

**Transaction Costs and Tradable Permit Markets:
The United States Lead Phasedown**

Suzi Kerr

*Motu Economic Research
New Zealand
Suzi.Kerr@motu.org.nz*

David Maré

New Zealand Department of Labour

30 June, 1998

Abstract

We econometrically estimate the effect of transaction costs on the cost-effectiveness of the tradable permit market created during the US lead phasedown. We develop a methodology to identify transaction costs in the absence of price data. We find that refineries generally trade efficiently. We also, however, find evidence that transaction costs affect trading. We find evidence that refineries are less likely to trade in cases where theory suggests they will face high transaction costs. The data were collected from 30 major oil companies and include the trading partners and quantities traded for all permit trades carried out by each of 87 refineries over 8 quarters.

1 Introduction

The allocative efficiency of markets is of interest in most applied fields of economics, including labor, health, and finance. In the regulation of environmental externalities, allocative efficiency is a major factor in instrument choice. Tradable permits can achieve the same level of cost-effectiveness as taxes if the permit market allocates permits perfectly. If the permit market operates poorly, the cost-effectiveness is closer to that of command and control.

In a permit market, firms theoretically trade until their marginal costs are equal, in order to achieve cost effectiveness. In reality, however, trading may be costly, so firms may choose to forgo some gains from trade to reduce their trading costs. This paper shows that real tradable permit markets can achieve high levels of cost-effectiveness. We offer the first econometric analysis of a tradable permit market using an original data set from the market created by the US Environmental Protection Agency (EPA) during the US lead phasedown. We look at why some refineries traded consistently while others did not. We investigate whether these patterns are explained primarily by differences in the potential gains from trading, or by differences in the transaction costs faced by refineries.

We use data for 87 refineries, from 30 different companies, over 8 quarters. Using an original methodology, we estimate the loss of cost effectiveness resulting from ‘first-trade’ transaction costs. ‘First-trade’ transaction costs are defined as the cost of making one trade rather than not trading. We find a loss on the order of 10 - 20 percent of potential gains from trade. This loss of cost-effectiveness comes from failure to trade, and from the direct cost of transactions. We also find consistent patterns of failure to trade. Refineries that are part of small companies, smaller refineries, and refineries that do not have other refineries to trade with within their company, more frequently choose not to trade.

The results have important implications for the design of regulations dealing with domestic environmental concerns, and with international problems such as global climate change. First, a permit market can be a cost-effective instrument. Thus it should be seriously considered as an alternative to taxes. Second, when most potential traders in a planned market are small, non-integrated, and unsophisticated, the market is less likely to be efficient. In these situations, an alternative instrument, such as a tax, could be more cost effective.

Since Coase's 1960 observation that individuals will deal efficiently with externalities, if property rights are fully allocated and there are no transaction costs, economists have been developing the idea of using tradable permit markets as a cost-effective solution to the problem of externalities. The theoretical literature on these markets has developed over the last 30 years and several markets have been created, including the US Lead Phasedown Permit Market and the US Acid Rain Program.¹ Simulations suggest that tradable permit markets can achieve significant efficiency gains relative to command and control.² These simulations, however, ignore the possible effect of transaction costs. Hahn and Hester (1989) identified some of the sources of transaction costs in three US environmental programs. They emphasized the importance of administrative requirements in both reducing the potential for trading, and raising the costs of trading. They show a relationship between the simplicity of the property rights and trading requirements, and the economic success of the program, by considering three case studies. Stavins (1995) shows the theoretical effect of transaction costs on cost-effectiveness in a simple model.

Our results are consistent with hypotheses proposed in earlier work on tradable permit markets. In particular, they provide empirical evidence to support Hahn and Hester's contention that the existence of well-established intermediate product markets in the refinery industry facilitated trading. Our results also provide a deeper understanding of their observations from aggregate data on the vitality of the market, and the difference in behavior between large and small refineries.

In the other major current contribution to the empirical analysis of transaction costs in permit markets, Foster and Hahn (1995) use transaction specific data to look at the effect of transaction costs on prices in the Los Angeles Emissions Trading market. They suggest that the administrative rules used by EPA have large effects on transaction costs. They provide evidence of the direct financial costs incurred by firms when they trade. They cannot provide strong evidence on the more indirect effects of regulations on transaction costs. The complexity of the

1 See Dales (1968), and Kneese and Schultze (1975) for the first discussions of tradeable permit markets. Montgomery (1972) showed conditions under which ambient permit systems are efficient. Tietenberg (1985) discusses a wide range of theoretical issues relating to the design of tradeable rights markets under different conditions. A recent branch of literature has used simulations to study the effects of sequential and bilateral trading on the efficiency of markets. For example see [Atkinson and Tietenberg \(1991\)](#). For discussion of the US Acid Rain program see Stavins (1998), Joskow and Schmalensee (1997) and Joskow et al (1996).

2 For a summary of some of the important simulation results see Tietenberg(1985). Predicted gains vary between

program and the illiquid market make rigorous modeling of trading decisions difficult. In addition, the permits are heterogeneous, and the data set used by Foster and Hahn is limited (154 trades over 8 years).³ In contrast, our paper deals with an extremely simple and relatively liquid market enabling us to create a reasonably complete model of trading decisions. We have a large data set (659 observations over 2 years) on trades of a homogeneous good. The transaction specific data is linked to data on firm characteristics.⁴ We can empirically isolate the effects of transaction costs.

The basic characteristics of the permit market are described in Section 2. Section 3 develops an optimization model for the refineries involved in the market. The model shows how the potential gains from trade are derived, and how the decision whether to trade is made in the face of transaction costs. We discuss the possible sources of these transaction costs, drawing on the theoretical literature and interviews with players involved in the market. We derive the econometric models from this optimization model. The econometric models represent a new approach to the study of market efficiency. The estimation procedure is based on observations of quantity traded rather than price, because price information is absent in our data set. Selection models are used to identify the parameters of the net supply of permits and the transaction cost function.

Section 4 describes the data set and variables used, and presents the two different methods of estimation. The data set is unique because it identifies the buyer, the seller, and the quantity traded in every trade, and matches this to the characteristics of each refinery. Section 5 presents the results. We conclude in Section 6 with some simulations of different measures of market efficiency based on the empirical estimates. We discuss some interpretations of the results and implications for future regulatory design.

2 The United States Lead Phasedown Tradable Permit Market

The lead phasedown tradable permit market was one part of an ongoing phasedown of lead in gasoline. The phasedown was a response to concerns about the effects of lead on

7% and 95% depending on the pollutant and the region.

³ They are also forced to deal with the problem that the regulatory changes affect the value of the permits as well as the transaction costs involved in trading them.

⁴ A promising new data set is being generated out of the SO₂ Allowance market. Some preliminary work has been done by Bailey (1996).

children's IQ and adult hypertension, and to the need to stimulate the provision of unleaded gasoline for use in cars with catalytic converters.⁵ The phasedown began in 1974 with a requirement for unleaded gasoline use in new cars. Regulations were introduced in 1980 that limited lead in gasoline to 4 grams per gallon on average over a refinery's production during one quarter. This standard was non-binding for most refineries. The rapid phasedown of lead per gallon of leaded gasoline, and the tradable permit market, began in November 1982.

Refineries use lead to raise the octane value of the gasoline they produce. Lead is a useful input, but is not essential; substitutes exist at a higher price. The costs of the substitutes depend on the characteristics of the refinery, and the outputs they produce. They vary among refineries and across quarters. These costs determine whether the refinery is a buyer or a seller of lead permits at the market price for each quarter.

The EPA allocated the permits to use lead in gasoline in a particular quarter on the basis of the amount of leaded gasoline produced by the refinery during that quarter. A refinery that produced 100m gallons of leaded gasoline during any quarter of 1983 or 1984 received 1.1 grams of lead permits per gallon, or 110m grams of lead permits to use or sell during that quarter.⁶ Permits expired at the end of each quarter. The rights were allocated to each refinery and were fully tradable between refineries in the same or different companies. EPA required no pre-trade approval.

The allocation procedure led to an incentive to over-produce leaded gasoline. This poor permit design lowers the efficiency of the regulation; the same environmental outcome requires a stricter standard per unit of output. In the lead phasedown this problem was limited by the fact that unleaded gasoline was steadily taking over the gasoline market. There may have been some over- production relative to an ideal system, but it was not visible in aggregate; leaded gasoline production continued to fall. A similar idea, 'output-based' allocation has been proposed for carbon permit markets where the efficiency losses would be much greater.

5 The ex ante cost-benefit analysis of this program is found in Schwartz et al (1985). Catalytic converters are used to reduce vehicle emissions. If leaded gasoline is used in a car with a catalytic converter it "poisons" the converter and makes it ineffective. Unleaded gasoline is considerably more expensive to produce than highly leaded gasoline and there are incentives to "misfuel" cars thus reducing the effect of air pollution control measures.

6 After 1984, the standard became more stringent and rights were able to be "banked" for use in future quarters. "Banked" permits could be used until the end of 1987 when the phasedown ended. We do not use data from the banking period. Trades after 1984 cannot be interpreted solely as responses to differences in marginal values; they may also be driven by speculation.

During 1983 and 1984, the market for lead permits was active. Between 300 and 400 potential traders existed, and between 167 and 345 trades occurred each quarter. Refineries bought between 10 and 20 percent of all the lead permits they used from other refineries. The number of trades increased over time, as did the total percentage of previously-traded lead permits used by refineries.⁷

2.1 Transaction Costs

The focus of this paper is on the role that transaction costs played in the operation of the lead phasedown tradable permit market. Transaction costs can make otherwise profitable trades unattractive, and thus reduce the amount of trading, and the gains from trade. In our empirical work, we identify potential gains from trading for refineries, and explain the failure of some refineries to trade in terms of transaction costs.

We define transaction costs as the costs incurred in making a trade, which would not occur if the trade did not take place.⁸ Our reasons for believing that transaction costs may have been an important feature of the lead phasedown tradable permit market are grounded on a knowledge of the specific features of the market, as well as on insights from the literature about the costs of market exchange.⁹ In this section, we discuss briefly some of the descriptive and anecdotal evidence of transaction costs in the market.¹⁰

One form of transaction cost is the cost of optimizing. One indication of optimization difficulties is that of the trading reports that the General Accounting Office sampled in 1986 around 20% contained errors that did not lead to commercial advantage and thus were probably not deliberate.¹¹

A second source of transaction costs is the cost of using the price mechanism, i.e. search costs (Coase 1937). The trader must find information about the distribution of market prices, then find a trading partner.¹² The market for lead permits was a decentralized confidential

7 These numbers come from aggregate statistics available from EPA from 1 July 1983 onward.

8 We are not concerned with government's administrative costs, or other compliance costs that are independent of whether the refinery trades that quarter.

9 This literature developed from Coase (1937). A summary of the major strands of the literature is found in Williamson and Winter (1993).

¹⁰ A fuller discussion of types of transaction cost is contained in Kerr & Maré (1996).

¹¹ Government Accounting Office (1986).

¹² If transaction costs are high we might expect specialized brokers to emerge to reduce these costs. A few brokers did gradually emerge but they faced the same hindrances to trading as refineries themselves, (e.g.: confidentiality and credibility) so were able to play only a limited role in the market. To a large extent it appears that most brokerage among refineries was done by the trading departments of large companies rather than by separate entities. One

market. Because the value of lead was commercially sensitive, refineries and others did not always advertise broadly to make trades. Large companies appear to have faced lower search costs, because they have models of refinery production for the US as a whole, which allow them to estimate equilibrium price, because they were able to rely on trading within the company, and because they had established trading networks for intermediate products.¹³

A third sort of transaction cost arises from private information about the validity of the rights being traded.¹⁴ The permits to use lead were created by the accounting of individual refineries. They were traded before the EPA verified them. Because of this, some problems arose with both deliberate and accidental use of invalid permits.¹⁵ Concerned about the possibility of buying invalid permits, refiners incurred costs checking the reputation of potential trading partners and checking the validity of the individual permits offered for sale. Compliance by major refiners was very good for the most part. Buyers were more likely to take precautions with small refineries and refineries without established reputations.¹⁶ Small, non-integrated refineries faced high credibility costs.

A fourth transaction cost is the cost of negotiation. Refineries find negotiating easier if they are existing trading partners, or come from similar companies. Large companies may prefer to deal with other large companies. It may be harder for a small company to gain access. The importance of negotiation is less in more liquid markets. In a very liquid market, the price will be the equilibrium price and traders will not be very concerned about the quantity traded. Negotiation costs will be lower for larger companies and in later years.

company, which appears to have played a brokerage role, both bought and sold externally in nearly every quarter of the program and was particularly active in the early years of the market when transaction costs could be expected to be highest.

13 (Telephone interview with ex-manager of lead trading in a large oil company, 29 June 1993.) In a medium sized company in contrast, each refinery traded separately with some centralized coordination. (Telephone interview with manager of medium sized company, 21 April 1993.) Internal trading was extremely important among refineries. In our sample, 67% of the quantity bought was bought within the company, and 70% of the quantity sold was sold within the company.

14 This is similar to the costs of certifying the existence of emission reduction credits in the EPA emissions trading program. Foster and Hahn (1995) find that banked credits which have already been certified, trade at a premium over non-certified credits where there is greater risk.

15 In 1986 the Government Accounting Office randomly sampled 374 reports and found that 35% contained errors. (Government Accounting Office (1986) p. 19.) Refiners made errors in 36.4 percent of cases, blenders in 49.2 percent, and importers in 14.4 percent. (Blenders are small players such as service stations which add alcohol to gasoline. This reduces the lead content and therefore allows them to claim permits.)

16 A broker in the market said that one of their main tasks was to establish the credibility of the permits they were selling on behalf of small sellers (Telephone conversation with Marilyn Herman of Herman and Associates).

The fifth transaction cost arises because trading lead permits involves the release of some confidential information about the likely outputs of the refinery. The difficulty others and we have experienced gaining access to information on this market may be indicative of the level of concern about confidentiality in this market. A manager in one company said that the price information was extremely sensitive and consequently 95% of their trading was done internally.¹⁷

Many people discuss transaction costs as fixed and variable costs. While this may represent markets with formal brokers reasonably well, we do not believe this is a useful or clear distinction in this market. Traders can carry out the activities that are necessary to trade the first time with different degrees of care. For example, if a trader plans to trade a small amount once, he will put a minimal effort into finding trading partners; if he plans to trade many times over many periods a more serious effort will be justified. Almost all transaction costs will be borne to a certain extent in all trades; the relative weight of each will vary with the quantity traded, number of trades, and length of trading relationship.

3 A Model of Refinery Decisions

In this section, we model the effect of transaction costs on a refinery's trading behavior. We first present an optimizing model of refinery behavior in the absence of transaction costs, and then discuss the refinery's response to transaction costs.

3.1 Optimization without Transaction Costs

We model each refinery, i , as a unitary, profit-maximizing actor. The revenue of refinery i is the cross product of the output vector (g_i, x_i') , where g_i is leaded gas production and x_i' is production of all other jointly produced outputs, and the output price vector (p_g, p) where p_g is the leaded gasoline price, and p is a vector of other output prices. If the refinery decides to trade, it can buy or sell permits Q_i (net permits sold) at the market price T . The constraint is that lead used within the refinery, l_i , plus the permits sold, cannot exceed the standard of 1.1 grams per leaded gallon. When less than 1.1 grams are used per gallon, the marginal value of lead is always greater than its cost, so the constraint will bind.¹⁸

The costs for refinery i are a function of inputs (l_i, f_i) , where f_i is a vector of all inputs other than lead, outputs, the current technology of the refinery, k_i , and exogenous refinery

¹⁷ Telephone Interview with Technical Advisor in a large oil company 21 April 1993; our data gives 94% internal trading for this company.

characteristics r_i , such as its location and the size of the company it belongs to.¹⁹ During 1983-84, the refinery's static optimization problem is:

$$\text{Max}_{g_i, x_i, l_i, f_i} \quad \Pi_i = g_i p_g + x_i' p + T_i Q_i - C_i(l_i, f_i; g_i, x_i, | k_i, r_i) \quad (3.1)$$

$$\text{s.t. } l_i + Q_i = 1.1 g_i$$

The first order conditions for (3.1) give the following equations.

$$(i) \quad p_{x_i} = \frac{\partial}{\partial x_i} C_i(l_i, f_i; g_i, x_i, | k_i, r_i) \quad \forall x_i \in x_i ;$$

$$0 = \frac{\partial}{\partial f_i} C_i(l_i, f_i; g_i, x_i, | k_i, r_i) \quad \forall f_i \in f_i \quad (3.2)$$

$$(ii) \quad p_g + 1.1T = \frac{\partial}{\partial g_i} C_i(l_i, f_i; g_i, x_i, | k_i, r_i)$$

$$(iii) \quad T = \frac{-\partial}{\partial l_i} C_i(l_i, f_i; g_i, x_i, | k_i, r_i)$$

$$(iv) \quad Q_i = 1.1 g_i - l_i$$

The sets of equations, (i), are the normal cost minimizing conditions. (ii) is the normal cost minimizing condition for leaded gasoline, except that due to the structure of the regulation an inefficient incentive to over-produce leaded gasoline exists. Equation (iii) shows that lead permits will be sold until the increase in cost from using one less unit of lead in production is equal to the market price of lead permits. Equation (iv) gives the net supply of lead permits to the market.

This is a non-linear system of equations. The non-linearities could be defined by engineering parameters to yield a system of linear equations that can be solved to find the optimal value of l_i , and hence Q_i , as well as x_i , g_i and f_i . Linear programming models of refineries, such as the Department of Energy REMS model, are used by large oil companies to solve such problems.²⁰ In a reduced form these yield an equation for the marginal value of lead, MV_i , in the production of refinery i .

18 From the aggregate EPA data we calculate that 96% of all lead permits are used.

19 Refinery characteristics are exogenous in the short run but not the long run. They are almost certainly exogenous with respect to lead demand which is a small aspect of overall refinery operations.

20 We investigated using this approach and abandoned it because we would have needed to predict the optimal inputs and outputs of each refinery as well as calculate the marginal value of lead. This was impossible to do with any accuracy given our knowledge and the information available.

$$MV_i = MV_i(p_g, p, g_i, x_i, Q_i, f_i, k_i, r_i) \quad (3.3)$$

We assume the marginal value of lead, MV_i , is linear in lead use, and hence lead permits sold. We define MV_i^0 as the marginal value at the no-trade solution and γ_i as the slope.²¹ (See Figure 3.1.)

$$MV_i = MV_i^0 + \gamma_i Q_i \quad (3.4)$$

MV_i is the cost of creating an additional lead permit to sell. Figure 3.1 shows the derived demand/supply of lead permits for a buyer j and a seller i . At the optimum, $T = MV_i$. Inverting this, we solve for the net supply of permits. This is the quantity that would be sold in the absence of transaction costs.

$$Q_i^* = \frac{T - MV_i^0}{\gamma_i} \quad (3.5)$$

The potential gains from trade are simply the consumer or producer surplus the refinery can receive in the market.

$$V_i^*(Q_i^*) = \frac{1}{2} Q_i^* (T - MV_i^0) \quad (3.6)$$

The potential surplus depends on two things. How far the refinery's marginal value of lead diverges from the market price of lead permits at the no-trade solution, $(T - MV_i^0)$, and the slope of their marginal product of lead function, γ_i . For simplicity we assume that all trades occur at the equilibrium price T .²²

3.2 Responses to Transaction Costs

When faced with transaction costs, a refinery will trade only if their potential surplus is at least as large as the transaction costs. There will therefore be some refineries with positive potential surplus from trade who do not trade. We define a refinery's 'first trade' as the highest net value trade available to them. This is the trade the refinery would make if it traded only once. We define V_{ij} as the surplus i will gain by trading with j , and TC_{ij} as the transaction cost to i from this trade. The net value from the first trade equals $\text{Max}_j \{V_{ij} - TC_{ij}\}$.²³ We can see that a refinery will not trade if $V_{ij} - TC_{ij} < 0$ for all j , i.e. if trade is unprofitable with every partner j

²¹ MV_i is the derived factor demand.

²² We also assume that no refinery or company can influence the average market price. This is not an unreasonable assumption, given the number of competitors in the gasoline, and hence lead permit, market.

²³ In reality, both parties must agree to trade. This affects each refinery's actual trading opportunities and generates a complex two-way matching problem that is beyond the scope of this paper.

even though $V_{ij} \geq 0$. Our empirical identification of transaction costs is based on identifying refineries that have positive potential surplus but do not trade.

In this paper, therefore, we are interested only in how refineries choose whether to trade at all. What probably matters most for cost-effectiveness is not the transaction costs of all potential trades, but the transaction cost of the ‘first trade’. The prohibitively high transaction costs associated with some potential trades will never be observed if the trader has a profitable alternative trading strategy. Trading yields diminishing marginal gains. If a refinery trades at least once in each quarter, it may not achieve all the possible gains from trade, but it will probably achieve a significant portion of the potential gains. We assume that gains from the first trade are proportional to total potential gains, and the ratio does not vary by observable refinery characteristics.²⁴

We model the transaction cost from the highest net value trade, TC_i , as a cost per quarter independent of the quantity traded or number of trades. The cost depends on the characteristics of the refiner and the quarter, X_i , but not the trading partner.

$$TC_i = TC(X_i) \tag{3.7}$$

The refinery will choose to trade (Trade = 1) at least once if the potential gains from trade exceed the first-trade transaction cost.²⁵ S^* is the latent selection variable.

$$\begin{aligned} \text{Trade} &= 1 && \text{if } S^* = V^*_i - TC(X_i) > 0 \\ &= 0 && \text{otherwise} \end{aligned} \tag{3.8}$$

3.3 *Econometric Model*

The aim of the econometric model is to estimate the size of first-trade transaction costs. We estimate potential surplus, and infer the size of transaction costs by observing that transaction costs must be greater than or equal to potential surplus for those refineries that we observe not trading and the converse for those we observe trading. Our estimate of potential surplus is based on an estimate of the quantity that a refinery would have traded, even if we do not observe them trading. We use observations of quantity traded from those refineries that do trade, and model the

²⁴ We only identify transaction costs as a percentage of surplus, so any fixed constant of proportionality has no effect on our results. For simplicity we normalise it to 1.

²⁵ In the first order conditions (3.2) we saw that refineries had an incentive to overproduce gasoline. This will increase the allocations of permits and lower the permit price. In the absence of transaction costs, it will have no effect on trading patterns. If transaction costs cause the allocation method to affect trading, we expect sellers with low transaction costs to overproduce and profit by selling more, and potential buyers with high transaction costs to over produce and avoid the need to buy. The effect on total gasoline was indiscernable at the aggregate level,

quantity as a function of covariates that are observed for all refineries. By controlling for selection into trading, we predict "latent" quantity traded for those refineries that do not trade.²⁶

Q_i^* is parameterized as a function of $Z_i'a$, where Z_i is a vector of refinery and quarter characteristics and 'a' is a vector of coefficients. The chemical engineering factors determining the marginal value of lead all operate on the response of lead per gallon of leaded gasoline, suggesting that it is appropriate to think of $Z_i'a$ as an index related to lead per gallon of leaded gasoline.²⁷ In addition, lead permits were allocated per gallon of leaded gasoline. It is therefore appropriate that the net supply of permits in equation 3.9 is scaled by g_i , the number of gallons of leaded gasoline produced by refinery i .

$$Q_i^* = g_i Z_i'a + \varepsilon_i \quad \varepsilon_i \sim N(0, \sigma_\varepsilon^2) \quad (3.9)$$

From our estimate of latent quantity traded, we wish to estimate the surplus that would have been obtained from that trade (latent surplus). As noted above, we have no data on marginal product or price so we cannot estimate the relationship between quantity traded and surplus. Thus we need to assume the form of the relationship between Q_i^* and V_i^* .

The imposition of a functional form on this relationship is unavoidable and, unfortunately, may not be innocuous. If the errors in our prediction of surplus are systematically related to the variables affecting transaction costs, we will get biased estimates of transaction costs. For example, if we tend to over-estimate the surplus received by small refineries, we will also tend to over-estimate the transaction costs faced by small refineries. Small refineries will appear not to trade despite large estimated surpluses.

To control partially for this problem we use two different models for MV_i and hence potential surplus. In both we use the construction of gains from trade in (3.6). The estimation of quantity is identical in the two models, but the surplus estimation differs. The two models can be illustrated with reference to Figure 3.1. Q_i^* is observed (for traders), but $(T-MV_i^0)$ is not, because of the absence of price data.

however, so we do not expect these effects to be important.

²⁶ The use of selection models such as this are commonly used in the analysis of labor supply. See Maddala (1983) for a fuller discussion of these techniques.

²⁷ These factors include the amount of lead per gallon, and the response to other available lead substitutes. The octane responses to lead and many lead substitutes are non-linear. In addition, for many substitutes, only limited amounts can be added per gallon because of their effects on other characteristics of the gasoline. A refinery with more and cheaper alternatives will replace more of the lead in each gallon of gasoline.

In Model A, surplus is proportional to the square of Q_i^* . This is achieved by setting $(T - MV_i^0)$ - the divergence of refinery i 's marginal product of lead from the market price for permits - equal to $Z_i'a$ and γ_i equal to $\frac{1}{g_i}$. Combining these assumptions with (3.5) we get the equation for net permit supply given earlier (3.9), and combining with (3.6) we get Model A surplus (the first term in 3.11A).

In Model B, surplus is modeled as a linear function of Q_i^* . We set $(T - MV_i^0) = \text{sign}(Z_i'a)$, i.e. 1 for sellers, and -1 for buyers. The slope of the MV curve, γ_i , is set equal to $|\frac{1}{g_i Z_i'a}|$. This gives us the same equation for net permit supply but a different surplus (the first term in 3.11B).

First-trade transaction costs are modeled as a latent variable that depends linearly on observable characteristics of the refinery and company.²⁸ b is a coefficient vector to be estimated.

$$TC_i = X_i'b + \mu_i \quad \mu_i \sim N(0, \sigma_\mu^2) \quad (3.10)$$

The latent first-trade transaction costs are identified through their effect on selection into trading, controlling for the latent surplus achievable through trading. Selection into trading depends on whether the latent surplus exceeds the latent transaction costs. Substituting for $(T - MV_i^0)$ and γ_i in the surplus equation (3.6), and using (3.10), we get the following two selection equations.²⁹

$$\begin{aligned} \text{Model A} \quad S_i^* &= \frac{1}{2} Z_i'a Q_i^* - TC_i \\ &= \frac{1}{2} g_i (Z_i'a)^2 - X_i'b + \frac{1}{2} Z_i'a \varepsilon_i - \mu_i \end{aligned} \quad (3.11A)$$

$$\begin{aligned} \text{Model B} \quad S_i^* &= \frac{1}{2} |Q_i^*| - TC_i \\ &= \frac{1}{2} |g_i(Z_i'a)| - X_i'b + \varepsilon_i - \mu_i \end{aligned} \quad (3.11B)$$

4 Data and Estimation

4.1 Data and descriptive statistics

The self-reporting requirements of the EPA lead-trading program generated the primary source of data in this paper.³⁰ Each quarter, each refinery reported to the EPA the names of the refineries they sold to (bought from), the quantity of lead they sold or bought in each trade, and

²⁸ We constrain transaction costs to be positive in the estimation by simply using $\text{POS}(TC_i)$.

²⁹ Note that the form of the trading surplus in Model A creates a specific form of heteroscedasticity in 3.11 A.

the quantity of leaded and unleaded gasoline they produced. EPA collected these data confidentially and they have not previously been available for analysis. We gained permission to use the data directly from the oil companies involved. Our data are for a sample of 87 refineries. Table 4.1 shows the characteristics of our sample of refineries relative to the total population. We have data on 27% of the companies, 48% of refineries and 65% of industry capacity. We observe 48% of all trades. No price data were collected. We use data on the years 1983 and 1984. Our refineries and companies are larger than the average and our results should be considered in that light.

For our econometric analysis we use a given refinery in a particular quarter as the unit of observation. That is, each refinery counts as an observation in each quarter where they were operating. We observe complete data for 659 “refinery quarters”. For each refinery we aggregate quantity traded over all trades within a quarter to get ΣQ_i as a proxy for Q_i^* .³¹ We do not use information on the number of trades per quarter. Figure 4.1 shows the number of observations where a refinery trades 0, 1, 2, ..., 24 times. Most trade once or not at all, but many trade several times in one quarter.

Table 4.2 gives some basic statistics on trading in our data set. Although only one refinery never trades, only 37 refineries out of 87 trade every quarter; seven sell permits in every quarter; five always buy permits. 58 refineries buy permits in some quarters and sell in others. Two thirds of both consistent traders and those that occasionally do not trade, swap sides of the market in this way. The lack of consistency in trading suggests that the demand for lead by a given refinery is very variable across quarters.

Variables in the "latent supply of permits" equation

We estimate the latent supply of permits as a function of observable variables. The production of leaded gasoline (LEADED), used to scale permit supply, is reported as part of the EPA trading program. Some refineries were explicitly created to produce solely gasoline, partly

30 US. Environmental Protection Agency, Form 40-CFR80.20.

31 We use ΣQ_i , which is observable, to estimate Q_i^* , the latent potential quantity variable. When the refinery makes at least one trade, this will be related to but not equal to the potential quantity Q_i^* defined in (3.5). ΣQ_i could be greater or less than Q_i^* . In order to focus on estimating the transaction cost of the first trade, we assume that Q_i^* is correlated with ΣQ_i when the refinery does trade, and that the difference between them does not depend on the observable characteristics of the refinery. The more liquid the market is, the more realistic is this assumption. In other words, we assume that refineries that make at least one trade choose to trade a level that is an unbiased

in response to the regulation of crude oil imports. Because of this specialization they may find it harder to adjust their production to use less lead. As a measure of the specialization of the refinery we use the ratio of total gasoline production (leaded and unleaded) to total crude input (GASRATIO). The EPA trading program recorded unleaded gasoline production as well as leaded. The crude capacity of each refinery is published annually in the US Department of Energy, Petroleum Supply Annual.

The technology of a refinery affects the availability of substitutes to replace lead. Three types of technology are significant, catalytic reforming, alkylation and isomerization. We combine these, and scale by the total capacity of the refinery to create a measure of the availability of alternatives (TECH). Data on the technological capacity of each refinery is published in the Petroleum Supply Annual. Unfortunately, the outputs and technology of a refinery are endogenous. The effect of technology differences on the availability of lead substitutes may be counteracted by unobservable output differences.

The lead phasedown required an adjustment in technology to produce gasoline without lead. In the presence of convex adjustment costs, the optimal dynamic response to a change in regulation will involve gradual investment. Refineries facing lower adjustment costs, sophisticated refineries and refineries in large companies, tend to adjust earlier.³² They could be expected to have more advanced technology relative to their outputs and hence sell more permits during the period when other refineries were still producing with older technology.

As general measures of the sophistication of the refinery and its engineers, we use the total crude input of the refinery, measured in barrels per calendar day (REFSIZE). Similarly, the engineering expertise available to the refinery, and its ability to alter production and technology rapidly in response to regulation, depend on the total amount of production in the company to which it belongs. This is measured as total company crude capacity per calendar day (COSIZE). The ownership of refineries was identified using the Petroleum Supply Annual and Pennwell, Worldwide Refining and Gas Processing.

Because of geographical, demographic and historical factors, prices, inputs and production patterns vary exogenously by location. To control for this variation, we use a series

predictor of their potential level. The difference between Q^*_i and ΣQ_i reduces the efficiency of the estimation but does not create bias.

32 For analysis of the technology response of refineries to the US lead phasedown see Kerr (1997).

of dummy variables for 9 regions. These are defined in Table 4.3. The region definitions were taken from the Bureau of Mines districts used by the Department of Energy, and aggregated consistent with the Department of Energy Refinery Model REMS. The Louisiana Gulf, which produces large amounts of petrochemicals, is the omitted region. Louisiana Gulf refineries were high demanders of lead because of the high octane levels petrochemical inputs require. Thus we might expect the other regions to supply relatively more permits.

Finally, we include a dummy variable for time to take account of changes in the dispersion of values for lead (YEAR84). It takes the values 0 in 1983, and 1 in 1984.

Variables in the "transaction costs" equation

We use four proxies for the first-trade transaction costs arising from optimizing costs, search costs, adverse selection, negotiation and loss of information. We use the size of the refinery (REFSIZE) as a proxy for on-site legal and accounting capacity. Larger refineries are also more likely to trade intermediate products. This means they may be able to draw on existing trading relationships and are more integrated in the market in terms of information.

We use the total capacity of the company (COSIZE) as a proxy for the skill of the traders. Large companies have a high level of sophistication in their modeling capacity. They have access to excellent legal and accounting assistance, and are highly integrated in the industry. Larger companies are more likely to trade in a centralized way through a previously existing trading department. They are generally more established, so are less likely to face credibility problems.

The number of refineries in the company (NREFINCO) is used as a measure of the refinery's access to internal trading. We expect internal trading to have lower transaction costs. Finally, we use the year (YEAR84) to account for learning and the development of relationships that reduce transaction costs.

Definitions, means and standard deviations of all variables are given in Table 4.3. Table 4.4 shows descriptive statistics on trading where the unit of observation is a refinery in a specific quarter. In 139 instances (20%) a refinery does not trade in a given quarter. It is unlikely that 20% have a marginal value of lead exactly equal to the permit price, suggesting, but not proving, the existence of transaction costs affecting the decision to trade. The non-trading refineries (weighted by the number of times they are observed not trading) have observably different average characteristics from traders. They are smaller refineries, belonging to smaller

companies, with fewer refineries in each company. This supports the argument that the first-trade transaction costs vary systematically across refineries, but does not control for variation in the gains from trade.

4.2 Estimation Method

In both Model A and Model B, only the sign of S^*_i is observed and the proxy for Q^*_i is observed only when $S^*_i > 0$. Q^*_i can be positive (selling) or negative (buying). Amemiya (1985) defines this model as Tobit Type II. The basic form of the likelihood function is,

$$\begin{aligned} L &= \Pi_0 P(S^*_i \leq 0) \Pi_1 f(Q^*_i | S^*_i > 0) P(S^*_i > 0) dS_i \\ &= \Pi_0 P(S^*_i \leq 0) \Pi_1 \int_0^{\infty} f(S^*_i, Q^*_i) dS^*_i \end{aligned} \quad (4.1)$$

The joint density can be written as the product of a conditional and a marginal density, that is, $f(S^*_i, Q^*_i) = f(S^*_i | Q^*_i) f(Q^*_i)$.

Because of unobserved output variation, it is difficult to distinguish empirically between firms that are likely to sell and those likely to buy. We therefore estimate the models separately on two sub-samples, sellers and non-traders, and buyers and non-traders, as well as on the total sample. We look at selection into selling in the first sub-sample, and selection into buying in the second. In these subsamples, we have to define the surplus as surplus from selling and buying respectively.

For the supply subsample the relevant surplus for selection is producer surplus. However $(Z_i'a)^2$ and $|g_i(Z_i'a)|$ in (3.11) are always positive even if Q^*_i is negative. We need to indicate whether producer surplus is positive or negative. A similar problem arises in the buyer subsample. w_i is used as an indicator of the sign of producer / consumer surplus. For the supply sample $w_i = 1$ if $Z_i'a \geq 0$, $w_i = -1$ if $Z_i'a < 0$. For the buyer sample the relevant surplus is consumer surplus, so the reverse is true, $w_i = -1$ if $Z_i'a \geq 0$, $w_i = 1$ if $Z_i'a < 0$. For the total sample, $w_i = 1$.

Model A: Surplus as a quadratic function of Q^*_i

We estimate the following model.

$$Q_i^* = g_i Z_i'a + \varepsilon_i \quad (4.2)$$

$$S^*_i = \frac{1}{2} w_i g_i (Z_i'a)^2 - X_i'b + [\frac{1}{2} w_i Z_i'a \varepsilon_i - \mu_i] \quad (4.3)$$

$$\Sigma Q_i = Q^*_i \text{ if } S^*_i > 0$$

$$= 0 \quad \text{if } S^*_i \leq 0$$

where $\{ \varepsilon_i, \mu_i \}$ are iid drawings from a bivariate normal distribution with zero means, variances σ_ε^2 and σ_μ^2 and covariance $\sigma_{\varepsilon\mu}$.³³

Following Amemiya, the conditional distribution $f(S^*_i | \Sigma Q_i)$ is normal with mean $\frac{1}{2} w_i g_i(Z_i'a)^2 - X_i'b + \sigma_{12i} \sigma_1^{-2}(\Sigma Q_i - g_i Z_i'a)$, and variance $\sigma_{2i}^2 - \sigma_{12i}^2 \sigma_1^{-2}$,

where: σ_1 = standard error on the quantity equation (4.2)

$$= \sigma_\varepsilon,$$

σ_{2i} = standard errors on the selection equation (4.3)

$$= (\frac{1}{2})^2 (Z_i'a)^2 \sigma_\varepsilon + \sigma_\mu^2 - 2w_i(\frac{1}{2}Z_i'a) \text{cov}(\varepsilon, \mu), \text{ and}$$

$$\sigma_{12i} = w_i \frac{1}{2} Z_i'a \sigma_\varepsilon - \text{cov}(\varepsilon, \mu)$$

Thus for Model A the likelihood function is:

$$\begin{aligned} L &= \Pi_0 [1 - \Phi\{w_i \frac{1}{2}g_i(Z_i'a)^2 - X_i'b\} / \sigma_{2i}] \\ &\times \Pi_1 \Phi\{ [w_i \frac{1}{2}g_i(Z_i'a)^2 - X_i'b] / \sigma_{2i} + \rho_i(\Sigma Q_i - g_i Z_i'a) / \sigma_1 \} [1 - \rho_i^2]^{-0.5} \\ &\times \sigma_1^{-1} \phi [(\Sigma Q_i - g_i Z_i'a) / \sigma_1] \end{aligned} \quad (4.4)$$

where: $\rho_i = \sigma_{12i} / \sigma_1 \sigma_{2i}$.

We estimate this function using the ML function in TSP. No consistent two step estimators exist for this model. Convergence difficulties arose with the buyer and total samples. We present, therefore, results on only sellers Model A.

Model B Surplus as a linear function of Q^*_i

We estimate the following model.

$$Q_i^* = g_i Z_i'a + \varepsilon_i \quad \varepsilon_i \sim N(0, \sigma_\varepsilon^2) \quad (4.5)$$

$$S^*_i = \frac{1}{2} w_i |g_i(Z_i'a)| - X_i'b + [\varepsilon_i - \mu_i] \quad \mu_i \sim N(0, \sigma_\mu^2) \quad (4.6)$$

$$\begin{aligned} \Sigma Q_i &= Q^*_i \text{ if } S^*_i > 0 \\ &= 0 \quad \text{if } S^*_i \leq 0 \end{aligned}$$

Model B has the same structure as the labor supply models estimated by Heckman and can be consistently estimated using a two step method.³⁴ It satisfies the identification

33 The errors from equations 4.2 and 4.3 are jointly distributed but, because the error term (in square brackets) in equation 4.3 is a function of ε_i, μ_i and values that vary by observation, the covariance matrix for this joint distribution has a complex form that depends on the values of covariates, as well as on the parameters of the bivariate normal distribution from which ε_i and μ_i are drawn.

34 Note that the error structure is much simpler than in Model A – the error in equation 4.6 is not a function of observation-specific values.

requirement that at least one variable in Z_i is not included in X_i .³⁵ The region variables, TECH and GASRATIO are expected to affect the quantity traded but not transaction costs. In addition, the variables in X_i enter without being interacted with g_i .

After the two step estimation, we backed out the parameters of the underlying model. The transaction cost coefficients and the parameters of the error structure are calculated according to the process outlined by Maddala (1983).³⁶

5 Results

The result tables 5.1 and 5.2 present the underlying structural parameters of models A and B. The top part of tables 5.1 and 5.2 reports the transaction cost coefficients, b , where a positive coefficient implies an increase in transaction cost. The lower part reports the quantity equation coefficients, a . Finally the estimated parameters of the error structure are reported. In Model B, the likelihood reported is the likelihood for the complete model, calculated by constructing the likelihood function and calculating the value at the estimated values of the coefficients. Thus it is comparable with Model A. The standard errors for the quantity coefficients are those for the second step of the two-part estimation, unadjusted for the selection. We do not report standard errors for the transaction cost coefficients in Model B because the coefficient estimates are non-linear functions of the estimated parameters.

5.1 Existence of Transaction Costs

We find clear evidence that transaction costs affect trading in this market. Comparing the likelihood of an unconditional OLS with either Model A (Table 5.1) or Model B for any sample (Table 5.2), using the Likelihood Ratio Test at the 99.5% level, clearly rejects the hypothesis that transaction costs are zero. Unconditional OLS is equivalent to Model B with the restriction that $TC_i = 0$ for all i . The likelihood value for OLS is -3160.8 in the total sample, -1784.6 in the seller sample and -1768.29 in the buyer sample.

Because costs are normalized in each model, the value of transaction costs is only meaningful in relation to the estimate of surplus. In Model A, for sellers, average transaction

35 See Maddala (1983) p. 231-232.

36 The variance of the probit is calculated using an average of the estimates given by each of the quantity equation variables excluding outliers.

costs are 31% of average surplus.³⁷ In Model B for sellers, average potential transaction costs are 33% of average potential surplus. For the total sample, in Model B they are estimated at 53%. Of course transaction costs are incurred only if trading occurs and they vary among refineries. These are an upper bound on efficiency losses from first-trade transaction costs.

5.2 *Quantity Equation (Equations 4.2 and 4.5)*

The following results are roughly consistent across models and samples. GASRATIO, reflecting the degree of specialization in gasoline production, significantly reduces the amount of lead sold as expected, in the Model A estimates. GASRATIO is of the expected signs though insignificant, in Model B. The region variables increase net selling in both models, when they are significant. This is consistent with the omission of the Louisiana Gulf, as noted above.

TECH, reflecting the sophistication of technology, is generally insignificant. Where it is significant, more specialized equipment for replacing lead results in higher amounts of lead being bought. The lack of significance of TECH, and the unexpected sign, is likely to be due to the endogenous relationship between technology and unobservable outputs.

An interesting and counter-intuitive result is that larger refineries and companies within our sample do not appear to sell more lead permits. Within our sample, the effects of size seem to be primarily on the volume not direction of trade. Larger refineries, REFSIZE, trade smaller quantities per gallon of leaded gasoline production; they both buy and sell less. This may be due to greater diversification within larger refineries. Diversification allows them to respond cheaply to shocks in lead demand through internal adjustments. In contrast, refineries within large companies, COSIZE, tend to trade more both when buying and when selling. This is likely to be a response to greater specialization by refineries within large companies. In the second year, YEAR, both buying and selling increases, possibly indicating greater variance in valuations of lead as some refineries adjust faster than others.³⁸

5.3 *Causes of Transaction Costs (Equation 3.10)*

The likelihood ratio test strongly suggests that transaction costs exist. Some evidence suggests they vary systematically across refineries. The number of refineries in the company,

³⁷ In the selling sample average transaction costs are 0.9 while average surplus is 2.9. In the selling sample for Model B, average transaction costs are 3 out of average surplus of 9.

³⁸ The larger trades by larger companies and in the second year also could reflect an effect of transaction costs on trading, that is beyond the scope of our model. Transaction costs may bias firms toward smaller trades, or only one

NREFINCO, increases the potential for internal trading. It significantly reduces estimated transaction costs in Models A and B, except where refinery size is excluded. An additional refinery reduces transaction costs by 55% of surplus. A one standard deviation increase in the number of refineries raises the probability of trading by 0.26.

YEAR has the expected negative sign in both models, except for the buyer subsample where there appears to be a small increase in transaction costs over time. Transaction costs generally fall in the second year, but the effect is mostly insignificant.

In Model A, and Model B for sellers, the size of the refinery, REFSIZE, has a negative, large and significant effect on transaction costs. In Model A, a one standard deviation rise in the size of a refinery reduces transaction costs by 183% of average surplus. The probability of trading rises by 0.37 due to the fall in transaction costs. However, large refineries tend to trade less when they do trade. Thus they tend to get smaller surpluses. This surplus effect lowers the probability of trading by 0.078 for a one standard deviation rise in refinery size. Despite this, nearly all large refineries trade in every quarter. The effect of refinery size on transaction costs in the total and buyer samples for model B is unexpectedly positive. The buyer and total samples had very unstable likelihood functions in both Models A and B so we regard these results as less robust.

The size of company, COSIZE, has the expected negative effect on transaction costs in most cases (not for buyers, or in one seller specification), but is not significant. Company size is correlated with refinery size and number of refineries in company. Overall, the evidence weakly supports our contentions that larger refineries, refineries in large companies and refineries in companies with many refineries face lower transaction costs.

6 Interpretation and Conclusion

6.1 Implications for cost effectiveness

The potential level of market efficiency is important to policy makers when designing environmental regulation. They need to choose between three major regulatory approaches, permits, taxes and command and control. The issues involved in this choice have been discussed

trade rather than multiple trades. Lower transaction costs may affect quantity traded as well as the probability of making the first trade.

extensively in the literature.³⁹ These choices involve considerations of cost effectiveness, distributional issues that affect political feasibility, and concerns about the certainty of the cost and environmental outcome.⁴⁰ Taxes have the advantage of low transaction costs and hence higher cost effectiveness. However, as well as leading to environmental uncertainty, taxes often suffer from the problem that the distribution of costs cannot be controlled by politicians, and the groups that will be taxed may be needy or powerful.⁴¹ If the regulator sets standards on technology of particular companies, he has considerable discretion in the distribution of costs as well advantages in monitoring, but standards impose serious efficiency costs. Tradable permit markets offer a way to achieve flexibility in the distribution of costs through the initial distribution of permits, while also achieving cost effectiveness.

Traditionally, when analysis is done to compare tradable permits to other forms of regulation it is assumed that the permit market will achieve the least cost outcome and hence be as cost effective as taxation. Our results suggest that this may be a fair approximation for some players but not all. The inclusion of transaction costs does improve the explanatory power of the model.

Tables 6.1, 6.2 and 6.3 give overall simulation results on market efficiency. In 6.1 we see that in the total sample, 79 percent of refineries trade. The ‘quantity’ measure, %Q, is the model estimate of quantity actually traded, compared with estimated total latent/potential trade, $\frac{\hat{Q}_i}{\hat{Q}_i^*}$. The ‘surplus’ measure, %Surp, is estimated realized surplus net of transaction costs as a percentage of estimated latent surplus, $\frac{1(\text{if } \hat{S}^* > 0)(\hat{V}_i^* - T\hat{C}_i)}{\hat{V}_i^*}$. For the total sample, Model B predicts quantity and net surplus efficiency of 85 percent and 72 percent respectively.⁴²

In Tables 6.2 and 6.3 we look at the selling and buying samples. 67 percent of the selling sample trade.⁴³ In the buyer and seller samples, trading efficiency is biased downward relative to

39 e.g. Stavins and Whitehead (1992).

40 See Weitzman (1974) and an extension in Stavins (1995).

41 For a discussion of the political and distributional issues as well as other issues related to the choice of regulatory instrument see Hahn and Noll (1990) and Hahn and Stavins (1991). For discussion in the context of US climate change policy see Cramton and Kerr (1998).

42 However the model under-predicts trading participation (70%) so these efficiency estimates are also under-estimates.

43 The models predict the proportion trading fairly accurately (74% for Model A and 69% for Model B).

the total sample because each sub-sample includes all non-traders. The models predict that around 90 percent of latent quantity is sold, and 80-90 percent of possible net surplus is achieved. In the buyer sample, 63 percent trade, we estimate that 95 percent of latent quantity is actually bought, and 76 percent of potential net surplus is realized.⁴⁴

The results from Model A are the most stable. The selling results in Model A are consistent in sign with the selling results in Model B. Both predict trading efficiency quite well. Therefore we have the most confidence in the results on the seller sample. Using the results from the seller sample suggests high levels of cost effectiveness in this market, between 80 and 90 percent. This is roughly consistent with the downward biased estimate of 72% in the total sample and 76% in the buyer sample. Thus our estimates imply a loss of cost effectiveness from non-trading of around 10 to 20 percent.

In our model, and hence in all of these estimates, we are considering only losses of efficiency from direct transaction costs and from refineries choosing not to trade. The estimates do not take into account inefficient trades due to poor choice of partner, or lower (or higher) quantity traded than optimal. For this reason these are overestimates of trading efficiency.

Our results suggest that the characteristics of the refinery and year affect transaction costs. Tables 6.1 - 6.3 give results of simulations showing how these findings affect overall trading efficiency for specific groups of refineries. Large refineries, large companies, and companies with many refineries, achieve high levels on all measures of efficiency. Transaction costs seem to most affect refineries in small companies or in companies with few refineries, with trading efficiency generally under 60 percent and dropping frequently below 50 percent. Small refineries seem to be less affected. Within our sample, many small refineries are part of large companies. The year of trading appears to have almost no effect on overall trading efficiency.

7 Conclusion

We have used selection models to identify ‘first-trade’ transaction costs (the cost of making one trade rather than none) and analyze the sources of such transaction costs. We have produced rough estimates of the cost-effectiveness impact of first-trade transaction costs in the US market for lead permits. We find significant evidence for the existence of first-trade transaction costs. Within our sample of generally larger companies, the loss of cost effectiveness

⁴⁴ The model predicts 62% trading.

from these costs is around 10 to 20 percent. We also find some evidence that first-trade transaction costs are dependent on the characteristics of traders and the market in theoretically predictable ways. The losses of cost effectiveness are highest for refineries in companies with few refineries and refineries in small companies.

Our results provide empirical support for the contention that tradable permit markets can be a cost-effective regulatory mechanism. Permit markets are quite efficient in a simple active market such as that in the lead-trading program. Markets probably work most efficiently where the information flows are good and where legal risk is minimized by good enforcement, for example among and within large companies.

On the other hand, our results can also be used to extrapolate to markets where trading conditions are not or will not be so favorable. The lead permit market was very close to a textbook tradable permit market. However, even in the lead permit market, some agents faced high transaction costs, and suffered significant efficiency losses. If a market were composed of small, unsophisticated players, or players with few existing connections among themselves, we might expect their behavior to reflect and exaggerate the behavior of the small players in our sample. This is consistent with the low trading levels observed in the EPA Emissions Trading program.⁴⁵

Tradable permit markets can be a cost-effective instrument for environmental regulation. They may, however, be less cost-effective when the potential traders are unsophisticated and poorly integrated, and where there is poor enforcement or under-developed brokerage and trading mechanisms.

45 For analysis of the EPA emissions program see Hahn and Hester (1989) and Foster and Hahn (1995).

References

- Amemiya, T. (1985) *Advanced Econometrics*. Harvard University Press, Cambridge MA.
- Atkinson, S., and Tietenberg, T. (1991) "Market Failure in Incentive Based Regulation: The Case of Emission Trading." *J. of Environmental Economics and Management*, 21
- Bailey, Elizabeth M. (1996) "Allowance Trading Activity and State Regulatory Rulings: Evidence from the U.S. Acid Rain Program" MIT-CEEPR 96-002 Working Paper, March
- Coase, R. H. (1937) "The Nature of the Firm." *Economica* N. S. 4,: 386-405.
- _____ (1960) "The Problem of Social Cost." *J. of Law and Econ.* 3
- Cramton, Peter and Suzi Kerr (1998) "A Tax-Cut Auction for the Environment: How and why to auction CO2 emissions permits" Resources for the Future Discussion Paper 98-34
- Dales, J. (1968) *Pollution, Property and Prices* (Toronto: University Press)
- Foster Vivien, and Robert W. Hahn (1995) "Designing More Efficient Markets: Lessons from Los Angeles Smog Control" *Journal of Law and Economics* 38 (1) April pp. 19 - 48
- Government Accounting Office, (1986) Report to the Chairman, Subcommittee on Oversight and Investigations, Committee on Energy and Commerce, House of Representatives. "Vehicle Emissions — EPA Program to Assist Leaded-Gasoline Producers Needs Prompt Improvement." GAO/RCED-86-182, August
- Hahn, Robert W. and Hester, Gordon L. (1989) "Marketable Permits: Lessons for Theory and Practice." *Ecology Law Quarterly* 16
- _____ and Roger G. Noll (1990) "Environmental Markets in the Year 2000" *Journal of Risk and Uncertainty* 3, pp. 351 - 367
- _____ and Robert N. Stavins (1991) "Incentive-based Environmental Regulation: A New Era from a Old Idea" *Ecology Law Quarterly* 18(1), pp. 1 - 42
- Joskow, Paul L., Richard Schmalensee, (1997) "The Political Economy of Market-Based Environmental Policy: The U.S. Acid Rain Program" *Journal of Law and Economics*, forthcoming.
- Joskow, Paul L., Richard Schmalensee, and Elizabeth M. Bailey (1996), "Auction Design and the Market for Sulfur Dioxide Emissions" Working Paper, MIT-CEEPR 96-007WP.
- Kerr, Suzi C. and David Maré (1996) "Efficient Regulation Through Tradable Permit Markets: The United States Lead Phasedown" Department of Agricultural and Resource Economics, University of Maryland at College Park, Working Paper 96-06

- Kneese, Allen V., and Charles L. Schultze (1975) *Pollution, Prices and Public Policy* (Washington D.C.: Brookings Institution)
- Maddala G.S. (1983) *Limited-Dependent and Qualitative Variables in Econometrics* Econometric Society Monograph No. 3 (Cambridge University Press: New York)
- Montgomery, David W. (1972) "Markets in Licenses and Efficient Pollution Control Programs." *J. of Econ. Theory*, 5
- Pennwell Publishing Company (1991) *Worldwide Refining and Gas Processing*
- Schwartz, J., Pitcher, H., Levin, R., et al. (1985) "Costs and Benefits of Reducing Lead in Gasoline: Final Regulatory Impact Analysis." EPA-230-05-85-006; U.S. Environmental Protection Agency: Office of Policy Analysis; pp. 1-6 - 1-12.
- Stavins, Robert. N. (1995) "Transaction Costs and Tradable Permit Markets." *Journal of Environmental Economics and Management*
- _____ (1995) "Correlated Uncertainty and Policy Instrument Choice" *Journal of Environmental Economics and Management*
- _____ (1997) "What Can We Learn from the Grand Policy Experiment?: Positive and Normative Lessons from SO₂ Allowance Trading," *Journal of Economic Perspectives*, forthcoming.
- _____ and Bradley W. Whitehead (1992) "Pollution Charges for Environmental Protection: A Policy Link between Energy and Environment" *Annual Review of Energy and the Environment* 17, pp. 187 - 210
- Tietenberg, Tom. (1985) *Emissions Trading: An Exercise in Reforming Pollution Policy* Washington: Resources for the Future
- US Department of Energy, *Oil and Gas Journal* 1984 -1987
- US Department of Energy, Energy Information Administration *Petroleum Supply Annual* 1982 - 1987.
- US Department of Energy, Energy Information Administration Office of Oil and Gas, Refinery Evaluation Modeling System (REMS) Model Documentation 1984.
- US EPA - Quarterly Reports on Lead in Gasoline.
- Weitzman, Martin L. (1974) "Prices vs. Quantities." *Review of Economic Studies* 41
- Williamson, Oliver E. and Winter, Sidney G. eds. (1993) *The Nature of the Firm: Origins, Evolution and Development* (Oxford University Press: New York).

8 Figures and Tables

Figure 3.1 Optimal Trading with No Transaction Costs

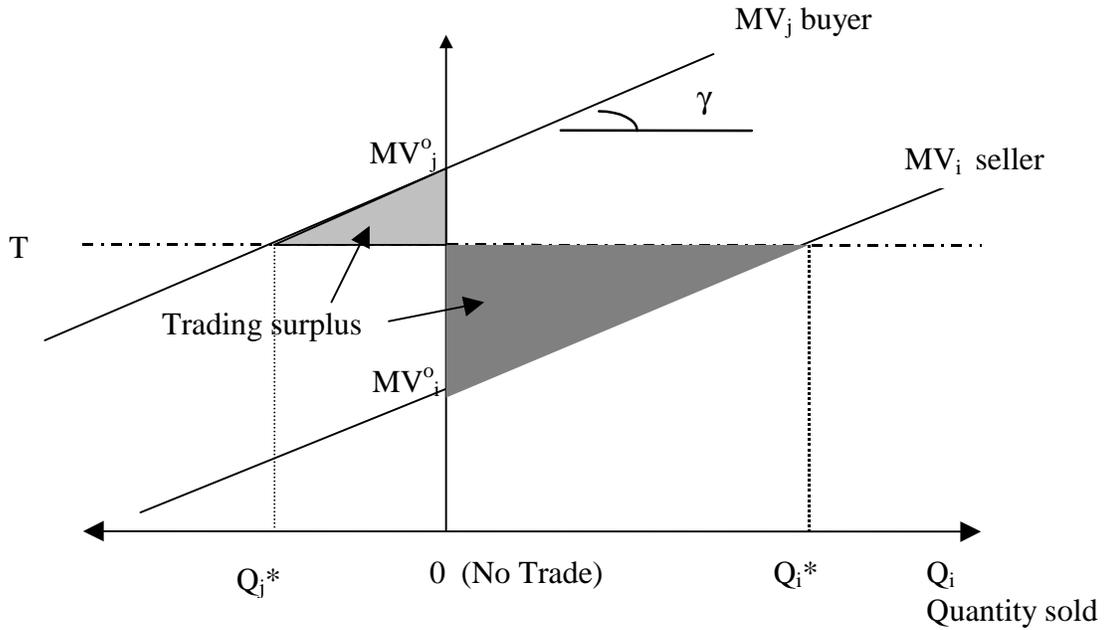


Figure 4.1 Distribution of Number of Trades Per Quarter

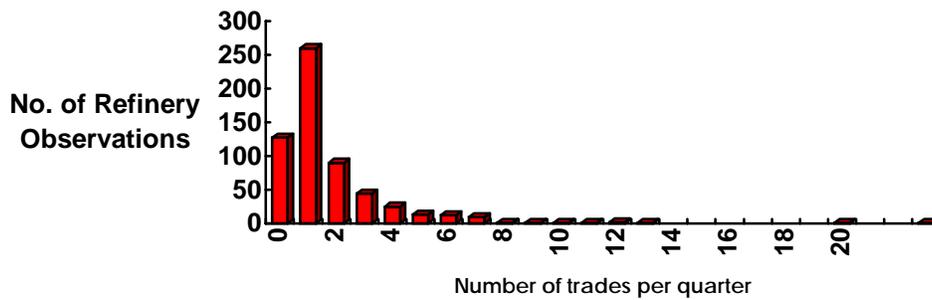


Table 4.1 Characteristics of Sample for years 1983-84⁴⁶

Characteristic	Sample	Aggregate	Percent
Number of Companies	30	110	27%
Number of Refineries	87	191	46%
Total Capacity	10.46 m b/cd ⁴⁷	16.2 m b/cd	65%
Average Capacity	0.13 m b/cd	0.085 m b/cd	
Average Number of Refineries in company	4.6	1.8	
<i>Jul. 1983 - Dec. 1984⁴⁸</i>			
Number of sellers ⁴⁹	36.5	71	51%
Total Quantity Sold ⁵⁰	5.5 b grams	9 b grams	61%
Number of buyers ³⁷	29	52	56%
Total Quantity Bought	3.8 b grams	8 b grams	47%
Total Market Number of Trades	700	1457	48%
Average Size of Trade	9.5 m grams	7 m grams	
Leaded Gasoline Production	36 b gallons	68 b gallons	53%

Table 4.2 Number of refineries observed not trading, buying, and selling

Never trade	1
Sell in every quarter	7
Buy in every quarter	5
Buy in some quarters, sell otherwise	25
Buy in some quarters, don't trade otherwise	6
Sell in some quarters, don't trade otherwise	10
Observed buying, selling, and not trading	33
TOTAL	87

46 These characteristics are based on data on the industry at 1 January 1986 from the Petroleum Supply Annual.

47 Million barrels per calendar day.

48 The period is limited here because aggregate data is available only from July 1983 onward. This includes only refiners with capacity over 5000 b/c.d.

49 This is the average number of refineries which sell (buy) per quarter.

50 Quantity sold is not equal to quantity bought because there are entities other than refineries in the market.

Table 4.3 Definition and Means of Variables

Variable	Definition	Total Sample	Seller Sample	Buyer Sample
		Mean (std dev)	Mean (std dev)	Mean (std dev)
ΣQ	Net permits sold (m grams)	0.28 (37.8)	15.16 (20.63)	-16.66 (38.81)
	Absolute value of quantity traded	19.19 (32.61)	15.16 (20.63)	16.66 (38.81)
LEADED	Quantity of leaded gasoline produced (m gallons)	88.82 (74.66)	81.80 (60.02)	86.66 (85.20)
REFSIZE	Refinery crude capacity (m b/cd)	0.12 (0.12)	0.11 (0.10)	0.12 (0.12)
COSIZE	Total Capacity of company (m b/cd)	0.66 (0.46)	0.63 (0.46)	0.60 (0.46)
NREFINCO	Number of Refineries in Company	4.61 (2.21)	4.40 (2.24)	4.42 (2.31)
YEAR 84	Dummy for 1984	0.50 (0.50)	0.51 (0.50)	0.49 (0.50)
Variables Interacted with Leaded Gasoline Production				
REFSIZE	Refinery crude capacity (m b/cd)	17.79 (29.83)	13.18 (20.99)	19.43 (34.25)
COSIZE	Total Capacity of company (m b/cd)	69.64 (85.08)	59.78 (68.72)	68.36 (96.27)
TECH	High Octane producing capacity	16.56 (14.01)	15.07 (11.38)	16.09 (15.68)
GASRATIO	Gasoline production relative to total capacity	180.27 (186.09)	173.42 (169.59)	175.42 (204.56)
YEAR84	Dummy for 1984	40.81 (62.61)	39.23 (56.48)	38.89 (65.26)
EAST	East Coast	8.28 (27.67)	8.15 (27.81)	8.33 (28.20)
MIDWEST	IL, IN, KY, MI, OH, TN	11.50 (43.14)	11.12 (42.20)	11.91 (42.32)
NTHCNTRL	MN, ND, SD, WI	3.03 (19.04)	3.65 (22.74)	1.41 (7.89)
SOUTH	AL, AR, LA Inland, MS	7.54 (27.11)	9.04 (30.28)	5.84 (21.68)
TXINLAND	NM, TX Inland	4.03 (18.32)	4.52 (19.43)	4.24 (18.56)
TXGULF	TX Gulf	20.22 (62.38)	9.47 (39.55)	27.17 (72.71)
LAGULF	LA Gulf (omitted dummy)	15.63 (38.65)	22.71 (45.19)	6.22 (24.33)
WEST	AK, AZ, CA, HI, NV, OR, WA	4.31 (14.59)	5.68 (17.10)	2.21 (10.22)
ROCKIES	CO, ID, MT, UT, WY	14.27 (51.54)	7.44 (32.29)	19.33 (62.60)

Table 4.4 Basic Trading Statistics for 1983/84 Sample

Unit of Observation	Total	Buyers	Sellers	Buy and Sell	Don't Trade
Refinery in a given quarter	666	252	311	36	139
Characteristics of Average Refineries					
REFSIZE (m b/cd)		0.14	0.14	0.16	0.091
COSIZE (m b/cd)		0.65	0.73	0.55	0.46
NREFINCO		4.79	4.85	4.42	3.60
GASRATIO		1.73	1.89	1.89	1.73
percentage in 1984		51.2	49.8	58.3	51.1

Table 5.1 Model A (Sellers and Non-Traders)

Transaction Costs	I ⁵¹	II	III
CONSTANT	8.2** (2.2)	7.73** (2.06)	6.32** (2.15)
NREFINCO	-1.6* (0.81)	-1.02* (0.47)	-1.35 (1.25)
YEAR84	-0.63 (0.93)	-0.77 (0.97)	-1.52# (1.16)
REFSIZE	-53** (21)	-52.7** (22.0)	
COSIZE	4.1 (3.6)		-9.01 (19.83)
Quantity Equation	<i>Dependent Variable = Net Quantity Supplied</i> <i>All variables are multiplied by leaded gasoline production</i>		
C	0.24** (0.054)	0.24** (0.054)	0.24** (0.055)
REFSIZE	-0.68** (0.11)	-0.68** (0.11)	-0.68** (0.12)
COSIZE	0.12** (0.03)	0.12** (0.03)	0.12** (0.034)
TECH	0.13 (0.19)	0.14 (0.19)	0.14 (0.19)
GASRATIO	-0.037** (0.012)	-0.037** (0.012)	-0.036** (0.013)
EAST	0.064* (0.031)	0.066* (0.031)	0.060* (0.031)
MIDWEST	-0.0035 (0.03)	-0.0052 (0.033)	-0.010 (0.033)
NTHCNTRL	0.26** (0.046)	0.25** (0.046)	0.25** (0.046)
SOUTH	0.089** (0.036)	0.088** (0.036)	0.081* (0.036)
TXINLAND	-0.029 (0.075)	-0.026 (0.076)	-0.033 (0.076)
TXGULF	0.044 (0.035)	0.042 (0.035)	0.041(0.035)
WEST	0.11** (0.028)	0.11** (0.028)	0.11** (0.028)
ROCKIES	0.17** (0.056)	0.16** (0.057)	0.15** (0.058)
YEAR84	0.022# (0.15)	0.022# (0.015)	0.022# (0.015)
σ_{ϵ}	17.48** (0.70)	17.49** (0.72)	17.45** (0.69)
σ_{μ}	4.89** (0.90)	5.21** (0.91)	5.00** (0.94)
Cov (ϵ, μ)	43.52* (20.1)	48.64** (20.78)	43.03* (21.05)
E(ρ_i)	-0.0049 (0.21)	-0.059 (0.20)	-0.0012 (0.20)
Log Likelihood	-1437.9	-1438.99	-1442.06
N	423	423	423

= Coefficient significant at 90%; * = Significant at 95%; ** = significant at 99%; ⁵²

51 The likelihood function was very “flat” near this solution and we were unable to make the function converge at a tolerance level of 0.001. It converged at 0.05.

52 These are not true standard errors because the observations on one refinery are not completely independent over time.

Table 5.2 Model B Heckman 2-Step Results

Transaction Costs	Total Sample	Sellers and Non-Traders	Buyers and Non-Traders
CONSTANT	19	20	12
NREFINCO	-5	-0.6	-4
YEAR84	-9	-4	0.1
REFSIZE	82	-24	82
COSIZE	-7	-13	1
Quantity Equation <i>All variables are multiplied by leaded gasoline production</i>			
Dependent Variable	Net Quantity Supplied	Net Quantity Supplied	Net Quantity Bought
C	-0.18* (0.11)	0.10 (0.087)	0.10 (0.15)
REFSIZE	-0.044 (0.15)	-0.72** (0.15)	-0.48** (0.18)
COSIZE	-0.15** (0.054)	0.18** (0.042)	0.35** (0.085)
TECH	-0.14 (0.26)	0.095 (0.22)	0.56# (0.35)
GASRATIO	0.0024 (0.024)	-0.0056 (0.020)	0.00063 (0.034)
EAST	0.32** (0.061)	0.062 (0.052)	-0.18* (0.090)
MIDWEST	0.34** (0.041)	-0.024 (0.043)	-0.32** (0.058)
NTHCNTRL	0.58** (0.079)	0.30** (0.053)	-0.44* (0.22)
SOUTH	0.37** (0.065)	0.095* (0.047)	-0.14 (0.13)
TXINLAND	0.31** (0.094)	-0.058 (0.067)	-0.20 (0.17)
TXGULF	0.23** (0.035)	0.037 (0.047)	-0.19** (0.047)
WEST	0.65** (0.044)	0.14** (0.041)	-0.31* (0.14)
ROCKIES	0.66** (0.10)	0.13* (0.067)	-0.34 (0.28)
YEAR84	0.0076 (0.025)	0.047* (0.020)	0.054* (0.032)
λ (Inverse Mills Ratio)	-1.39 (5.29)	2.37 (2.95)	4.51 (5.51)
σ_{ϵ}	32.07	17.48	31.53
σ_{μ}	45.6	23.3	35.9
$E(\rho_i)$	-0.001	0.01	0.01
Log Likelihood	-2967.5	-1512.6	-1440.3
Total N. (N. in OLS)	659 (521)	423 (285)	374 (236)

= Coefficient significant at 90%; * = Significant at 95%; ** = significant at 99%;

Table 6.1 Trading Efficiency in Total Sample (Model B Results)

	N	% Trade	%Q	%Surp
TOTAL	659	79	85+	72+
REFSIZE < 0.05	215	73	79	74
REFSIZE ≥ 0.12	217	90	90	74
NREFINCO < 3	151	58	51	23
NREFINCO ≥ 5	357	88	95	85
COSIZE < 0.15	159	57	36	17
COSIZE ≥ 0.64	377	88	93	83
YEAR = 1983	329	80	79	64
YEAR = 1984	330	79	91	81

+ indicates estimate is biased downward.

Table 6.2 Trading Efficiency in Selling Sample (Models A and B)

	N	%Trade	%Q Model A/ B	%Surp Model A/ B
TOTAL	423	67+	90/92	88/81
REFSIZE < 0.05	127	54+	74/81	71/66
REFSIZE ≥ 0.12	123	82+	100/99	100/91
NREFINCO < 3	113	42+	52/60	54/37
NREFINCO ≥ 5	221	80+	100/100	97/96
COSIZE < 0.15	118	41+	28/31	21/5
COSIZE ≥ 0.64	227	79+	100/100	97/97
YEAR = 1983	208	67+	90/91	88/78
YEAR = 1984	215	67+	89/92	88/83

+ indicates estimate is biased downward.

Table 6.3 Trading Efficiency in Buying Sample (Model B)

	N	%Trade	%Q	%Surp
TOTAL	374	63+	95	76
REFSIZE < 0.05	147	60+	100	100
REFSIZE ≥ 0.12	116	81+	95	72
NREFINCO < 3	103	37+	41	11
NREFINCO ≥ 5	180	76+	99	80
COSIZE < 0.15	111	37+	56	14
COSIZE ≥ 0.64	198	76+	97	78
YEAR = 1983	189	64+	96	76
YEAR = 1984	185	62+	94	75

+ indicates estimate is biased downward.