Breakdown of the Equi-Marginal Principle in Permit Markets Involving Multiple Pollutants and Exogenous Caps

Carson J. Reeling
Department of Economics
Environmental and Sustainability Studies Program
Western Michigan University
carson.reeling@wmich.edu

Richard D. Horan
Department of Agricultural, Food, and Resource Economics
Michigan State University
horan@msu.edu

Cloé Garnache
Department of Agricultural, Food, and Resource Economics
Michigan State University
garnache@msu.edu

Abstract: Pollution permit trading programs typically focus on individual pollutants, yet many environmental management problems involve multiple pollutants. Hence, potential benefits may arise from developing multi-pollutant strategies that take a more comprehensive approach to environmental management. There is interest in expanding market-based approaches that traditionally involve intra-pollutant trading, or trading “like” pollutants, to allow inter-pollutant trading, or trading across imperfectly substitutable pollutants. We examine the design of a market involving both inter- and intra-pollutant trades when some firms generate multiple pollutants. Our focus is on choosing trade ratios for both types of trades. We also extend prior work on market-based approaches for multi-pollutant problems by examining the second-best design of inter- and intra-pollutant trading ratios when permit caps are inefficient. This is important since permit caps are typically set exogenously in offset programs. We demonstrate analytically that one-for-one intra-pollutant trading is not generally optimal for uniformly mixed pollutants in this setting, regardless of whether inter-pollutant trading is allowed, because the equi-marginal principle is not satisfied. This contrasts with prior work that analyzes inter-pollutant trade ratios when one-to-one intra-pollutant trading rates are assumed due to uniform mixing of like pollutants. We illustrate our analytical results using a numerical example of nitrogen trading in the Susquehanna River Basin in Pennsylvania.

JEL Codes: Q50, Q53, Q54, Q58

Keywords: equi-marginal principle; instrument design; multi-pollutant markets; pollution permit markets; second-best
Pollution permit trading offers the potential to improve the efficiency of pollution control by reallocating abatement effort towards lower marginal cost emitters. Pollution markets have been successfully applied to manage various air pollutants (Burtraw et al. 2005), and efforts to apply markets to water pollutants have been increasing (Fisher-Vanden and Olmstead 2013). While programs have historically focused on individual pollutants, many problems involve multiple, linked pollutants; for example, firms may emit multiple pollutants and/or multiple pollutants may contribute to a common environmental problem (US EPA 2015, 2011; NRC 2004).

The presence of pollutant linkages implies potential benefits from developing multi-pollutant strategies that take a more comprehensive approach to management (NRC 2004; Lutter and Burtraw 2002), including multi-media efforts to protect both air and water (EPA 1997; Aillery et al. 2005). Multi-pollutant management is considered to be a key facet of next-generation pollution control efforts, with market-based approaches offering an important mechanism for implementation (EPA 2011; NRC 2004). For instance, the U.S. Cross-State Air Pollution Rule is a multi-pollutant approach, albeit based on distinct markets for each pollutant. The European Union Emissions Trading System (EU ETS) allows trades among CO₂, nitrous oxide (N₂O), and perfluorocarbons (PFCs) by converting all pollutants into CO₂ equivalents (EC 2013). There have been calls for expanding market-based approaches that traditionally involve intra-pollutant trading, or trading “like” pollutants, to allow inter-pollutant trading, or trading across imperfectly substitutable pollutants that cannot easily be converted into equivalent units.

---

1 In the U.S., for example, atmospheric emissions of NOₓ and SO₂ have been regulated via separate pollution trading markets under the Acid Rain Program and Cross-State Air Pollution Rule. Greenhouse gas emissions are governed in some regions by carbon trading programs, including California’s Cap and Trade Program (GWSA 2006) and the Regional Greenhouse Gas Initiative in nine Northeastern States. In the case of water markets, the focus has largely been on point-nonpoint nutrient trading whereby high-cost point source abatement is exchanged for low-cost nonpoint source abatement.
(e.g., due to heterogeneous and nonlinear environmental impacts). Realizing the gains from multi-pollutant management requires careful policy design (NRC 2004; EPA 2011).

We examine the design of a market involving both inter- and intra-pollutant trades when some firms generate multiple pollutants. Our particular focus is on choosing the rates at which pollutants are traded, or trade ratios, for both types of trades. This contrasts with the limited prior work on multi-pollutant markets (Lutter and Burtraw 2002; Montero 2001), which analyzes inter-pollutant trade ratios when one-to-one intra-pollutant trading rates have simply been assumed due to uniform mixing of like pollutants.\(^2\) Although one-to-one intra-pollutant trades are first-best in single-pollutant markets with uniform mixing, this result has not been shown in a multi-pollutant setting where optimality requires choosing the policy variables jointly.

We also extend prior work on market-based approaches for multi-pollutant problems by examining the second-best design of inter- and intra-pollutant trading ratios when permit caps are inefficient. This is important since permit caps are typically set exogenously in the sense that they are not chosen based on cost-benefit analysis and therefore are unlikely to reflect the efficient level of emissions (Tietenberg 2005). Moreover, policy tools for managing pollutants are often designed using a piecemeal approach (Yaffee 1997; Lutter and Burtraw 2002). Trading rules may be established after—not jointly with—emissions regulations, particularly for the offset programs that have been used to address nonpoint source water pollution and carbon emissions (Woodward 2011). These programs allow previously-regulated point sources to

---

\(^2\) In related work, Muller (2012) models co-pollutants as being produced as perfect complements, such that firms make a single abatement choice that implicitly reduces emissions of greenhouse gases and co-pollutants (e.g., SO\(_2\), NO\(_x\), and PM\(_{2.5}\)) to produce both global and local benefits. With only a single choice being made, trades simultaneously involve tradeoffs among both like pollutants (i.e., the uniformly-mixed, global pollutant CO\(_2\)) and dissimilar pollutants (the co-pollutants, which have distinct local impacts), as does the single trade ratio. Absent co-pollutants, all trades could be considered intra-pollutant trades and would occur at a one-for-one rate. Hence, the trade ratio differs from unity because intra- and inter-pollutant trades occur simultaneously due to co-pollutants. A number of other papers deal with multi-pollutant problems (particularly greenhouse gases), but not in the context of trading (e.g., Moslener and Requate 2007; Nordhaus 2000; Repetto 1987).
purchase offsets from nonpoint sources to improve the cost-effectiveness of water quality management (Wainger and Shortle 2013; Ribaudo and Nickerson 2009; Fisher-Vanden and Olmstead 2013) or carbon management (California’s Global Warming Solutions Act of 2006).

The impact of inefficient caps on the design of both inter- and intra-pollutant trading ratios has yet to be addressed. Lutter and Burtraw (2002) examine inter-pollutant trading in this context using an ad hoc trade ratio. Others explore single markets, with no inter-pollutant trading, when multiple pollutant markets exist and are interdependent since some firms produce multiple pollutants (e.g., Woodward 2011; Stranlund and Son 2015). These studies investigate design rules apart from intra-pollutant trading ratios, as these are set at a one-to-one rate. For instance, Woodward (2011) considers whether it is more efficient to allow previously unregulated firms to sell abatement in only one or several pollution offset markets.

We demonstrate analytically that one-for-one intra-pollutant trading is not optimal for uniformly mixed pollutants in our framework, and hence the equi-marginal principle (i.e., that effective marginal abatement costs should be equated) is not satisfied, regardless of whether inter-pollutant trading is allowed. This is in contrast to textbook pollution markets, which promote efficiency by replicating the equi-marginal principle. The sub-optimality of using market-based approaches to replicate the equi-marginal principle in the current framework means trading is not even second-best.

We proceed by developing a model of multi-pollutant abatement in the next section. Section 2 derives conditions for efficient abatement, and Section 3 explores the outcome of several pollution permit trading scenarios. We highlight our analytical results using a numerical example of nitrogen trading in the Susquehanna River Basin (SRB) in Pennsylvania in Section 4. Section 5 concludes.
1. A MODEL OF MULTI-POLLUTANT ABATEMENT

Consider a pollution problem whereby two environmental media (air and water) are being harmed by two pollutants. For expositional purposes, we focus on two different forms of nitrogen pollution: nitrogen emissions (e.g., N₂O) that contribute to atmospheric pollution, and nitrogen loadings that pollute aquatic ecosystems (where loadings are defined as emissions that are delivered to a particular water body). This pollution is produced in three sectors, with one sector contributing pollution to both media: an industrial sector (indexed by \( I \)) that produces point source emissions, denoted \( e_I \); a wastewater treatment sector (indexed by \( W \)) that produces point source loadings, denoted \( r_W \); and an agricultural sector (indexed by \( A \)) that produces nonpoint pollution in the form of both emissions, \( e_A \), and loadings, \( r_A \).

We assume emissions from each source are uniformly mixed in the atmosphere so that they are perfect substitutes in creating harm within a particular environmental medium. Likewise, our definition of loadings implicitly accounts for the spatial effects on delivered nutrients, which means we can consider loadings from the various sources to be uniformly mixed and therefore perfect substitutes in creating environmental harm. Finally, we assume emissions and loadings from each source are deterministic. These assumptions allow us to focus on aggregate decisions at the sector level rather than focusing on individual firms within each sector. More importantly, these assumptions allow us to illustrate how the multi-pollutant case differs from textbook models of permit markets, which focus on a single pollutant involving deterministic and uniformly-mixed emissions.³

³ We concentrate on optimal market design under somewhat pristine conditions (apart from exogenous permit caps) and ignore factors (e.g., uncertainty) discussed in prior work that complicate the analysis but do not help illustrate the present findings. Making the more realistic assumption that agricultural emissions are stochastic makes it more difficult to parse out the effects of multiple pollutants. This is because permit markets in such settings (based on
Moving forward, it will be simpler to work with abatement rather than pollution levels. Define abatement by the industrial sector as $a_{el} = e_{0} - e_{f}$, where the subscript “0” denotes initial emissions prior to abatement. The industrial abatement cost function is $C_{I}(a_{el})$, where $C_{I}(0) = 0$ and $C_{I}', C_{I}'' > 0$. Likewise, wastewater treatment sector abatement is $a_{rw} = r_{w0} - r_{w}$, with the increasing, convex abatement cost function $C_{W}(a_{rw})$, where $C_{W}(0) = 0$. Finally, agricultural emissions and loadings abatement are $a_{eA} = e_{0} - e_{A}$ and $a_{rA} = r_{0} - r_{A}$, respectively, with the increasing, convex abatement cost function $C_{A}(a_{eA}, a_{rA})$, where $C_{A}(0,0) = 0$.\(^4\)

Finally, suppose pollution abatement prevents economic damage that varies with the aggregate level of each pollutant. Let $E_{e} = e_{0} - a_{el} + e_{0} - a_{eA}$ and $E_{r} = r_{w0} - a_{rw} + r_{0} - a_{rA}$ represent the ambient concentration of the air and water pollutants, respectively. Aggregate abatement of these pollutants is then $a_{e} = E_{e0} - E_{e}$ and $a_{r} = E_{r0} - E_{r}$, where the subscript “0” denotes unregulated pollution (i.e., with zero abatement). Abatement benefits—expressed as avoided economic damage—are then denoted $B_{s}(a_{s}) = D_{s}(E_{s0}) - D_{s}(E_{s}) = D_{s}(E_{s0}) - D_{s}(E_{s0} - a_{s})$ and $B_{s}(a_{s}) = D_{s}(E_{s0}) - D_{s}(E_{s0} - a_{s})$, where $D_{s}(\cdot)$ represents economic damages from pollutant $s \in \{e, r\}$. We assume $D_{e}', D_{r}'' > 0$, and hence $B_{e}'(a_{e}) > 0$ and $B_{r}'(a_{r}) < 0$.

Under the framework presented here, the only potential linkage between the air and water pollution problems arises through the agricultural abatement cost function. We assume trades of estimated or mean agricultural emissions) can only be second-best when damages are nonlinear, and generally involve a number of complex design elements (e.g., uncertainty trading ratios to adjust for risk) that significantly alter these markets relative to textbook markets (Shortle and Horan 2001). Our focus on the effects of multiple pollutants when emissions are deterministic offers a clearer comparison to textbook markets. The potential impact of stochastic agricultural pollution is described in the Discussion section.

\(^4\) Abatement cost functions are defined as follows. Let $\pi_{i}(z_{i})$ be profits for sector $i \in \{I, W, A\}$, $g_{s}(z_{i})$ be the sector’s emissions or loadings of pollutant $s \in \{e, r\}$, and $z_{i}$ be a vector of the sector’s production and pollution control choices. Absent abatement activity, sector $i$’s maximized choice vector is $z_{0i}$. Hence, abatement costs are defined as the reduction in profits, $\pi_{i}(z_{0i}) - \pi_{i}(z_{i})$. The abatement cost function for sector $i \in \{I, W\}$, $C_{i}(a_{s})$, is obtained by choosing $z_{i}$ to minimize $\pi_{i}(z_{0i}) - \pi_{i}(z_{i})$ subject to $a_{s} \leq g_{s}(z_{0i}) - g_{s}(z_{i})$. Agriculture’s abatement cost function $C_{A}(a_{eA}, a_{rA})$ is obtained by choosing $z_{i}$ to minimize $\pi_{i}(z_{0i}) - \pi_{i}(z_{i})$ subject to $a_{el} \leq g_{el}(z_{0i}) - g_{el}(z_{i})$ and $a_{rA} \leq g_{rA}(z_{0i}) - g_{rA}(z_{i})$.\
agricultural abatement costs are not linearly separable, i.e., \( \frac{\partial^2 C_A}{(\partial a_{eA} \partial a_{rA})} \neq 0 \), to ensure the problems are linked. Otherwise, the pollution management problem could be treated as two independent problems—one for each environmental medium.\(^5\)

2. FIRST-BEST CONTROL

We first characterize the efficient, or first-best, allocation of pollution control effort as a benchmark for comparison with market outcomes. The efficient outcome is defined as an allocation of pollution control effort that maximizes social net benefits

\[
\max_{a_{el}, a_{rW}, a_{eA}, a_{rA}} V = B_e (a_{el} + a_{el}) + B_r (a_{rW} + a_{rW}) - C_I (a_{el}) - C_W (a_{rW}) - C_A (a_{eA}, a_{rA}).
\]

Assuming an interior solution, the first-order conditions (FOCs) for problem (1) are

\[
\begin{align*}
\frac{\partial V}{\partial e_I} &= 0 \Rightarrow C_I' = B_e', \\
\frac{\partial V}{\partial r_W} &= 0 \Rightarrow C_W' = B_r', \\
\frac{\partial V}{\partial e_A} &= 0 \Rightarrow \frac{\partial C_A}{\partial a_{eA}} = B_e', \\
\frac{\partial V}{\partial r_A} &= 0 \Rightarrow \frac{\partial C_A}{\partial a_{rA}} = B_r'.
\end{align*}
\]

The FOCs (2)–(5) state the familiar result that, at the first-best abatement levels \( a_{el}^*, a_{rW}^*, a_{eA}^*, a_{rA}^* \), each sector’s marginal abatement costs equal the marginal benefits from abatement.

Additional insight can be had by manipulating (2)–(5) to yield the following modified equi-marginal condition

---

\(^5\)Air and water pollution linkages could also arise through economic damages if, for example, a fraction of all air emissions are deposited into the water resource rather than contributing to \( a_r \). We do not analyze such linkages here.
so that the effective marginal cost of abatement—measured by the marginal cost normalized by the marginal avoided damages from abatement—is equalized across all sources under the efficient outcome. The first and third equalities in (6) imply the conventional equi-marginal condition: marginal abatement costs should be equated within each environmental medium. The second equality extends the equi-marginal principle across media by normalizing costs in a way that treats abatement benefits in the different media as fungible or perfectly substitutable. The final equality says marginal costs equal marginal benefits in each case. While not a surprising result, this modification contrasts with current regulatory approaches (e.g., distinct markets for different pollutants) that treat abatement of distinct pollutants as non-substitutable even when those pollutants arise from the same source.

3. MARKET TRADING SCENARIOS

We now consider the outcome under various trading scenarios that differ along two dimensions. First, pollution trading can occur either in distinct markets for each pollutant that allow only intra-pollutant trading (reflecting current market-based approaches) or in an integrated, multipollutant market that allows both intra- and inter-pollutant trading. Second, point source pollutant caps can either be chosen optimally or they can be set exogenously relative to the market. We begin by defining the sectors’ market responses in a general model of trading, as these responses will be used when constructing each of the various trading scenarios.

3.1. Market Responses in a General Model of Pollution Trading

We adopt a very general model of trading by assuming pollution permits are defined for each
pollutant and for each sector. Agricultural permits are denoted \( \hat{e}_A \) and \( \hat{r}_A \), with initial allocations \( \hat{e}_{A0} \) and \( \hat{r}_{A0} \), and permit prices, \( p_{el} \) and \( p_{rl} \). The agricultural sector is not initially regulated, i.e., \( \hat{e}_{A0} = e_{A0} \) and \( \hat{r}_{A0} = r_{A0} \), and so farmers have initial rights to pollute. Point source permits are denoted \( \hat{e}_I \) and \( \hat{r}_W \), with associated initial allocations or permit caps \( \hat{e}_{I0} \) and \( \hat{r}_{W0} \), and permit prices \( p_{el} \) and \( p_{rl} \). Point sources initially face binding regulations, i.e., \( \hat{e}_{I0} < e_{I0} \) and \( \hat{r}_{W0} < r_{W0} \), but they may purchase permits (or offsets) from other sources to pollute more and reduce costs.

We examine two types of market structures. First is an integrated market that allows both intra- and inter-pollutant trading. Industrial emissions are defined as the numeraire pollutant, and trades are guided by trading or exchange ratios that define the number of permits that must be purchased for an industrial source to increase emissions by one unit. Three ratios are required: an intra-pollutant ratio for emissions, \( \tau_{eAcl} = \left| \frac{\hat{d}e_A}{\hat{d}e_I} \right| \), and two inter-pollutant ratios,

\[
\tau_{rAcl} = \left| \frac{\hat{d}r_A}{\hat{d}e_I} \right| \quad \text{and} \quad \tau_{rWcl} = \left| \frac{\hat{d}r_W}{\hat{d}e_I} \right| .
\]

These ratios can be used to define the trade ratios for the remaining combinations of potential permit trades: an intra-pollutant ratio for loadings,

\[
\tau_{rAcl} = \left| \frac{\hat{d}r_A}{\hat{d}e_A} \right| = \tau_{rAcI} / \tau_{eAcI},
\]

and the two remaining inter-pollutant ratios,

\[
\tau_{rAcl} = \left| \frac{\hat{d}r_A}{\hat{d}e_I} \right| = \tau_{rAcI} / \tau_{eAcI} \quad \text{and} \quad \tau_{rAcI} = \left| \frac{\hat{d}r_A}{\hat{d}r_I} \right| = \tau_{rAcI} / \tau_{rAcI} .
\]

The market clearing condition for this case is

\[
\frac{\hat{e}_{I0} + \hat{e}_{A0} / \tau_{eAcI} + \hat{r}_{W0} / \tau_{rWcl} + \hat{r}_{A0} / \tau_{rAcI}}{Q} = \frac{(e_{I0} - a_{el}) + (e_{A0} - a_{el}) / \tau_{eAcI}}{Q} + \frac{(r_{W0} - a_{rl}) / \tau_{rWcl} + (r_{A0} - a_{rl}) / \tau_{rAcI}}{Q}
\]

(7a)

where the left hand side (LHS), denoted as \( Q \), represents the effective aggregate permit cap expressed in terms of industrial emissions.
Now consider the case of distinct markets, defined here as an emissions market and a loadings market in which no inter-pollutant trades are allowed—only intra-pollutant trading occurs within each market. Point source permits serve as a numeraire in their respective markets, with one intra-pollutant trade ratio required for each market: $\tau_{el,el} = |d\hat{e}_l / d\hat{e}_l|$ and $\tau_{rW,rW} = |d\hat{r}_W / d\hat{r}_W|$. There is also a distinct market clearing condition for each market:

\[ (7b) \quad \frac{\hat{e}_{t0} + \frac{\hat{e}_{A0}}{\tau_{el,el}}}{Q_e} + \frac{(e_{t0} - a_{el}) - \frac{(e_{A0} - a_{el})}{\tau_{el,el}}}{Q_e} = \frac{\hat{r}_{w0} + \frac{\hat{r}_{A0}}{\tau_{rW,rW}}}{Q_r} + \frac{(r_{w0} - a_{rW}) - \frac{(r_{A0} - a_{rW})}{\tau_{rW,rW}}}{Q_r}. \]

where $Q_e$ and $Q_r$ represent the effective aggregate permit caps for each market.

Now consider each sector’s decisions. Each sector chooses abatement to minimize abatement costs plus the cost of purchasing permits. We show in the Appendix that these problems can be written as follows, regardless of whether we are dealing with an integrated market or distinct markets:

\[ (8) \quad \min_{a_{el}} C_i(a_{el}) + p_{el}[e_{t0} - a_{el} - \hat{e}_{t0}] \]

\[ (9) \quad \min_{a_{rW}} C_r(a_{rW}) + p_{rW}[r_{w0} - a_{rW} - \hat{r}_{w0}] \]

\[ (10) \quad \min_{a_{el},a_{rW}} C_i(a_{el},a_{rW}) + p_{el}[e_{A0} - a_{el} - \hat{e}_{A0}] + p_{rW}[r_{A0} - a_{rW} - \hat{r}_{A0}] \]

with the following FOCs for interior solutions

\[ (11) \quad C'_i = p_{el} \]

\[ (12) \quad C'_r = p_{rW} \]

\[ (13) \quad \frac{\partial C_A}{\partial a_{el}} = p_{el} \]

\[ (14) \quad \frac{\partial C_A}{\partial a_{rW}} = p_{rW} \]
Conditions (11)–(14) simply state that, at the optimum, each sector’s marginal abatement cost equals the permit price of the abated pollutant. Additionally, we show in the Appendix that additional market equilibrium conditions relate the trade ratios to permit price ratios, consistent with prior work involving markets for a single pollutant (e.g., Malik et al. 1993). Using these relations from the Appendix, along with (11)–(14), the market equilibrium conditions for an integrated market are:

\[
(15a) \quad \tau_{eA,el} = \frac{p_{eA}}{p_{eA}} \frac{C'_I}{\partial C_A / \partial a_{eA}}, \quad \tau_{rA,el} = \frac{p_{rA}}{p_{rA}} \frac{C'_I}{\partial C_A / \partial a_{rA}}, \quad \text{and} \quad \tau_{rW,el} = \frac{p_{rW}}{p_{rW}} \frac{C'_I}{C'_W}.
\]

The market equilibrium conditions for distinct markets are

\[
(15b) \quad \tau_{eA,el} = \frac{p_{eA}}{p_{eA}} \frac{C'_I}{\partial C_A / \partial a_{eA}}, \quad \tau_{rA,rW} = \frac{p_{rW}}{p_{rW}} \frac{C'_W}{\partial C_A / \partial a_{rA}}.
\]

The abatement choices that solve the relevant market equilibrium conditions (15) and the relevant market-clearing condition (7) can be expressed as responses to the policy variables, 

\[ a_{el}(\tau, \hat{\tau}), a_{el}(\tau, \hat{\tau}), a_{el}(\tau, \hat{\tau}), \text{ and } a_{rW}(\tau, \hat{\tau}), \] where \( \tau \) and \( \hat{\tau} \) represent the policy variables.

Specifically, \( \tau = [\tau_{rA,el} \tau_{rW,el} \tau_{eA,el}] \) in the integrated market scenario and \( \tau = [\tau_{eA,el} \tau_{rA,rW}] \) in the distinct markets scenario, whereas \( \hat{\tau} = [\hat{\tau}_{eA,el} \hat{\tau}_{rW,el}] \) in both scenarios. Note that each sector’s behavior depends on the policy variables associated with both air and water pollution. It is obvious that this should be the case for an integrated market, but it is also true for distinct markets due to the agricultural source participating in both markets; its non-separable abatement costs imply its decisions are not independent across markets. It is analytically intractable to identify the signs of the comparative statics results for the abatement response functions, as this requires the application of Cramer’s Rule to evaluate a \( 4 \times 4 \) matrix. However, prior work on trading a single pollutant across two sectors indicates that a larger ratio \( \tau_{y,z} \) generally makes it
more expensive for sector $z$ to trade to reduce abatement (Horan and Shortle 2015).

### 3.2. Market Design with Endogenous Permit Caps: First-Best Markets

We begin with the case in which permit caps are endogenously chosen, as this case most directly relates to prior work on market design (e.g., Montero 2001). The planner’s problem, after substituting the behavioral responses into $V$, is

$$
\max_{\tau, e} \ V = B_e(a_{el}(\tau, \hat{e}) + a_{et}(\tau, \hat{e})) + B_r(a_{rw}(\tau, \hat{e}) + a_{ra}(\tau, \hat{e})) - C_I(a_{el}(\tau, \hat{e})) - C_I(a_{rw}(\tau, \hat{e})) - C_I(a_{el}(\tau, \hat{e}), a_{ra}(\tau, \hat{e}))
$$

(16)

Problem (16) is written generally to reflect both market scenarios. The FOC for any relevant policy parameter $u$, defined as a scalar element of the policy vectors, is

$$
\frac{\partial V}{\partial u} = [B'_e - C'_I]\frac{\partial a_{el}}{\partial u} + [B'_e - \frac{\partial C_I}{\partial a_{et}}]\frac{\partial a_{et}}{\partial u} + [B'_r - C'_I]\frac{\partial a_{rw}}{\partial u} + [B'_r - \frac{\partial C_I}{\partial a_{ra}}]\frac{\partial a_{ra}}{\partial u} = 0 \ \forall u.
$$

(17)

Comparing (17) with the efficient conditions (2)–(5), it is clear that an efficient market design is one that causes each of the four bracketed terms in (17) to vanish. Ensuring such an outcome will generally require four instruments, which are available in both the integrated and distinct market scenarios.\(^6\)

For the integrated market scenario, FOC (17) can be written in terms of three trading ratios by using the sectors’ FOCs along with the market equilibrium relations (15a):

$$
[B'_e - \frac{p_{el}}{\tau_{et}}]\frac{\partial a_{el}}{\partial u} + \left[B'_e - \frac{p_{el}}{\tau_{et}}\right]\frac{\partial a_{et}}{\partial u} + \left[B'_r - \frac{p_{el}}{\tau_{et}}\right]\frac{\partial a_{rw}}{\partial u} + \left[B'_r - \frac{p_{el}}{\tau_{et}}\right]\frac{\partial a_{ra}}{\partial u} = 0 \ \forall u.
$$

(18a)

---

\(^6\) As depicted in problem (16), the integrated market scenario involves five policy choices to manage four abatement levels. In this case it is equivalent to transform the problem slightly by treating $Q$ from the market clearing condition (7a) as the policy variable rather than individual permit levels. Then the abatement response functions would take the form $a_{el}(\tau, Q)$, $a_{et}(\tau, Q)$, $a_{rw}(\tau, Q)$, and $a_{ra}(\tau, Q)$, and the choice of $Q$ along with the three trading ratios would be sufficient to manage the four abatement levels. Optimizing (16) with respect to $Q$ and $\tau$ yields the first-best levels $Q'$ and $\tau'$. Then the initial permit allocation $\hat{\tau}_{i0}$ and $\hat{p}_{w0}$ may be set at any combination to satisfy $Q' = \hat{\tau}_{f0} + \hat{\tau}_{a0}/\tau_{et, f0} + \hat{p}_{f0}/\tau_{et, f0} + \hat{p}_{w0}/\tau_{et, w0}$.
The bracketed terms in (18a), and hence in (17), will vanish to yield an efficient outcome when
\[ \tau_{rW,el} = \tau_{rA,el} = \frac{B_e}{B_r} \quad \text{and} \quad \tau_{eA,el} = \frac{B_e}{B_r} = 1 \] and \( \hat{e} \) is set such that \( p_{el} = B_e^* \) (see footnote 6). Using the market equilibrium conditions, we can use the trading ratio results to derive the remaining equilibrium prices \( p_{rW} = \frac{p_{el}}{\tau_{rW,el}} = B_r^* \), \( p_{el} = \frac{p_{el}}{\tau_{eA,el}} = B_e^* \), and \( p_{rA} = \frac{p_{el}}{\tau_{rA,el}} = B_r^* \). This solution ensures equivalent intra-pollutant trading ratios for emissions, \( \tau_{eA,el}^* = \frac{p_{el}}{p_{eA}} = 1 \), and loadings, \( \tau_{rA,rW}^* = \frac{p_{rW}}{p_{rA}} = 1 \), as well as equivalent inter-pollutant trading ratios, \( \tau^* = \tau_{rA,el}^* = \tau_{rW,el}^* = \frac{B_e^*}{B_r^*} \). The optimal intra-pollutant ratios imply one-to-one permit trading within each environmental medium, consistent with current approaches for uniformly-mixed pollutants. The optimal inter-pollutant trade ratio, based on relative economic benefits, contrasts with the standard approach in current multipollutant markets of setting the ratio according to the pollutants’ relative physical or chemical qualities. However, the efficient ratio is consistent with prior economic research that says efficient trades should occur at the marginal rate of substitution of damages (e.g., Schmalensee 1993; Lutter and Burtraw 2002).

For the distinct markets scenario, FOC (17) can be written in terms of two trading ratios by using the sectors’ FOCs along with the market equilibrium relations (15b):

\[(18b) \quad \left[ B_e' - p_{el} \frac{\partial a_{el}}{\partial u} + \left[ B_e' - \frac{p_{el}}{\tau_{eA,el}} \right] \frac{\partial a_{eA}}{\partial u} + \left[ B_r' - p_{rW} \right] \frac{\partial a_{rW}}{\partial u} + \left[ B_r' - \frac{p_{rW}}{\tau_{rA,rW}} \right] \frac{\partial a_{rA}}{\partial u} \right] = 0 \quad \forall u \]

The bracketed terms in (18b), and hence in (17), will vanish to yield an efficient outcome when
\[ \tau_{eA,el} = \frac{B_e^*}{B_r^*} = 1, \quad \tau_{rA,rW} = \frac{B_e^*}{B_r^*} = 1, \quad \hat{e}^* \] is set such that \( p_{el} = B_e^* \), and \( \hat{r}_{e0}^* \) is set such that

---

7 For example, different types of GHGs are traded based on their global warming potential, denominated in units of “carbon dioxide equivalents.” Trades governed by these types of trade ratios are unlikely to be cost-effective as they ignore the economic characteristics of pollution that vary across pollutant species (Schmalensee 1993; Muller 2012).
Using the market equilibrium conditions, we can use the trading ratio results to derive the remaining equilibrium prices $p_{eA} = \frac{P_{eA}}{\tau_{r_{eA}r_{W}}} = B^*_{r}$ and $p_{rA} = \frac{P_{rA}}{\tau_{r_{A}r_{W}}} = B^*_{r}$. This solution ensures the point-nonpoint emissions and loadings trading ratios are $\tau^*_{e_{A}e_{A}} = \tau^*_{r_{A}r_{W}} = 1$.

There is one important caveat to these results for both the integrated market and the distinct markets. Let $E^*_e = (e_{0e} - a^*_{eA}) + (e_{0A} - a^*_{eA})$ and $E^*_r = (r_{0W} - a^*_{r_{W}}) + (r_{0A} - a^*_{r_{A}})$ be the efficient levels of total emissions and loadings, respectively. Assuming agricultural sources are not initially regulated so that they have implicit initial permit caps of $e_{0A}$ and $r_{0A}$, then the efficient permit caps for the emissions and loadings by the industrial and wastewater treatment sectors, respectively, are $\hat{e}_{0A} = E^*_e - e_{0A}$ and $\hat{r}_{0W} = E^*_r - r_{0A}$. Note that $\hat{e}_{0A} \leq E^*_e$ and $\hat{r}_{0W} \leq E^*_r$ are required to obtain the efficient outcome with $\hat{e}_{0A} = e_{0A}$ and $\hat{r}_{0W} = r_{0A}$. Otherwise, initial nonpoint source emissions and/or loadings are so large that the first-best outcome cannot be attained simply by regulating point source emissions and loadings. In such instances, agriculture must be regulated, $\hat{e}_{0A} < e_{0A}$ and/or $\hat{r}_{0W} < r_{0A}$, to obtain the first-best outcome.

The results for the two market scenarios imply the following result:

**Result 1.** Suppose point source permit caps are endogenously chosen while nonpoint source caps are essentially set at unregulated levels. Either the integrated or distinct markets can be efficient in this case (provided unregulated agricultural pollution is not too great) because the tradeable permit markets are able to replicate the first-best equi-marginal principle. Moreover, one-to-one trading is optimal for intra-pollutant trades involving uniformly-mixed pollutants, consistent with the traditional equi-marginal principle.
Result 1 indicates that we could obtain an efficient outcome by integrating permit markets, but there is no need for this integration provided the point source permit caps are chosen optimally to reflect the linkages created by the agricultural sector (or, more generally, any sector that pollutes in multiple markets). Moreover, neither the linkages nor the chosen market scheme affect the standard result of one-to-one trading for intra-pollutant trades involving uniformly-mixed pollutants.

In practice, market caps are not typically chosen efficiently, but instead are set outside the market to meet environmental or human health standards. We now turn to the more realistic case of exogenously-defined pollution caps to examine trading in a second-best setting.

3.3. Market Design with Exogenous Permit Caps

Consider the optimal choice of \( \tau \) given that \( \hat{e} \) has already been exogenously specified, likely at a sub-optimal value. The objective function for this new problem, in which we simplify the notation by suppressing \( \hat{e} \), is

\[
\max_\tau V = B_e (a_{el} (\tau) + a_{el} (\tau)) + B_r (a_{ew} (\tau) + a_{el} (\tau))
- C_f (a_{el} (\tau)) - C_w (a_{ew} (\tau)) - C_A(a_{el} (\tau), a_{el} (\tau))
\]

The first order conditions are analytically equivalent to (17), but now \( u \) is only defined as a scalar element of the vector \( \tau \). This means there are no longer four policy instruments—in either type of market—to ensure that each of the bracketed terms in (17) vanishes to produce the efficient outcome; the resulting market solutions therefore cannot be first-best.

First consider the case of integrated markets. We show in the Appendix that the optimal trading ratios are defined implicitly by the following relations:
As described in the Appendix, the term \( \rho_{j}(\tau_{y,z}) \) is indicated to be a function of \( \tau_{y,z} \) to reflect the derivatives arising within this term, but actually the term depends on all of the trade ratios.

---

\[ \begin{align*}
(20a) \quad \tau_{W,el} &= \frac{B'_{e}}{C'_{W}} + \frac{B'_{e} \left[ \rho_{el}(\tau_{W,el}) - 1 \right] + \rho_{el}(\tau_{W,el})}{C'_{W}} + \frac{B'_{e} \rho_{A}(\tau_{W,el})}{C'_{W}} + \frac{B'_{e} - C'_{W} \rho_{W}(\tau_{W,el})}{C'_{W}} \\
(20b) \quad \tau_{el,el} &= \frac{B'_{e}}{\frac{\partial C_{el}}{\partial a_{el}}} + \frac{B'_{e} \left[ \rho_{el}(\tau_{el,el}) - 1 \right]}{\frac{\partial C_{el}}{\partial a_{el}}} + \frac{B'_{e} \rho_{W}(\tau_{el,el})}{\frac{\partial C_{el}}{\partial a_{el}}} + \frac{B'_{e} - \partial C_{el} / \partial a_{el} \rho_{el}(\tau_{el,el})}{\frac{\partial C_{el}}{\partial a_{el}}} \\
(20c) \quad \tau_{rA,el} &= \frac{B'_{e}}{\frac{\partial C_{el}}{\partial a_{el}}} + \frac{B'_{e} \left[ \rho_{el}(\tau_{rA,el}) - 1 \right] + \rho_{el}(\tau_{rA,el})}{\frac{\partial C_{el}}{\partial a_{el}}} + \frac{B'_{r} \rho_{W}(\tau_{rA,el})}{\frac{\partial C_{el}}{\partial a_{el}}} + \frac{B'_{r} - \partial C_{el} / \partial a_{el} \rho_{el}(\tau_{rA,el})}{\frac{\partial C_{el}}{\partial a_{el}}} 
\end{align*} \]

where \( \rho_{j}(\tau_{y,z}) = \frac{(\partial a_{j} / \partial \tau_{y,z}) / \tau_{y,z} - (\partial a_{k} / \partial \tau_{y,z}) / \tau_{k,z}}{\tau_{j,z}} \) for \( j, y, z \in \{ el, eA, rW, rA \} \) and \( \tau_{y,z} = 1 \).

The first RHS term in (20) represents the ratio of marginal benefits of abatement in sector \( z \) relative to the marginal cost of abatement in sector \( y \), which is of the same form as the first-best trade ratio \( \tau_{y,z} \) presented above (given that sector \( y \)’s marginal benefits and marginal costs of abatement are equated in the first-best outcome).

The remaining terms in (20) are adjustments to address inefficiencies from having too few instruments to perfectly control each source’s abatement. The second and third terms reflect the economic impacts of inefficient behavioral responses to \( \tau_{y,z} \) outside of sector \( y \). These responses are manifested through the \( \rho_{j}(\tau_{y,z}) \) terms, which represent sector \( j \)’s behavioral response to the trade ratio \( \tau_{y,z} \) relative to the abatement responses in all sectors other than \( y \) (with all effects being denominated in terms of sector \( z \)’s permits). The final RHS term in (20) reflects an inefficient deviation in the marginal benefits and marginal costs of abatement in sector \( y \), arising from an inability to adjust permit levels to equate these values. We expect \( \rho_{y}(\tau_{y,z}) \) to be
negative, which means $\tau_{y,z}$ is optimally decreased (increased) to incentivize more (less) abatement in sector $y$ whenever the inefficiencies increase (decrease) the marginal benefits of abatement in sector $y$ relative to the marginal abatement costs. The first-best forms of the trade ratios (i.e., ratios of marginal benefits) arise when permit levels are set to ensure the behavioral responses $\rho_j(\tau_{y,z}) = 0 \ \forall j \neq z,y$, such that $\rho_z(\tau_{y,z}) = 1$ (i.e., a change in $\tau_{y,z}$ for $y \neq z$ only guides trades between sectors $y$ and $z$, having no effect on other sectors’ abatement responses at the margin), and $\rho_y(\tau_{y,z}) = -\tau_{y,z} \ \forall j \neq z,y$. With too few instruments, however, sub-optimal abatement responses to a larger $\tau_{y,z}$ occur in all sectors, with beneficial (adverse) responses yielding a larger (smaller) $\tau_{y,z}$.

A key result from equation (20) is that a one-to-one rate of intra-pollutant trading (i.e., $\tau_{eA} = 1; \tau_{rA} = \tau_{rW} = 1$) is unlikely to be optimal, even though pollution is uniformly mixed within each environmental medium. Note that this result holds even in the special case where agricultural abatement costs are linearly separable, due to the fact that the integrated market links behaviors affecting both environmental media.

Another key result is that the optimal permit market for the integrated scenario is not even second-best in the current setting. To see this, denote the level of benefits associated with an optimal policy in this setting as $B^{**} = B^{**}_e + B^{**}_r$. As shown in the Appendix, the least-cost (or most efficient) method of achieving these benefits satisfies the equi-marginal principle

$$
\frac{C'_e}{B'_e} = \frac{\partial C_e}{\partial a_{eA}} \neq 1, \quad \frac{C'_W}{B'_W} = \frac{\partial C_W}{\partial a_{rA}} \neq 1, \quad \frac{C'_r}{B'_r} = \frac{\partial C_r}{\partial a_{rW}} \neq 1,
$$

where the inequality holds except in the special case where point source permit levels are set at the first-best values. We refer to this outcome as the second-best outcome. In contrast, the FOC for the intra-pollutant trade ratio can be written as
(22) \[ \frac{C'_e}{B'_e} = \left[ 1 - \frac{\partial C_e}{\partial a_{el}} \right] \rho_{el}(\tau_{el,el}) + \rho_{el}(\tau_{el,el}) + \frac{B'_e}{B'_e} \left[ \rho_{el}(\tau_{el,el}) + \rho_{el}(\tau_{el,el}) \right] \]

(with the other FOCs being analogous). Clearly, relations (21) and (22) generally differ: the trading outcome with inefficient permit levels cannot replicate the equi-marginal principle, and therefore does not represent the second-best approach as we have defined it. Rather, trading in this situation can only be third best.\(^9\) Nevertheless, pollution markets may remain more desirable than other policy instruments, such as taxes and standards, since markets can coordinate efforts across sectors that are typically managed separately but that are responsible for multiple, linked pollutants.

Now consider the case of distinct markets. We begin by noting this scenario has one less policy tool than the integrated market scenario (two ratios here versus three in the integrated market).\(^10\) This means the distinct market scenario cannot yield larger benefits than the integrated market case by Le Chatelier’s Principle (Samuelson 1947). Even so, the optimality conditions are still given by (17). Using a process analogous to that described above, the optimal trade ratios for this case are (20b) and

\[(20d) \quad \tau_{r,rrW} = \frac{B'_r}{\delta C_A / \delta a_{rA}} + \frac{B'_r}{\delta C_A / \delta a_{rA}} \left[ \rho_{el}(\tau_{el,el}) + \rho_{el}(\tau_{el,el}) \right] + \frac{B'_r}{\delta C_A / \delta a_{rA}} \left[ \rho_{el}(\tau_{el,el}) + \rho_{el}(\tau_{el,el}) \right] \]

The interpretation of this intra-pollutant ratio is the same as above, as is the result that intra-pollutant trading at a one-to-one rate is unlikely to be optimal. The only difference is that the behavioral results \( \rho_z(\tau_{y,z}) \neq 1, \rho_y(\tau_{y,z}) \neq -\tau_{y,z} \), and \( \rho_j(\tau_{y,j}) \neq 0 \) \( \forall j \neq z,y \) now arise simply due

\(^9\) Beavis and Walker (1983) use the terminology “third-best” to describe their solution to the pollution control problem, which satisfies the environmental constraint but is inefficient.

\(^10\) Recall this was also true when permit levels were endogenous, although there were no efficiency implications in that case since there were sufficient numbers of controls in each market scenario to attain a first-best outcome (see footnote 6). Making permit levels exogenous results in two fewer controls in each market scenario, imposing efficiency-reducing restrictions for each scenario.
to the fact that agricultural abatement costs are not linearly separable, whereas explicit market linkages also contributed to this result in the integrated markets scenario. Suppose agricultural abatement costs were linearly separable, such that the markets were truly separate with

$$\rho_z(\tau_{y,z}) = 1, \quad \rho_y(\tau_{y,z}) = -\tau_{y,z}, \quad \text{and} \quad \rho_j(\tau_{y,z}) = 0 \quad \forall j \neq z, y.$$  

Then setting $$\tau_{e_d,e_d} = \tau_{r_d,r_d} = 1$$ yields the least cost allocation associated with the chosen permit levels.

The results of this section are summarized as follows:

**Result 2.** Suppose point source permit caps and nonpoint source caps are set exogenously relative to the trading program. We find that one-to-one intra-pollutant trading is generally sub-optimal in this case, even when pollutants within a particular medium are uniformly-mixed. More generally, the equi-marginal principle does not generally apply when permit markets are designed optimally for this setting. This means integrated markets can only be third best, whereas distinct markets are likely to be even less efficient.

In practice, caps are set exogenously for individual pollutants, and then distinct markets are implemented with trading occurring on a one-to-one basis for uniformly mixed pollutants. Result 2 indicates that such a market design is not even third best.

**4. NUMERICAL MODEL: MULTI-POLLUTANT TRADING IN THE SUSQUEHANNA RIVER BASIN**

We now illustrate the theory using a model of multipollutant trading in the Pennsylvania portion of the SRB. The SRB is the Chesapeake Bay’s largest drainage basin, contributing about 60 percent of the total streamflow and nearly 46 percent of the nitrogen loads to the Bay (US EPA
2010). Most of the SRB is in Pennsylvania, which is the major source of the SRB’s nutrient inputs. Pennsylvania established a nutrient water quality trading program in 2005 under its 2004 Chesapeake Bay Tributary Strategy (Shortle 2012). Trading activity, while sparse initially, has recently increased due to stringent new caps imposed by the 2010 Chesapeake Bay Total Maximum Daily Load (TMDL). The Pennsylvania portion of the SRB also features numerous point and nonpoint sources of greenhouse gases (GHGs). Pollution in the form of greenhouse gas emissions and water quality loadings poses a major threat to environmental quality and economic activities in the Chesapeake Bay and the surrounding airshed (Birch et al. 2011).

4.1. Model Specification and Calibration

We begin by specifying and calibrating a numerical model of abatement costs and benefits for the SRB. Let $C_i(a_{el}) = \psi a_{el}^3$ be the industrial sector’s abatement cost so that marginal abatement costs are $C_i'(a_{el}) = 3\psi a_{el}^2$. Assuming a marginal abatement cost of $35 \text{ (for a one metric ton increase in CO}_2 \text{ equivalent emissions [mtCO}_2\text{e])}$ at an abatement level of $a_{el} = 19.8 \text{ million mtCO}_2\text{e (RGGI 2014; US EPA 2013)}$, we solve for $\psi = 8.93 \times 10^{-14}$.

Let $C_w(a_{rw}) = \phi a_{rw}^3$ be the wastewater treatment sector’s abatement costs so that marginal abatement costs are $C_w'(a_{rw}) = 3\phi a_{rw}^2$. Kaufman et al. (2014) estimate the marginal abatement cost for point sources in the Pennsylvania portion of the SRB to be $33,000 \text{ (for a one metric ton increase in total nitrogen [mtN])}$ at an abatement level $a_{rw} = 2,267 \text{ mtN}$. Substituting this information into the marginal abatement cost relation and solving for $\phi$ yields $\phi = 0.0064$.

Finally, let $C_A(a_{el}, a_{ra}) = (\alpha / 3)a_{el}^3 + (\beta / 3)a_{ra}^3 - \gamma a_{el}a_{ra}$ be the agricultural sources’ abatement cost function. Following Woodward (2011), we assume $\gamma > 0$, which means agricultural emissions and loadings abatement are complements, with $\gamma$ representing the degree
of complementarity. We are unaware of any empirical measurement of the complementarity between agricultural emissions and loadings abatement, so we assume $\gamma = 5 \times 10^{-6}$. This is small enough that marginal abatement costs do not become negative over reasonable ranges of abatement.\textsuperscript{11} The parameters $\alpha$ and $\beta$ are calibrated by substituting values for marginal abatement costs and abatement levels into the marginal cost relations. For loadings, we use Kaufmann et al.’s (2014) marginal abatement cost value of $66,000/\text{mtN}$ at an abatement level of 9,525 mtN. Using these values in the marginal abatement cost relation $\partial C/A / \partial a_e = \alpha a_e^2 - \gamma a_e$, we solve for $\alpha = 1.82 \times 10^{-11}$.

We use the agricultural loadings abatement value to determine the corresponding level of emissions abatement of $a_e = 734,451 \text{mtCO}_2\text{e}$.\textsuperscript{12} Assuming a marginal abatement cost of $10/\text{mtCO}_2\text{e}$ (Golub et al. 2009), we use the marginal abatement cost relation $\partial C/A / \partial a_e = \beta a_e^2 - \gamma a_e$ to solve for $\beta = 0.00073$.

We assume the damage function for emissions takes the form $D(A_e) = \varepsilon E_e$ since GHGs are a globally mixed pollutant and the emissions from the Pennsylvania portion of the SRB represent a small portion of the world’s GHG emissions. We set $\varepsilon = 14/\text{mtCO}_2\text{e}$ in accordance with Tol’s (2005) median estimate of the marginal damage value of CO\textsubscript{2} emissions.

Finally, the damage from loadings is assumed to take the form $D(r(E_r)) = \nu E_r^2$, so that marginal damages are $MD(r(E_r)) = 2\nu E_r^2$ where $\nu > 0$ is a parameter. Kaufman et al. (2014) estimate marginal damages for total N loadings in the Chesapeake Bay to be $7,414/\text{mtN}$ at an aggregate emissions level of $E_r = 45,000 \text{mtN}$. Substituting these values into the marginal

\textsuperscript{11} We performed sensitivity analysis and found that our numerical results are largely insensitive to $\gamma$.

\textsuperscript{12} Specifically, we used the initial loadings abatement value along with transport coefficients from the USGS SPARROW model (Ator et al. 2011) to calculate the average change in applied nitrogen for cropland in the SRB, assuming all abatement was due to changes in nitrogen application. We then follow the approach of Reeling and Gramig (2012) of using the DAYCENT model (NREL 2011) to estimate emissions abatement associated with this change in nitrogen applications.
damage relation, we solve for \( \nu = 0.082 \). The model parameters are presented in Table 1.

4.2. Simulation Results

We simulate a variety of trading scenarios for the SRB using Mathematica 7.0 (Wolfram Research, Inc. 2008); the results are presented in Table 2.\(^\text{13}\) The first scenario, which we refer to as the efficient outcome, is the solution to the social planner’s problem (1) or from efficiently designed integrated or distinct pollution markets (i.e., where the initial emissions and loading caps, \( \hat{e}_{f0} \) and \( \hat{r}_{W0} \), are chosen to ensure the market outcome yields \( p_{el} = p_{e4} = B_e^* \) and

\[
p_{rW} = p_{rA} = B_r^* , \text{ given that nonpoint sources have an implicit right to pollute}.
\]

The social net benefits in the efficient outcome total $141.5 million. The efficient inter-pollutant trade ratio is small (\( \tau_{rW,el} = \tau_{rA,el} = 0.0023 \)) and suggests that, at the first-best outcome, the marginal benefits (and hence the marginal costs) from abating loadings are more than 430 times larger than those from abating emissions. Accordingly, the small ratio encourages more abatement in the loadings sector at the margin. The optimal infra-marginal trade ratios are unity, as is expected in a first-best setting. Point sources are optimally responsible for the majority of emissions abatement, whereas agriculture optimally abates the majority of loadings. The difference arises due to the agricultural sector’s relatively high marginal costs of abating emissions, but relatively low marginal cost of abating loadings.

We compare the efficient outcome to the more realistic integrated and distinct market scenarios in which the initial permit caps for each pollutant have been set exogenously and sub-optimally, e.g., by different agencies regulating point sources in each sector. These sub-optimal

\(^{13}\) Notice that three trade ratios are presented for each scenario: one inter-pollutant trade ratio and two intra-pollutant trade ratios. In the analytical section above, we optimized over the inter-pollutant ratio \( \tau_{rA,el} \) rather than the intra-pollutant ratio \( \tau_{rA,W} \). This is of no consequence, since any three ratios give rise to all other ratios. We focus on the intra-pollutant ratio in our numerical results to make comparisons with the case in which one-to-one intra-pollutant trade ratios are imposed exogenously.
point source permit caps represent the only differences in initial regulations relative to the efficient outcome, as we have assumed the agricultural sector is not initially regulated in either case. Therefore, to facilitate comparison with the first-best case, let the point source permit caps be \( \hat{e}_{j0} = \zeta_s \hat{e}_{j0}^* \) and \( \hat{r}_{w0} = \zeta_s \hat{r}_{w0}^* \), where \( \zeta_s \) (for \( s \in \{ e, r \} \)) is a parameter indicating the degree of regulation relative to the efficient case. A value of \( \zeta_s = 1 \) represents an efficient cap on point sources of pollutant \( s \), whereas \( \zeta_s < 1 \) (\( > 1 \)) implies an inefficiently strict (lax) cap on point sources of pollutant \( s \). Note that an efficient cap in only one sector (e.g., \( \zeta_e = 1, \zeta_r \neq 1 \) or vice versa) will not yield an efficient allocation of either type of pollutant since agricultural choices respond to incentives in both markets. Efficiency is only obtained when \( \zeta_e = \zeta_r = 1 \). Because our interest is in highlighting how traditional results change when caps are set exogenously, rather than in exploring every qualitative combination of \( \zeta_e \) and \( \zeta_r \) relative to unity, we set \( \zeta_e = \zeta_r = \zeta \).

Table 2 illustrates results for cases where \( \zeta = 1 \pm 0.05 \). First consider the case of excessively strict caps \( (\zeta = 0.95) \) where intra-pollutant trading ratios are chosen optimally. Market integration results in greater social net benefits than distinct markets, although the benefits in both integrated and distinct markets are, respectively, nine to fourteen percent smaller than those in the efficient scenario.

Consider how the trade ratios are set to reallocate abatement in these scenarios. For the integrated market scenario, the inter-pollutant trade ratio involving point source emissions and wastewater treatment plant loadings \( (\tau_{rW,el} = 0.0089) \) is larger than the efficient ratio, reducing the incentives to reallocate abatement to wastewater treatment facilities. In contrast, the inter-pollutant trade ratio involving point source emissions and agricultural loadings \( (\tau_{rA,el} = 0.0009; \) computed based on other ratios in Table 2) is smaller than the efficient ratio to encourage an exchange of industrial emissions abatement for agricultural loadings abatement. Likewise, the
small intra-pollutant trade ratio in the loadings sector \( (\tau_{rA,W} = 0.101) \) encourages an exchange of wastewater treatment abatement for agricultural loadings abatement. The small intra-pollutant trade ratio in the emissions sector \( (\tau_{eA,eI} = 0.369) \) also encourages more abatement by agriculture by taking advantage of the complementarities in agricultural abatement costs. Indeed, total agricultural loadings abatement increases by 56 percent relative to the efficient scenario, while industrial emissions abatement is reduced slightly even though it faces more stringent caps in this scenario. Agriculture’s abatement of emissions do increase relative to the first-best case, but this is economic due to abatement cost complementarities.

Similar results arise for the case of distinct markets, although the lack of inter-pollutant trading means that taking advantage of agricultural abatement cost complementarities is the only mechanism for reallocating abatement from emissions to loadings. This explains the smaller intra-pollutant ratio for emissions relative to the integrated market scenario.

Further insight is obtained by examining the case where one-to-one intra-pollutant trade ratios are exogenously imposed. Social net benefits decline significantly relative to the scenarios where these ratios are chosen optimally: a twenty percent decline occurs under integrated markets, and a 34 percent decline occurs under distinct markets. Moreover, integrated markets perform 27 percent better than distinct markets in this setting. For the integrated market scenario, the single inter-pollutant trade ratio \( (\tau_{rW,eI} = \tau_{rA,eI} = 0.0005) \) is smaller than any of the other ratios examined thus far, as this is now the only mechanism for reallocating abatement to the loadings sector. But note that, unlike the case with differentiated ratios, it is no longer possible to target reallocations towards abatement in agricultural loadings. Instead, the small inter-pollutant trade ratio encourages more loadings abatement by both agriculture and wastewater treatment plants. At the same time, the small ratio encourages less emissions.
abatement by agriculture. The result is larger control costs. The distinct market scenario is even more limited in reallocating abatement efforts when the intra-pollutant ratios are set at unity, because now policy makers have no tools available to guide the market. Consequently, this is the least efficient outcome overall.

The welfare rankings for the case of excessively lax caps (ζ = 1.05) are the same as those for overly strict caps, except now the integrated market with one-to-one intra-pollutant trading outperforms the distinct markets with optimally chosen trade ratios. Table 2 indicates the optimal choices of trade ratios, and hence the allocation of abatement efforts, are largely opposite of the outcomes arising when ζ = 0.95, as might be expected.

Figure 1 illustrates social net benefits under each of the trading scenarios described above for a wider range of values of ζ.\textsuperscript{14} The figure shows the ranking of trade scenarios is preserved for other values of ζ, with the welfare deviation expanding the further is ζ from unity, at which point all scenarios converge.

5. DISCUSSION AND CONCLUSION

Environmental managers engaged in permit market design decisions may face many practical constraints, including the inability to efficiently set permit caps. We demonstrate that these constraints have important consequences for market design choices in the context of multi-pollutant problems. In particular, the use of inefficient permit caps causes the equi-marginal principle to break down in the permit market, which in turn affects the optimal values of both inter-pollutant and intra-pollutant trade ratios. The standard rule of one-to-one intra-pollutant trading of uniformly-mixed pollutants is no longer optimal when permit caps are inefficient and

\textsuperscript{14} An emissions cap greater than 107 percent of the efficient level results in initial permit allocations greater than initial industrial emissions.
abatement costs are non-separable. This is true even when inter-pollutant trading is not allowed. Rather, both types of trade ratios are optimally adjusted to reflect behavioral responses. This result has potentially wide-ranging implications, given the interest in offset markets, the fact that many polluters generate multiple pollutants, and that permit caps are unlikely to be set efficiently in practice. Moreover, the result is likely to hold in settings involving non-uniformly mixed pollutants: optimal trade ratios for non-uniformly mixed pollutants are unlikely to equal the relative marginal benefits from abating each pollutant, in contrast with prior work.

Failure of the equi-marginal principle to hold means that the level of damages arising under an optimal trading program could be achieved at lower cost under an alternative program: pollution markets fail to even be second-best in this case. Still, we demonstrate that environmental markets may have the potential to perform well, but only if they are designed adequately. We examined two choices that can enhance efficiency and found larger efficiency gains arise for more constrained (i.e., sub-optimal) permit levels: (i) integrating permit markets to allow inter-pollutant trading, and (ii) setting all trading ratios to optimally reflect behavioral responses. In particular, market integration enhances efficiency because the inter-pollutant trade ratio offers an additional policy tool that can help compensate for the inefficient permit levels.

Finally, our analysis ignores stochastic emissions and uncertainties about abatement effectiveness that are often characteristic of nonpoint pollution problems. Recent numerical work on nutrient water quality problems in the SRB (Horan and Shortle 2015) finds that intra-pollutant trade ratios in this single-pollutant setting may be less than one due to nonpoint source risks, especially when permit caps are exogenous. These results suggest that risk may further enhance our numerical results that optimal intra-pollutant ratios involving nonpoint sources are less than one. However, including cost uncertainty is unlikely to affect our qualitative results.
since we are comparing market designs within a pollution permit trading framework rather than comparing different policy instruments such as price and quantity instruments. We leave a formal accounting of these uncertainties for future research.

REFERENCES


National Research Council (NRC), Committee on Air Quality Management in the United States, 2004. *Air Quality Management in the United States*, National Academy of Sciences,


Appendix

Deriving Sector-Level Cost Minimization Problems

Consider the case of an integrated market. Following the approach of Horan and Shortle (2005), suppose the wastewater sector may hold permits sold by any other sector. This sector’s initial emissions permit holdings are $\hat{r}_W^0$, and denote its purchases of permits from other sources by $\hat{e}_W$, $\hat{e}_A$, and $\hat{r}_A$. The sector’s costs of abatement and permit purchases are $C_W(a_r) + p_{eA}\hat{e}_W + p_{eA}\hat{e}_A + p_{eA}\hat{r}_A$. Moreover, this sector is constrained in that its total emissions cannot be greater than its permit holdings, $r_W \leq \hat{r}_W + \hat{e}_W / \tau_{eA,W} + \hat{e}_A / \tau_{eA,W} + \hat{r}_A / \tau_{rA,W}$, where the final three RHS terms represent the emissions that the wastewater treatment sector can generate based on permits obtained from other sectors. Assuming the emissions constraint is satisfied as an equality, then $\hat{r}_A$ can be eliminated as a choice variable so that total costs are

\begin{equation}
A1 \quad C_W(a_r) + p_{eA}\tau_{eA,W} [r_W - \hat{r}_W - \hat{e}_W / \tau_{eA,W} - \hat{e}_A / \tau_{eA,W}] + p_{eA}\hat{e}_W + p_{eA}\hat{e}_A
= C_W(a_r) + p_{eA}\tau_{eA,W} [r_W - a_r - \hat{e}_W / \tau_{eA,W} - \hat{e}_A / \tau_{eA,W}] + p_{eA}\hat{e}_W + p_{eA}\hat{e}_A
\end{equation}

Here, the choice variables are $a_r$, $\hat{e}_W$, and $\hat{e}_A$. The FOCs associated with $\hat{e}_W$ and $\hat{e}_A$ are

\begin{align}
A2 \quad p_{eA} &= p_{eA}\frac{\tau_{eA,W}}{\tau_{eA,W}}
\end{align}

\begin{align}
A3 \quad p_{eA} &= p_{eA}\frac{\tau_{eA,W}}{\tau_{eA,W}}
\end{align}

where the equalities in equations (A2)–(A3) emerge in a competitive market equilibrium. Divide (A2) by (A3) to yield
(A4) \[ \frac{p_{el}}{p_{el}} = \frac{\tau_{el, rW}}{\tau_{el, rW}} = \frac{d\hat{e}_t}{d\hat{e}_t} = \tau_{el, rW}. \]

The cost minimization problems for the other sectors are analogous, with FOCs ultimately yielding the other two primary trade ratios

(A5) \[ \tau_{rA, el} = \frac{p_{el}}{p_{rA}} \]

(A6) \[ \tau_{rW, el} = \frac{p_{el}}{p_{rW}}. \]

Given these relations, the \( \hat{e} \) terms in (A1) cancel, so we can write the cost function for a particular sector restricted on only that sector’s pollution: \( C_I(a_{el}) + p_{el}[r_{i0} - a_{el} - \hat{e}_{i0}], \)

\( C_W(a_{rW}) + p_{rW}[r_{W0} - a_{rW} - \hat{r}_{W0}], \) and \( C_A(a_{eA}, a_{rA}) + p_{eA}[e_{A0} - a_{eA} - \hat{e}_{A0}] + p_{rA}[r_{A0} - a_{rA} - \hat{r}_{A0}] \).

**Deriving Second-Best Trade Ratios**

Consider the integrated market, with the first-order conditions given by (17). In particular, the condition for the choice of \( \tau_{rW, el} \) is

\[
\left[ B'_e - C'_I \right] \frac{\partial a_{el}}{\partial \tau_{rW, el}} + \left[ B'_e - \frac{\partial C_A}{\partial a_{eA}} \right] \frac{\partial a_{eA}}{\partial \tau_{rW, el}} + \left[ B'_e - C'_W \right] \frac{\partial a_{rW}}{\partial \tau_{rW, el}} + \left[ B'_r - \frac{\partial C_A}{\partial a_{rA}} \right] \frac{\partial a_{rA}}{\partial \tau_{rW, el}} = 0
\]

which can be rearranged and written as

\[
B'_e \left[ \frac{\partial a_{el}}{\partial \tau_{rW, el}} + \frac{\partial a_{eA}}{\partial \tau_{rW, el}} \right] + B'_e \left[ \frac{\partial a_{rW}}{\partial \tau_{rW, el}} + \frac{\partial a_{rA}}{\partial \tau_{rW, el}} \right] = C'_I \frac{\partial a_{el}}{\partial \tau_{rW, el}} + \frac{\partial C_A}{\partial a_{eA}} \frac{\partial a_{eA}}{\partial \tau_{rW, el}} + C'_W \frac{\partial a_{rW}}{\partial \tau_{rW, el}} + \frac{\partial C_A}{\partial a_{rA}} \frac{\partial a_{rA}}{\partial \tau_{rW, el}}
\]

Next, use the market equilibrium conditions (15) to write (A7) as
Using a process analogous to that described above, the optimal trade ratios for this case are
\begin{align}
B_e' \left[ \frac{\partial a_{el}}{\partial \tau_{rW,el}} + \frac{\partial a_{el}}{\partial \tau_{rW,el}} \right] + B_r' \left[ \frac{\partial a_{rW}}{\partial \tau_{rW,el}} + \frac{\partial a_{rW}}{\partial \tau_{rW,el}} \right] \\
= C_i' \frac{\partial a_{el}}{\partial \tau_{rW,el}} + C_i' \frac{\partial a_{el}}{\partial \tau_{rW,el}} + C_i' \frac{\partial a_{rA}}{\partial \tau_{rW,el}} + C_w' \frac{\partial a_{rW}}{\partial \tau_{rW,el}} \\
= C_i' \left[ \frac{\partial a_{el}}{\partial \tau_{rW,el}} + \frac{1}{\tau_{rA,el}} \frac{\partial a_{el}}{\partial \tau_{rW,el}} + \frac{1}{\tau_{rA,el}} \frac{\partial a_{rA}}{\partial \tau_{rW,el}} \right] + C_w' \frac{\partial a_{rW}}{\partial \tau_{rW,el}}.
\end{align}

We can solve (A8) for
\begin{align}
\tau_{rW,el} = \frac{C_i'}{C_w'} = \frac{B_e'}{C_w'} \left[ \rho_{el}(\tau_{rW,el}) + \rho_{el}(\tau_{rW,el}) \right] + \frac{B_r'}{C_w'} \rho_{rW}(\tau_{rW,el}) + \frac{B_r'}{C_w'} \rho_{rA}(\tau_{rW,el})
\end{align}
where the first equality stems from the polluters’ market equilibrium condition (15a), and where
\[
\rho_j(\tau_{y,z}) = \frac{(\partial a_j / \partial \tau_{y,z}) / \tau_{j,z}}{\sum_k (\partial a_k / \partial \tau_{y,z}) / \tau_{k,z}}, \text{ for } y,z \in \{ eL, eA, rW, rA \} \text{ and } \tau_{j,y} = 1.
\]
The term \( \rho_j(\tau_{y,z}) \) is indicated to be a function of \( \tau_{y,z} \) to reflect the derivatives arising within this term, but actually
the term depends on all trade ratios. Likewise, we can derive
\begin{align}
\tau_{el,el} &= \frac{B_e'}{\partial C_A / \partial a_{el}} \rho_{el}(\tau_{el,el}) + \left[ B_e' - \frac{\partial C_A}{\partial a_{el}} \right] \rho_{el}(\tau_{el,el}) \\
&\quad + \frac{B_r'}{\partial C_A / \partial a_{el}} \left[ \rho_{rW}(\tau_{el,el}) + \rho_{rA}(\tau_{el,el}) \right]
\end{align}
\begin{align}
\tau_{rA,el} &= \frac{B_r'}{\partial C_A / \partial a_{el}} \left[ \rho_{el}(\tau_{rA,el}) + \rho_{el}(\tau_{rA,el}) \right] + \left[ B_r' - \frac{\partial C_A}{\partial a_{el}} \right] \rho_{rA}(\tau_{rA,el}) \\
&\quad + \frac{B_r'}{\partial C_A / \partial a_{el}} \rho_{rW}(\tau_{rA,el})
\end{align}

Now consider the case of distinct markets. The optimality conditions are still given by
(17). Using a process analogous to that described above, the optimal trade ratios for this case are
(A10) and
\[
\tau_{r, e, r} = \frac{B'_r}{\partial C_A / \partial a_{r,d}} \rho_{r, w}(\tau_{r, e, r}) + \left[ \frac{B'_r - \partial C_A / \partial a_{r,d}}{\partial C_A / \partial a_{r,d}} \right] \rho_{e, t}(\tau_{r, e, r}) \\
+ \frac{B'_r}{\partial C_A / \partial a_{r,d}} \left[ \rho_{e, t}(\tau_{r, e, r}) + \rho_{e, t}(\tau_{r, e, r}) \right].
\]

(A12)

**Efficiency in Attaining** \(B^{**}\)

Consider the problem of trying to attain the aggregate benefit level \(B^{**}\), as defined in the main text, in the most efficient way possible. Since benefits are fixed, we can write the problem as a cost minimization problem

\[
\min_{a_{e,t}, a_{r,d}, a_{r,d}} C_j(a_{e,t}) + C_w(a_{r,d}) + C_A(a_{e,t}, a_{r,d})
\]

\(s.t. \quad B_e(a_{e,t} + a_{r,d}) + B_r(a_{r,d} + a_{r,d}) \geq B^{**}\)

(A13)

Defining \(\lambda\) as the Lagrangian multiplier for this problem, the FOCs are \(C'_{j} = \lambda B'_{e}\), \(C'_{w} = \lambda B'_{r}\), \(\frac{\partial C_A}{\partial a_{e,t}} = \lambda B'_{e}\), and \(\frac{\partial C_A}{\partial a_{r,d}} = \lambda B'_{r}\) (along with the constraint in (A13)). These conditions give rise to condition (21).
<table>
<thead>
<tr>
<th>Sector</th>
<th>Parameter</th>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Industrial</td>
<td>$\psi$</td>
<td>Marginal cost parameter, industrial sector</td>
<td>$8.93 \times 10^{-14}$</td>
</tr>
<tr>
<td></td>
<td>$e_{I0}$</td>
<td>Initial industrial emissions (million mtCO\textsubscript{2}e)</td>
<td>153.5</td>
</tr>
<tr>
<td>Wastewater</td>
<td>$\phi$</td>
<td>Marginal cost parameter, wastewater treatment sector</td>
<td>0.0064</td>
</tr>
<tr>
<td></td>
<td>$r_{W0}$</td>
<td>Initial wastewater treatment loadings (thousand mtN)</td>
<td>12</td>
</tr>
<tr>
<td>Agricultural</td>
<td>$\alpha$</td>
<td>Marginal cost parameter, agricultural sector</td>
<td>$1.82 \times 10^{-11}$</td>
</tr>
<tr>
<td></td>
<td>$\beta$</td>
<td>Marginal cost parameter, agricultural sector</td>
<td>$7.28 \times 10^{-4}$</td>
</tr>
<tr>
<td></td>
<td>$\gamma$</td>
<td>Complementarity parameter, agricultural sector</td>
<td>$5 \times 10^{-6}$</td>
</tr>
<tr>
<td></td>
<td>$e_{A0}$</td>
<td>Initial agricultural emissions (million mtCO\textsubscript{2}e)</td>
<td>5.54</td>
</tr>
<tr>
<td></td>
<td>$r_{A0}$</td>
<td>Initial agricultural loadings (thousand mtN)</td>
<td>28</td>
</tr>
<tr>
<td>Damages</td>
<td>$\varepsilon$</td>
<td>Marginal damage from emissions ($/\text{mtCO}_2\text{e}$)</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>$\nu$</td>
<td>Marginal damage parameter, loadings</td>
<td>0.08</td>
</tr>
</tbody>
</table>
Table 2. Simulation Results from Pollutant Trading Scenarios

<table>
<thead>
<tr>
<th>Scenarios</th>
<th>Social net benefits ($ million)</th>
<th>Inter-pollutant ratios</th>
<th>Intra-pollutant ratios</th>
<th>Emissions abatement (millions of mtCO₂e)</th>
<th>Loadings abatement (thousands of mtN)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Efficient (first-best)</td>
<td>141.5</td>
<td>0.0023</td>
<td>1</td>
<td>12.52</td>
<td>0.88</td>
</tr>
<tr>
<td>Exogenous caps: strict (ζ = 0.95)b</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No trading</td>
<td>53.8</td>
<td>—</td>
<td>—</td>
<td>19.5</td>
<td>0</td>
</tr>
<tr>
<td>Optimally chosen trade ratios</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Integrated market</td>
<td>128.5</td>
<td>0.0089</td>
<td>0.369</td>
<td>12.07</td>
<td>1.39</td>
</tr>
<tr>
<td>Distinct markets</td>
<td>122.3</td>
<td>—</td>
<td>0.212</td>
<td>11.9</td>
<td>1.81</td>
</tr>
<tr>
<td>One-to-one intra-pollutant ratios</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Integrated market</td>
<td>103.1</td>
<td>0.0005</td>
<td>1</td>
<td>12.23</td>
<td>0.86</td>
</tr>
<tr>
<td>Distinct markets</td>
<td>81.0</td>
<td>—</td>
<td>1</td>
<td>19.07</td>
<td>1.34</td>
</tr>
<tr>
<td>Exogenous caps: lax (ζ = 1.05)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No trading</td>
<td>74.6</td>
<td>—</td>
<td>—</td>
<td>5.50</td>
<td>0</td>
</tr>
<tr>
<td>Optimally chosen trade ratios</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Integrated markets</td>
<td>124.3</td>
<td>0.0005</td>
<td>1.380</td>
<td>9.30</td>
<td>0.56</td>
</tr>
<tr>
<td>Distinct markets</td>
<td>100.8</td>
<td>—</td>
<td>2.59</td>
<td>6.29</td>
<td>0.28</td>
</tr>
<tr>
<td>One-to-one intra-pollutant ratios</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Integrated market</td>
<td>103.2</td>
<td>0.001</td>
<td>1</td>
<td>6.50</td>
<td>0.46</td>
</tr>
<tr>
<td>Distinct markets</td>
<td>98.89</td>
<td>—</td>
<td>1</td>
<td>5.98</td>
<td>0.42</td>
</tr>
</tbody>
</table>

a WWT = Wastewater treatment sector.
b The efficient emissions and loadings caps are of 145.6 million mtCO₂e and 36180 mtN, respectively. The strict emissions and loadings caps (ζ = 0.95) are set at 138.4 million mtCO₂e and 34,371 mtN; while the lax emissions and loadings caps (ζ = 1.05) are set at 152.88 million mtCO₂e and 37,989 mtN, respectively. Baseline emissions are 159.04 million mtCO₂e and baseline loadings are of 40,000 mtN.(Table 1).
Figure 1. Social net benefits from abatement in the SRB under separate and integrated markets when permit caps are set exogenously at the fraction $\zeta$ of efficient levels.