only with the MCD office, which then accepts the responsibility of obtaining the necessary offsets from farmers. Buyers in this market face none of the transaction costs associated with negotiation, aggregation of sellers, or monitoring and enforcement. This effect is very important to the program's success given that transaction costs have often noted as a problem in WQT programs (Jarvie and Solomon. 1998).

Related to the previous point, but even more innovative in the MCD program was the coordination of the bids with the county-level SWCD offices throughout the watershed. Farmers wishing to be paid to reduce their load by implementing best management practices (BMPs) submit bids to the reverse auctions run periodically by the MCD. In this first-price reverse auction, the winners are those who submit the lowest bids in dollars per pound.[2] Farmers do not, however, interact directly with the MCD program managers. Instead, SWCD agents are the sole contact of the farmers with the program. These agents actively promote the program within their county to help recruit farmers, provide advice on bidding and other information on submitting project applications, disburse funds if a bid is accepted, and monitor the projects for compliance. As discussed by Breetz et al. (2005), SWCD agents have been used by other WQT programs and the longstanding relationships that these agents offer have played an important role in obtaining farmer participation. In other words, these agents have the trust and frequent contact with farmers, which would be costly to establish for any newly established WQT program. The MCD program has taken advantage of these ties to a greater extent than any other program, and we believe this is a central reason that for the large number and variety of farmers who have submitted bids to the program.

An additional important innovation in the MCD program has been the use of a standardized Spreadsheet Tool for Estimating Pollutant Load (STEPL, USEPA Region 5, 2005). As in all WQT programs involving nonpoint sources, the nutrient reductions achieved by a change in agricultural practices is difficult to measure for each farm individually. Hence, the expected nutrient reductions for a given practice are estimated using standardized formulas. The STEPL spreadsheet, which is completed by the SWCD agent, gives estimates of the nutrient reductions that can be achieved through the implementation of a particular management practice. Separate spreadsheets are available for tillage practices, gully stabilization, bank stabilization, and manure management. The spreadsheet uses geographically specific soil characteristics, slopes, and other physical characteristics, in order to estimate the pounds of nutrients that will enter waterways before and after a practice is implemented. This creates reasonable estimates of the load reductions that can be achieved through the program, and a consistent model that is used by SWCD agents throughout the watershed. Inevitably, the spreadsheet is not a perfect predictor of mean runoff and there is some degree of subjectivity in its completion. Further, it represents overall annual estimates, and therefore, it ignores the inherent temporal and spatial variability in nutrient loads that economists have noted repeatedly is an important factor in designing an optimal trading program (Shortle, 1987; Hennessy and Feng. 2008). Finally, there does not appear to be any attempt to calculate site-specific transfer rates to equate the environmental impact of loads throughout the watershed. Despite these limitations, the spreadsheet does represent a practical means of calculating the nutrient reductions that a farmer can provide.

There are two other features of the MCD project applications that had potential to yield valuable information, but do not appear to have worked in practice. First, the applications filed with each bid include space for detailed information on the costs that a farmer would face in implementing his or her practice. If completed and accurate, this would have provided program managers and

---

[2] Total pounds are calculated by simply summing up the pounds of nitrogen and phosphorus and

analysts valuable information about the underlying cost of reducing nutrient loadings. The applications also ask about ancillary benefits that might be provided by a practice change. In both case, however, these variables were not used in making funding decisions and at least some SWCD agents were aware of this fact. As a result, these data are unreliable and of little value in assessing the program.

IV. Preliminary findings on the MCD program

In this section we summarize our preliminary findings in our evaluation of the MCD program. Our observations are based on two sources: the complete set of project applications that have been submitted to the program and semi-structured interviews with program managers, SWCD agents, and farmers. In the fall of 2009 we will be conducting a survey of SWCD offices and farmers in Ohio to provide a stronger statistical foundation for our analysis and explore additional issues.

The traditional model of a farmer's willingness to implement a conservation practice assumes that the reservation price is primarily equal to the change in a farmer's profits that is expected to result from the practice plus a risk premium if the practice changes the distribution of returns over time (Kurkalova et al. 2006). If this holds and a large number of bids are obtained, then theory predicts that bids will tend to be an increasing function of the cost of implementing the conservation practice (Latacz-Lohmann and Van der Hamsvoort, 1997). The reverse auction process, therefore, can be an important step in the achievement of a cost-minimizing outcome.

In practice, we have found that bidding in the MCD program is only loosely tied to the actual cost to the farmers of the associated programs. Our observations to date focus on three main issues that we have found to be particularly interesting about the MCD program: the interaction between the program and existing Farm Bill programs, the importance of county-level institutions, and the extent to which learning has occurred over time.

---

multiplying by the duration of the proposed BMP.

A. Water quality trading in the context of Farm Bill programs

The first reason why bids to the MCD program do not appear to simply reflect the underlying costs of implementation is because of the interaction with the USDA Farm Bill programs. In the MCD program, no practice receiving funding from other sources can also generate credits saleable to the MCD program.[3] This is an important restriction since it helps ensure that credits generated reflect additional nutrient reductions that would not be achieved without the payment. As we have learned through our interviews of SWCD agents, however, because of this, the MCD program is not viewed in isolation but as an option to be added to the regular programs with federal funding, especially the Environmental Quality Incentive Program (EQIP) and the Conservation Reserve Program (CRP). The interviews gave us the strong impression that the SWCD agents view their job as primarily one of achieving conservation objectives while at the same time serving the farmers in their counties. While it certainly holds that a farmer's willingness to adopt a conservation practice will depend upon the cost of that practice, they are also guided by the SWCD agents about which program will be their best option. Farmers likely to qualify for higher payments from the MCD program, therefore, are the most likely to apply.

For example, the continuous CRP program[4] funds small scale conservation activities such as grass waterways to reduce erosion and nutrient run-off within a farm. When farmers can qualify for this program, the CRP program pays 90% cost share and rental payment to the farmers. As one SWCD agent said, "CRP is gold." Similarly, in the Miami watershed, EQIP pays farmers to implement manure management. Since these programs offer an alternative funding source, farmers are unlikely to place MCD bids that differ much from the costs-share rates available through those programs, even if the farmers' actual costs are far below those rates. Further, even

---

[3] In the first year of the program this restriction was not established at the time bids were solicited. However, when bids were evaluated, it was decided that already funded projects could not qualify for funding and this restriction was clarified in later rounds of bidding.

[4] The continuous CRP, which funds small-scale conservation and has a continuous source of funding, is commonly used in the watershed. In the watershed in recent years there have been no opportunities for farmers in the Great Miami watershed to participate in the general signup CRP program, which has a competitive bidding process and makes payments for the retirement of cropland production.

for farms who do not qualify for any of the USDA programs, the government programs provide an initial price signal that may influence the bids.

There are several reasons that were mentioned during our interviews for why the MCD program can be a preferred source of funding to the Farm Bill conservation programs. First, the MCD program often provides greater flexibility. For example, consider the case of a farmer who enters a five-year contract for adopting conservation tillage and then wants to till one year. The Farm Bill programs consider this a violation, whereas the MCD program would just pay for four of the five years of conservation tillage.

Second, another advantage of the MCD program is that the Farm Bill programs may have insufficient funds. In particular, the EQIP program has a backlog due to the low acceptance rate, and therefore, the MCD program is a viable alternative. Third, there are some criteria that make landowners ineligible to compete for the Farm Bill program. For example, land is not eligible for participation in the continuous CRP program, according to program rules, if it was not in the crop rotation during the period 1996-2001. Additionally, some conservation practices are only permitted in the MCD program, such as retiring cropland to establish hayfield with some livestock grazing. We have been told these reasons directly during interviews. Our upcoming survey of a much larger group of farmers will help to document the relative importance of these and other reasons.

B. County-level support and guidance

As discussed above, the MCD program relies heavily on county-level SWCD agents. In this section we will discuss how these agents have played central roles in the bidding process. With the exception of Breetz and her colleagues (2005, 2007), the economics literature on water quality trading has seen intermediaries as playing a relatively minor role in WQT programs. In policy discussions of water quality trading programs there is often mention of the role that might be played by “aggregators” or brokers to reduce the transaction costs of identifying the farmers who might be interested in participating. Woodward et al. (2002) emphasize that an institution might play an intermediary role as a “clearing house” for trading in which a single price is established for credits.

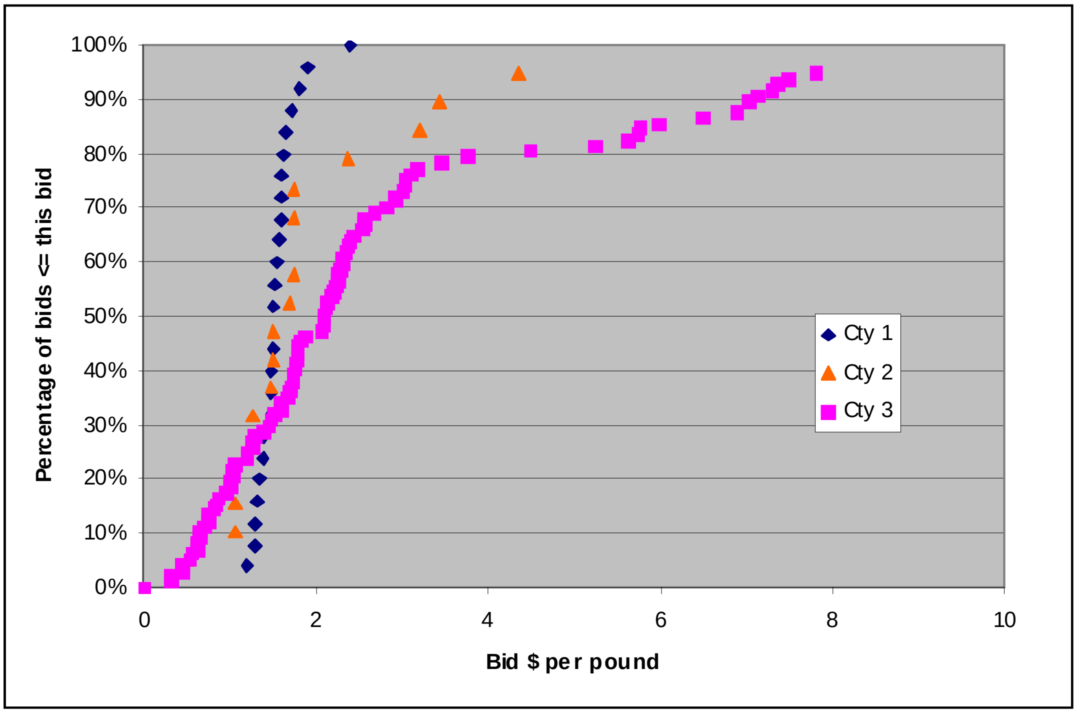
In terms of its ability to reduce transaction costs, we believe that the MCD program has been quite successful. The county-level SWCD agents have long-term relationships with farmers and know well the conservation practices that have the best potential to reduce nutrient runoff. Hence, by relying on SWCD agents to act as intermediaries, the MCD program has been able to generate a large number of bids from farmers and had a credible monitoring and enforcement.

In our conversations with SWCD agents, however, we have found that the agents have done more than just reduce transaction costs. In contrast with the well established and larger Farm Bill programs, most farmers are not even aware of the MCD program. Therefore, the SWCD agents are critical for disseminating information about the purpose and rules of the new MCD program. Through newsletters to farmers and direct conversions either in the SWCD office or during field visits, they have also played a critical role in recruiting potential participants. They have an active role in the project application, often providing advice on bidding strategies and assessing the likelihood of acceptance in Farm Bill versus MCD program. As a result, the bids do not reflect just the preferences and costs of the farmers; they are filtered by the information and help that the SWCD agents provide. These agents, seeing themselves as primarily serving the interests of the landowners in their district, will guide the farmers toward the program with greater success and bids that maximize the payments that they can obtain from the program.

Further, some of the counties have been more aggressive in promoting the MCD program than others. The majority of the applications (83%) have come from only three counties. On the other hand, three of the ten participating counties have submitted three or fewer applications each. Five counties in the watershed did not submit any applications in the first four rounds (Auglaize, Champaign, Hamilton, Hardin, and Logan), though for most of the nonparticipating counties only a small percentage of the county's area lies within the Miami watershed (See Figure **1**).

The SWCD agents' influence on bidding can be seen in Figure 2 where we display the distributions of the bids from the three counties that have submitted the most bids to the program. The bids from county 1 are remarkably concentrated around the mean value of $1.55 with a standard deviation of only $0.23, all falling within a very narrow band between $1.20 and $2.40 per pound. In contrast, County 3's bids averaged $4.92, had a standard deviation of $7.13 and ranged from $0.31 per pound to as high as $65 per pound with about 5% of the 97 bids from that county exceeding $10 per pound. Although there is spatial variation across the watershed in soils, farm types and conservation practices being proposed, it is highly unlikely that this is the source of the striking variation across the counties as seen in the figure.

Figure 2: Distribution of bids from three counties[5]



source: MCD Application data set.

We also found evidence for the SWCD influence on bidding in our interviews with the agents and farmers. In county 1, for example, the agent who helped most of the bidders had the farmers choose their bids from a range of *prices per pound* of nutrients reduced, with clear knowledge of the price that has been paid in prior rounds of the MCD program. In contrast, the agents in county 3 had farmers choose practices and specified set *prices per acre*, regardless of soil characteristics that influence the emission reductions that can actually be achieved. The role played by SWCD agents indicated that they view their role in the MCD program differently than in the Farm Bill conservation programs. In the MCD program an agent is actively helping his or

[5] County names are not identified here to diminish the extent to which this report exposes details of the program that will influence future bidding.

her farmers compete with the farmers of the other counties. In contrast, when working with the Farm Bill programs the agents either are simply technical assistants of applications (e.g., continuous CRP has eligibility criteria but no competitive bidding) or help generate a pool of applicants that compete at the county level for a fixed amount of money (e.g., EQIP provides a fixed budget for each county office).

C. Learning over time

Five rounds of bids have taken place in the MCD program, providing an opportunity to determine whether bidders and the SWCD agents who assist in the bidding have learned over time. There is some evidence of learning as seen in Figure 3.[6] The most evident pattern is the sharp contrast between the first and second rounds. These two rounds make up 82% of all the bids made. There is a sharp shift in the distribution of bids from round 1 to round 2. The average bid in round 1 was \$1.65 per pound-year, but this increased to \$4.57 in round 2. The difference between these two periods can largely be attributed to a change in the MCD's eligibility criteria. Prior to round 1, MCD had indicated that bids that also received conservation funding from government cost-share programs could participate. However, such applications were not accepted in round 1 and the MCD clarified that policy change to the SWCD offices prior to round 2. It appears that the SWCD offices then expected the MCD to pay higher prices in round 2. In fact, the highest price per credit that the MCD paid in round 2 was 3¢ lower than the highest price paid in round 1.

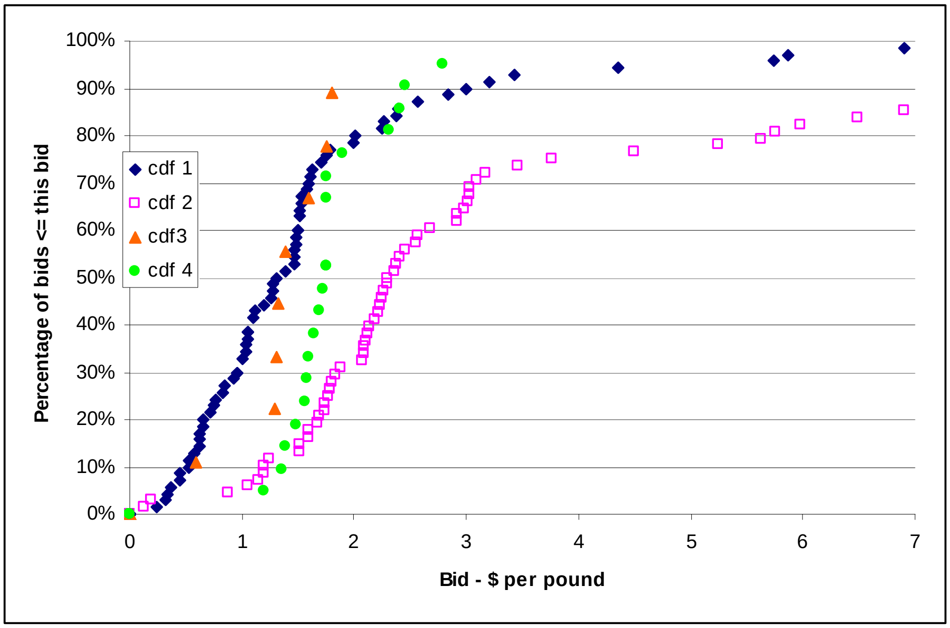
The experience in rounds 1 and 2 seem to have had two main effects. First, we see that far fewer bids were offered in rounds 3 and 4. Some of the counties that were the most active participants in the first two rounds actually dropped out completely in the later two rounds. The differences across counties seen in Figure 2 is to some extent a result of the fact that different counties are participating in later rounds, after some learning has occurred, than were active in rounds 1 and 2. Secondly, we see in round 3 and 4 more concentrated bids. In round 1 there was a \$5.38 gap between the bids at the 5^{th} and 95^{th} percentiles, and this gap increased to \$12.46 in round 2. By

---

[6] For a variety of reasons, only two bids were submitted in round five. Hence, we exclude it from this figure.

rounds 3 and 4, the bids had become much more concentrated with the centrally 90% spread falling to $1.20 and $1.60 respectively. Clearly, by round 3 bidders were getting much better information about how much to bid. This was accompanied by a sharp decline in the number of bidders and the county that submitted the most bids in rounds 1 and 2 did not solicit any bids after that point. To our surprise, we found that there are very few (perhaps one or two) farmers who have offered repeat bids in the data. Hence, we believe that the changes seen in Figure 3 reflect learning by the SWCD agents, not individual farmers.

Figure 3: Distributions of bids in rounds 1 - 4



source: MCD Application data set.

Based on these data and our interviews with SWCD agents, there appear to be two main types of agents. The first type attempted to participate during the first two rounds in the MCD program. In the early rounds there was some uncertainty and confusion among the agents due to novelty of the program, the evolving rules and bidding, and the fact that the MCD's ranking based on dollars per pound as opposed to Farm Bill programs that are tied to practices and land area. Some of the agents that failed in those rounds, despite substantial effort, became discouraged and declined to participate in later rounds. The second type of agent was able to learn quickly about the MCD program criteria or began participating after the bugs had been worked out. They experienced higher acceptance rates for their farmers and have continued to promote the MCD program and recruit farmers. This is seen in Table 1, where we see that the number of counties participating in the MCD program fell sharply after the second round, while the acceptance rate went up sharply.

Table 1: Counties Participating and Acceptance Rates

Round	**1**	**2**	**3**	**4**
Number of counties with an application	9	6	3	4
Acceptance Rate	18%	22%	89%	67%

While there certainly is variation across the watershed, it seems unlikely to us that there are no cost-effective opportunities to reduce nutrient loading in 3/4 of the counties in the watershed. If this is true, then for reasons that are entirely unrelated to the cost of achieving water quality improvements, participation in the program is concentrated in some counties and not in others. Hence, this process of learning has important implications for program efficiency. Where SWCD have stopped participating, a substantial portion of the watershed, landowners are not even aware of this WQT program and certainly not submitting bids.

D. Practices across counties

One of the features of the MCD program that distinguishes it from most other WQT programs to date is that it has involved a wide range of pollution reducing practices. Two other programs that have paid a large number of farmers to implement practices were established for point sources in Minnesota: the Rahr Malting Company and the Southern Minnesota Beet Sugar Cooperative. In both cases, the range of management practices that could apply for funding was relatively narrow. In contrast, a wide range of practices could apply for funding in the MCD program. As indicated above, the nutrient load reductions are calculated using four separate spreadsheets, however, each of these offers many combinations and variations of practices that might be proposed. Aggregating into five main practice types, we see that 40% of the applications have proposed some sort of tillage or field management change, and this makes up almost exactly

40% of all of the accepted projects. In contrast, projects that have proposed funding of a technological way to reduce runoff, mostly in the form of variable rate fertilizer applications, were rejected in all cases. The most successful category is hayfield and sod establishment, where 9 of the 11 proposals have been accepted. These projects were relatively small, averaging only 2,400 pounds per project, when compared to other practices such as livestock management projects, which averaged 50,790 pounds per project. In fact, livestock project represent XX% of the total nutrient reductions accepted by the MCD program.

Table 2: Applications by practice type

Practice Type	Applications	% Accepted	Average Lbs/Project	Average bid/lb
Grass waterways	49	10%	1,449	1.54
Livestock	18	56%	50,970	1.40
Tillage	67	31%	8,253	1.22
Hayfield & Pasture	12	75%	2,436	1.51
Bank Stabilization	7	71%	7,005	2.10
Fertilizer Management	16	0%	7,727	1.12
All Applications	169*	30%	10,316	1.35

* Includes only projects in rounds 1-4, excluding the two proposals submitted in round 5
Source: MCD Application data set.

Across the watershed there is substantial variation in the types of agricultural operations. According to the 2002 Agricultural Census, 88% of the agricultural revenue was derived from crops in counties 7 and 10, whereas, county 1 is livestock dominated with only 17% of the agricultural revenue was from crops (Table 3). Hence, it is not suprising that there is also variation of in the practices that are proposed across the watershed. We find that counties tend to specialize in one or two practice types. For example, 13 of the 25 applications from county 1 specifically related to livestock, consistent with the strong presence of the livestock sector in that county. However, not all of the variation can be explained as a reflection of the agricultural sector in the counties. The applications from counties 2 and 3, for example, seem out of proportion with the distribution of revenue for these counties. Some of this can be explained by changes in program rules over time, in particular the fact that livestock practice were not allowed until the third round of bidding, due to delays in formulating the nutrient worksheet to evaluate livestock projects. But, consistent with our meetings with SWCD agents, there also appears to be specialization by county offices that does not reflect only the agricultural sector in the county.

Table 3: Practice types in applications by county
(Predominant practice for each county in bold face)

		Management Practices in Project Proposals						
County*	% of Agriculture Revenue From Crops**	Bank Stabilization	Tillage	Fertilizer Management	Grass Waterways	Hayfield & Pasture	Livestock	Total MCD Applications
County 1	17		5		1	6	**13**	25
County 2	42	**7**	1	5	5		1	19
County 3	56		**46**	11	40			97
County 4	78		**5**		1			6
County 5	78					**5**		5
County 6	57		**3**			1	1	5
County 7	88		**4**					4
County 8	68		1		**2**			3
County 9	13						3	3
County 10	88		**2**					2
County 11	43							0
County 12	74							0
County 13	83							0
County 14	82							0
County 15	36							0
All Counties	40	7	**67**	16	49	12	18	169

* County numbers are the same as in Figure 2.
** Agricultural Revenue data is from the 2002 Census of Agriculture. The remaining portion of the revenue is from livestock.
Source: MCD Application data set.

V. Conclusions on the efficiency of bidding in the MCD program

We have argued that the MCD program is probably the most interesting point-nonpoint source WQT program to be developed to date. In sheer volume of trades it is among the largest WQT programs and in terms of variety of traders and prices paid, it is unsurpassed. The MCD *looks* like a real market with the MCD playing the central role of a clearinghouse.   Since many of the payments made by point sources were made in anticipation of TMDL restrictions that have not yet materialized, ex post, it is not possible to document whether there have been any cost savings. However, as TMDLs are finalized in the coming years, the watershed has the advantage of an institutional capacity that will make it possible to use trading as an alternative to wastewater plant upgrades.

However, there are three aspects of the MCD program that require attention and have may have inhibited the program's efficiency to date. First, we find evidence that the prominent role of the SWCD agents has introduced some distortionary impacts on the bidding. Second, although the program has not witnessed a great deal of repeat bidding by landowners seeking to "game the system," there is evidence that county-level representatives have learned over time and this is leading to bids that are near to the price being paid by MCD. To some extent it appears that much of the learning was done at the expense of counties that submitted a large number of bids in the early rounds. Counties that began bidding in later rounds were able to set their bids more precisely to increase the chance of having a bid accepted. Finally, our interviews have indicated that bids in the program have also been influenced by the options that farmers have to take advantage of conservation funding in the USDA Farm Bill programs. It is too early to tell if these effects will diminish over time. But if future programs use the MCD program as a model, we believe it would be useful to structure the program in a way that provides a common body of information and guidance.

Together, the distortionary role of SWCD agents in setting bids and the alternative available to farmers to obtain funding from other sources mean that bids in the MCD program do not reflect simply the cost of implementing a project. We do not know the magnitude of this difference and future work will attempt to measure it to the extent possible. But if there is a gap, then farmers in the MCD program will not tend to reveal in their bids their true reservation price for implementing a BMP, which is the principal motivation for using a reverse auction type structure (Latacz-Lohmann and Van der Hamsvoort, 1997).

VI. References

Breetz, Hanna L. et al. 2004. Water Quality Trading and Offset Initiatives in the U.S.: A Comprehensive Survey. Working Paper: Dartmouth College, Hanover, New Hampshire. http://www.dartmouth.edu/~kfv/waterqualitytradingdatabase.pdf (Accessed on 9/26/2006)

Breetz, Hanna L, and Karen Fisher-Vanden. 2007. Does Cost-Share Replicate Water Quality Trading Projects? Implications for a Possible Partnership. *Review of Agricultural Economics* 29(2):201-15.

Breetz, Hanna L, Karen Fisher-Vanden, Hannah Jacobs, and Claire Schary. 2005. Trust and Communication: Mechanisms for Increasing Farmers' Participation in Water Quality Trading. *Land Economics* 81(2):170-90.

Dales, John H. 1968. "Land, Water, and Ownership." *The Canadian Journal of Economics*, 1(4):791-804.

David, M., W. Eheart, E. Joeres, and E. David. "Marketable Permits for the Control of Phoshporus Effluent into Lake Michigan." *Water Resources Research* 16(April 1980):23-63-70.

Hennessy, David A, and Hongli Feng. 2008. When Should Uncertain Nonpoint Emissions Be Penalized in a Trading Program?. *American Journal of Agricultural Economics* 90(1):249-55.

Hoag, Dana L, Bruce A Babcock, and William E Foster. 1993. Field-Level Measurement of Land Productivity and Program Slippage. *American Journal of Agricultural Economics* 75(1):181-89.

Horan, Richard D., Shortle, James S., and Abler, David G.. 2004. The Coordination and Design of Point-Nonpoint Trading Programs and Agri-Environmental Policies. *Agricultural and Resource Economics Review* 33 (1): 61-78.

Jarvie, Michelle and Barry Solomon. 1998. Point-Nonpoint Effluent Trading in Watersheds: A Review and Critique. Environmental Impact Assessment Review 18(2):135-57.

Kieser & Associates. 2004. "Preliminary Economic Analysis of Water Quality Trading Opportunities in the Great Miami River Watershed, Ohio. " Prepared for The Miami Conservancy District. Kalamazoo, Michigan.

Kurkalova, Lyubov, Catherine Kling, and Jinhua Zhao. 2006. Green Subsidies in Agriculture: Estimating the Adoption Costs of Conservation Tillage from Observed Behavior. *Canadian Journal of Agricultural Economics* 54(2):247-67.

Latacz-Lohmann, Uwe, and Carel Van der Hamsvoort. 1997. Auctioning Conservation Contracts: A Theoretical Analysis and an Application. *American Journal of Agricultural Economics* 79(2):407-18.

Latacz-Lohmann, Uwe, and Carel Van der Hamsvoort. 1997. Auctioning Conservation Contracts: A Theoretical Analysis and an Application. *American Journal of Agricultural Economics* 79(2):407-18.

Miami Conservancy District. 2005. *Great Miami River Watershed Water Quality Credit Trading Program Operations Manual* Dayton, OH: The Miami Conservancy District.

Ohio EPA. 2006. *Ohio TMDLs in Progress* http://www.epa.state.oh.us/dsw/tmdl/OhioTMDLs_InProgress.html

Shabman, Leonard, Kurt Stephenson, and William Shobe. 2002. Trading Programs for Environmental Management: Reflections on the Air and Water Experiences. *Environmental Practice* 4(3):123 - 175.

Shortle, James S. 1987. "Allocative implications of comparisons between the marginal costs of point and nonpoint source pollution abatement." *Northeastern Journal of Agricultural and Resource Economics* 16(1):17-23.

Stavins, Robert N. 2008. Addressing Climate Change with a Comprehensive US Cap-and-Trade System. *Oxford Review of Economic Policy* 24(2):298-321.

Sutherly, Ben. 2008. "EPA: Stillwater River Pollution Overstated." *Dayton Daily News*, April 28, 2008, A6.

United States Environmental Protection Agency (USEPA). 1994. "President Clinton's Clean Water Initiative: Analysis of Costs and Benefits. EPA-800-R-94-002. Washington, DC: Office of Water.

USDA. 2002. *2002 Census of Agriculture*. Online data available at http://www.agcensus.usda.gov/

United States Environmental Protection Agency, Office of Water. 2003. Final Water Quality Trading Policy, issued January 13, 2003. Available http://www.epa.gov/owow/watershed/trading/tradingpolicy.html (Accessed on 9/25/2006)

USEPA. 2007. *Water Quality Trading Toolkit for Permit Writers*. (August 2007) http://www.epa.gov/owow/watershed/trading/WQTToolkit.html (accessed on 4/1/2009)

USEPA. 2008. *Water Quality Trading Evaluation: Final Report* (October 2008) http://www.epa.gov/evaluate/wqt.pdf (accessed on 4/1/2009)

USEPA Region 5, 2005. Spreadsheet Tool for Estimating Pollutant Load (STEPL). Available at http://it.tetratech-ffx.com/stepl/ accessed on 9/14/2006.

Woodward, Richard T. "Lessons about Effluent Trading from a Single Trade." *Review of Agricultural Economics* 25(2003):235-45.

Woodward, Richard T., Ronald A. Kaiser , and Aaron-Marie B. Wicks. 2002. The Structure and Practice of Water Quality Trading Markets. *Journal of the American Water Resources Association* 38 (4): 967-979.

Background

- US EPA (2004) *Water Quality Trading Assessment Handbook*, EPA 841-B-04-001.
- King, Dennis (2005) "Crunch Time for Water Quality Trading," Choices, 20(1), pp. 71-75.
  - Why are there so few WQ trading success stories
  - Why aren't the point and nonpoint sources who are supposed to benefit from WQ trading more supportive
  - What can be done to improve the situation

Woodward and Newburn

- Success factors of the Great Miami Program
  - Inspired leadership
  - MCD’s intermediary role as a clearinghouse
  - Role of SWCD offices and agents
    - Standardized Spreadsheet tool for Estimating Pollutant Load (STEPL)

To Increase Usefulness to EPA: Quantification

- Are the SWCD Agents worth the cost
  - Working for the NPs
  - Working harder in some counties than others
  - Is there gaming or not, and will it be measurable
- CRP and EQIP interfere (slippage)
  - Interference measurable
- If NPs are not revealing their reservation price
  - How far off are they
- If TMDLs are drivers and TMDLs don’t exist
  - To what extent is this lack of a regulatory stick reducing the efficiency of the program

To Increase Usefulness to EPA: Action

- "If this is true [interference by SWCD and other programs] and general, then farmers in the MCD program will not tend to reveal their reservation price in their bids, which is the principal motivation for using a reverse auction type structure."
  - Principal yes, but there are coordination features the auction adds
  - What is the alternative [if this is a successful WQT program, what does this augur] or
  - How can this be fixed
    - Can the role of the SWCD agents be modified
    - Can some information be withheld (as it is in the CRP's EBI) to increase competition

In General

- "Hence, despite its use of a reverse-auction structure, the MCD program has not pushed farmers to reveal their direct cost of implementing a conservation practice. Overall, the MCD experience demonstrates how effective design can improve the prospects for WQT, but it also demonstrates some of the limitations this approach has in achieving a **cost-minimizing allocation** of pollution reduction efforts."
  - This doesn't bother me if it can be quantified

Market Mechanisms and Incentives: Applications to Environmental Policy

A Workshop sponsored by U.S. Environmental Protection Agency's National Center for Environmental Economics (NCEE) and National Center for Environmental Research (NCER)

Resources for the Future
1616 P Street, NW, Washington, DC 20036
(202) 328-5000

Wednesday, April 29, 2009

Disclaimer

These proceedings are being distributed in the interest of increasing public understanding and knowledge of the issues discussed at the workshop. The contents of this document may not necessarily reflect the views of the U.S. Environmental Protection Agency and no official endorsement should be inferred.

U.S. Environmental Protection Agency (EPA) National Center for Environmental Economics (NCEE) and National Center for Environmental Research (NCER)

Market Mechanisms and Incentives: Applications to Environmental Policy

Resources for the Future
1616 P Street, NW, Washington, DC 20036
(202) 328-5000

Wednesday, April 29, 2009

8:30 a.m. – 9:15 a.m.	**Registration**	
9:15 a.m. – 9:30 a.m.	**Introductory Remarks:** *Julie Hewitt, Chief, Economic and Environmental Assessment Branch, Office of Water*	
9:30 a.m. – 11:20 a.m.	**Session I: Fuel Economy and Gasoline Prices**	
	Session Moderator:	*Cynthia Morgan, EPA, National Center for Environmental Economics*
	9:30 a.m. – 10:00 a.m.	Imperfect Competition, Consumer Behavior, and the Provision of Fuel Efficiency in Vehicles *Carolyn Fischer, Resources for the Future*
	10:00 a.m. – 10:30 a.m.	New Vehicle Characteristics and the Cost of the Corporate Average Fuel Economy Standard *Thomas Klier, Federal Reserve Bank of Chicago, and Joshua Linn, University of Illinois at Chicago*
	10:30 a.m. – 10:40 a.m.	Discussant: *Gloria Helfand, University of Michigan, and EPA, Office of Transportation and Air Quality*
	10:40 a.m. – 10:50 a.m.	Discussant: *Winston Harrington, Resources for the Future*
	10:50 a.m. – 11:20 a.m.	Questions and Discussion
11:20 a.m. – 12:30 p.m.	**Lunch** (On Your Own)	
12:30 p.m. – 1:30 p.m.	Panel Discussion: Role of Market Mechanisms and Incentives to Climate Change	
	Moderator: *Dick Morgenstern, Resources for the Future*	
	Panelist: *Joe Aldy, Special Assistant to the President for Energy and the Environment* *David McIntosh, Senior Counsel in the Office of Congressional and Intergovernmental Relations* *Brian Murray, Duke University*	
1:30 p.m. – 1:40 p.m.	**Break**	
1:40 p.m. – 3:30 p.m.	**Session II: Applications of Environmental Trading Programs**	
	Session Moderator: *Will Wheeler, EPA, National Center for Environmental Economics*	
	1:40 p.m. – 2:00 p.m.	An Experimental Analysis of Compliance in Dynamic Emissions Markets: Theory and Experimental Design *John Stranlund, University of Massachusetts - Amherst, James Murphy, University of Alaska – Anchorage, and John Spraggon, University of Massachusetts - Amherst*

	2:00 p.m. – 2:30 p.m.	Can Markets for Development Rights Improve Land Use and Environmental Outcomes? *Virginia McConnell, Elena Safirova, Margaret Walls, and Nick Magliocca, Resources for the Future*
	2:30 p.m. – 2:35 p.m.	Discussant: *Heather Klemick, EPA, National Center for Environmental Economics*
	2:35 p.m. – 3:05 p.m.	Preliminary Findings and Observations on Ohio's Great Miami River Water Quality Credit Trading Program *Richard Woodward, Texas A&M University*
	3:05 p.m. – 3:10 p.m.	Discussant: *Hale Thurston, EPA, National Risk Management Research Laboratory*
	3:10 p.m. – 3:30 p.m.	Questions and Discussion
3:30 p.m. – 3:40 p.m.	**Break**	
3:40 p.m. – 5:30 p.m.	**Session III: Winners and Losers in Cap and Trade** Session Moderator: *Charles Griffiths, EPA, National Center for Environmental Economics*	
	3:40 p.m. – 4:10 p.m.	Paving the Way for Climate Policy: Compensation for Electricity Consumers and Producers Under a CO_2 Cap and Trade Policy *Karen Palmer, Dallas Burtraw, and Anthony Paul, Resources for the Future*
	4:10 p.m. – 4:40 p.m.	When Does Cap-and-Trade Increase Regulated Firms' Profits? *Dave Evans, EPA, National Center for Environmental Economics; Ian Lange, University of Stirling; and Joshua Linn, University of Illinois at Chicago*
	4:40 p.m. – 4:50 p.m.	Discussant: *Ann Wolverton, EPA, National Center for Environmental Economics*
	4:50 p.m. – 5:00 p.m.	Discussant: *Terry Dinan, Congressional Budget Office*
	5:00 p.m. – 5:30 p.m.	Questions and Discussion
5:30 p.m.	**Adjournment**	

DISCUSSION PAPER

July 2008 ■ RFF DP 08-25

Compensation for Electricity Consumers under a U.S. CO_2 Emissions Cap

Anthony Paul, Dallas Burtraw and Karen Palmer

1616 P St. NW
Washington, DC 20036
202-328-5000 www.rff.org

RESOURCES
FOR THE FUTURE

Compensation for Electricity Consumers Under a U.S. CO_2 Emissions Cap

Anthony Paul, Dallas Burtraw, and Karen Palmer

Abstract

Policies to cap emissions of carbon dioxide (CO_2) in the U.S. economy could pose significant costs on the electricity sector, which contributes roughly 40 percent of total CO_2 emissions in the U.S. Using a detailed simulation model of the electricity sector, we evaluate alternative ways that emission allowances can be allocated. Most previous emissions trading programs have allocated the major portion of allowances for free to incumbent firms. In the electricity sector this approach would lead to changes in electricity price that vary by region primarily based primarily on whether prices are market-based or determined by cost-of-service regulation. Allocation to customers, which could be achieved by allocation to local distribution companies (retail utilities) would recover symmetry in the effect of free allocation and lead to signficiantly lower overall electricity prices. However, this form of compensation comes with an efficiency cost that will increase the overall cost of climate policy.

Key Words: emissions trading, allowance allocations, electricity, air pollution, auction, grandfathering, cost-effectiveness, greenhouse gases, climate change, global warming, carbon dioxide, asset value, compensation

JEL Classification Numbers: Q2, Q25, Q4, L94

Contents

Compensation for Electricity Consumers Under a U.S. CO_2 Emissions Cap

Anthony Paul, Dallas Burtraw, and Karen Palmer *

1 Introduction

A crucial decision in the design of a cap-and-trade program for CO_2 is the initial distribution, or allocation, of emission allowances. The creation of a market for CO_2 emissions would involve the largest distribution and enforcement of new property rights in North America in over a century, and the decision about allocation has efficiency and distributional consequences. The economics literature finds significant efficiency advantages to the use of an auction rather than free distribution of emission allowances. One reason is that an auction is administratively simple and precludes regulated parties from seeking a more generous future allocation. Another is that free allocation in competitive markets, like some markets for electricity in the United States, can move consumer prices away from the marginal cost of production and therefore distort resource allocation in the wider economy away from the efficient optimum. Compared with other approaches, an auction helps maintain transparency and the perception of fairness, and it leads to more efficient pricing of goods in the economy, which reduces the cost of the policy. These are important principles for the formation of a new market for an environmental commodity.

Most previous programs have relied on free distribution rather than an auction. Generally speaking, free allocation of allowances gives interested parties strong incentives to argue for an ever-increasing share of emissions allowances. In contrast, many authors suggest that auctions reduce rent-seeking, which occurs when regulated parties invest resources in trying to affect the outcome of an administrative process that distributes allowances freely. One particularly insidious aspect of free allocation is the adjustment to allocation rules for new emissions sources and for old sources that retire. The sulfur dioxide (SO_2) trading program in the United States has no adjustments for these sources, which is a virtue because it does not create incentives for investment

---

* This research was supported in part by Mistra's Climate Policy Research Forum (Clipore). Erica Myers and Susie Chung proivided technical assistance. All errors and opinions are the responsibility of the authors. Please send comments to paul@rff.org. This paper was prepared for the conference *Reforming Rules and Regulations: Laws, Institutions and Enforcement*, 18-19 July, 2008,Venice, Italy. All errors and opinions are the responsibility of the authors.

behavior to deviate from what is otherwise efficient. However, most other trading programs have such adjustments. In the NO_x budget program in the United States, for example, individual states determine the allocation of allowances; most have set-asides for new sources, and sources that retire lose their allocations. Adjustments also are ubiquitous in the EU Emission Trading Scheme. The problem with such adjustments is that they alter the incentives for investment and retirement in a way that can lead to unintended consequences. For instance, there is evidence that as a result of adjustments to allocation rules for new sources in the EU, the value of emissions allowances can bias investment toward higher-emitting generating sources. This bias can result from the value of the subsidy embodied in free allowance allocations. Furthermore, the removal of allocations from sources that retire provides a financial incentive to continue the operation of existing facilities that are often inefficient and that otherwise would retire, except for the value of the allowances that they earn by remaining in operation. The use of an auction avoids this predicament entirely.[1]

The second, and equally forceful, reason that economists favor the use of auctions is that they generate funds that can be used to help reduce the cost of policy. For the purposes of minimizing the cost of climate policy on the economy and promoting economic growth, the economics literature has focused on dedicating the use of revenue from an auction to reduce preexisting taxes. Like any new regulation, climate policy imposes a cost on households and firms; that cost acts like a virtual tax, reducing the real wages of workers. This hidden cost can be especially large under a cap-and-trade program because the price placed on the scarcity value of carbon is reflected in the cost of goods that use carbon in their production, which are ubiquitous in the economy. However, the revenue raised through an auction (or an emissions tax), if dedicated to reducing other preexisting taxes, can reduce this cost. This so-called revenue recycling would have substantial efficiency advantages compared with free distribution.[2]

A compelling justification *for* free distribution of emission allowances is that public policy should do "no direct harm" through changes in government rules and regulations.[3] This justification has been invoked to argue for free allocation to firms, in order to soften the impact of the policy. However, consumers rather than firms or their

---

[1] Åhman et al. 2007; Åhman and Holmgren 2006.

[2] Bovenberg and Goulder, 1996; Bovenberg and de Mooij, 1994; Golder et al. 1999; Parry et al. 1999. Smith et al. 2002.

[3] Schultze, 1977.

shareholders may be the most adversely affected by climate policy. Consequently, compensation for consumers has become a central element of the political dialogue about climate policy in the United States, which is made even more salient by recently increasing fuel prices. This paper looks at the effect on consumers from climate policy and approaches to compensating consumers.

We focus exclusively on the electricity sector. Although the electricity sector is responsible for about 40 percent of CO_2 emissions in the United States, most models indicate that under a cost-effective program, two-thirds to three-quarters of emission reductions in the first couple decades of climate policy are likely to come from this sector. Consequently the electricity sector is a very special case; it constitutes the heart of any proposal to implement market-based approaches to achieving CO_2 emission reductions. All of the important existing trading programs include the electricity sector, and usually they exclude other sources. In the United States, the sulfur dioxide (SO_2) trading program, which began in 1995, and the nitrogen oxide (NOx) trading program, which began in 1999, each has a pool of emission allowances with annual value of $1–3 billion and focus on the electricity sector almost exclusively. The EU Emission Trading Scheme (ETS), which began in 2005, includes major point sources, of which the electricity sector constitutes the most significant portion. In addition, the ten-state Regional Greenhouse Gas Initiative (RGGI), which will be the second mandatory cap and trade program in the world for CO_2 beginning in 2009, covers just the electricity sector.

Previous analysis of the electricity sector relying on detailed simulation modeling indicates that on an industry-wide basis only 6 percent of the allowance value created within the electricity sector (2.5 percent overall) is sufficient to hold the industry harmless because the majority of costs are recovered by changes in product prices.[4] General equilibrium models with less information about the structure of costs and production within the sector have found results that are broadly comparable. One study found that most of the economic effect of climate policy would be felt in the oil, gas, and coal industries, which could be compensated with just 19 percent of allowance value.[5] That paper found that the most important downstream industry to be compensated is the electricity sector, but that it would be much less affected than the primary fuel sectors.

[4] Burtraw and Palmer, 2008.

[5] Bovenberg and Goulder (2001) considered the effect of a constant $25 allowance value sufficient to achieve emissions reductions of 18 percent in the long run.

Another study estimated that the reduction in equity value in the electricity sector would be equivalent to 6 percent of the total allowance value.[6]

Although harm to producers may be concentrated and visible to the politicians, consumers in the electricity sector would incur a loss approximately eight times as great as that of producers (Burtraw and Palmer 2008, U.S. EIA 2005). Consequently, the political economy of climate policy in the United States invites some form of compensation for consumers, at least as a transition to full implementation of CO_2 allowance auctions.

The obvious way in which compensation for electricity consumers can be achieved is through free allocation of emission allowances. Emissions allowances represent enormous economic value—tens of billions of dollars annually under a federal carbon policy—that arises due to the value placed on emissions within a cap-and-trade system. Paltsev et al. (2007) put the possible annual auction revenue at $130–$370 billion by 2015, an amount equivalent to $1,600 to $4,900 per family of four. The initial distribution of just a portion of the valuable emissions allowances represents a significant potential source of compensation. The enormous value of the allowances makes this high-stakes issue perhaps the greatest political challenge in designing climate policy.

This paper highlights the important role that market organization and regulatory institutions in the electricity sector play in affecting the efficacy of climate policy. Specifically, the regulatory setting plays a crucial role in determining whether free allocation will effectively deliver compensation to its intended recipients. The U.S. electricity sector is split so that about one-third of the electricity consumed from the power grid is sold at market-based competitive prices and the other two-thirds are sold under cost of service regulation.[7]

As mentioned above, one virtue of an auction is the possibility to direct revenues to purposes that reduce overall cost. Another virtue applies specifically to the electricity sector. In regulated regions, compared with free allocation, an auction approach tends to reduce the difference between price and marginal production cost for electricity generation—a source of inefficiency that is endemic to the electricity industry.[8] Within a

---

[6] Smith et al. (2002) estimated the effects of a 14 percent decrease in emissions to be achieved by 2010, and a 32 percent decrease by 2030.

[7] As of April 2007, the following jurisdictions had deregulated electricity markets: ME, NH, MA, CT, RI, NY, NJ,PA, DE, MD, DC, OH, MI, IL, and TX (EIA 2007c). In 2006 these states and the District of Columbia consumed 36 percent of all retailed electricity in the lower 48 states (EIA 2007b).

[8] Beamon et al. 2001. Burtraw et al. 2001. Burtraw et al. 2002. Parry 2005.

partial equilibrium model, the efficiency gains from using an auction in regulated settings can be at least as great as the gains from revenue recycling in a general equilibrium context.[9]

This paper incorporates the mechanisms of electricity price formation under competitive and regulated electricity markets in a detailed simulation model to investigate the magnitude of the effects that can be anticipated from alternative methods of allowance allocation within the electricity sector. We examine the effects on consumers under an auction of allowances, and under grandfathering – free distribution to incumbent electricity-generating firms. We contrast these approaches with three allocation schemes that are primarily aimed at compensating consumers. These all involve allocation to local distribution companies, the retail companies that deliver electricity to customers. The prices that these entities charge for electricity distribution are regulated throughout the United States and local distribution companies have been identified in legislative proposals as potential trustees to act on behalf of customers with respect to the allocation of emissions allowances. Various proposals have suggested allocation to local distribution companies be done on the basis of population, emissions or consumption.

Free allocation of emissions allowances to consumers or generators diverts revenues that otherwise could be dedicated to general tax relief, which offers efficiency gains and forms broad-based compensation for the diffuse effects of the policy on households. Free allocation also diverts revenues from other purposes, such as research initiatives or energy efficiency programs linked to climate policy. In the electricity sector free allocation also moves electricity price in regulated regions further away from marginal cost. Policymakers need to be cognizant of likely impacts on all affected parties, and they may want to limit and narrowly target free distribution of emissions allowances to better address a broader set of efficiency and compensation goals.

2 Analysis of the Electricity Sector

The electricity sector deserves special attention not only because of the emission intensity of its product, but also because of the long-lived nature of capital and the idiosyncratic way in which electricity markets are organized. Regulation of generation and retail services are generally left to states. However, because electrons flow freely

[9] Burtraw et al. 2001; Parry 2005.

over the wires and across state lines, the transmission grid is regulated by the federal government. The way that most states choose to regulate generation and retail services typically differs, and the choices vary across the states.

2.1 Institutions

Economists tend to think most markets are fundamentally competitive, at least in the long run. As a general principle, in competitive markets free allocation to firms will not benefit consumers because the economic value of a commodity in a competitive market is determined by its scarcity. Emissions allowances are a valuable asset, and as long as there is a liquid allowance market, a firm can sell allowances at the market price instead of using them for its own compliance responsibilities. The firm will recognize the lost opportunity for revenue from the sale of an allowance each time it uses the allowance itself for compliance. So in most markets economists would not expect to see consumers receive the benefit from free allocation to firms. Instead the value of emission allowances would be captured by shareholders who, in turn, would recognize their opportunity cost in production decisions.

The fact that a firm in a competitive market will charge its customers for the use of an asset that the firm has received for free is often a difficult idea for people to grasp, but it is wholly consistent with economic theory and it is in general what is observed in empirical studies. Indeed, sometimes economists seek evidence of noncompetitive behavior and "market power" by looking for instances when the price of a good differs from the cost of factor inputs used in its production. An emissions allowance in a cap-and-trade program is one such factor. If a firm did not pass through the cost of an allowance in the pricing of its product, it would be *prima facie* evidence of a noncompetitive market—and of possible market manipulation.

However, a substantial portion of the electricity sold in the United States is not traded in competitive markets, but instead sold in markets that are subject to cost-of-service regulation. In these cases regulators set prices to allow firms to recover their costs, which are usually calculated on an original-cost basis. If allowances are received for free by regulated electricity generators, then the addition to the cost basis for the purpose of cost recovery is zero. Roughly speaking, this situation applies to about two-thirds of the electricity customers in the country. In these areas the benefit of free allocation to emitters or producers can be expected to be passed on to consumers.

The contrast between regions with market-based electricity prices and regulated prices could yield asymmetric changes in retail electricity prices under free allocation to firms. These asymmetric effects on electricity consumers, in which free allocation to producers benefits consumers in regulated regions of the country, but not those in regions with market-based prices, introduce a challenging dilemma to climate policy.

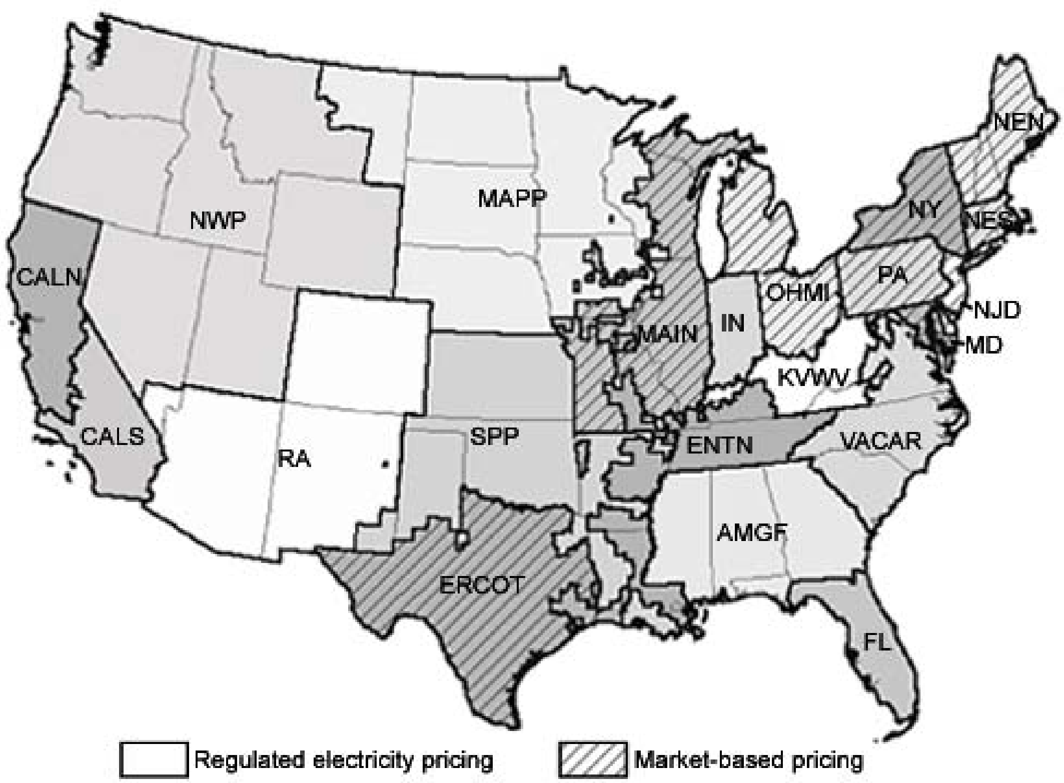
An alternative approach to compensation is allocation to local distribution companies (LDCs), the retail electricity companies that deliver electricity to customers and that could be directed to act as trustees on behalf of consumers. Although retail companies would see the cost of power in the wholesale power market increase under a cap-and-trade program, they would have substantial allowance value to rebate to consumers, and this would reduce the cost impact for their customers in competitive and regulated regions alike. Virtually the entire country is regulated in retail services, and some recent proposals, including the Lieberman–Warner climate proposal (SB 2191), would allocate some fraction of allowance value to LDCs for compensation to electricity consumers through rate reductions. This approach is expected to have the advantage of maintaining symmetry on a regional basis in the electricity sector.

2.2 Model

Several features of the market determine the relationship between CO_2 allowance price and the electricity price (Reinaud 2007). The most important are the fuel use, the portfolio of generation technologies, the nature of economic regulation and market structure, and the approach to allocation. To analyze these relationships we rely on a detailed simulation model of the electricity sector known as the Haiku Electricity Market Model, which is maintained by Resources for the Future. Haiku is a deterministic, highly parameterized model that calculates information similar to the National Energy Modeling System used by the Energy Information Administration, and the Integrated Planning Model developed by ICF Consulting and used by the U.S. Environmental Protection Agency (EPA). The Haiku model is distinguished from these models by its capacity to evaluate policy effects on consumer and producer surplus in the electricity sector and express these as a measure of economic welfare within the national and regional electricity markets.

The Haiku model simulates equilibrium in regional electricity markets and interregional electricity trade with an integrated algorithm for emission control technology choices for SO_2, NO_x, mercury and CO_2. The composition of electricity supply is calculated using a fully integrated algorithm for capacity planning and retirement coupled with system operation in temporally and geographically linked electricity markets. The model solves for electricity market equilibrium in 21 Haiku market regions (HMRs) within the continental United States. Each of the 21 HMRs is classified by its electricity pricing regime as having either competitive or regulated electricity pricing, as shown in Figure 2.2-1. Electricity markets are assumed to maintain their current regulatory status throughout the modeling horizon; that is, regions that have already moved to market-based pricing of generation continue that practice, and those that have not made that move remain regulated. The price of electricity to consumers does not vary by time of day in any region, though all customers in competitive regions face prices that vary from season to season.

Figure 2.2-1 Haiku Market Regions and Electricity Pricing



Each year is subdivided into three seasons (summer, winter, and spring-fall) and each season into four time blocks (superpeak, peak, shoulder, and base). For each time block, demand is modeled for three customer classes (residential, industrial, and commercial). Supply is represented using model plants that are aggregated according to their technology and fuel source from the complete set of commercial electricity generation plants in the country. Investment in new generation capacity and the retirement of existing facilities is determined endogenously in a dynamic framework, based on capacity-related costs of providing service in the future ("going forward costs"). Operation of the electricity system ("generator dispatch") in the model is based on the minimization of short-run variable costs of generation.

Equilibrium in interregional power trading is identified as the level of trading necessary to equilibrate regional marginal generation costs net of transmission costs and power losses. These interregional transactions are constrained by the level of the available interregional transmission capability as reported by the North American Electric Reliability Council (2003a, 2003b).[10] Factor prices, such as the cost of capital and labor, are held constant. Fuel prices are benchmarked to the forecasts of the Annual Energy Outlook 2007 for both level and elasticity (U.S. EIA 2007). Coal is differentiated along several dimensions, including fuel quality and content and location of supply; and both coal and natural gas prices are differentiated by point of delivery. The price of biomass fuel also varies by region depending on the mix of biomass types available and delivery costs. Other fuel prices are specified exogenously.

Emissions caps in the Haiku model, such as the Title IV cap on national SO_2 emissions, EPA's Clean Air Interstate Rule caps on emissions of SO_2 and NO_x, and the Regional Greenhouse Gas Initiative (RGGI) cap on CO_2 emissions, are imposed as constraints on the sum of emissions across all covered generation sources in the relevant region. Emissions of CO_2 from individual sources depend on emission rates, which vary by type of fuel and technology, and total fuel use at the facility. The sum of these emissions across all sources must be no greater than the total number of allowances available, including those issued for the current year and any unused allowances from previous years when banking is permitted. In this analysis, banking for CO_2 is not enabled. Rather, year-specific emission targets are taken from the Energy Information Administration analysis described below, to which the simulations are calibrated. This

[10] Some of the HMRs are not coterminous with North American Electric Reliability Council (NERC) regions and therefore NERC data cannot be used to parameterize transmission constraints. Haiku assumes no transmission constraints among OHMI, KVWV, and IN. NER and NEO are also assumed to trade power without constraints. The transmission constraints among the regions ENTN, VACAR, and AMGF, as well as those among MAACR, MD, and PA, are derived from version 2.1.9 of the Integrated Planning Model (EPA 2005). Additionally, starting in 2014, we include the incremental transfer capability associated with two new 500-KV transmission lines into and, in one case, through Maryland, which are modeled after a line proposed by Allegheny Electric Power and one proposed by PEPCO Holdings (CIER 2007). We also include the transmission capability between Long Island and PJM made possible by the Neptune line that began operation in 2007.

approach allows for a more transparent comparison of the effects of different approaches to allocation because the quantity of emissions in each year is held constant.

3 Scenarios

One way that electricity consumers can be compensated directly is for each citizen to receive allowance value directly. This approach has recently been described as "cap and dividend" because the allowance value would be refunded as a dividend on a per capita basis. This approach would be among the most progressive in its distributional consequences (Burtraw et al. 2008; Boyce and Riddle 2007). The Center on Budget and Policy Priorities (2007) identify another approach that would take advantage of information about household income to target the most disadvantaged households using just a portion of the allowance value.

Environmental advocates typically take a different view, however, aiming to direct auction revenue to complementary initiatives to reduce emissions. For example, the Model Rule for the 10 northeastern U.S. states in RGGI specifies that each state must allocate at least 25 percent of its budgeted allowances to a consumer benefit or strategic energy purpose. These "consumer benefit" allowances are to be sold or otherwise distributed to promote energy efficiency, to directly mitigate electricity ratepayer impacts, or to promote lower-carbon-emitting energy technologies. (Most of these ten states have indicated their intention to auction nearly 100 percent of their budgeted allowances.) Ruth et al. (2008) found the dedication of 25 percent of the allowance value to investments in end-use efficiency would offset any increase in retail electricity price from the policy. A similar plan to direct a portion of allowance value to strategic energy purposes is part of the European Commission's proposal for moving to an auction in the EU ETS beginning in 2013. The merits of this strategy rest on the belief that there exist market barriers that prevent the realization of opportunities for improving efficiency in the end-use of energy or to bringing renewable energy sources to market. The merits rest as well on the ability to design institutions that can use allowance value effectively to overcome these barriers. Other claims for allowance value are based on the need to accelerate the adaptation to climate change. Atmospheric scientists tell us that we are already at the point where some climate warming is inevitable and that adaptation will be necessary. Adaptation will involve significant investment by the private and public sectors. An auction provides revenues that could be directed to this variety of purposes.

The scenarios we model are limited to the electricity sector, but they capture the heart of the debate regarding treatment of allocation for that sector in the United States.

The allowance distribution plan for Lieberman–Warner (S 2191) dedicates 22 percent of the allowances in the year 2012 to states in one fashion or another.[11] One major portion is directed to electricity local distribution companies (9 percent) and natural gas distribution companies (2 percent). These allocations are intended to address a variety of purposes including promotion of investment in end-use efficiency or direct rate relief for customers. Other proposals have been even more aggressive.[12] The Jeffords bill in 2002 would have allocated two-thirds of emissions allowances to the states for determination of ultimate allocation by trustees. It would be plausible for this decision to be left to the state public utility commissions, who would act as trustees on behalf of consumers.

3.1 Alternative Methods for an Initial Distribution of Emissions Allowances

The Lieberman–Warner proposal includes a cap-and-trade system for the entire economy with point of compliance at upstream fuel supply in almost every case. In general, the policy would require fuel suppliers to surrender allowances equal to the carbon content of the fuel and byproducts that they sell or consume in their refining and manufacturing processes. The exception is coal-fired power plants, which would have compliance responsibility at the point of consumption. For natural gas use in the downstream electricity sector the cost of the cap-and-trade system would be perceived as a change in the relative cost of fuel. Fuel with relatively high carbon content would be expected to have a higher price because of the opportunity cost of emissions allowances that fuel suppliers would have to surrender to bring that fuel to market. For coal-fired power plants the cost of power would depend on the method of allocation and the type of regulation in place.

In general, it is vital to recognize that the point of allocation of emissions allowances is distinct from the point of compliance. We evaluate several methods for the initial distribution of emissions allowances (Sterner and Muller 2007). One alternative is **upstream allocation**, with all emissions allowances distributed initially to fuel suppliers and with no allowances distributed to the electricity sector. Within the electricity sector, this approach is equivalent to an **auction** regardless of how the allowances are actually distributed to fuel suppliers because electricity generators purchase their emissions allowances bundled along with their fuel through an increase in the price of fuel.

---

[11] The remainder are allocated using a mix of free allocation to industry and an auction.

[12] The National Association of Regulatory Utility Commissioners (April 21, 2008) has called for 100 percent of the allowances to be distributed for free in the electricity sector to be given to LDCs.

As one alternative to an auction, we consider free distribution of allowances to incumbent firms in the electricity sector on the basis of historic measures of electricity generation. This approach is often called **grandfathering** because it distributes allowances without charge to incumbents in the industry. Another approach, which we do not explore here, is to regularly **update** the calculation underlying the allowance distribution based on current- or recent-year data. Like distribution based on historic data, an updating approach distributes allowances free of charge and also could distribute them according to various measures, such as the share of electricity generation or heat input (a measure related to fuel use) or a share of emissions at a facility (Burtraw et al. 2001; Fischer and Fox 2004; Rosendahl 2008). An updating approach leads to lower electricity prices than an auction or historic approach and is expected to have greater social costs because it does not provide the same incentive through higher prices for consumers to improve the efficiency of energy use (Burtraw et al. 2006).

The focus of this analysis is the modeling of allocation to local distribution companies, the retail companies that directly serve customers. This approach is described as "**allocation to load"** or **"load-based allocation**" because in one form or another it would allocate to customer demand for electricity (load). We model this at the level of 21 regions in our model, and based on three different metrics. One is the portion of electricity consumption in each region, a second is the portion of population and the third is the portion of emissions. These metrics are calculated on a one-time basis, drawing on the baseline model run for each simulation year in the model. The value of emissions allowances under allocation to load is used to offset the average cost of electricity directly, for example by offsetting the transmission and distribution charge. Although electricity prices vary by customer class because of varying time profiles of demand and different shares of transmission charges assigned to each class, we assume a uniform distribution of the value of allowances in reducing electricity price across all customers.

3.2 Baseline

The baseline scenario is constructed to incorporate all major federal legislation governing airborne emissions from the electricity sector including Title IV and CAIR for SO_2 emissions, the annual and ozone season caps on emissions of NO_x under CAIR, and CAMR for mercury emissions. Also included are some state level legislation, including RGGI, and other policies that are specific to individual states. For nuclear capacity additions, Haiku uses the regional output of the EIA National Energy Modeling System for 2007 as capacity limits on new construction of nuclear plants. All of these potential capacity additions are east of the Mississippi River (U.S. EIA 2007).

Two of the most important baseline scenario assumptions are the treatment of Federal Renewable Energy Production Tax Credit (REPTC) and of state level Renewable Portfolio Standards (RPS) in several western states, including California. The REPTC provides a production tax credit of $19/MWh to new wind, geothermal, and dedicated biomass generators, and a credit of $9.50/MWh is available to new landfill gas and non-dedicated biomass generators.[13] Since the federal REPTC has repeatedly been renewed just prior to lapsing and has actually lapsed three times for a total of 16 months (over the 15 years since it was initially instituted) before being reinstituted, it is modeled in perpetuity in Haiku as a tax credit that is received with 90 percent probability. The state level RPS mandates within the Western Electricity Coordinating Council (WECC) region require substantial increases in renewables generation in the coming years. The resulting capacity additions are not modeled endogenously within Haiku. Instead, we force new renewable capacity into our model in order to meet these standards in the western states[14] according to forecasts provided by Energy and Environmental Economics, Inc (E3).[15] These forecasts of renewable resource additions include the planned capacity additions for wind, geothermal and biomass reported by the Transmission Expansion Planning Policy Committee (TEPPC) along with additional resources that E3 forecasts would be needed to meet RPS standards. These renewable policies are significant because of their potency in reducing emissions, but also because by including them in the baseline it reduces the cost of achieving a specific emissions cap under the policy scenario.

3.3 Policy Scenario

The emissions reduction targets that we model are taken from the U.S. Energy Information Administration modeling of the Lieberman–Warner proposal (U.S. EIA 2008). From that modeling we take the CO_2 emissions determined at the national level as given, and we assume it is not affected by small changes in the electricity sector that result under the variations of policies we model. We do not allow inter-annual banking in the runs of our model, although it is implicit in the quantity targets we take from EIA.

---

[13] All values are reported in 2004$ unless indicated.

[14] The western states for which we forced in new renewables capacity include California, Arizona, Montana, Colorado, New Mexico, Utah, Nevada and Wyoming.

[15] "Electricity and Natural Gas GHG Modeling: Methodology and Key Revisions," Slides 38-39, April 21, 2008; <http://www.ethree.com/GHG/E3_CPUC_GHG_21April08.pdf>.

Investment and operational decisions in our model respond to this fixed emission target. In reality (as opposed to in the model), the electricity sector decisions would play a role in the determination of the electricity sector's share of national emissions that obtain under each policy scenario, but we maintain the fixed quantity to achieve comparability across scenarios. Since the emission quantity is the same and the models are different, our model will result in a different level of allowance price and electricity price across scenarios and different from that obtained in the EIA exercise.

In addition to the no-policy baseline, five policy scenarios are modeled. These are comprised of an allowance auction, free allowances to incumbent generating firms (grandfathering), and the three allocation-to-load scenarios described above, based on consumption, population and emissions.

4 Results

The effect of climate policy on electricity consumers depends on several factors that vary by region of the country including the fuel mix and technology used for generating electricity, economic regulation and the approach to allocation. Moreover these factors interact. For example, the portfolio of generation technology determines the fuel that is used at the margin at different times of day and year, and the economic regulation in the region determines whether changes in average or marginal cost determine changes in electricity price. This analysis focuses on the role of allocation, but highlights the important interactions among all these factors, particularly how different approaches to allocation can have different effects depending whether markets are regulated.

If allowances are allocated upstream or auctioned to electricity producers, the opportunity cost of the allowances would be reflected in the price of electricity in both regions with competitive electricity markets and regions subject to cost-of-service regulation. We find that if allowances are allocated for free to generators on the basis of historic generation, the effect on electricity prices and thus on consumers would depend on whether electricity markets are regulated or not. Allocating allowances to local distribution companies would largely erase these inter-regional differences based on regulation, with remaining differences across regions being largely a function of the mix of resources used to generate electricity within a region.

This modeling exercise was performed for a horizon beginning in 2010 and ending in 2025. For expositional simplicity this paper focuses on the results obtained for the year 2020.

4.1 Allowance Allocation and Consumers

The quantitative effect of different approaches to allowance allocation on average electricity prices is shown in Table 4.1-1. The table reports price effects in each Haiku market region as well as aggregate price effects aggregated into regions defined by geography and by regulatory institutions. The auction has the biggest effect on electricity prices in both types of regions and nationwide. Prices in competitive regions increase by $8.50 per MWh with an auction, while in regulated regions the increase in average price due to the policy is closer to $6 per MWh. The difference is related to the differences in resource mix between the two types of regions and the difference in regulation. Under cost-of-service regulation, allowance costs from an auction are passed through to consumers in a way that makes the generators whole and thus these costs are averaged over all MWh sold. However, in competitive regions, the allowance cost to the marginal generator is passed through in the market-determined price that is charged for all electricity, which may be generated with an emitting technology or a non-emitting technology at any point in time. Many generators, particularly those with substantial reliance on non-emitting technologies like nuclear or hydro, will earn rents from this type of pricing (Burtraw and Palmer, 2008). Nationwide the price increase under an auction averages $7 per MWh.[16]

When allowances are grandfathered to generators based on historic generation the inevitable effect of the policy is an electricity price reduction in regulated regions relative to the auction scenario. This reduction leads to increased consumption of electricity, more power imported from neighboring competitive regions and a resultant increase in the electricity price in the competitive regions relative to the auction. Relative to the baseline, competitive regions will see an increase in electricity prices of nearly $10 per MWh under grandfathering, while regulated regions will actually experience a small decline in prices of $1 per MWh. The decline in prices is made possible by the disconnect in

[16] This price effect is difficult to compare directly with the EIA analysis. US EIA (2007) models a mixed allocation of some auction, some free allocation. Also they do not model the continuing availability of the REPTC, even on a probabilistic basis. Further, the EIA analysis of an economy-wide policy will have broader effects in primary fuel markets, especially with respect to the change in demand for natural gas and resultant price change. The demand response we capture is only that pertaining to the electricity sector. US EIA finds the change in electricity price to be $4.7/MWh (2004$).

regulated regions between marginal and average costs. Marginal costs will rise by nearly $10 per MWh, but average costs will fall as a result of the displacement of relatively carbon intensive generation resources in the baseload part of the supply curve with less carbon intensive resources, especially subsidized renewable resources. This also allows for the profitable export of grandfathered allowances. Nationwide the price increase under grandfathering will be roughly one-third of what it will be with an auction.

Table 4.1-1. Change in Electricity Price by Region and Allocation Method in 2020 (2004$/MWh)

Region	Auction	Grandfathering	Load-Based (population)
Regulated Regions	6.1	(1.0)	(0.0)
Competitive Regions	8.5	9.9	1.8
National	7.0	2.7	0.6

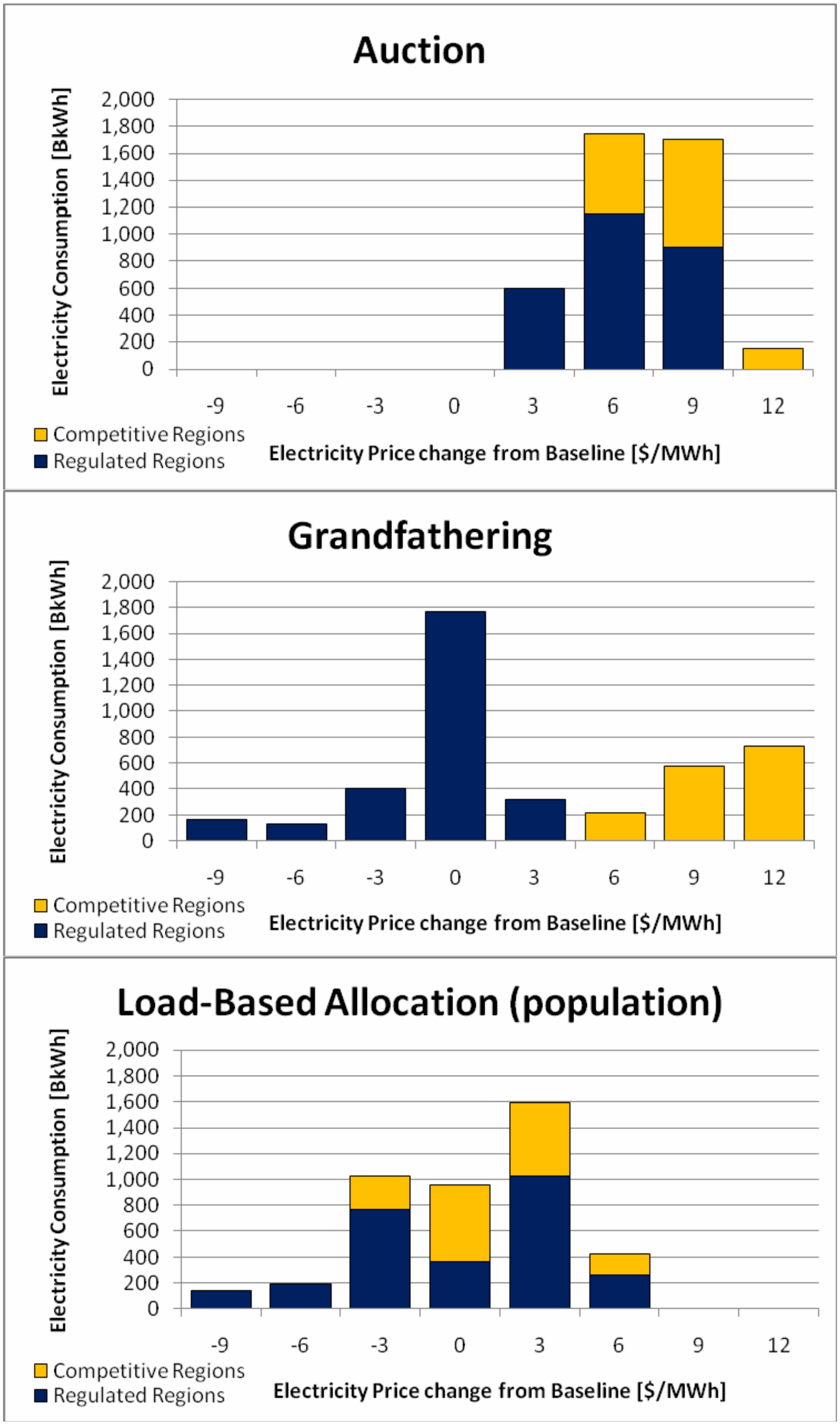
The asymmetric consequence of grandfathering emission allowances is illustrated in the first two panels of Figure 4.1-1, which illustrates the distributions of price changes across regions under different allocation approaches. The graphs in this figure provide histograms of the frequency of various levels of electricity price change resulting from the cap and trade policy under different allocation approaches. The horizontal axis in each graph indicates the size of the change in electricity price while the vertical axis indicates the number of billion kWh of electricity consumption that experience each level of price change from the policy. Changes associated with regulated and competitive regions are indicated in contrasting shading.

The top panel of the figure shows how electricity prices change in 2020 in response to a climate policy with an allowance auction. Under the auction we see that the average change in electricity price nationwide is roughly $7 per MWh with impacts in particular regions varying between $1.80 and $10.60 per MWh. This graph shows that there is much overlap between price impacts in regulated and competitive regions under this policy. The main differences in the distribution of the change in price result from the fuels and technologies used to generate electricity in each region. There remain important differences between regulated and competitive regions in the rates at which compliance costs are passed through to customers as changes in electricity prices because of the difference in average and marginal cost pricing. However, the difference among regions under an auction, and hence the explanation for the distribution of the change in costs across regions, is fundamentally driven by the change in the emission intensity of electricity generation. The middle panel of the graph shows the distribution of price effects across regions under grandfathering. This figure illustrates the dramatic difference between regulated regions and competitive regions. In regulated regions, consumers benefit from a grandfathering approach and price increases are much reduced. As shown in appendix Table 1, in two regions, Indiana and a region spanning Kentucky, West Virginia and a small part of Virginia, prices actually fall by roughly $7 and $9 per MWh, respectively, while several other regions experience price declines of between $0.20 and $1.80 per MWh. In competitive regions the distribution of price impacts actually shifts slightly to the right, reflecting the increase in generation costs associated with the increase in demand (due to the lower price) in regulated regions, compared to an auction. The biggest price increases occur in Pennsylvania (PA) and the region that includes Ohio and Michigan (OHMI).

The asymmetry in price effects under grandfathering between regions under different pricing regimes has posed one of the major political challenges to the design of climate policy in the United States. While a point of departure for policy design from the perspective of the electricity industry has been free allocation of emission allowances, analysis has informed industry members of their opposing interests, depending on what kind of region they are in, and increasingly consumer interests have taken notice. The emerging attention being given to load-based allocation is potentially one way for the industry to navigate these issues.

Compared to an auction, load-based allocation attenuates price increases from climate policy in both competitive and regulated regions. As a point of departure, we consider first LBA on the basis of population. In practice this would be implemented by initially apportioning allowances to the service territories of individual retail utilities, or more probably to states and charging state public utility commissions to complete the apportionment to service territories. In our model, this is implemented by apportioning allowances among the 21 market regions according to population. Table 4.1-1 indicates that in competitive regions the price increase from the policy falls from $8.50 per MWh with an auction to $1.80 per MWh with the load-based allocation approach. In regulated regions, the average price of electricity is unchanged under the climate policy with load-based allocation compared to the baseline. Nationwide electricity price increases by $0.60 per MWh. In general, load-based allocation dramatically reduces the effect of the policy on consumers in both competitive and regulated regions relative to the auction scenario.

Figure 4.1-1. Distribution of Electricity Price Effects in 2020



The bottom panel of Figure 4.1-1 illustrates this effect in a histogram that can be compared with the other approaches to allocation. To a rough approximation, load based allocation restores the symmetry in the price impacts on regulated and competitive regions that would be observed under an auction, albeit at much lower levels. This has made load-based allocation increasingly popular to overcome political opposition to the effect of climate policy on electricity prices.

Giving allowances away for free will help to soften the impact of the policy on consumers, but it will do so at a cost. By reducing prices, the load-based allocation approach and, for consumers in regulated regions, the grandfathering approach, mute the incentive to conserve electricity that exists with an auction. The effective subsidy to electricity consumption causes generation to be higher, leading to a higher demand for CO_2 allowances. This results in an increase in allowance price, which will spread throughout the economy under an economy-wide cap-and-trade program. Allowance prices in 2020 under the different allocation approaches are reported in Table 4.1-2. The table indicates that allowance price in 2020 rises from $14.10 per ton CO_2 under an auction to $15.30 under grandfathering and even higher, to $15.80, under a load-based approach.

Table 4.1-2. National CO_2 Allowance Price in 2020 (2004$/ton)

Auction	Grandfathering	Load-Based (population)
14.1	15.3	15.8

4.2 Different Flavors of Load Based Allocation

In a second set of simulation runs we consider three alternative bases for determining a region's load-based allocation of emission allowances. Previously we considered total population in the region, and in addition we consider total electricity consumption and total emissions from electricity generation.[17] Under load-based allocation based on population, heavily populated regions would receive a greater share of the allowance value than less populated ones. In comparison, the consumption-based approach would tend to favor consumers who reside in regions where electricity

[17] The National Association of Regulatory Utility Commissioners (April 21, 2008) call for load-based allocation on the basis of historic emissions.

consumption per capita is higher and the emissions-based approach would favor consumers in coal-intensive regions. The emissions-based approach introduces a more prominent role for the resource mix in the determination of each region's share of emission allowances, as occurs with grandfathering, except in this case the benefits of free allocation in both regulated and competitive regions accrue to consumers instead of to generators.

Varying the basis for allocation to local distribution companies from population to another measure will have different effects on electricity prices in different regions. Appendix Table 2 shows some of the factors underlying those differences. For example, both Northern (CALN) and Southern California (CALS) have low consumption per capita and conversely relatively high population per unit of consumption relative to many other regions in the model, and thus would receive a larger share of emissions allowances under a population-based allocation than a consumption-based approach. A population-based approach would produce a more substantial price reduction, especially in Northern California, which has a low rate of electricity consumption per capita.[18] An emissions-based approach also would not be favorable for California, which has a relatively clean portfolio of generators. In contrast, a coal-intensive region such as the one including much of Kentucky, part of Virginia and West Virginia (KVWV) has relatively high CO_2 emissions per capita and thus consumers in that state are expected to find an emissions-based approach to allocation to be more favorable.

To provide a summary of these different regional effects we aggregate the 21 Haiku market regions into six regions:

- Northeast states in the Regional Greenhouse Gas Initiative (RGGI)
- Southeast
- Midwest and Appalachia
- Plains
- Rocky Mountains and Northwest
- California

The composition of each region is shown in the map displayed in Figure 4.2-1.

[18] Note that the measures of electricity consumption per capital reported in the table include total consumption by all classes of customers in the state divided by total population.

Figure 4.2-1 Aggregated Regions

Table 4.2-1 reports the regional electricity price changes in 2020 under the three different approaches to load-based allocation. The last row of the table reports the average price changes for the nation as a whole and shows that varying the basis for load-based allocation does not have much impact on average electricity price nationwide. Under all three load-based scenarios national average electricity price rises by less than a dollar per MWh, which is substantially less than the price difference resulting under an auction. However, this apparent similarity masks some substantial differences across regions.

The biggest difference in price effects across the different load-based approaches occurs in California, where allocation to local distribution companies based on population yields an average electricity price that is $8.50 below baseline levels (e.g. price actually falls under the climate policy) while an allocation based on emissions yields a price that is $3.10 above baseline levels, on par with the price increase experienced under grandfathering. If allocation to local distribution companies is based on electricity consumption, the average price is also lower than in the baseline, by $3.60.

Table 4.2-1. Change in Electricity Price by Region and Approach to Load-Based Allocation in 2020 (2004$/MWh)

Region	Load-Based (population)	Load-Based (consumption)	Load-Based (emissions)
RGGI	(1.4)	2.0	5.3
Southeast	0.6	(0.8)	(0.3)
Midwest and Appalachia	4.2	3.5	0.5
Plains	2.3	1.5	0.0
California	(8.5)	(3.6)	2.9
Rockies and Northwest	(2.6)	(2.2)	(2.1)
Competitive	1.8	2.7	2.6
Regulated	(0.0)	(0.6)	(0.7)
National	0.6	0.6	0.4

The second biggest differences in price effects across the three load-based allocation scenarios are found in RGGI. Electricity consumers in the RGGI states would clearly be better off when allocation is based on population and average price is $1.40 per MWh lower than in the absence of a climate policy. Under the consumption-based allocation to local distribution companies, prices in the RGGI states increase by $2 per MWh. Consumers in this region are least well off under an emissions based approach, which produces average price increases under the policy of $5.30 per MWh. Some parts of this region experience substantially higher price increases; in northern New England, which relies heavily on hydro and nuclear power, average electricity prices will increase by $9.20 under the emissions-based approach.

The only region to experience price increases under all three load-based approaches is the Midwest and Appalachia. In this region average price rises by $4.20 in the per capita scenario, driven in large part by even bigger increases in coal-rich Kentucky and West Virginia. When emissions are used as the basis for allocation to load, the average electricity price rises by only $0.50. However, this increase masks large declines in price in the states of Kentucky, West Virginia and Indiana that are offset by price increases in the region that combines Illinois and Wisconsin, which has a fair amount of nuclear generation. Customers in the plains states also experience price increases larger than the national average under the population and consumption based approaches. However, average price remains virtually on par with baseline levels under the emissions based approach.

Our findings suggest that electricity consumers in the western region excluding California should be largely indifferent between the three approaches to load-based allocation, all of which produce price drops in 2020 of roughly $2.00. Closer examination of the results for the two Haiku market regions that comprise this larger region suggest slightly greater differences in price effects between the consumption-based and the emissions-based approaches. In particular the consumption based approach leads to slightly larger price drops in the northern part of this region, which is rich in hydro resources as well as other types of renewables. Consumers in the southern part of this region fare better under an emissions based approach as coal plays a greater role in the generation mix there.

4.3 Efficiency Consequences

The beneficial effects of load-based allocation accrue to electricity consumers through lower electricity prices; however, the downside of a load-based approach is that it yields a higher allowance price than would prevail under an auction. This effect on allowance price is essentially invariant with respect to the basis on which allowances are allocated to local distribution companies. Table 4.3-1 shows the CO_2 allowance price under all three approaches and reveals that allowance prices are little changed across the three scenarios, and they each lead to significant differences in allowance price compared to an auction.

Table 4.3-1. National CO_2 Allowance Price in 2020 (2004$/ton)

Load-Based (population)	Load-Based (consumption)	Load-Based (emissions)
15.8	16.0	16.0

From a sector-specific perspective, the difference in allowance price is not significant, but within the broader economy it signals that greater use of resources and greater cost would be required to achieve emission reduction goals. Any kind of free allocation including grandfathering will raise the allowance price, but the load-based approach does so most importantly. While grandfathering is generally intended to compensate the owners of incumbent facilities that will be regulated by climate policy, in contrast, free allocation to load is a subsidy to consumers of electricity. We find it would raise the price of allowances by about 12 percent compared to an auction. As a consequence, within a nationwide cap-and-trade program, other actors in the economy such as industries that use natural gas, or households that use fuels for home heating, or people who drive cars, would pay for this subsidy to electricity by higher prices for the use of energy elsewhere in the economy.

The subsidy to electricity consumption has the effect of reducing the incentive for consumers to make investments in end-use efficiency. Recent analyses (McKinsey 2007, Nadel et al. 2004) suggest that there are substantial opportunities to improve the efficiency with which electricity is used in the economy. While government programs and standards may contribute to the realization of these efficiency gains, another important component is the capital purchase decisions of individuals. If electricity price rises less due to free allocation to electricity consumers then those consumers will have less incentive to purchase efficient air conditioners, refrigerators, etc., causing other sectors to do more work to achieve overall emission reductions.

Electricity consumers, and the industry that supplies them with power, have a parochial interest in trying to lessen the impact of climate policy on prices and in capturing the value for the electricity sector associated with placing a scarcity value on CO_2 in the economy. However, there is no economic logic why the value of emission allowances should be reserved for a sector just because it has historically been the source of emissions. The parochial assignment of allowance value to any one sector of the economy could lead to different marginal costs for emission reductions throughout the economy, and it could lead to some sectors having to achieve greater emission reductions than would be efficient overall, which offers the prospect of raising the cost from a nationwide perspective.

5 Conclusion

It is noteworthy that precisely because the cost of climate policy is large, a good way to achieve broad-based compensation is to reduce the overall social cost of the policy. Recycling revenue raised under an allowance auction to reduce preexisting taxes, helps achieves efficiency goals and these achievements are compounded since this approach reduces the overall cost of climate policy, thereby lessening the impact on households overall. However, it would not succeed in compensating lower income households who spend a larger portion of their income on energy than wealthier households who would benefit the most from revenue recycling. Burtraw et al. (2008) find that a proportional reduction in labor income taxes would be highly regressive.

One approach to compensating households that is embodied in current legislative proposals is free allocation to electricity customers, to be achieved by free allocation to local distribution companies. This approach seems appealing because customers may have relatively little opportunity to reduce electricity use in the short term, at least until they have the opportunity to make new capital investments in appliances, home weatherization, etc.

From the national perspective, we find significant benefits for electricity consumers from free allocation to local distribution companies. In addition, this approach reconciles the important regional differences that would emerge with a grandfathering approach that distributed allowances for free to incumbent emitters. However, the benefits that emerge at the national level mask important differences across regions that depend on how the allocations are determined. Allocation to local distribution companies based on population could yield electricity prices in 2020 for populous regions with relatively clean sources of electricity generation that are actually below prices in the absence of climate policy. Consumers residing in regions that rely heavily on coal will tend to fare better under an approach that uses emissions to determine allocation.

We also find free allocation to local distribution companies comes with an important efficiency cost, not just in a general equilibrium context stemming from foregone revenue but also due to the market dynamics in the regulated industries. When electricity customers do not see the increase in retail electricity prices, they do not have an incentive to reduce electricity consumption. Across the sector, this effect would lead to more electricity consumption, and under an economy-wide program, it would lead to more emissions from the electricity sector, requiring more reductions from other sectors. This is expected to raise the overall cost of achieving climate goals. However, the political virtue of this approach is that using allocation to load provides a mechanism in the short run to avoid sudden changes in electricity prices for consumers. Because free allocation to customers has the political virtue of lessening the price effect, it may provide for a useful transition path to phasing in a full auction in the electricity sector.

Appendix

Appendix Table 1. Change in Electricity Price from by Region in 2020 (2004$/MWh)

Electricity Pricing Regime	Haiku Market Region	Auction	Grandfathering	Load-Based (population)	Load-Based (consumption)	Load-Based (emissions)
Competition	NEN	9.2	10.5	0.7	3.9	9.2
	NES	5.4	6.0	(3.6)	(0.6)	1.7
	NY	9.0	10.0	(2.0)	2.9	7.1
	NJD	6.4	7.0	(0.6)	1.6	5.3
	MD	8.4	9.9	0.8	3.3	4.7
	PA	10.6	12.1	4.6	4.6	0.8
	OHMI	10.1	11.5	3.4	3.4	1.4
	MAIN	9.7	10.6	3.6	3.6	3.5
	ERCOT	5.8	7.9	1.2	1.1	0.5
Regulation	FRCC	4.1	0.7	(2.6)	(2.5)	(1.1)
	AMGF	6.6	(0.2)	2.1	0.3	(0.2)
	VACAR	4.8	(0.4)	0.0	(1.8)	0.3
	KVWV	9.5	(9.3)	5.7	2.6	(4.1)
	IN	9.2	(7.4)	4.8	3.2	(3.3)
	ENTN	7.5	(1.1)	2.7	1.0	(0.7)
	SPP	8.0	0.2	3.6	1.5	(0.6)
	MAPP	7.9	(1.8)	2.7	1.8	(0.1)
	NWP	5.5	(0.1)	(2.3)	(2.7)	(0.6)
	RA	5.8	(1.7)	(2.9)	(1.4)	(4.2)
	CALN	1.8	2.6	(9.8)	(5.0)	2.0
	CALS	2.7	4.1	(7.5)	(2.6)	3.6
	RGGI	7.6	8.5	(1.4)	2.0	5.3
	Southeast	5.8	(0.3)	0.6	(0.8)	(0.3)
	Midwest and Appalachia	9.9	5.4	4.2	3.5	0.5
	Plains	7.1	3.1	2.3	1.5	0.0
	California	2.3	3.5	(8.5)	(3.6)	2.9
	Rockies and Northwest	5.7	(0.8)	(2.6)	(2.2)	(2.1)
	Competitive Regions	8.5	9.9	1.8	2.7	2.6
	Regulated Regions	6.1	(1.0)	(0.0)	(0.6)	(0.7)
	National	7.0	2.7	0.6	0.6	0.4

Appendix Table 2. Baseline Regional Characteristics in 2020

Electricity Pricing Regime	Haiku Market Region	Aggregate Region	Population	CO2 Emissions Rate (tons/MWh)	Electricity Consumption per Capita (MWh/person)	CO2 Emissions per Capita (tons/person)
Competition	NEN	RGGI	3,624,102	0.13	8.8	1.5
	NES	RGGI	11,685,426	0.44	9.1	3.8
	NY	RGGI	19,576,920	0.35	7.6	2.4
	NJD	RGGI	10,905,384	0.51	10.4	3.3
	MD	RGGI	6,497,626	0.68	10.3	5.1
	PA	Midwest and Appalachia	12,787,354	0.61	13.0	12.7
	OHMI	Midwest and Appalachia	21,998,225	0.85	13.2	11.1
	MAIN	Midwest and Appalachia	22,050,673	0.62	13.1	8.7
	ERCOT	Plains	25,040,400	0.63	15.0	9.7
Regulation	FRCC	Southeast	22,140,641	0.61	12.9	6.7
	AMGF	Southeast	19,883,364	0.62	17.9	12.2
	VACAR	Southeast	22,940,469	0.50	16.1	7.8
	KVWV	Midwest and Appalachia	6,976,401	1.01	21.6	27.2
	IN	Midwest and Appalachia	6,627,008	1.07	18.0	21.6
	ENTN	Southeast	15,470,035	0.68	17.0	13.5
	SPP	Plains	12,287,603	0.73	18.1	15.0
	MAPP	Plains	13,039,828	0.69	15.0	12.2
	NWP	Rockies and Northwest	18,976,707	0.34	14.2	6.6
	RA	Rockies and Northwest	18,902,845	0.67	10.7	9.8
	CALN	California	17,379,247	0.12	7.8	0.8
	CALS	California	24,827,496	0.21	7.5	0.7
		RGGI	52,289,458	0.42	8.9	3.2
		Southeast	80,434,509	0.60	15.8	9.7
		Midwest and Appalachia	70,439,660	0.79	14.4	13.2
		Plains	50,367,830	0.67	15.8	11.6
		California	42,206,743	0.16	7.6	0.7
		Rockies and Northwest	37,879,552	0.48	12.5	8.2
		Competitive Regions	134,166,109	0.61	11.8	7.5
		Regulated Regions	199,451,643	0.61	13.8	9.0
		National	333,617,752	0.61	13.0	8.4

References

Åhman, M., J. Kruger, and L. Zetterberg 2007. "A Ten-Year Rule to Guide the Allocation of EU Emission Allowances," *Energy Policy* 35 (3):1718-1730.

Åhman, Markus, Kristina Holmgren, 2006. "New entrant allocation in the Nordic energy sectors: incentives and options in the EU ETS." *Climate Policy* 6: 423-440.

Beamon, J.A., Leckey, T., and Martin, L. 2001. "Power Plant Emission Reductions Using a Generation Performance Standard." Draft. Washington, DC: U.S. Department of Energy, Energy Information Administration.

Bovenberg, A., de Mooij, R. 1994. "Environmental Levies and Distortionary Taxation." *American Economic Review* 84: 1085–89.

Bovenberg, A.L., and Goulder, L.H. 1996. "Optimal Environmental Taxation in the Presence of Other Taxes: General Equilibrium Analyses." *American Economic Review* 86: 985–1000.

Bovenberg, A.L., and Goulder, L.H. 2001. "Neutralizing the Adverse Industry Impacts of CO2 Abatement policies: What Does it Cost?" In C. Carraro and G. Metcalf (eds.), *Behavioral and Distributional Effects of Environmental Policy*. Chicago: University of Chicago Press.

Boyce, James K., and Matthew Riddle. 2007. Cap and Dividend: How to Curb Global Warming while Protecting the Incomes of American Families. Working paper No. 150. University of Massachusetts, Amherst, MA: Political Economy Research Institute.

Burtraw, D., R. Bharvirkar, A. Paul and K. Palmer, 2001. "The Effect of Allowance Allocation on the Cost of Carbon Emission Trading," Resources for the Future Discussion Paper 01-30.

Burtraw, D., Palmer, K., Bharvirkar, R., and Paul, A. 2002. "The Effect on Asset Values of the Allocation of Carbon Dioxide Emission Allowances." *The Electricity Journal* 15(5): 51–62.

Burtraw, D., Kahn, D., Palmer, K., 2006. CO2 allowance allocation in the regional greenhouse gas initiative and the effect on electricity investors. *Electricity Journal* 19(2), 79–90.

Burtraw, D. and K. Palmer 2008. "Compensation Rules for Climate Policy in the Electricity Sector," *Journal of Policy Analysis and Management,* (forthcoming).

Burtraw, D., R. Sweeney and M. Walls 2008. "Where You Stand Depends on Where You Sit: The Regional Incidence of Climate Policy," Resources for the Future.

Center on Budget and Policy Priorities. 2007. Climate-Change Policies Can Treat Poor Families Fairly and be Fiscally Responsible. Washington D.C.: Center on Budget and Policy Priorities.

Fischer, Carolyn, and Alan Fox. 2004. Output-Based Allocations of Emissions Permits: Efficiency and Distributional Effects in a General Equilibrium Setting with Taxes and Trade. RFF Discussion Paper 04-37. Washington, DC: Resources for the Future.

Goulder, L.H., Parry, I.W.H., Williams III, R.C., and Burtraw, D. 1999. "The Cost-Effectiveness of Alternative Instruments for Environmental Protection in a Second-Best Setting." *Journal of Public Economics* 72(3): 329–360.

McKinsey and Company. 2007. *Reducing U.S. Greenhouse Gas Emissions: How Much at What Cost?* US Greenhouse Gas Abatement Mapping Initiative, Executive Report, December.

Nadel, Steven, Anna Shipley and R. Neal Elliot. 2004. "The Technical, Economic and Achievable Potential for Energy Efficiency in the U.S. – A Meta-Analysis of Recent Studies", *Proceedings of the 2004 ACEEE Summer Study on Energy Efficiency in Buildings*.

Paltsev, Sergey, et al. 2007. *Assessment of U.S. Cap-and-Trade Proposals*. MIT Joint Program on the Science and Policy of Global Change, Report No. 146 (April).

Parry, I.W.H., Williams, R.C., and Goulder, L.H. 1999. "When Can Carbon Abatement Policies Increase Welfare? The Fundamental Role of Distorted Factor Markets." *Journal of Environmental Economics and Management* 37(1): 52–84.

Parry, I.W.H. 2005. "Fiscal Interactions and the Costs of Controlling Pollution from Electricity." *Rand Journal of Economics* 36(4): 850–70.

Reinaud, Julia. 2007. CO_2 Allowance & Electricity Price Interaction. International Energy Information Paper (February). Paris: OECD/IEA.

Rosendahl, Knut Einer. 2008. Incentives and Prices in an Emission Trading Scheme with Updating. *Journal of Environmental Economics and Management*, 56(1):69-82 (July).

Ruth, Matthias, Steven Gabriel, Karen Palmer, Dallas Burtraw, Anthony Paul, Yihsu Chen, Benjamin Hobbs, Daraius Irani, Jeffrey Michael, Kim Ross, Russell Conklin, Julia Miller, 2008. "Economic and Energy Impacts from Participation in the Regional Greenhouse Gas Initiative: A Case Study of the State of Maryland. Energy Policy, 36: 2279-2289.

Schultze, C.L. 1977. *The Public Use of Private Interest*. Washington, DC: The Brookings Institution.

Smith, A.E., Ross, M.T., and Montgomery, W.D. 2002. "Implications of Trading Implementation Design for Equity-Efficiency Trade-Offs in Carbon Permit Allocations." Washington, DC: Charles River Associates.

Sterner, Thomas, and Adrian Muller. 2007. Output and Abetment Effects of Allocation Readjustment in Permit Trade. *Climatic Change*, (forthcoming).

U.S. Energy Information Administration. 2005. *Impacts of Modeled Recommendations of the National Commission on Energy Policy*. SR/OIAF/2005-02, April.

U.S. Energy Information Administration. 2008. *Energy Market and Economic Impacts of S.2191, The Lieberman–Warner Climate Security Act of 2007*. SR/OIAF/2008-01, April.

U.S. Energy Information Administration 2007. *Annual Energy Outlook 2007*, DOE/EIA-0383(2007).

PRELIMINARY AND INCOMPLETE: PLEASE DO NOT CITE

When Does Cap-and-Trade Increase Regulated Firms' Profits?

David A. Evans, Ian Lange and Josh Linn

DRAFT: May 28, 2009

Many recent studies have concluded that cap-and-trade programs for greenhouse gases with free allocation of permits would increase regulated firms' profits, chiefly for the electric power sector. It is likely that future greenhouse gas programs would induce reductions in abatement costs and the costs of generating electricity from low-emissions sources, such as wind power. Past research has not considered the effects of endogenous or subsidized cost reductions on incumbents' profits in detail, however. This paper investigates these effects and calculates the share of permits that would need to be grandfathered while keeping profits constant. A simple analytical model illustrates that technological change or subsidies to abatement equipment or renewables would likely decrease the profits of regulated firms. The simulation results of a detailed partial equilibrium model of the electric power sector confirm these conclusions and suggest that cost reductions are likely to decrease profits.

Keywords: Cap-and-Trade; Abatement Costs; Climate Change Policy; Compensation

JEL codes: Q58, Q40, Q52

PRELIMINARY AND INCOMPLETE: PLEASE DO NOT CITE

When Does Cap-and-Trade Increase Regulated Firms' Profits?

David A. Evans, Ian Lange and Josh Linn[ψ]

1 Introduction

As the policy debate over reducing greenhouse gas emissions has intensified, particular attention has been devoted to the effect of regulation on incumbent firms' profits. Numerous modeling studies predict that a carbon dioxide cap-and-trade program in the U.S. with 100% free allocation of permits would increase the profits of regulated firms, chiefly those firms in the electric power sector (e.g., Goulder and Bovenberg, 2001 and Burtraw *et al*., 2002). Evidence from the European Union's Emissions Trading System (EU ETS) for carbon dioxide (CO_2) appears to show that profits of electric power producers have increased (Sijm *et al*., 2006). The seemingly counter-intuitive result arises from the fact that a large share of the emissions reduction would be due to a reduction in output. With a relatively price-inelastic demand, the increase in revenue more than offsets the increase in costs provided that the right to emit is grandfathered to emitters.

Policy makers have responded to these results by proposing to auction a share of the permits to ensure that regulated firms' profits do not increase. For example, the European Parliament approved a post-2012 EU ETS with auctioned permits (100% in the electric power sector) in order to prevent the average profits of regulated firms from increasing.[1]

[ψ] Evans (evans.davida@epa.gov) is an Economist in the National Center for Environmental Economics, U.S. Environmental Protection Agency. Lange (i.a.lange@stir.ac.uk) is a Lecturer in the Department of Economics, University of Stirling. Linn (jlinn@uic.edu) is an Assistant Professor in the Department of Economics, University of Illinois at Chicago.

The views expressed in this paper are those of the author(s) and do not necessarily represent those of the U.S. Environmental Protection Agency. In addition, although the research described in this paper may have been funded entirely or in part by the U.S. Environmental Protection Agency, it has not been subjected to the Agency's required peer and policy review. No official Agency endorsement should be inferred.

[1] The auction revenue represents a transfer from the regulated firms to the government, provided that the auction revenues are not refunded to the affected firms. The regulator can control total emissions by setting the cap, and can influence the effect of the regulation on profits through the fraction of permits it provides freely.

The modeling studies cited above are essentially static in nature, however. In particular, they do not consider the effect of cap-and-trade programs on costs, which in turn affect incumbents' profits. For two reasons, it is highly likely that cap-and-trade programs for greenhouse gases cause reductions in the cost of abatement technologies and in the cost generation technologies with low emission rates, such as renewables. First, there is considerable evidence that cap-and-trade programs cause innovation and technology adoption (e.g., Popp, 2002, Bellas and Lange, 2005 and Linn, 2008). Regulation-induced innovation and technology adoption could reduce the cost of abatement technologies, such as carbon capture and storage (CCS), and the cost of renewable energy sources, such as wind and solar. Second, in recent CO_2 cap-and-trade programs and proposals for future programs, a share of the auction revenue is used to directly subsidize abatement technology and renewables generation. Similarly to the effect of regulation-induced technological change, these subsidies and research programs would reduce abatement costs and the cost of producing electricity from renewable sources.[2] Understanding the effect of cap-and-trade on profits therefore requires a more flexible approach that accounts for endogenous reductions in costs or reductions in cost attributable to direct incentives associated with the cap-and-trade program.

This paper provides a framework for understanding how cap-and-trade programs affect profits in a dynamic setting. Our analysis is strictly positive in nature; we are interested in how the value of existing capital may change as result of policy-induced reductions in the cost of lowering emissions intensity. We are not making claims that additional compensation is required of firms affected by a cap-and-trade policy, although the analysis may inform the political economy surrounding the debate regarding the level and appropriate form of legislation and regulation to control greenhouse gases.

[2] The two uses of auction revenues differ in their desirability from the perspective of economic efficiency. If there is a market failure in that investment in research and development is undersupplied, then using auction revenues to support this research may be welfare-improving. However, unless the direct subsidy to renewable production allows these technologies to overcome some other market failure, then using the auction revenues in this manner is inefficient.

We first use a simple analytical model in which firms can reduce total emissions under a cap-and-trade program by reducing output or by using direct abatement technology to reduce the average rate of emissions per unit of output. Direct abatement technologies include those controls that can be retrofit to an existing facility or added to a new facility to reduce the emission rate. We show that if additional direct abatement is more costly, i.e. the rate at which the marginal direct abatement cost increases is higher, the greater is the reduction in output and consequently the increase in revenues. On net, provided that a sufficient number of allowances are grandfathered to emitters, the regulated firms are better off. Alternatively, considering an increase in regulatory stringency – a decrease in the cap – the greater is the magnitude of the elasticity of total output to emissions; the greater is the increase in profits. Several numerical examples illustrate the main intuition and provide insight into the likely effect of decreasing costs on the profits of incumbent firms. The main predictions are, first, that a decrease in the cost of direct abatement technologies, either because of regulation-induced technological change or subsidy, may reduce profits if firms are less likely to reduce output. Second, decreases in the cost of low-emission generation sources are also likely to reduce profits because they decrease the elasticity of total output to emissions.

These predictions are confirmed by simulations of a detailed market model of the U.S. electricity sector. The baseline simulation reproduces the findings of the previous literature that a carbon dioxide cap-and-trade program increases profits. However, we also find that profits are increasing in the elasticity of output to emissions. Moreover, the effect of abatement costs and the costs of renewables on profits is quantitatively large, implying that endogenous cost reductions are likely to have a significant effect on profits. The paper concludes by discussing the implications of these results for auction design. In particular, using auction revenue to subsidize new abatement technology or renewables substantially reduces the profits of existing generating units.

2 Background

This section briefly summarizes firms' compliance behavior in cap-and-trade programs, which is modeled more formally below, and summarizes the recent literature. The imposition of a cap-and-trade program leaves firms in the regulated industry with essentially two compliance options: reduce output or reduce the rate of emissions per unit of output. The emissions rate can be reduced by switching to a cleaner fuel, such as substituting natural gas for coal, or by installing abatement equipment, such as carbon capture and storage (CCS). Regulated firms in a competitive market choose the profit-maximizing compliance behavior, based on the output price and permit price; the output price represents the cost of reducing output and the permit price represents the benefit of reducing output as well as the benefit of reducing the emission rate. If firms primarily respond to the regulation by reducing output and the demand curve for output is relatively inelastic (as in the case of electricity), then the regulation causes profits to increase. In this case the the limit on emissions acts as a limit on output and firms behave similarly to a cooperative oligopoly in which firms restrict output to increase profits (Mansur, 2007).

The other option is to reduce the rate of emissions per unit of output, either by installing abatement equipment or by switching to fuels with low emission rates. In this paper, direct abatement refers to a reduction in the emission rate obtained by installing capital equipment such as CCS. Fuel switching under a cap-and-trade program is more costly when the price of permits is arbitraged into the price of the cleaner fuels (Keohane and Busse, 2007). If firms comply entirely by reducing their emissions per unit of output, profits must decrease because production costs increase and there is no effect on the output market. Therefore, whether firms abate by reducing output or reducing the emission rate influences the effect of a cap-and-trade program on profits.

The incentive for firms to choose a decrease in output versus a decrease in the emissions rate depends crucially on the availability of low-cost direct abatement technology and low-emissions generating entrants. These two factors essentially decrease the elasticity of abatement costs with respect to (w.r.t) output. If the elasticity is high, this can be interpreted as a case where emissions and output can be unattached and the cap-and-trade program will no longer act as an output constraint. If the elasticity is low, emissions and output are attached and the cap-and-trade program will act as an output constraint allowing firms to reduce production, increase prices and profits.

Many potential greenhouse gas policies include substantial auctioning of permits with plans to use the auction revenue to subsidize direct abatement technology (e.g., CCS) and low-emissions generating entrants (e.g., renewables).[3] The post-2012 EU ETS calls for full auctioning in the electricity sector in 2012 and full auctioning in other sectors by 2020. The auction revenues will be used to fund CCS demonstration projects and renewable technologies. Every recent greenhouse gas legislative proposal in the 110th US Congress called for some auctioning of emissions allowances, while a share of allowances is provided to incumbent emitters in part to keep their profits "whole" in the face of the cap. For example, the Lieberman-Warner Climate Security Act specifies that 21% of permits be auctioned in 2012, increasing to 69% in 2031 and staying at 69% until 2050. Approximately 13% and 18% of auction proceeds would be used to subsidize renewable energy (Title IV-Subtitle D-Section 4406) and research into sequestration technology (Title IV-Subtitle D-Section 4403), respectively. The Regional Greenhouse Gas Initiative, a cap-and-trade program in the Northeast, requires that 25% of permits be auctioned and the proceeds used to fund investment in energy efficiency or low-emissions generation. The states can decide how to allocate the remaining 75%, although most states have announced they will auction off a large percentage of their permits.

Maloney and McCormick (1982) first explore the possibility that environmental regulations may cause firms to reduce production, which would increase the output price and perhaps profits. They note the importance of free allocation and entry costs in determining the effect of regulation on profits.

Dinan (2003) provides a broad overview of the literature related to firm and consumer compensation with the imposition of a cap-and-trade regulation. Goulder and Bovenberg (2001) use a general equilibrium model of the U.S. to determine the percentage of allowances that should be grandfathered (i.e., allocated freely to existing units) to hold profits constant under an economy-wide cap-and-trade program for carbon dioxide (CO_2). They conclude that for many sectors, especially electricity and fossil fuel, full grandfathering would increase profits, so that less than 100 percent of the permits should be grandfathered to leave profits unaffected.

---

[3] Other compelling reasons for auctioning abound. For a survey of these arguments, see Burtraw and Evans (2008).

Cameron *et al*. (2006) uses a Cournot model to determine the percentage of grandfathered permits needed to hold industry profits constant under a CO_2 program. Using UK and European industry data for electricity, cement, steel, and newspaper generation, they find that these industries require no more than 65% of permits to be grandfathered to ensure that the industries' profits are constant; the electricity industry requires no more than 6%. Burtraw and Palmer (2007) use a detailed partial equilibrium model of the US electricity sector to compare alternative allocation methods.[4] Results show that grandfathering permits to electricity generators leaves virtually all of them better off and the largest firms substantially better off.

The conclusion that the imposition of a cap-and-trade program may lead to increased profits is not limited to carbon dioxide. Burtraw and Palmer (2003, 2004) show that grandfathering allowances under cap-and-trade programs for sulfur dioxide, nitrogen oxides and mercury would also lead to an increase in total profits for the electricity sector. Parry (2004) addresses a similar question from a different perspective, focusing on the effect of cap-and-trade programs on household wealth. For small reductions in emissions, households in the top wealth quintile, who are most likely to own shares of firms affected by the regulation, are made better off with the introduction of emissions caps on sulfur dioxide, nitrogen oxides and carbon dioxide when allowances are fully grandfathered. Parry (2004) uses an aggregated partial equilibrium model of the US electricity sector with a linear marginal abatement cost function and argues that a convex function would exacerbate the regressivity of permits. Mansur (2007) uses the imposition of a tradable permit scheme for nitrogen oxides (NO_x SIP Call) to show that firms reduced output more than in a competitive market in an effort to raise prices.
While many models of the effect of potential greenhouse gas legislation on firms' profits find they will be overcompensated, other modeling (Farrell *et al*., 1999) and empirical (Linn, 2006)

---

[4] Burtraw and Palmer (2007) and Burtraw *et.al* (2002) are part of a series of studies that look at the effect of different allocation schemes for cap-and-trade programs in the electricity sector on producer and consumer surplus using the model used in this paper. Generally they find that incumbent firms prefer grandfathering, followed by auctioning of allowances and finally allocating allowances based on recent production levels (output-based updating) regardless of the pollutant being regulated. Furthermore, all of the allocation methods yield to higher total profits for all firms in the industry (although some a are certainly worse off). In part this is due to the nature in which electricity is priced, as described below.

studies of cap-and-trade programs for conventional pollutants have found that these programs reduce profits of affected firms. Farrell *et al*. (1999) uses a detailed partial equilibrium model of the US electricity sector to investigate the effect of a cap-and-trade program for nitrogen oxides. Linn (2006) finds that the stock prices of regulated firms decreased in anticipation of a regional cap-and-trade program for nitrogen oxides.

3 The Relationship Between Profits, Abatement Costs, Entry Costs and Baseline Emissions

3.1 *Increasing Profits with a Cap-and-Trade Program*

We begin by demonstrating the possibility that profits can increase with the imposition of a cap-and-trade program. The analysis focuses on a representative firm. Let q be the total production/consumption of the output in the market subject to the cap and trade regulation. Uncontrolled emissions equal q*z, where z is the uncontrolled emissions per unit of q (i.e., the emissions intensity of q), while actual emission are represented by e. Abatement is then $a \equiv qz - e$. The abatement technology is represented by the cost function: $f(qz-e)$, where $f'(qz-e) \geq 0$ and $f''(qz-e) \geq 0$. Marginal-willingness to pay for output is represented by $P(q)$ while the cost of producing q net of emissions control costs is $C(q)$ and these functions are assumed to have the usual shape ($P'(q) \leq 0$, $C'(q) \geq 0$, and $C''(q) \geq 0$).

Each firm in this market is subject to an emissions cap represented by A, so that compliance requires that $A \geq e$. We assume that the regulated firms do not have market power, and therefore take the output price, P, as given. The profit maximization for the firm is:

$$\max_{q,e} \ \pi \equiv Pq - C(q) - f(zq-e) \qquad s.t. \quad A \geq e \tag{1}$$

The first order conditions of this problem (assuming the allowance constraint binds) are:

$$P - C'(q) - zf'(zq - e) = 0$$
$$f'(zq - e) - \lambda = 0$$
$$A - e = 0$$

where λ is the shadow value on the emission constraint and is therefore the permit price.

The effect on profits of tightening the cap (i.e. reducing *A*,) depends on how *A* affects production, emissions and the allowance price. After substituting *P=P(q)* in the first order conditions, it can be shown that:

$$\frac{dq}{dA} = \frac{zf''(qz-e)}{SOC} \geq 0$$
$$\frac{de}{dA} = 1$$
$$\frac{d\lambda}{dA} = \frac{[P'(q) - C''(q)]f''(qz-e)}{SOC} \leq 0$$

where $SOC = -P'(q) + C''(q) + z^2 f''(qz - e) \geq 0$ and is the second-order condition of equation (1). The first equation shows that the production of *q* increases as the cap is loosened. Critical to the hypothesis being explored here, the increase in production from an increase in the cap is larger if the rate of increase in the cost of directly controlling emissions rises (i.e., notionally $dq^2 / dAdf''(.) \geq 0$). Production is more responsive to the level of allowable emissions if direct emissions control is more expensive. Therefore, we see that the reduction in production is greater if *A* decreases and direct abatement is increasingly costly. The second and third equations show that, unsurprisingly, emissions rise in proportion to *A*, and the allowance price falls with *A*.

Total profits to the industry can be expressed as:

(2) $$\Pi \equiv P(q)q - C(q) - f(zq - e) - \lambda(1-s)A$$

The expression $\lambda(1-s)A$ represents the value of the allowance rent that the government retains where the share of grandfathered allowances to emitters is *s*. The expression can be thought of as the total payment paid by the industry to purchase the share of the allowances that is auctioned.[5] This term was not included in the exercise above because we assume that firms take the allowance price as given (also note that if *s*=1 the regulated firms capture the entire scarcity rent of the cap). Therefore, profits equal industry revenues minus production, abatement costs, and the share of the scarcity rent from the allowances that is not allocated to industry.

Taking the total derivative of (2) with respect to the cap, we can demonstrate the ambiguity of industry profits to a decline in the cap. Applying the profit maximization relationship that price equals marginal production cost, and that the cap binds such that the change in allocation equals the change in the cap, we obtain:

$$-\frac{d\Pi}{dA} \equiv \underbrace{-\left[P'(q)q\right]\frac{dq}{dA}}_{(+)} \underbrace{-s\lambda\frac{de}{dA}}_{(-)} \underbrace{+A(1-s)\frac{d\lambda}{dA}}_{(-)} \tag{3}$$

The first term of this expression shows that profits increase from reducing the cap due to the contraction of production that in turn raises inframarginal revenues.[6] The more responsive demand is to a change in production (i.e, the bigger the absolute value of $P'(q)$), the greater is this first term. The second term captures the decrease in the value of the allowances allocated to firms as a result of the cap being smaller. As the cap falls, the allowance price will also rise, and these additional allowances must be purchased from the government. This effect is represented by the last term in (3). If all allowances were allocated freely (s=1), then the entire expression would simply equal the change in inframarginal revenues from tightening the cap minus the

---

[5] We are assuming that the method of auctioning allowances yields an allowance price that equals the shadow value of the allowance constraint.

[6] This expression would not be present in the case of regulating a monopolist. This is because the monopolist is already accounting for the inframarginal change in revenues as supply is contracted. If the supply side of the market were characterized by monopoly pricing, then the amount of compensation required to keep profits neutral will be higher than if the market is characterized by marginal cost pricing.

marginal cost of the direct control technology.[7] If *s*=1, it is straightforward to show that if the marginal cost of the direct control technology is increasing rapidly ($f''(.)$ is large), then firms are made better off by tightening the cap.

3.2 Extending the Model to Multiple Producer Types

The preceding analysis provides important intuition, and this section extends the analysis to include two important features of electricity markets: heterogeneous producers with different uncontrolled emission rates. It is assumed that there is marginal cost pricing and particular functional forms are chosen for electricity demand, generation costs, and the direct abatement cost function. There are two types of electric generating units, which have different fixed and marginal costs, where all fixed costs are assumed to be entirely sunk. Base load units, denoted with the subscript b, have relatively high fixed costs, and low marginal generating costs, equal to mc_b. Peaking units have lower fixed costs but higher marginal costs of $mc_p > mc_b$. Each base load unit has a maximum generating capacity of $\bar{q}_b$ and each peaking unit has a maximum generating capacity of $\bar{q}_p$. There are n_b and n_p base load and peaking units in the market, meaning that they have all paid their sunk costs. The electricity market is competitive, and each unit operates at full capacity as long as the price exceeds its marginal operating costs.[8] The demand curve for electricity has a constant elasticity of substitution (CES), and is given by:

$$P = \gamma(n_b q_b + n_p q_p)^{1/\delta},$$

where γ and δ are constants and q_b and q_p equal the output of each base load and peaking unit; δ is the own-price elasticity of demand.

---

[7] Note that the change in profits from lowering the allocation is linear in *s*. As *s* increases the previously grandfathered allowances have a bigger influence on the change in the value of those allowances. However, the effect of the higher allowance price on the share that must be purchased from the government declines.

[8] In this model, high and low cost units operate in a competitive market and we abstract from time-varying electricity demand.

Figure 1 illustrates the market equilibrium. The supply curve is horizontal at mc_b until all base load units are operating at full capacity, at which point the supply curve is horizontal at mc_p. The supply curve is vertical when peaking units are also operating at full capacity. The figure shows the long run equilibrium, where all units earn zero profits. Price exceeds marginal costs for all units, and operating revenue exactly covers fixed costs (recall that the base load units have higher fixed costs).

Starting from long run equilibrium, the government unexpectedly announces that there will be a cap-and-trade program for carbon dioxide emissions. The cap is set at C and permits are grandfathered. Peaking units have zero emissions, while base load units have uncontrolled emission rates of $z > 0$. However, base load units may reduce their emissions according to the function:

$$x = f\left(qz - e\right) = \alpha(qz - e)^{\beta},$$

where α and β are positive constants, with $\beta > 1$. The variable x equals abatement expenditure and e is the unit's emissions. The abatement cost function has positive and increasing marginal abatement costs, and the elasticity of emissions with respect to abatement expenditure is constant.

Base load units are price-takers in both the electricity and permit markets. They choose the profit-maximizing level of output and abatement expenditure. Under the cap-and-trade program, they must hold enough permits to cover their emissions. Their optimization problem is given by:

$$\max_{q_b, x_b} q_b(p - mc_b) - f\left(qz - e\right) - \pi(zq_b - (x/\alpha)^{1/\beta} - a_b)$$
$$s.t. q_b \leq \overline{q}_b,$$

where π is the equilibrium permit price and a_b is the number of permits allocated to the unit. The first order condition for output is:

$$p - mc_b - \pi z - \mu = 0,$$

where μ is the multiplier on the production constraint. The marginal operating cost of the base load unit now includes the marginal cost of generating electricity, plus the product of the permit price and the uncontrolled emission rate. The first order condition illustrates a central result: as long as the permit price is sufficiently low, the constraint binds and base load generating units operate at full capacity. However, high permit prices may cause the base load units to reduce their output below full capacity.

The optimization problem of a peaking unit is similar, except for the fact that peaking units have zero emissions and therefore do not choose abatement.

$$\max_{q_p} q_p (p - mc_p) + a_p \pi$$
$$s.t. q_p \leq \overline{q}_p$$

Because peaking units have zero emissions, the permit price does not appear in their first order condition. As a result, peaking units always operate at full capacity.[9] The equilibrium electricity and permit prices, as well as the effect of the cap on profits, depend on the cap and other parameter values. The following subsections investigate the effect on generating units' profits of the cap and the elasticity of abatement costs.

3.2.1 Effect of the cap on profits

[9] The assumption that peaking units have zero emissions is for simplicity, however, and does not affect the main results.

Figures 2a and 2b show how the equilibrium varies with the emissions cap, A, for particular parameter values. In both figures, the horizontal axis shows the required abatement under the cap, which is equal to the difference between unregulated emissions and A. The figures show as functions of total abatement the equilibrium output and abatement expenditure of base load units; and the permit price, electricity price and total industry output. The figures correspond to the short run equilibrium, in which entry of new generating units is not possible.[10]

Figure 2a shows that as total abatement increases from zero, abatement expenditure and the permit price increase. However, the permit price never becomes sufficiently high that base load units reduce output (base load generating capacity is set equal to 4 in these simulations). Figure 2b shows that as total abatement increases, industry profits decrease. The reason is that under this set of parameter values, increasing total abatement (i.e., decreasing A) does not affect the electricity market, as all generating units continue to operate at full capacity; as the figure shows, the price of electricity does not change with the cap. As total abatement increases, however, base load units must increase their abatement expenditure to reduce emissions sufficiently to meet the cap, which reduces profits. Note that although permits are allocated freely to generating units, total profits are lower than if there were no cap. We next show that this result arises because regulation does not affect the equilibrium level of electricity production. That is, profits decrease when output and emissions are uncorrelated in equilibrium.

3.2.2 Effect of the abatement function on profits

We now demonstrate one of the main results of this section: as abatement becomes more costly, profits may increase. Subsequently, we will interpret this result in the context of recent policy proposals to support innovation in abatement technologies, which would reduce their costs. For the simulations, total abatement is fixed at 50. The value of β varies from 1 (the value used in

---

[10] Focusing on the short run simplifies the exposition, but allowing for long run entry would not significantly affect the main conclusions.

Figure 2) up to 5. Figures 3a and 3b plot as functions of β the equilibrium output and abatement expenditure of base load units; and the permit price, electricity price and total industry output. An increase in β represents an increase in the elasticity of abatement costs w.r.t output. That is, higher values of β reflect more costly abatement.

For low and moderate values of β , increasing the elasticity of abatement costs operates similarly to an increase in total abatement, which was shown in Figure 2. When the elasticity increases, for a given A, abatement expenditure increases and the permit price increases. Electricity production is not affected, and profits decrease.

A striking change occurs when β is greater than about 3: an increase in β causes industry profits to increase, even though abatement is more expensive. The first order condition for base load output provides the intuition. For low values of β, the permit price is low and the output price exceeds the marginal operating cost for base load units, which therefore operate at full capacity. As β increases, however, the permit price increases as well, and at some point the price of electricity no longer exceeds marginal operating costs. Therefore, base load plants reduce output, which causes the electricity price to rise. There are therefore two effects of increasing β, while holding A constant. The first is that abatement costs increase, which reduces profits. On the other hand, electricity production also decreases, which increases profits because electricity demand is price-inelastic. Under the particular parameter values chosen for these simulations, the second effect is greater in magnitude, and an increase in β leads to higher profits when β is sufficiently large. Note that when β exceeds 3, total industry profits are greater than when there is no cap on emissions; this case is commonly referred to as overcompensation.

In the extreme case where base load units cannot reduce their emissions rates at all, which is roughly the case for carbon dioxide regulation, all abatement occurs via curtailing base load generation. That is, all abatement occurs via fuel switching, in which the share of generation of high-marginal cost peaking units increases at the expense of base load units. In this case, even a very large A could significantly increase firms' profits. On the other hand, when abatement is relatively inexpensive, as with nitrogen oxides or sulfur dioxide, there is likely to be a much smaller effect on electricity production, and a small or moderate cap may decrease profits.

3.2.3 Entry of Zero Emission Rate Generating Units

We now discuss the effect of the entry of generating units with zero emission rates, such as wind or solar, which will be referred to as renewable generating units. It is assumed that the renewable technology has fixed costs of F_r, zero marginal costs and a maximum generating capacity of $\overline{q}_r$; renewable generators enter as long as the output price exceeds $F_r / \overline{q}_r$. There is an unlimited number of renewable generating units that could potentially enter the market. Therefore, once the cap-and-trade program is implemented, renewable generators will enter the market until the output price equals $F_r / \overline{q}_r$. That is, the renewable technology amounts to a cap on the price of electricity.

In practice, F_r might decrease because of technological change or because of government subsidies (e.g., an investment tax credit). The first case to consider is the one in which the capacity constraint does not bind. A decrease in F_r reduces the price of output and profits unambiguously fall. The more interesting case is hen the capacity constraint does not bind, i.e., as in the right section of Figure 3.

For this case, Figure 4 plots the equilibrium outcomes as functions of the price of electricity. That is, as F_r decreases, the electricity price decreases, moving from right to left in the diagram. The simulations use the baseline parameter values from Figure 3, setting $\beta = 5$, and varying F_r. In the diagram, the minimum price is the price of electricity if there were no cap-and-trade program, and the maximum price is the price from Figure 3 when $\beta = 5$.

Consider first the case in which F_r is sufficiently high that there is no entry of renewable in equilibrium. As Figure 3 shows, base load units comply with the cap by abatement expenditure and reducing output. The output constraint does not bind and the price of electricity exactly equals operating costs $mc_b + \pi z$ .Now consider a decrease in F_r large enough that renewable units enter the market. The price of output decreases, but the permit price must also decrease so that the first order condition for base load output continues to be satisfied. Therefore, abatement expenditure decreases. To comply with the cap, base load units must decrease output. Thus, there are two effects on industry output. Base load units decrease production while renewable units increase production. The latter effect is greater in magnitude, so the output price falls and industry profits decrease.

Recall that the simulations reported in Figures 2-4 assume that there is no entry of base load or peaking units. The results in Figure 4 indicate that allowing for such entry generally offsets any increase in profits, to the extent that entering units have a lower emission rate than the average unit.

3.2.4 Emissions Rates of Peaking Units

One of the reasons for distinguishing base load and peaking units is that we can investigate the importance of relative (uncontrolled) emission rates of the two classes of units. Figure 5 shows the effect of increasing the emission rate of peaking units from the baseline value of zero to 0.4. It is assumed that the direct abatement technology is not available to peaking units, so that they cannot reduce their emission rate. This assumption which seems reasonable given that abatement technology is typically characterized by having high fixed costs and low marginal costs; i.e., peaking units will not generally find it optimal to install abatement equipment even when base load units do install such equipment.

Intuitively, increasing the emission rate of peaking units amounts to increasing total abatement or decreasing the cap. Figure 5 shows that for low emission rates of peaking units, increasing the emission rate has the same effect as increasing total abatement, which was depicted in Figure 2. Base load units comply by reducing their emission rate instead of decreasing output, which leaves the output price unaffected and decreases profits.

When the emission rate of peaking units increases above about 0.1, the permit price is sufficiently high that base load units reduce output in addition to reducing their emission rate. As a result, the output price and profits increase, and the cap-and-trade program results in greater profits.[11]

3.3 Summary of Modeling Results

The numerical results suggest that the greater the reduction in total industry output, the greater are profits. That is, the greater the (negative) elasticity of output to emissions, the greater the effect on profits. Therefore, a reduction in abatement costs may decrease profits because firms are less likely to reduce output in response to a cap. Greater entry of renewables or a reduction in emissions rates of non-base load units would have a similar effect. The next section shows that these results hold in a detailed model of the electric power sector and quantifies the magnitude of these effects.

4 Demonstrating the Importance of Abatement Opportunities using a Detailed Electricity-Sector Model

4.1 Model Setup

The previous discussion describes the relationship between the availability of techniques to reduce emissions intensity, either through direct abatement or cleaner technologies, and the elasticity of demand on the profits of incumbents. In particular, we see how the availability of cleaner technologies impacts the profits of incumbent firms qualitatively. What is not yet clear is how sensitive forecasts of the change in firm profits as a result of a particular policy are to

[11] This result implies that if Figure 2 were extended to show the effect of greater total abatement, there would be a similar inflection point, when base load units begin restricting output and profits begin to increase. Furthermore, if the emission rate of peaking units becomes sufficiently high, they may begin to reduce output. This causes profits to increase at an even greater rate.

assumptions about critical parameters discussed in the previous section. To address this issue we use a detailed partial equilibrium model of the U.S. electricity sector to explore how incumbents' profits are affected by a reduction in the cost of renewable generation in the presence of a CO_2 emissions cap. We use the Haiku electricity sector model developed at Resources for the Future (Paul et al, 2008). Future work will use the model to show the affect on profits from changes in direct abatement costs and the emission rate of non-base load units under a climate policy.

While we will identify reasonable bounds of the induced reduction in the cost of wind generation in forthcoming analysis, the analysis reported herein is intended to demonstrate the influence of the cost of abating and relatively clean new production technology in the abstract. That is, we are not claiming that our different cost sensitivities fall into bounds that are empirically informed, and therefore are knowingly relevant for policy. However, if a significant effect on profits cannot be shown at the extremes considered in this analysis, then they presumably would not be important within a policy-relevant range. Also, as discussed below, we model policy-directed reductions in the cost of wind generation relative to incumbent generators. In the Lieberman-Warner climate proposal 4% of annual auction revenues are set aside to subsidize renewable generation. While it requires some assumptions about exactly how exactly these allowances will be distributed, we model the effect of this use of auction revenues on the profit of incumbent generators. Therefore, this portion of the analysis can be viewed as empirically informed.[12]

The Haiku electricity market model has been used primarily to study national and regional environmental and market structure policies that affect this sector. Haiku solves via an iterative tâtonnement algorithm in the prices for both electricity production and major fuel inputs,

[12] In both of these cases other potentially important sources of relative cost changes are admittedly being ignored. We recognize that the source of any reduction in the cost of new wind generation may spillover or be applicable to other generating technologies (e.g., a general improvement in the cost of turbine production). However, these effects may only exacerbate the effect of induced innovation on the profits of incumbent generators. We also note that other uses of the asset value of allowances, such as provisions for bonus allowances for adopting facilities with carbon capture and storage, may also influence the profitability of incumbent firms. However, we can not exhaustively consider all of the possible direct uses of auction revenues that may meaningfully influence the relative costs of generation and pollution control in the electricity industry.

including coal and natural gas, as well as the price of emissions allowances. Welfare includes both consumer and producer surplus in the electricity market adjusted by any changes in government revenues. The model allows for dynamic investment and compliance behavior. For the purposes of this study the model is solved for five evenly-spaced simulation years over a 20-year time horizon.

The model divides the contiguous states of the U.S. into 21 regional electricity markets roughly defined by historical regional electricity reliability council boundaries, plus some further spatial disaggregation. Electricity demand is price responsive and distinguished by consumer class (residential, industrial, commercial), season (summer, winter, spring/fall), and time of day (base load, shoulder, peak, and super-peak); there are 756 distinct retail electricity markets in each simulation year. Interregional trade is modeled endogenously subject to transmission capacity constraints between regions. The wholesale electricity market is assumed to be competitive so that prices are based on the relationship between marginal generating costs in different regions. Each of the 21 regional retail markets can be characterized by either average (cost-of-service) or marginal cost pricing.[13] Whether a region is treated as an average cost or marginal cost pricing region depends on the pricing structure faced by the majority of consumers in that region.

Electricity generators are represented by "model plants" defined by nine criteria: location, vintage (existing or new), prime mover, fuel, relative operating cost (for those generators using natural gas or nuclear fuel), and, for coal-fired boilers, coal demand region, capacity, and the expected presence of SO_2 and NO_X post-combustion abatement controls in 2010. Generator dispatch at any time is determined by a model plant's short-run operating cost. Capital stock investment and retirement, as well as investment in pollution control technologies, are determined by the expected profitability of generation assets over time. Assumptions regarding the performance of new generation capacity and pollution control technologies are drawn from a variety of sources.

[13] Marginal cost pricing refers to "average marginal costs" over time blocks in a particular season, as opposed to time-of-day pricing.

The possibility that firms are subject to average-cost pricing adds an additional nuance to the analysis that is not captured in the simple model in the previous section. Emissions allowances in average-cost pricing regions are assumed to be treated in the rate base at their original costs, i.e. zero. If the entire market is characterized by average cost pricing it is assumed that full grandfathering allowances will have no direct effect on the output price (Burtraw et al, 2002). Furthermore, auctioning allowances has been shown to increase welfare relative to grandfathering allowances because average-cost pricing causes the price of electricity to differ from marginal generating costs, auctioning increases total welfare by reducing the difference. In a perfectly competitive markets this would not be the case; total welfare would be the same with both auctioning and grandfathering and the only difference would be the wealth transfer between the government and the regulated firms.[14] That all said, the focus of our analysis is on the profits of incumbent firms in deregulated regions. While a plant in a cost-of-service region may earn profits, the firm that owns that plant may not. On net, the model requires that the net return on all assets in the regulated regions are zero but for the covering the cost of capital.

The Haiku model has been used in other studies exploring different dimensions of the question of the effect of cap-and-trade programs on the profits of incumbent firms, including the composition of the portfolio of generating technologies of firms and the method of allocation allowances (Burtraw et al. 2002, Burtraw and Palmer, 2007). Furthermore, outputs from this model have been used to calibrate more aggregate models of the electricity sector used in the literature on the incidence of environmental regulations (Parry, 2004 and Bovenberg *et al*. 2005).

The emissions cap used in this analysis mimics the required economy-wide percentage emissions reduction in the early simulation periods (through 2020) required by the Lieberman-Warner Climate Security Act of 2007, but then is about 25% tighter than the Lieberman-Warner restrictions in the later periods. However, there are important differences between the cap we model and the recent economy-wide caps that are the foundation of recent legislative proposals. One of the differences is obvious: that we are using a sector specific approach which may over

[14] Subject to the caveat that the receiving grandfathered allowances is not conditional on operating the plant, so that there is no distortion in the entry/exit decisions of affected firms.

estimate any allowance price effects from changes in the cost of abatement in the electricity sector as other sectors of the economy may absorb emissions or provide reductions. Another difference is that we do not allow emissions banking for carbon emissions for modeling convenience. However, we pick an allowance cap path intended to mimic a Hotelling-price path for allowances.

4.2 The Cost of Renewable Generation on Incumbent's Profits

In this experiment we explore how the profits of incumbent firms under a national emissions cap on CO_2 are influenced by the capital and operating cost of new wind generators.[15] The cost and maximum potential capacity of new wind generators in Haiku is drawn from the EIA's NEMS model. For each region the total available wind capacity is divided into five capital cost "bins". The cost of installing wind capacity is 20%, 50%, 100% and 200% for each cost step relative to the first step and the bins are not of equal size. The accelerating cost of adding new wind capacity is meant to reflect the decline in the suitability of sites (due to terrain, distance from the grid, etc.) to install new wind turbines.[16] Under the two cost sensitivities, each capital cost "step" is decreased by 30% or 50%. The 30% cost reduction was chosen because this is roughly the per MW value of the subsidy paid to renewable generators under a 20% renewable portfolio standard policy in 2020 using Haiku, and therefore the results may be roughly compared to that analysis (Blair et al., 2009). The 50% cost sensitivity was chosen because it is likely a lower bound on the possible reduction in cost. Both of these cost reductions are assumed to occur immediately after the policy is adopted.[17]

---

[15] While it is possible for existing plants to improve their efficiency and therefore emissions per unit of production, Haiku does not model this possibility due to a lack of information on the cost and effectiveness of these efficiency improvements.

[16] All of the assumptions about the cost, availability and performance of wind in Haiku are adopted from the version of Energy Information Administration's (EIA) National Energy Modeling System used for the Annual Energy Outlook 2008.

[17] Future work will consider the possibility of a gradual or future reduction in costs, which seems more likely, and perhaps also less apt to influence incumbent firm profits because these cost reductions will occur after their meaningful life.

When the cost of a new generating technology is changed in the model it would affect both conditions in the policy scenario and conditions in the baseline. That is, the introduction of less expensive wind would outcompete at least some existing generation regardless of whether or not a policy is in place. For this reason, existing sensitivity analyses consider cost changes in both the baseline and in the policy case. While our interest is in situations where costs may differ between the baseline and the policy, we also model the case where wind cost are lower absent the policy to identify what portion of the affect on profits is attributable to the lower cost of new wind generation and what is attributable to the policy. The comparison between the profits in case with and without the policy with different assumed wind costs is independently informative. As explained in further below, the metric we use to measure the change in incumbent firm profits is as a share of auction revenues. Few existing analyses report the sensitivity of this statistic to model assumptions, with Bovenberg et al. 2005 being a notable exception.

As described above we also look at the case where a share of the auction revenue is provided to new wind generation. We consider two levels. In the first, 4% of the auction revenue is provided to new wind generation (built after 2010) in the form of a subsidy per kWh generated. As discussed above, under the Lieberman-Warner proposal 4% of the auction revenue is to support renewable generation, although the bill is not explicit about the form of the support, although we think it is reasonable to assume that it would be in the form of a production subsidy.[18] Also, under the Lieberman-Warner proposal the share of auction revenues allocated to renewables is based on an economy-wide cap, whereas we are only modeling a sector specific cap. For this reason, we consider the case where 10% of the auction revenue is provided to wind generators. We chose 10% because the electricity sector comprises approximately 40% of U.S. greenhouse gas emissions.[19] Of course, this share has a larger effect on the allowance price than would occur

[18] We retain the native assumption in Haiku that wind receives the 1.9 cent/kWh (2005$) tax credit as provided by the Energy Policy Act of 2005. The tax credit is earned on generation in the first 10 years of operation of a facility. The credit is discounted to 90% to reflect the fact that over time it has lapsed on previous occasions, but not for long spells since it was originally introduced in 1992. It might be argued that the allocation of allowances to wind energy will be treated as a replacement for the tax credit on wind energy, which must be renewed.

[19] Although, because of the availability of relatively low cost controls in the electricity sector, it is forecast to emit an even lower share of emissions under a cap-and-trade program.

under an economy-wide cap, but one can view the effect of the 4% and 10% cases as roughly bounding the effect that would occur had we modeled and economy-wide program. Another notable difference is that unlike the Lieberman-Warner approach where the share of allowances allocated to renewables falls to 1% in 2031, we assume that wind receives a 4% or 10% share indefinitely in that the terminal year of our analysis, 2030, is modeled to represent all future periods.

We deviate from an important native assumption in Haiku, which is to restrict the total amount of new construction of wind generation to a certain percentage of existing capacity in the region. For some regions these constraints bind, although without much consequence, in the baseline case. We remove them because as we move to lower wind costs under the policy case, they bind further and the consequence of less expensive wind no longer has a marginal effect on electricity prices and composition of fuels used to generate electricity (the cost steps described above have a similar effect as we will see).

Table 1 summarizes the change in the national electricity price, which is a simple consumption-weighted index of prices in all regional markets, the allowance price and the CO_2 emissions level under the different policy and wind cost scenarios. The cap is binding with 100% wind costs in all years where there is a cap, but that is not the case with lower wind costs (recall that we assume that there is no banking).[20] The findings comport with out expectations; the price of electricity rises as a result of the policy, but falls as wind costs fall.

Table 2 shows the difference in national generation of coal, natural gas and wind energy between the policy case and the baseline. Overall generation changes little as a result of the policy or the cost of wind generation because of the very low elasticity of demand for electricity. We see that coal generation falls in the presence of the cap, while natural gas and wind generation rise. Wind is a large share of total generation in the model, even in the baseline at 100% cost relative to other model forecasts (e.g., the EIA's Annual Energy Outlook). This is because we removed the

[20] With a very low allowance price in 2015 emissions do not quite converge to the cap under the 70% wind cost scenario.

restrictions on the amount of new wind capacity that can be installed. That all said, these results must be taken as preliminary. Through the course of our analysis we discovered a small, but consequential, coding error in Haiku which lowered the fixed operating cost of all new capacity. Not only does this reduce the overall cost of new generation, but it also puts high-fixed operating cost generators, like wind, at an advantage relative to low-fixed operating cost generators.

One interesting side-observation arises from reviewing Table 2. With a lower cost of wind generation in the presence of a cap, coal generation is rises in the later years of the analysis. That is, despite there being a cap and a lower cost competitor in the form of wind, the lower cost competitor actually displaces other higher costs forms of emitting generation, thus effectively loosening the cap and encouraging more coal generation to remain. This finding is similar to the observation that “green serves the dirtiest” when subsidies to renewable generation are imposed in the presence of a cap-and-trade system (Böhringer and Rosendahl, 2009).

Figure 4a shows the share of total auction revenue over the period from 2010 to 2030 discounted at the cost of capital (*r* in the expression that follows) required to make keep the profits (producer surplus) of incumbent firms whole over this period given the different wind cost assumptions.[21] The calculation is:

$$\% = \frac{\sum_{t=1}^{20} (1+r)^{-t} \left(PS_t^{NC} - PS_t^{C} \right)}{\sum_{t=1}^{20} (1+r)^{-t} \left(P_t^{A} * A_t \right)}$$

Where *PS* is the producer surplus either with the cap (*NC*) or without (*C*), *P* is the price of allowances and *A* is the total allocation of allowances, all scripted by time (*t*). The bar on the left is calculated assuming that the cost of wind is the same in both with and without the policy. The bar on the right assumes that the cost of wind only falls with the policy. Under these two approaches the bars are necessarily the same height under the assumption that wind costs are 100% of the native assumption in Haiku.

[21] Recall that the model is simulated at five year intervals. Producer surplus is linearly interpolated between simulation years.

As we can see, the share of allowances that would make incumbent generators operating in competitive markets whole is insensitive to the cost of wind. This finding is consistent with other studies that have looked at the sensitivity of model assumptions to the compensation required to keep regulated firms' profits whole. However, we also see that this share rises considerably if the cost of wind generation falls as a consequence of the policy. The share required make incumbent firms whole more than doubles under the lower wind cost cases. This is true despite the fact that wind generation only displaces about 6% of other sources of generation nationally as the cost of wind is 30% lower and 10-11% if it is 50% lower.

Admittedly an induced and immediate 30% or 50% reduction in the cost of new wind capacity realized at the outset of a policy is dramatic. However, the policy itself has provisions to make wind preferable to dirtier existing sources of generation other than the cap itself, such as the share of auction revenues provided to wind generation. The total value of this inducement rises overtime as the price rises relative to the decline in the cap. The total value per unit generated depends on whether the asset value of allowances rises faster than the generation of wind. We might expect that it does as the supply of wind steepens as total wind generation increases, thus limiting the amount of wind generation and in turn contributes to a higher allowance price. Figure 4b presents the preliminary findings regarding the share of auction revenue that would need to be provided to incumbent firms operating in wholesale markets to keep their profits whole. Despite the lower allowance price and increased wind generation as the share of the auction revenue going to firms rises, the share required to make incumbent producers whole *falls*. This finding clearly warrants further analysis as it contradicts the previous findings reported and our expectations based on the modeling section 3. In part it may be where new wind is locates given the subsidy. But if it holds up to further scrutiny or if the magnitude does not change significantly, perhaps importantly from a political economy perspective the effect of this particular use of the asset value of allowances is not consequential to the profits of incumbent firms.

5 Conclusion

The effect of a cap-and-trade program on profits depends crucially on the availability of low cost abatement options and generating sources with low emissions. These two conditions essentially determine the relationship between output and emissions; the greater the elasticity of output to emissions, the greater the increase in profits. A simple model of the electricity industry demonstrates this point, which is confirmed in the simulations of the Haiku model.

An important policy implication of this result is that technological change or economic incentives for renewable generating units reduces profits of existing generating units. Such changes therefore reduce the share of permits that must be auctioned to prevent these units from earning positive profits under regulation. Extensions to this analysis include exploring how the level of the cap effects compensation, the importance of the emission rate of the marginal producer of the final good, and the pricing regime in the market (i.e., whether regulated sources face marginal cost or cost-of-service pricing as is common in the electricity sector), as well as the elasticity of demand.

Future work may focus on how the allocation of allowances to clean units will affect the compensation to dirty firms, and that R&D policy will over time affect the relationship between output and emissions. Of course, the latter needs to be considered in the context of the lifetime of capital; in a partial equilibrium model it is only the incumbent firms that may realize positive profits from the cap (entry of all others is subject to a zero-profit condition). If the benefits of R&D are only realized in the long run, after a full turnover of the existing stock of capital, then it may not have an influence on the profits of incumbent firms. However, the incentive to bank allowances under a cap that declines over time may fall, and there may still be a welfare effect on incumbents.

References

Blair, Nate and others. 2009. *Renewable Energy & Efficiency Modeling Analysis Partnership (REMAP): How Different Energy Models Addressed a Common High Renewable Energy Penetration Scenario in 2025.* National Renewable Energy Laboratory.

Böhringer, C. K. E. Rosendahl. 2009. Green Serves the Dirtiest. On the Interaction between Black and Green Quotas. Discussion Papers 581 - Statistics Norway, April 2009.

Bovenberg, A. Lans, Lawrence Goulder and Derek Gurney. 2005. Efficiency Costs of Meeting Industry-Distributional Constraints Under Environmental Permits and Taxes *RAND Journal of Economics* 36(4), 951-971.

Burtraw ,Dallas and David A. Evans. 2008. Tradable Rights to Emit Air Pollution. Resources for the Future Discussion Paper 08-08. March 2008.

Burtraw, Dallas and Karen Palmer. 2007. Compensation Rules for Climate Policy in the Electricity Sector. RFF Discussion Paper 07-41.

Burtraw, Dallas, Karen Palmer, Ranjit Bharvirkar, and Anthony Paul. 2002. The Effect on Asset Values of the Allocation of Carbon Dioxide Emission Allowances. *The Electricity Journal* 15(5), 51–62.

Carlson Curtis P., Dallas Burtraw, Maueen Cropper, Karen Palmer. 2000. "SO_2 control by electric utilities: What are the gains from trade?" *Journal of Political Economy*. 108(6), 1292–326.

Ellerman, A. Denny, Paul Joskow, Richard Schmalensee, Juan-Pablo Montero, and Elizabeth Bailey. 2000. *Markets for Clean Air: The U.S. Acid Rain Program*, Cambridge University Press, New York

European Union. http://ec.europa.eu/environment/climat/emission/ets_post2012_en.htm

Farrell Alex E., Carter R, R. Raufer. 1999. The NOx budget: market-based control of tropospheric ozone in the northeastern United States. *Resource and Energy Economics.* 21(2), 103–24

Goulder, Lawrence and A. Lans Bovenberg. 2001. "Neutralizing the Adverse Industry Impacts of CO_2 Abatement Policies: What Does It Cost?" in C. Carraro and G. Metcalf, eds., *Behavioral and Distributional Effects of Environmental Policies*, University of Chicago Press.

Hepburn, Cameron, John Quah, and Rober Ritz. 2006. Emissions Trading and Profit-Neutral Grandfathering. Oxford Department of Economics Working Paper 295.

Keohane Nathaniel O. & Meghan Busse. 2007. Market Effects of Environmental Regulation: Coal, Railroads, and the 1990 Clean Air Act. *RAND Journal of Economics*, 38(4), 1159-1179.

Lange, Ian and Allen Bellas. 2005. Technological Change for Sulfur Dioxide Scrubbers Under Market-Based Regulation. *Land Economics* 81(4), 546-556.

Linn, Joshua. 2006. The Effect of Cap-and-Trade Programs on Firms' Profits: Evidence from the Nitrogen Oxides Budget Trading Program. MIT CEEPR Working Paper 06-011

Linn, Joshua. 2008. Technological Modifications in the Nitrogen Oxides Permit Program. *The Energy Journal*, 29(3), 153-176.

Maloney, Michael and Robert McCormick, 1982. A Positive Theory of Environmental Quality Regulation. *Journal of Law and Economics.* 25(1), 99-123

Mansur, Eric. 2007. Do Oligopolists Pollute Less? Evidence from a Restructured Electricity Market. *Journal of Industrial Economics*, 55(4), 661-689

Palmer, Karen, and Dallas Burtraw. 2004. Distribution and Efficiency Consequences of Different

Approaches to Allocating Tradable Emission Allowances for Sulfur Dioxide, Nitrogen Oxides, and Mercury. Preliminary Manuscript. Washington, DC: Resources for the Future.

Parry, Ian. 2004. Are Emissions Permits Regressive? *Journal of Environmental Economics and Management* 47(2), 364-387.

Paul, Anthony, Dallas Burtraw and Karen Palmer. 2009. *Haiku Documentation: RFF's Electricity Market Model Version 2.0.* Resources for the Future Report.

Popp, David. 2002. Induced Innovation and Energy Prices. *American Economic Review* 92(1), 160-180

Sijm, Jos, Karsten Neuhoff, and Yihsu Chen, CO2 Cost Pass-Through and Windfall Profits in the Power Sector. *Climate Policy* 6(1), 2006, 49-72.

Tables and Figures

Table 1: Electricity Price, Allowance Price, and Emissions under Different Wind Cost and CO_2 Policy Assumptions.

		2010	2015	2020	2025	2030
National Electricity Price (2005$/MWh)	No Cap, 100% Wind	$ 81.10	$ 77.86	$ 80.27	$ 81.24	$ 82.53
	No Cap, 70% Wind	$ 79.79	$ 76.62	$ 79.03	$ 79.75	$ 81.42
	No Cap, 50% Wind	$ 78.49	$ 75.33	$ 77.38	$ 78.54	$ 80.42
	Cap, 100% Wind	$ 81.63	$ 82.95	$ 89.17	$ 93.76	$ 99.61
	Cap, 70% Wind	$ 80.04	$ 77.68	$ 85.85	$ 89.89	$ 94.33
	Cap, 50% Wind	$ 78.80	$ 76.09	$ 81.43	$ 86.53	$ 91.12
Allowance Price 2005$/ton	No Cap, 100% Wind	$ -	$ -	$ -	$ -	$ -
	No Cap, 70% Wind	$ -	$ -	$ -	$ -	$ -
	No Cap, 50% Wind	$ -	$ -	$ -	$ -	$ -
	Cap, 100% Wind	$ -	$ 7.09	$ 15.20	$ 23.54	$ 30.39
	Cap, 70% Wind	$ -	$ 0.13	$ 13.60	$ 20.66	$ 25.07
	Cap, 50% Wind	$ -	$ 0.00	$ 8.69	$ 19.05	$ 23.96
Total CO2 Emissions (Million Tons)	No Cap, 100% Wind	2562	2670	2724	2859	2966
	No Cap, 70% Wind	2187	2328	2464	2544	2644
	No Cap, 50% Wind	2016	2149	2259	2282	2323
	Cap, 100% Wind	2470	2210	1770	1570	1380
	Cap, 70% Wind	2120	2180	1770	1570	1380
	Cap, 50% Wind	2010	2080	1770	1570	1380

Table 2: Total, Coal, Natural Gas, and Wind Generation under Different Wind Cost and CO_2 Policy Assumptions.

		2010	**2015**	**2020**	**2025**	**2030**
National Generation	No Cap, 100% Wind	4150	4390	4686	5007	5337
billion kWh	No Cap, 70% Wind	4166	4424	4709	5055	5391
	No Cap, 50% Wind	4200	4469	4767	5092	5442
	Cap, 100% Wind	4136	4284	4458	4655	4873
	Cap, 70% Wind	4169	4398	4563	4752	5017
	Cap, 50% Wind	4190	4457	4669	4875	5130
Coal Generation	No Cap, 100% Wind	1984	2038	2128	2341	2499
billion kWh	No Cap, 70% Wind	1708	1814	1941	2056	2214
	No Cap, 50% Wind	1564	1668	1778	1840	1907
	Cap, 100% Wind	1885	1625	1301	1098	901
	Cap, 70% Wind	1637	1658	1325	1142	952
	Cap, 50% Wind	1568	1584	1339	1153	977
Natural Gas Generation	No Cap, 100% Wind	787	906	845	725	641
billion kWh	No Cap, 70% Wind	686	747	753	714	625
	No Cap, 50% Wind	668	724	712	662	595
	Cap, 100% Wind	825	968	832	963	1215
	Cap, 70% Wind	701	803	791	834	1048
	Cap, 50% Wind	646	783	737	797	989
Wind Generation	No Cap, 100% Wind	164	205	426	494	543
billion kWh	No Cap, 70% Wind	584	650	762	868	922
	No Cap, 50% Wind	859	951	1072	1238	1454
	Cap, 100% Wind	176	377	699	798	862
	Cap, 70% Wind	637	713	916	1085	1217
	Cap, 50% Wind	856	955	1161	1314	1448

Figure 4a: Percentage of Allowance Revenue Required to keep Incumbent Firms' Profits Whole under Different Wind Cost Assumptions.

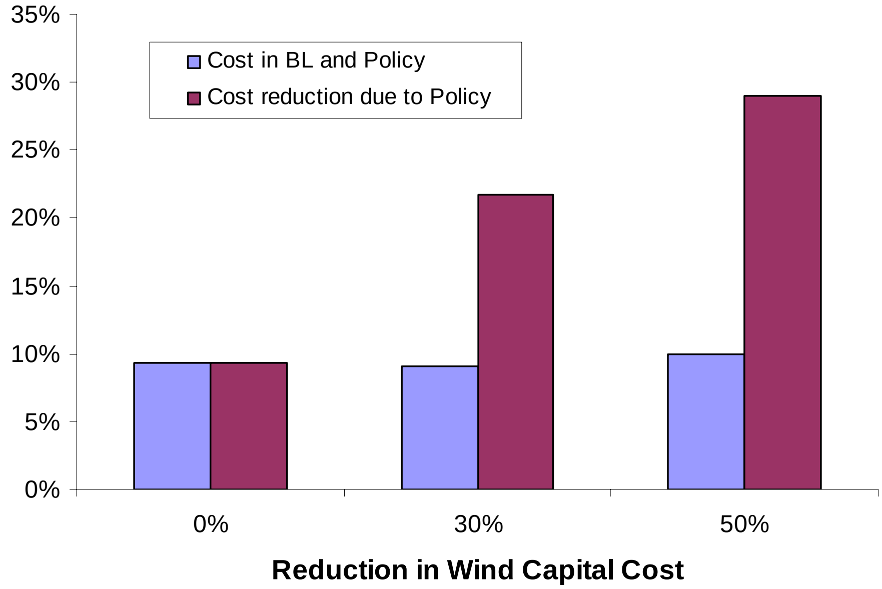


Figure 4b: Percentage of Allowance Revenue Required to keep Incumbent Firms' Profits Whole under Different Assumptions Regarding the Share of Allowance Auction Revenue Provided to Win.

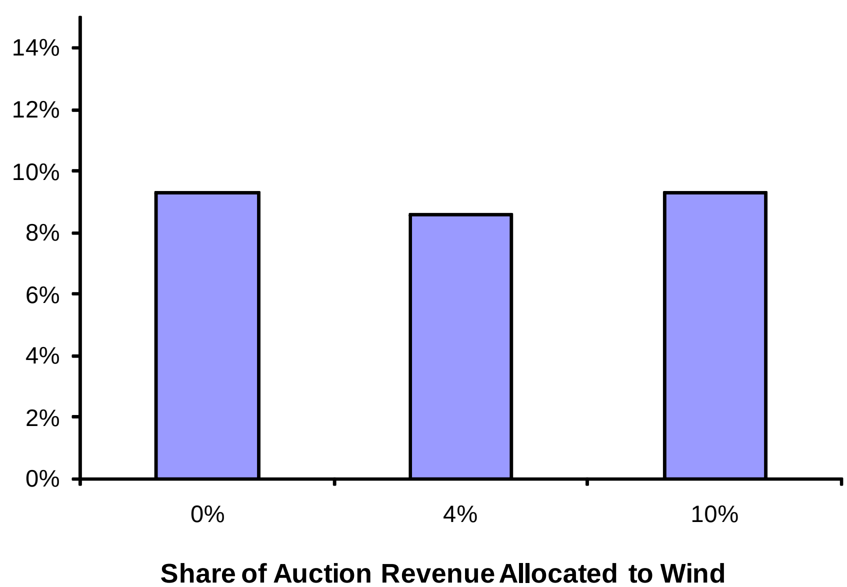


Figure 1: Electricity Market Equilibrium

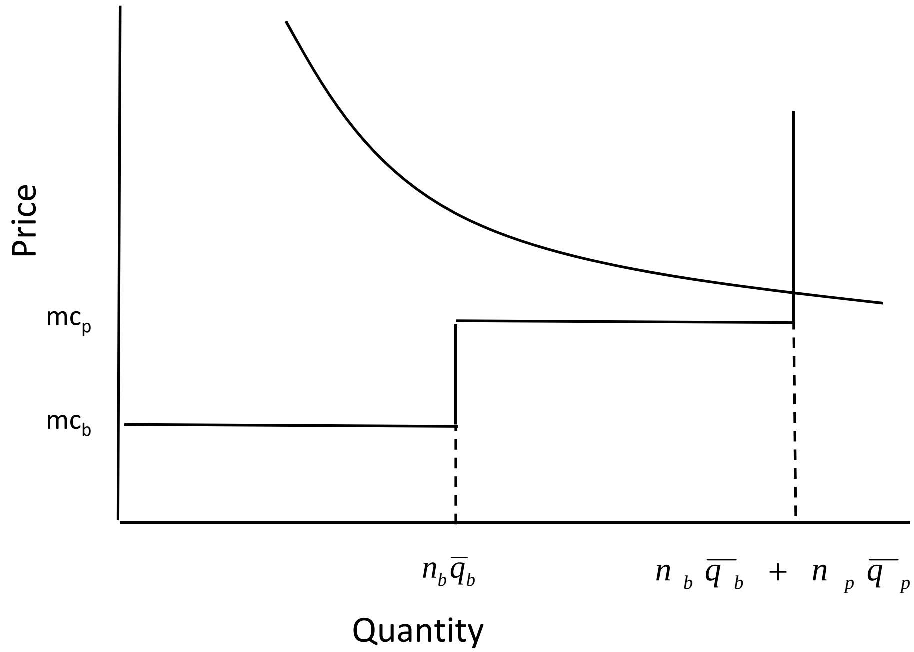


Figure 2a: Baseline Scenario, Output, Abatement Expenditure and Permit Price vs. Total Abatement

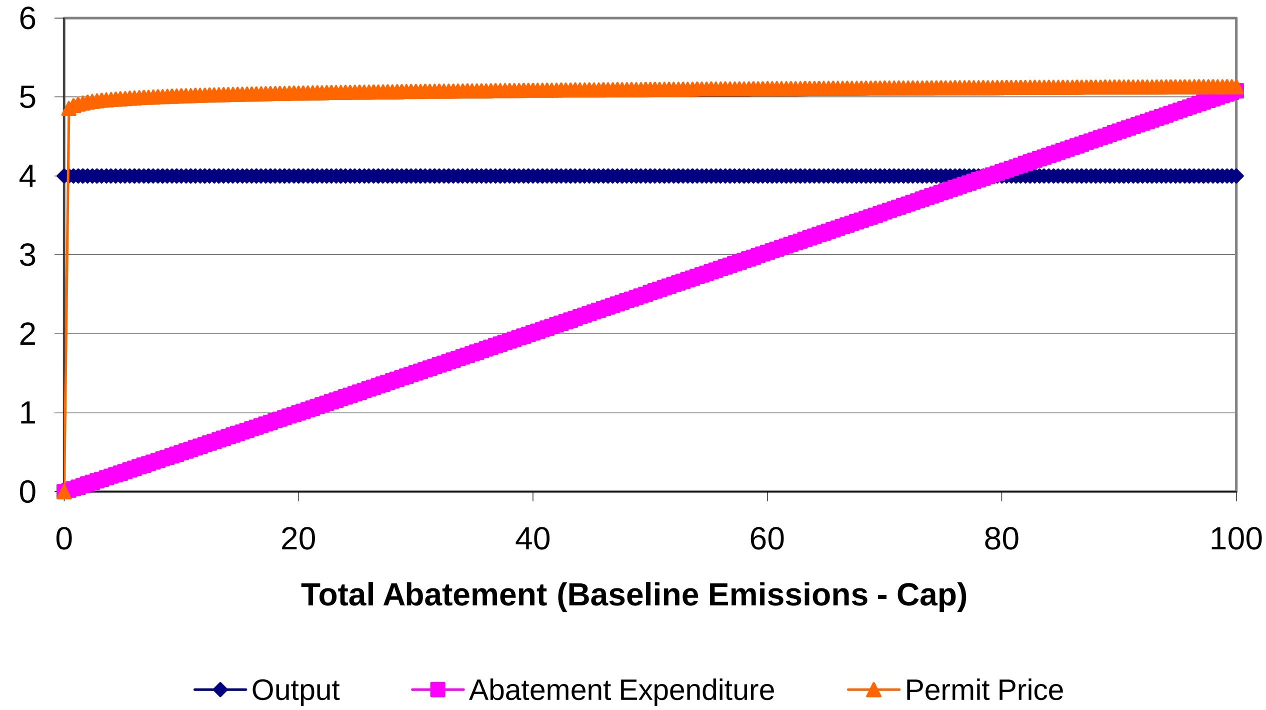


Figure 2b: Baseline Scenario, Output Price and Profits vs. Total Abatement

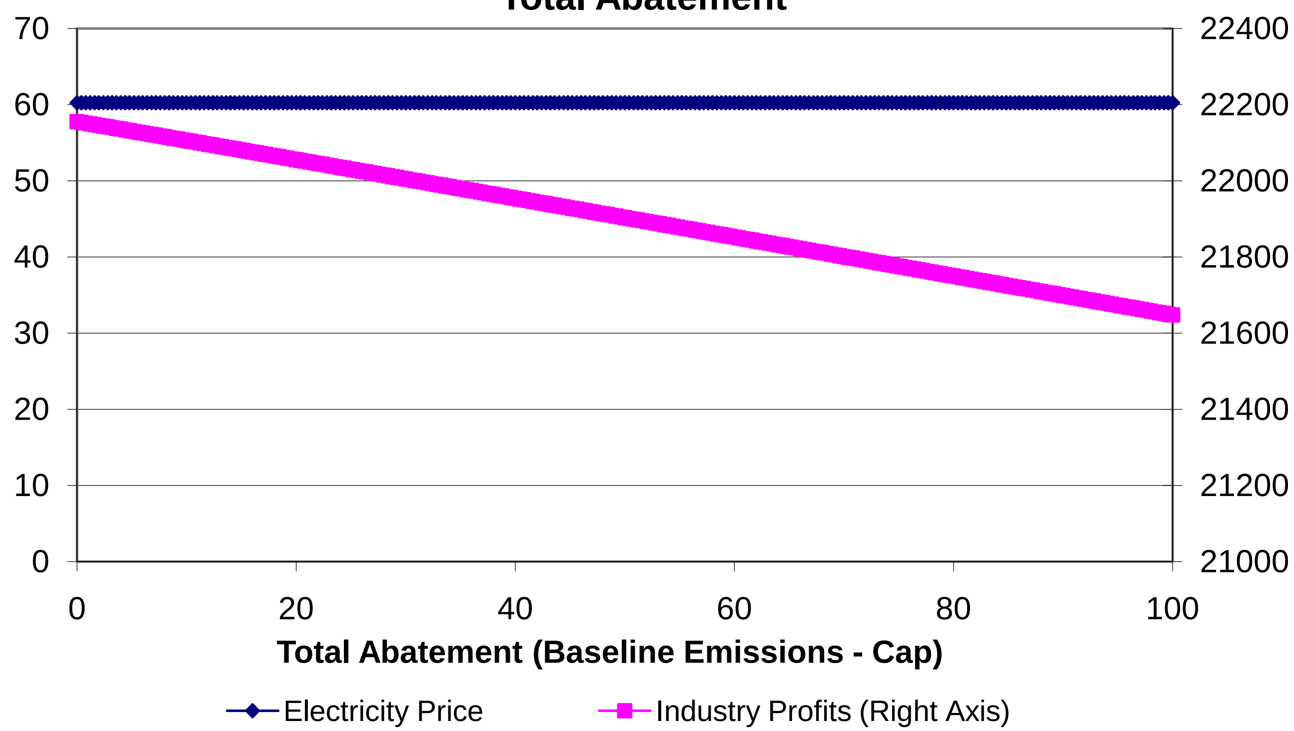


Figure 3a: Output, Abatement Expenditure and Permit Price as Functions of Beta

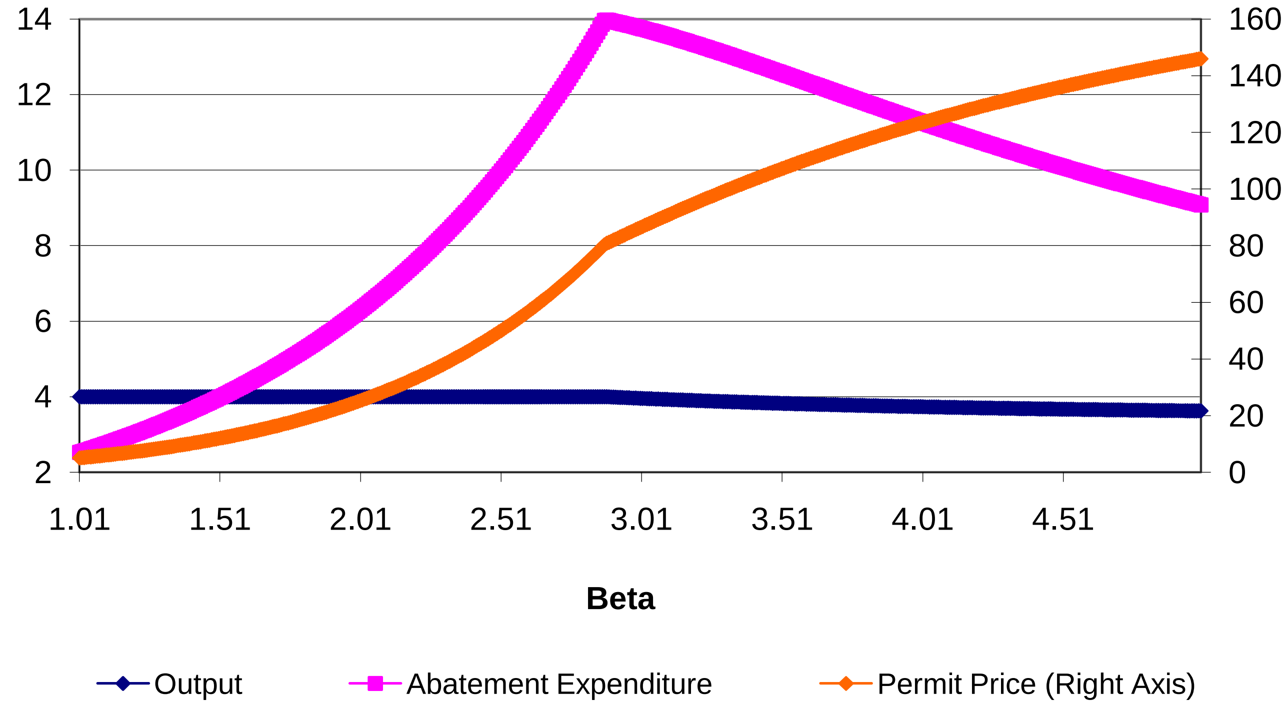


Figure 3b: Output Price and Profits as Functions of Beta

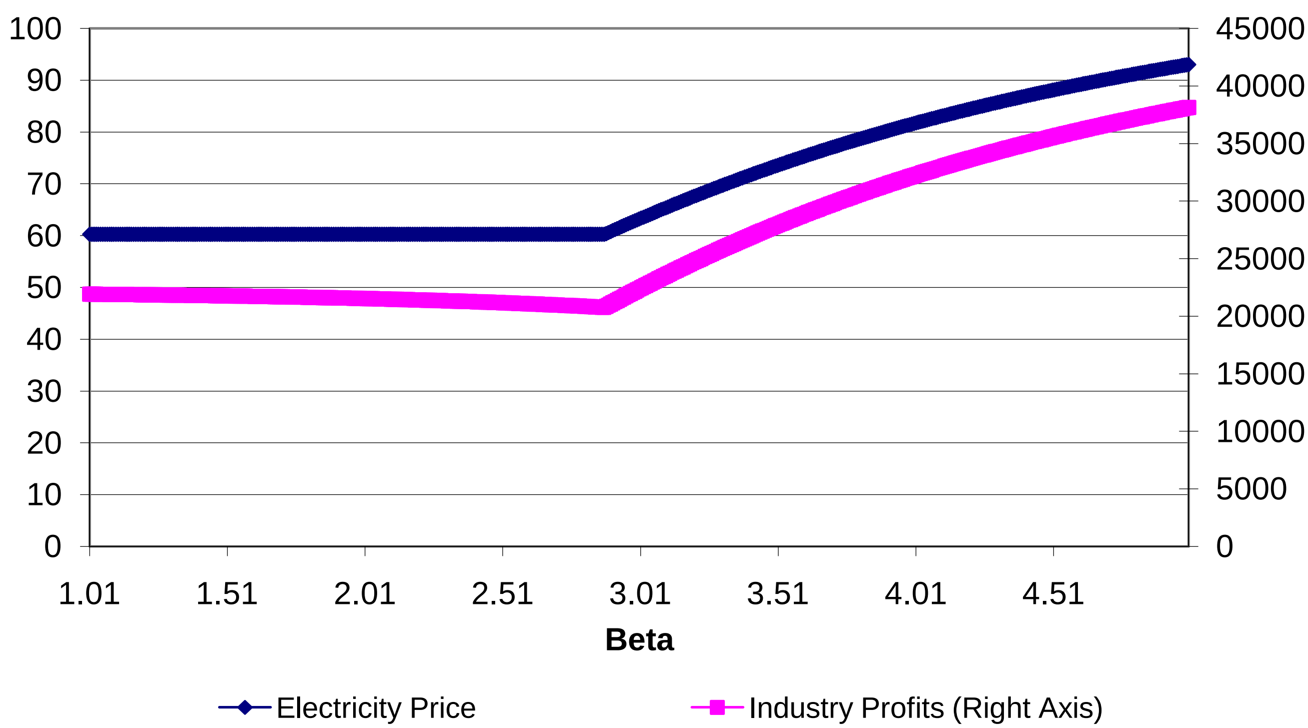


Fig 4a: Abatement and Permit Price as Functions of the Output Price Cap

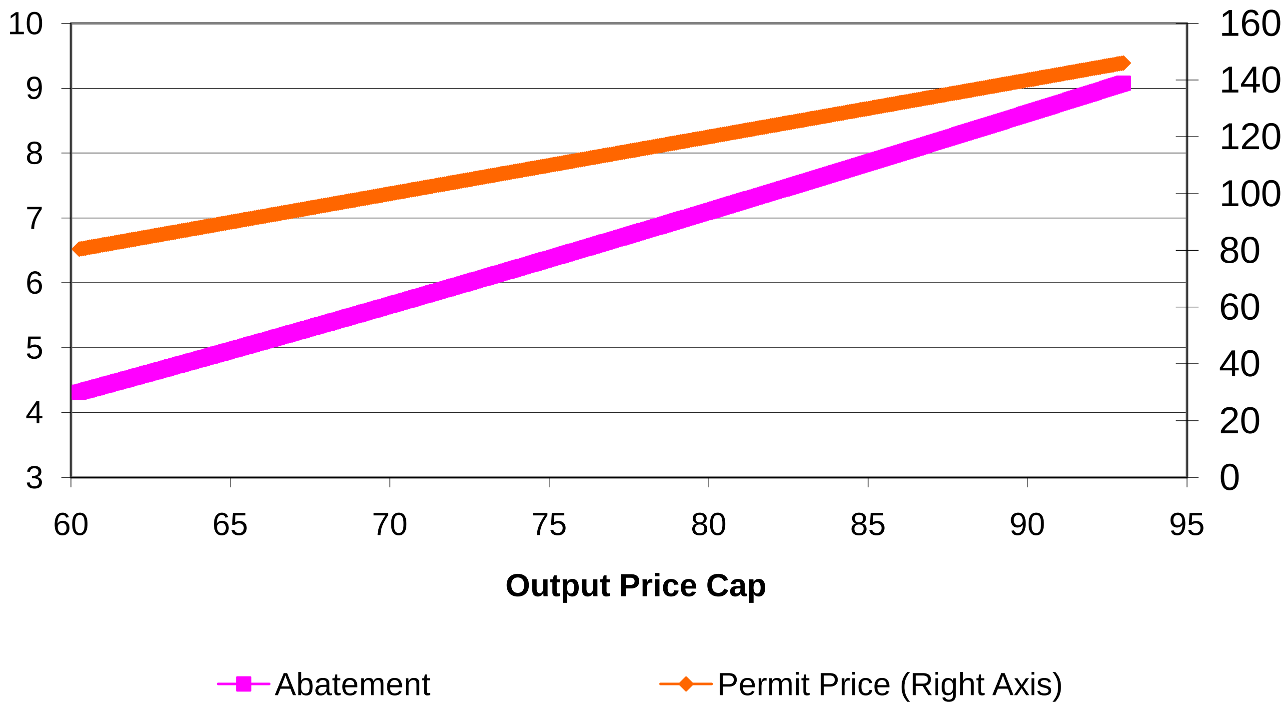


Fig 4b: Output and Profits as Functions of the Output Price Cap

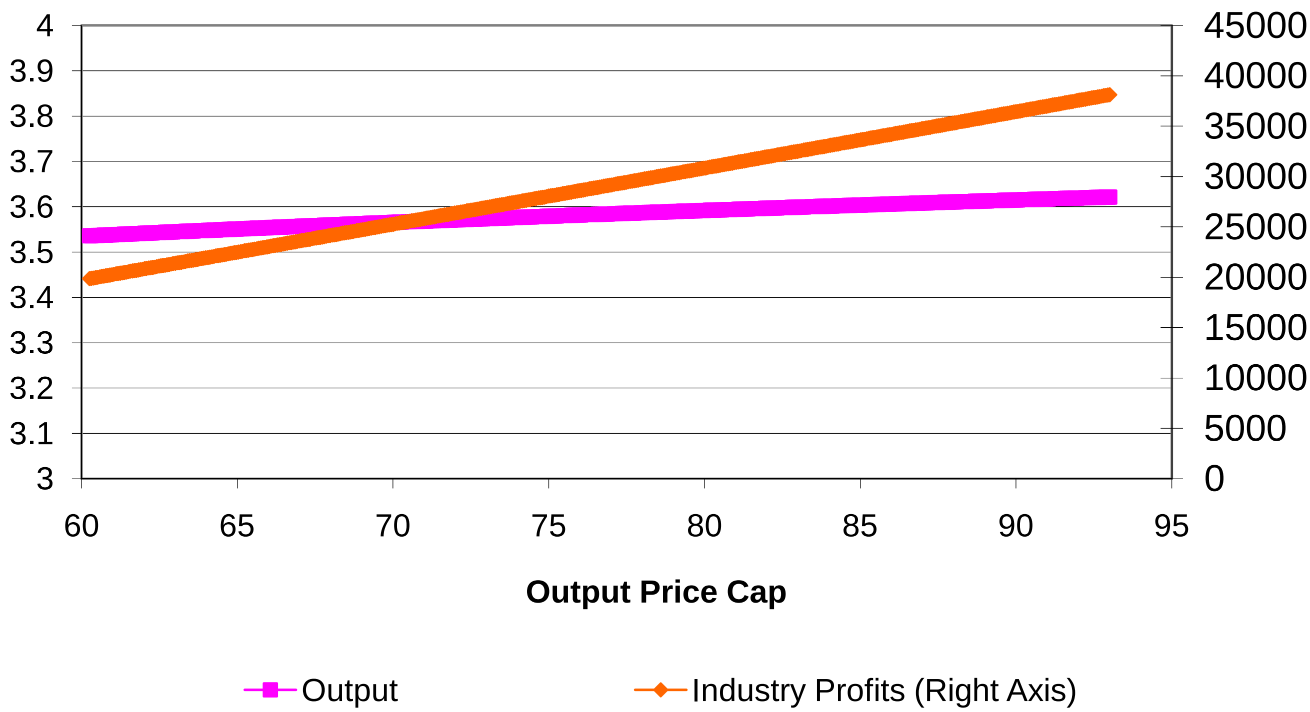


Fig 5a: Baseload Output, Permit Price and Peaking Output as Functions of Peaking Unit Emission Rate

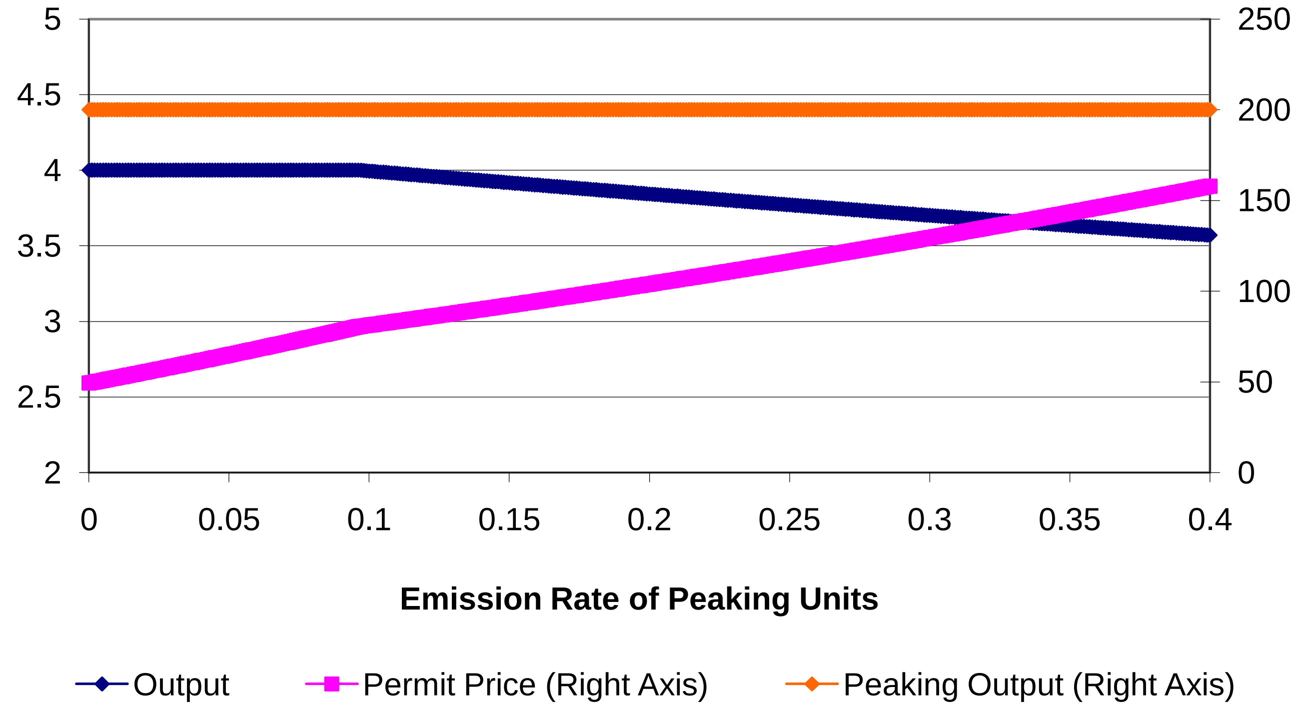
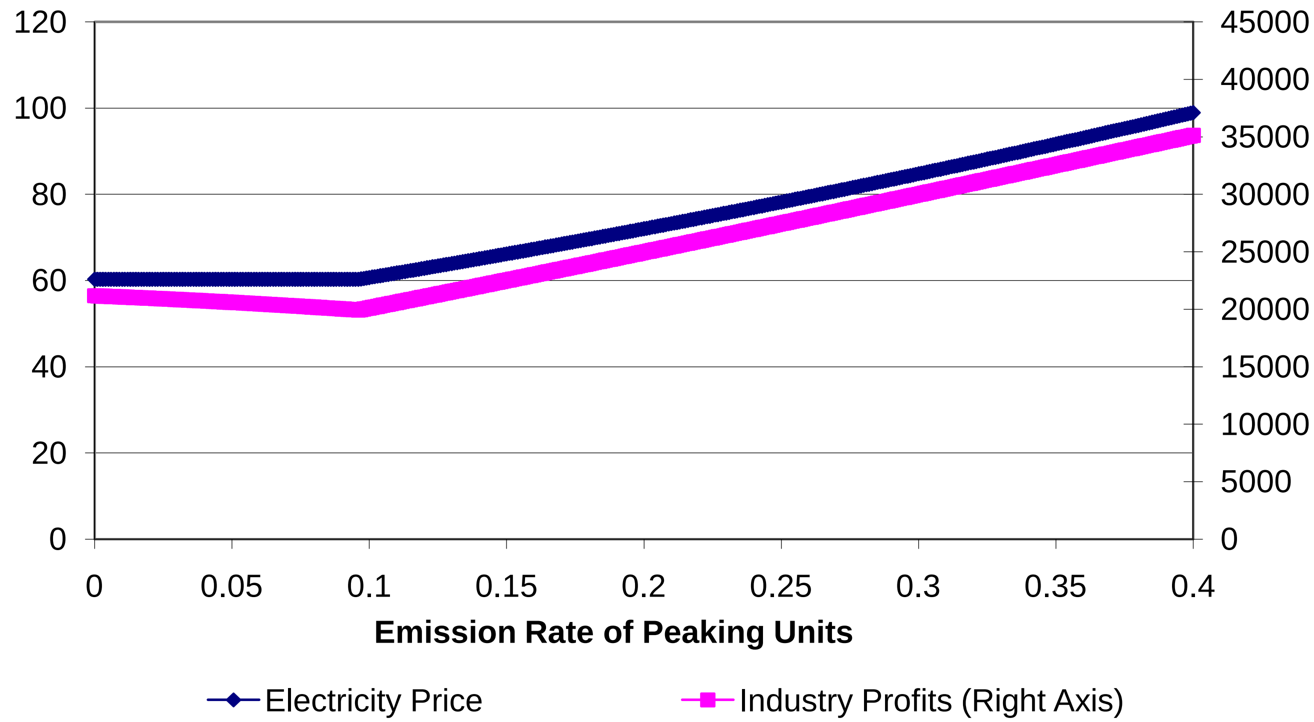


Fig 5b: Output Price and Profits as Functions of Peaking Unit Emission Rates



Session III: Winners and Losers of Cap and Trade

Discussant Comments

Dr. Ann Wolverton

National Center for Environmental Economics, US EPA

Evaluating potential trade-offs between auctioning and giving away allowances for free

- Giving allowances away for free is often discussed as a way to compensate those that incur the cost of regulation
  - Much of the policy debate has been about how to adequately compensate firms
  - But firms pass most of the costs of regulation onto consumers in the form of higher prices.
  - Within the electricity sector, consumers are estimated to incur a loss approximately eight times greater than producers when allowances are auctioned.
- Auctioning creates revenue that can be used to:
  - Offset pre-existing tax distortions, which increases economic efficiency and reduces the overall cost of the policy
  - Compensate households for increases in the prices of energy intensive goods
- In cost-of-service regions, auctioning allowances results in increased electricity prices which improves economic efficiency since it moves these utilities closer to marginal cost pricing.

- Given the regulated nature of the electricity sector, the question examined by both studies is how many allowances would you have to give away for free to fully compensate utilities and does it effectively target the intended recipients?
  - Palmer et al also evaluate alternative compensation schemes where allowances are distributed directly to consumers via local electricity distribution companies
  - Evans et al evaluate the question of firm compensation in the face of endogenous cost reductions and its effect on firm profit

The method of allowance allocation matters in the electricity sector

- Normally, how allowances are allocated should not affect prices
  - Even if allocated for free, firms recognize the opportunity cost associated with an allowance and incorporate that value into decision-making
  - Instead the consequences are usually distributional – who captures the value of the allowances?
- Why is the electricity sector the exception?
  - Not all markets are competitive
  - Some are cost-of-service areas, which use average cost pricing.

The method of allowance allocation matters in the electricity sector

- In the electricity sector, pricing in competitive markets remains unaffected by the method of allocation, but cost-of-service regions do not: the method of allowance allocation affects production decisions.
  - Auctioned allowances or free allocation upstream of the electricity sector means that the opportunity cost of allowances will be passed along to consumers in the form of a higher price of electricity in both cost-of service and competitive regions.
  - Free allocation to electricity producers benefits consumers in cost-of-service regions (prices faced by the consumer do not rise), but benefits shareholders in competitive regions (costs are passed through to the consumer).
    - Palmer et al find that the lack of a price increase in cost-of-service regions leads to increased consumption of electricity, greater imports from neighboring competitive regions, and further increases in the price of electricity in competitive regions relative to an auction.
  - Free allocation to consumers via local electricity distribution companies largely erases the differences in burden caused by differences in regulatory structure.
    - Because allowances are allocated on the basis of electricity use, electricity prices are kept low.
    - This form of allowance allocation benefits consumers but foregoes the price signal that induces consumers to conserve electricity, making it more expensive to achieve given emission reductions (the permit price will rise).

Palmer and Evans papers

- Both papers focus on the impacts of a cap-and-trade policy.
- They examine the electricity sector and evaluate the effects of different compensation rules in a dynamic framework
- Both papers are successful in conveying that the story is more complicated than what you might initially anticipate due - in part - to the regulatory structure of the electricity sector.
- They stay agnostic with regard to who should be compensated.
- Instead, they concentrate on the distributional and sometime efficiency implications of different compensation schemes given a more careful treatment of the electricity sector.

Many policy questions arise with regard to compensation

- How much weight should we put on compensating “losers” versus other revenue uses (e.g., funding R&D, offsetting pre-existing tax distortions)?
- Should we compensate firms or consumers?
  - The literature presents evidence of diffuse but greater impacts on consumers relative to producers
  - The literature finds that some energy-intensive industries are likely to be affected, at least in the short run but sometimes also in the longer run.
  - Evans et al show that firms in the electricity sector respond to a cap-and-trade policy by reducing output, which increases profits (due to inelastic demand).
- With regard to firms, how does one determine negative impacts from the policy? At the industry level? At the firm level? Short run or longer run?
  - Palmer et al show that the level of aggregation matters in determining negative impacts.
  - Evans et al demonstrate that because abatement costs are expected to diminish in the longer run, firm profits gained through the policy will also diminish relative to initial predictions. This is because firms are then less likely to reduce output but also may imply higher levels of compensation in the long run.
- Can we minimize strategic behavior by firms to qualify for compensation?

Many policy questions arise with regard to compensation

- How do we minimize overcompensation of producers for their loss in value?
  - Palmer et al find that allocating allowances for free is a “blunt instrument” for compensation since it rewards both winners and losers of the policy. LBA may help with this but it also has the already discussed efficiency impacts on the electricity sector by foregoing demand reductions in cost-of-service regions.
  - Evans et al show that using auction revenue to subsidize new abatement technology or renewables substantially reduces profits of existing generating units.
- How do we evaluate trade-offs between efficiency and equity? How close can we get to an efficient outcome using different compensation rules?
  - A perfect example of this is illustrated with load-based allocations: the trade-off between keeping electricity prices low and foregone energy efficiency improvements, but these trade-offs exist in other forms:
  - Auctioning allowances and then compensating consumers via a reduction in the payroll tax improves economic efficiency of the policy but can be highly regressive.
  - Auctioning allowances and then compensating consumers via a cap-and-dividend policy foregoes the economic efficiency improvement but is more progressive in nature.

Comments On Two Papers on Compensation

Terry Dinan

April 29, 2009

These are my own views and should not be interpreted as those of the Congressional Budget Office

Evans, Lang and Linn: Findings

- **Find that decreases in compliance costs (renewable or abatement) *decrease* existing generators' profits under free allocation**
  - Relevant for technological innovations and subsidies
  - **Could** be used to justify giving a larger fraction of allowances to existing generators
- **Previous literature finds small fraction of allowances necessary to offset losses to existing producers when measured at industry level:**
  - Bovenberg and Goulder find less than 20% of allowances needed to hold fossil fuel suppliers and energy-intensive manufacturers harmless
  - Burtraw and Palmer find only 6% of allowances (economy wide system) necessary to hold electricity sector harmless
  - Those estimates based on given set of technologies.

Thoughts on Evans, Lang and Linn

- **Interesting that cost-lowering innovations can increase "need" for firm compensation**
  - Underscores need to think carefully about subsidizing renewable energy technologies or abatement (if not justified by other market failures)
- **How meaning full are industry-level loss estimates?**
  - Mask individual firms' gains and losses
    - Burtraw and Palmer estimate that compensating industry level losses for electricity-sector cap-and-trade would take 6% of allowance value; making losers whole would take 11%.
  - Over-compensating some firms and under-compensating others is likely to be inevitable.
    - Burtraw and Palmer's 11% estimate could climb to over 40% due to overcompensation.
- **Identifying a loss does not necessarily imply a need to compensate because not all losses can be offset. Trade-offs will be inevitable**

Paul, Burtraw and Palmer Findings

- **Previous literature:**
  - No previous analyses of load-based allocations
  - Burtraw and Palmer find free allocation to utilities benefits consumers of regulated utilities only, creates efficiency cost
- **Paul et al. highlights several points:**
  - Load based allocations suppress price increases to *all* electricity consumers.
  - Relative to direct allocation to utilities:
    - Potentially “more fair” – customers of both regulated and unregulated utilities benefit
    - Increases efficiency cost
  - Regional Distributional results vary depending on method of load based allocation
    - Population vs. emissions vs. consumption

Thoughts on Paul, Burtraw and Palmer

- **Allocations to LDCs blunt price increases:**
  - Efficiency cost
    - Creates need for higher prices elsewhere in the economy
    - Magnitude of efficiency cost depends on elasticity of demand in electricity sector
      - Sensitivity analysis could be helpful
  - Is policy more “fair”?
    - Policy subsidizes electricity industry
      - Possible to measure effect on industry profits?
    - Consumers face disproportionate burdens in other ways
      - Implications for regional and income-related effects?
- **Broader coverage decreases ability to protect low-income households**
  - Giving all households some protection for some price decreases available funds for focusing on most vulnerable households
  - LDCs lack information to target relief to low-income households.

Allocation Involves Hard Choices

- **Full Compensation is impossible: More losers than winners**
  - Based on CBO illustrative example (2000) allowance value could be roughly 85% of total policy costs
    - Based on conservative assumption that cost of tax-interaction cost same as resource cost.
  - Some of the allowance value will be used to offset government costs (roughly 25%)
- **Inability to avoid over-compensation reduces fraction of costs that can be offset**
  - E.g., generator costs total to 11% of allowance value but need significantly higher fraction (20-40 %) when overcompensation is taken into account.
  - Uses of allowance value considered by CBO and others over-compensates some households and under-compensates others

Papers Highlight Trade-Offs: No "Right" Answer as to How to Distribute Allowances

- **Producer losses**
  - Very hard to identify individual firm losses and avoid overcompensation
  - Description of producer losses as "concentrated" not accurate if shareholders have diversified portfolios
    - Shareholder losses typically borne by higher-income households
- **Worker losses**
  - Extremely concentrated on workers and communities
  - Very little analysis of the magnitude or of compensation methods
- **Consumer losses**
  - Wide-spread losses across all households
    - Some protection through COLA-adjusted incomes (Soc. Security)
  - Regressive effects
  - Compensation through local distribution companies
    - Provides widespread compensation for fraction of costs
    - Raises overall costs (if not lump sum)
    - Can't target low-income households
    - Has different regional effects depending on allocation method