Agglomeration Economies and Efficient Environmental Regulation

JunJie Wu
Oregon State University
Department of Applied Economics
213 Ballard Extension Hall
Corvallis, OR 97331
Email: junjie.wu@oregonstate.edu

and

Kathleen Segerson
The University of Connecticut
Department of Economics
365 Fairfield Way, U-1063
Storrs, CT 06269-1063
Email: kathleen.segerson@uconn.edu

First draft, February 28, 2018

Abstract: Agglomeration economies have long been a focal point of regional and urban economics; however, little research has explored their implications for environmental policy design. This paper first develops a model to analyze firms' location decisions in an economy with agglomeration economies and then uses it to examine the interaction effect between agglomeration economies and environmental regulation. Results suggest that consideration of agglomeration economies and firm relocations can change some of the classic results from traditional Pigouvian marginal analysis. Perhaps the most striking result is that a performance or technology standard commonly used for pollution control can lead to a higher aggregate ambient pollution level, an increased pollution concentration, and a larger amount of total pollution damage when firms can relocate in response to environmental regulation and agglomeration economies. Adoption of cleaner technology makes it less costly for firms to agglomerate, which leads to increased concentration of firms and higher pollution exposures. Agglomeration economies enhance the effect of a regulatory standard when production cost is sufficiently sensitive to it, making it more likely to be counterproductive. An emission tax can also lead to greater concentration of pollution and more pollution damage, even if the aggregate ambient pollution level may be lower under the tax.

JEL classification: Q51, Q58, R38

Key Words: Agglomeration economies, firm relocation, environmental regulation

1. Introduction

This paper explores two under-studied challenges for environmental policy design. One is related to relocations of firms in response to environmental policy. Firm relocations can pose both challenges and opportunities for environmental management. Because firms can move, at least in the long run, governments can design policy to induce or force heavily polluting firms to move from a densely populated area to a sparsely populated area. This is exactly what the Chinese government has been doing in recent years to improve air quality in its major cities such as Beijing. But the outcome has been disappointing so far. As heavily polluting firms were forced to move from Beijing to surrounding provinces such as Hebai and Henan, air quality in Beijing improved, but air quality in the surrounding provinces deteriorated. In addition, air quality in Beijing has not improved as much as hoped by the government. One reason is that more people and firms have moved to Beijing due to improved air quality and job opportunities (Burkitt and Spegele, 2013). The large increase in the volume of economic activity and the associated congestion and pollution have significantly impeded air quality improvement in Beijing.

This Chinese example is not unique; many studies in environmental economics have documented firm relocations in response to environmental regulation, focusing on the so-called pollution haven hypothesis.[1] In addition, many studies consider spatial dimensions of polluting activities in environmental policy analysis (Montgomery, 1972; Tietenberg, 1980; Atkinson and Tietenberg, 1982; Krupnick et al., 1983; Baumol and Oates, 1988; Klaassen et al., 1994; Farrow et al., 2005; Muller and Mendelsohn, 2009). For example, in a classic paper Montgomery (1972) shows that although an ambient-based tradeable permit system that takes into account the diffusion from sources to receptor points can lead to the least-cost solution, an emission-based permit system cannot because it ignores the diffusion process of emissions. Muller and Mendelsohn (2009) demonstrated that the sulfur dioxide allowance trading program in the US would generate much larger social welfare if trading ratios are based on spatially differentiated marginal damages. Although it has been well recognized that firms may relocate in response to environmental regulation, previous studies on spatial policy design typically take firm locations as given. Relatively little research has explored the implications of firms' reactive relocations for environmental policy. One noticeable exceptions is Markusen et al. (1993), who show that the

[1] For a review of this literature, please see Jeppesen et. al (2002), Brunnermeier and Levinson (2004), or Levinson and Taylor (2008).

implications for optimal environmental policy differ significantly from those suggested by traditional Pigouvian marginal analysis when firms can choose the number and the regional locations of their plants in response to policy.

Another under-studied challenge for environmental policy design is related to agglomeration economies. Agglomeration economies have been a focal point of urban and regional economics, but have virtually been ignored by environmental economists. Broadly speaking, agglomeration economies are gains that occur when a large number of firms or people are located in close proximity to one another. Agglomeration economies can form through several mechanisms or channels. For example, a large market allows for more efficient sharing of workers with similar skills (Krugman, 1991; Fujita and Thisse, 2002); better matching between employers and employees, or buyers and sellers (Ellison et al., 2010); and more efficient learning from each other (Porter, 2003; Lucas, 1988; Feldman and Audretch, 1999). Previous studies have also identified various benefits from agglomeration (Rosenthal, 2004). For example, the classic work by Marshall (1920) emphasizes transport costs – costs of moving goods, people, and ideas – that can be saved by agglomeration. Recent empirical evidence shows that agglomeration may increase firms' productivity by facilitating information exchanges and cross-fertilization of ideas (Moretti, 2004; Shapiro, 2006; Henderson, 2003; Rosenthal, 2003).

Agglomeration economies have been identified as a major driver of economic growth and the spatial distribution of economic activity in modern economies (Glaeser et al., 1992; Glaeser and Gottlieb, 2008; Duranton and Puga, 2014; Artz et al., 2016). For this reason, they can have important environmental implications. For example, agglomeration may lead to more people exposed to high concentrations of air pollution in urban areas, and therefore may affect the environmental impact of economic activity. Agglomeration may also affect the effectiveness of environmental regulation by making enforcement more difficult or by changing firms' compliance costs. Because firms in a larger market have easier access to information and technologies, they may face lower compliance costs. Environmental regulation may also affect the magnitude of agglomeration economies. For example, by reducing pollution, environmental regulation may reduce the cost of agglomeration, leading to increased concentration of firms. Such interaction effects may in turn affect the impact of environmental regulation (see figure 1). But surprisingly, agglomeration economies have rarely been considered in environmental policy analysis.

In fact, there is a disconnection between environmental economics and urban economics. Urban economics tends to focus on centripetal forces for concentrations and centrifugal forces for decentralization and the effect of these forces on the spatial distribution of economic activity and economic growth. Environmental economics, in contrast, tends to focus on the effect of environmental regulation on firms' location decisions. Few studies, to the best of our knowledge, have evaluated the interaction effects between agglomeration economies and environmental regulation. This disconnection is not only surprising, but also costly for research and problem-solving because key feedbacks are lost in both fields and real-world problems may be addressed in a suboptimal manner.

A primary objective of this paper is to demonstrate the importance of considering agglomeration economies and firm relocations in environmental policy analysis. The specific research questions are: How do agglomeration economies and firm relocations affect the efficiency of environmental regulation? How does environmental regulation affect the magnitude of agglomeration economies? To address these question, we first develop a model to analyze firms' location decisions in an economy with agglomeration economies and then use the model to analyze the interaction effect between agglomeration economies and environmental regulation commonly used in practice.

The analysis generates several striking results. First, we find that consideration of firms' reactive relocations and agglomeration economies can change some of the classic results from traditional Pigouvian marginal analysis. Perhaps the most striking result is that a performance or technology standard commonly used for pollution control can lead to a higher aggregate ambient pollution level, an increased pollution concentration, and a larger amount of total pollution damage when firms can relocate in response to environmental regulation and agglomeration economies. Adoption of cleaner technology makes agglomeration less costly, which leads to increased concentration of firms and higher pollution exposures. Agglomeration economies enhance the effect of an environmental standard when production cost is sufficiently sensitive to the regulation, making it more likely to be counterproductive.

Second, an emission tax can also lead to a greater concentration of pollution and more pollution damage, even if the aggregate level of ambient pollution may be lower under the tax. In particular, when an emission tax and an emission standard are designed to achieve the same reduction in emission intensity (emission per unit of output), an emission tax will lead to a larger

increase in pollution concentration and damage than an emission standard when production generates a large amount of stationary pollution that leads to increasing marginal pollution damage.

It may be unreasonable to expect an emission tax to "fix" the problem in the presence of other distortions. But we should expect that a policy would not make things worse (i.e., to reduce social welfare). Unfortunately, an emission tax may not even pass this lower bar for some polluting industries. Specifically, for highly polluting industries with large agglomeration economies, firms would be overly concentrated and produce more than the socially optimal level of pollution. In such circumstances, an emission tax may lead to a larger improvement in environmental quality in the more concentrated region, which will attract more people and firms to the region. The increased concentration of firms will offset the effect of the emission tax on total emissions and may render it ineffective in reducing the total amount of ambient pollution, especially when the transport cost is minimal. In the worst case scenario, an emission tax can lead to more pollution damage because of the increased pollution concentrations. For industries that generate relatively low levels of pollution, the equilibrium output prices will be above the optimal prices because firms' desire to under-produce, associated with market power, is greater than the tendency to overproduce, due to their failure to consider pollution externalities. In such cases, an emission tax would cause even a larger deviation from the optimal prices and thus would reduce social welfare.

2. The Model

2.1. Basic Setup

Consider an economy with two regions, East (E) and West (W), and two sectors, manufacturing (M) and agriculture (A). The manufacturing sector produces differentiated goods, and the agricultural sector produces a homogenous, numéraire good. There are two types of individual consumers in the economy: laborers and entrepreneurs. As in Ottaviano et al. (2002) and others, we assume laborers are immobile and are evenly dispersed across the two regions.[2]

[2] The assumption of an immobile input captures the more general idea that some inputs (such as land) are immobile while some others (such as low-skilled workers) have a very low spatial mobility.

Laborers provide labor as a variable input for the two sectors in exchange for wages.[3] L is the total mass of laborers in the economy and is assumed to be large enough to satisfy firms' demand for labor. For notional simplicity, L is set equal to one.[4] The agricultural sector A is perfectly competitive in both regions, and it has a constant return to scale technology that requires one laborer to produce one unit of A. This, coupled with the choice of sector A's good as the numéraire and equilibrium in the labor market, implies that in both regions the wage paid to each laborer employed in either the agricultural or the manufacturing sector is equal to one.

Entrepreneurs are owner/operators of manufacturing firms. Each entrepreneur owns a single firm that produces a unique manufactured good (i.e., a unique variety). Thus, the total number of manufactured goods equals the total mass of the firms, which is also normalized to one. The production process is identical for firms in a given region, and the total cost of producing q units of any variety is given by $C(q) = K + wq$, where w is the (constant) marginal cost of production, which reflects wage payments to laborers;[5] and K is the fixed cost, such as capital investment and overhead cost. To capture agglomeration economies that occur in a larger market, such as those from more efficient exchange of information, services, and ideas, we assume the fixed cost is lower as more firms agglomerate in the region: $K = k(1 - \phi\lambda)$, where λ is the share of entrepreneurs located in the region, and k and ϕ are positive parameters. In addition to the agglomeration economies, the mere presence of a fixed cost implies that the technology exhibits economies of scale. It is well-known that economies of scale provide an incentive for each entrepreneur to produce in a limited number of locations and that, in the absence of concerns about environmental quality, if transportation of goods is costly, the preferred sites will be those where demand is relatively larger, thereby providing a potential rationale for agglomeration (Krugman, 1991). Henceforth, k is referred to as the economies of scale parameter, and ϕ the agglomeration economies parameter.

Up to this point, our model follows the general specifications used in the agglomeration literature. Our primary interest is in extending the model to recognize the environmental impacts

[3] This is in contrast to Ottaviano et al. (2002), who assume that the immobile factor is only used as an input in the agricultural sector. This allows us to examine how an emission tax affects pollution because an emission tax corresponds to an increase in w.
[4] This assumption is made to simplify notations and does not change general results below.
[5] This is similar to the assumption made in Krugman (1991). In contrast, Ottaviana et al. (2002) assume zero marginal cost of production for manufactured goods, and a constant K.

of production. To do this, we assume that production of manufactured goods generates a fixed amount of emissions of a given pollutant per unit of output (given by z), which is identical for all varieties of the manufactured good (i.e., all varieties are equally polluting). A key aspect of emissions is that different types of pollutants have different levels of mobility. For example, greenhouse gas emissions may be more easily transmitted across regions than industrial waste. To capture this, we assume that, of the z units of the pollutant emitted per unit of manufacturing output, $\frac{(z+s)}{2}$ ends up in the home region where the firm is located, and $\frac{(z-s)}{2}$ drifts to the other region, where $0 \leq s \leq z$. Thus, s measures the stationarity of emissions, with $s = z$ indicating emissions remain local and s=0 indicating that the emissions are is perfectly mobile, i.e., the location of the emissions is unimportant.

As noted, a key objective here is to examine the impact of pollution on location decisions. We assume that entrepreneurs are mobile and can choose the region in which to live and produce,[6] taking into consideration the profit they will earn from production, the prices of consumption goods produced in both regions, and the level of ambient pollution in the two regions.

Finally, we modify standard assumptions about consumer preferences used in previous studies to incorporate the disutility of pollution. In particular, we assume that all consumers (laborers and entrepreneurs) have identical preferences defined by a quadratic utility function:[7]

$$U = \alpha \int_0^1 q(i)di - \frac{\beta}{2} \int_0^1 q^2(i)di + q_A - (Z + \delta Z^2), \tag{1}$$

where $q(i)$ is the consumption of the i-th good, q_A is the consumption of the numéraire good, and Z is the level of ambient pollution in the region where the consumer lives (which could differ from emissions in that region if pollution is mobile). α, β, and δ are positive parameters. α measures the intensity of consumer preferences for manufactured goods (relative to the agricultural good) (Ottaviano et al., 2002). $\beta > 0$ captures consumers' preference for variety, as well as the substitutability between varieties, with a higher value of β implying less

[6] We do not consider the possibility of an entrepreneur living in one location and producing in another. Thus, our entrepreneurs should be interpreted as "hands on" owners of their firms.

[7] We adopt a simplified version of the quadratic utility function of Ottaviano et al. (2002) to simplify the model and notations. The analysis can be carried out in a similar fashion with the more general specification without changing the fundamental results of this paper.

substitutability.[8] $D(Z) = Z + \delta Z^2$ is the pollution damage function, with $\delta \geq 0$ indicating that the marginal pollution damage is non-decreasing as pollution level Z increases in the region. All consumers (including entrepreneurs) view the level of ambient pollution in each region as a public bad, i.e., they do not consider the impact of their consumption decisions on overall pollution levels.

2.2. Consumer and Producer Decisions

Consumers choose the consumption level of each good to maximize their utility subject to a budget constraint:

(2) $$Max_{q(i),q_A} \quad U \qquad s.t. \int_0^1 p(i)q(i)di + q_A = Y + Y_0,$$

where $p(i)$ is the price of the *i*-th variety of the manufactured good, *Y* is the consumers income (from wages for laborers and from profits for entrepreneurs), and Y_0 is the initial endowment of the numéraire good, which is assumed to be large enough to ensure some consumption of that good. Utility maximization yields the following demand functions for the different varieties of the manufactured good and for the agricultural good:

(3) $$q(i) = \frac{\alpha}{\beta} - \frac{1}{\beta}p(i)$$

(4) $$q_A = (Y + Y_0) - \frac{\alpha}{\beta}\left(\int_0^1 p(i)di\right) + \frac{1}{\beta}\left(\int_0^1 p^2(i)di\right).$$

Substituting (3)–(4) into (1) yields the indirect utility function:

(5) $$V = \frac{\alpha^2}{2\beta} - \frac{\alpha}{\beta}\left(\int_0^1 p(i)di\right) + \frac{1}{2\beta}\left(\int_0^1 p^2(i)di\right) + Y + Y_0 - (Z + \delta Z^2),$$

which shows that prices, income, and environmental quality all affect a consumer's utility.

[8] To see this, suppose an individual consumer consumes a total mass of *Q* of the differentiated manufacturing goods. If consumption is uniform on [0, x] and zero on (x, 1), then density on [0, x] is *Q/x*. Utility for this consumption pattern is $U = \alpha Q - \frac{\beta}{2x}Q^2 + q_A - (Z + \delta Z^2)$, which is strictly increasing in the number of variety *x* and is maximized at *x* =1 when $\beta > 0$.

Each firm faces a downward-sloping demand curve and chooses its price to maximize profits. We assume that the number of firms is large enough that each firm can ignore its influence on, and reactions from, other firms. Thus, each firm solves the following problem:

(6) $$Max_{p_{rr},p_{ro}}\ (p_{rr} - w)q_{rr}M_r + (p_{ro} - w - t)q_{ro}M_o - K_r,$$

where $r, o = E, W$ and $r \neq o$. p_{rr} and q_{rr} are, respectively, the price and quantity of a manufactured good produced in region *r* and sold in region *r*; p_{ro} and q_{ro} are, respectively, the price and quantity of a good produced in region *r* and sold in the other region; *t* is the cost of transporting one unit of any variety from one region to the other; and M_E and M_W are the total masses of population in the East and West, respectively. Let λ denote the share of entrepreneurs living in the east. Then $M_E = \lambda + 0.5$ and $M_W = (1 - \lambda) + 0.5$. Note that $\lambda = 1/2$ corresponds to a totally dispersed economy, while $\lambda = 0$ or $\lambda = 1$ implies total agglomeration of manufacturing into one of the two regions.

First-order conditions of the firm's maximization problem result in the following pricing strategy:

(7) $$p_{rr} = \frac{\alpha + w}{2},\ \ p_{ro} = \frac{\alpha + w + t}{2}\ .$$

Substituting p_{ro} into $(p_{ro} - w - t)$, we see that the firm's markup over the variable and transport cost is positive if and only if $t < (\alpha - w)$. The same condition must hold for consumers in one region to buy products produced in the other region. When $t \geq (\alpha - w)$, there will be no inter-region trade between the two regions. Therefore, throughout the analysis below we assume $0 < t < (\alpha - w)$ to ensure inter-region trade, as in Ottaviano et al. (2002).

Substituting equations (7) into (6) gives the markup over variable and fixed costs, which is the profit that goes to the entrepreneur as returns to his entrepreneurial skill. Thus, for a given λ (and hence M_r and M_o), the income for an entrepreneur living and producing in region *r* equals:

(8) $$Y_r = \frac{1}{4\beta}[(\alpha - w)^2 M_r + (\alpha - w - t)^2 M_o] - K_r.$$

The difference in income for entrepreneurs located in the two regions equals:

(9) $$Y_E - Y_W = \left[\frac{t(\alpha - w - 0.5t)}{\beta} + 2k\phi\right](\lambda - 0.5),$$

which is positive if and only if $t > 0$ and $\lambda > 0.5$. In the absence of transportation costs, entrepreneurs will earn the same income regardless of where firms are located. However, when transportation is costly, an entrepreneur will earn more profit when its firm is located in the larger market. The larger the agglomeration economies (i.e., a larger ϕ) and the economies of scale (k), the larger the income difference.

2.3. Ambient Pollution

Production/consumption decisions, emissions per unit of production, the location of firms, and the mobility of emissions all combine to determine total ambient pollution in the two regions, given by:

(10a) $$Z_E = \frac{(z+s)}{2}\lambda(q_{EE}M_E + q_{EW}M_W) + \frac{(z-s)}{2}(1-\lambda)(q_{WE}M_E + q_{WW}M_W)$$

and

(10b) $$Z_W = \frac{(z+s)}{2}(1-\lambda)(q_{WE}M_E + q_{WW}M_W) + \frac{(z-s)}{2}\lambda(q_{EE}M_E + q_{EW}M_W).$$

The first term in (10a) is the amount of the emissions generated in the East that remains in the East, and the second term in (10a) is the amount of emissions generated in the West that drifts to the East. The interpretation for (10b) is analogous. Equations (10a) and (10b) show that both the total ambient pollution in each region, and hence the marginal pollution damage (when $\delta > 0$), are influenced by firms' location decisions.

We can derive the total amount of ambient pollution in the two regions by substituting the demand functions into (10a) and (10b), which yields:

(11) $$Z_E + Z_W = z \cdot (Q_E + Q_W) = \frac{z}{\beta}[t(\lambda - 0.5)^2 + (\alpha - w - 0.5t)],$$

where Q_r is the total quantity of the manufactured good produced in region *r*. Equation (11) implies that, in the presence of transportation costs, the total amount of ambient pollution is minimized when the polluting firms are perfectly dispersed between the two regions (i.e., $\lambda = 0.5$), increases as firms become more concentrated, and is maximized when all firms fully agglomerate in one region (i.e., $\lambda = 0$ or 1). Intuitively, when all firms locate in one region, the overall price is lowest due to the saving of transport costs. As a result, the total demand and

hence aggregate production is highest when all firms are located in one region. Thus, purely from the perspective of minimizing aggregate ambient pollution, dispersing production is optimal.

Similarly, the difference in pollution between East and West can be derived:

(12) $$Z_E - Z_W = \frac{s}{\beta}(2\alpha - 2w - 0.5t)(\lambda - 0.5).$$

This difference is positive when $\lambda > 0.5$ and $s > 0$, which suggests that the region with more polluting firms will have a higher level of ambient pollution when emissions are not completely mobile (i.e., $s \neq 0$).

The damages from pollution depend not only on ambient pollution levels but also on the distribution of the population exposed to those damages. Across the two regions, total damages from pollution equal (see the derivation in (A5b) in the appendix):

(13) $$TD = (Z_E + \delta Z_E^2)M_E + (Z_W + \delta Z_W^2)M_W$$

$$= (Z_E + Z_W) + \frac{\delta}{2}[(Z_E + Z_W)^2 + (Z_E - Z_W)^2]$$

$$+(\lambda - 0.5)(Z_E - Z_W)[1 + \delta(Z_E + Z_W)]$$

which is also minimized when the polluting firms are perfectly dispersed between the two regions (i.e., $\lambda = 0.5$) because both $(Z_E + Z_W)$ and $(Z_E - Z_W)$ are minimized at $\lambda = 0.5$.

3. Equilibrium Outcomes

3.1. Equilibrium Distribution of Firms

A distribution of firms between the two regions is in equilibrium if no firm has incentives to move to the other region. This can occur only if $V_E = V_W$ unless all firms have already moved to the region that offers higher utility. Formally, a distribution of firms between the two regions λ^* is an equilibrium if and only if

(14) $$\begin{array}{ll} V_E = V_W & if\ 0 < \lambda^* < 1 \\ V_E \leq V_W & if\ \lambda^* = 0 \\ V_E \geq V_W & if\ \lambda^* = 1 \end{array}.$$

To study the stability of equilibrium, we follow a well-established tradition in the migration literature by assuming that people are attracted to regions that offer higher utility, with the power of attraction increasing with a region's size (Tabuchi and Thisse, 2006; Fujita et al., 1999). Formally, we study stability through replicator dynamics defined by:

(15) $$\dot{\lambda} = \lambda(1-\lambda)(V_E - V_W),$$

where $\dot{\lambda}$ denotes the time derivative of λ. Like Tabuchi and Thisse (2006), we assume when firms move, all markets adjust instantaneously.

As derived in the appendix (see the proof of proposition 1), the utility difference for entrepreneurs located in the two regions equals:

(16) $$V_E - V_W = C[s\delta zt(\lambda - 0.5)^2 - \theta](\lambda - 0.5),$$

where

$$C = -\frac{2\alpha - 2w - 0.5t}{\beta^2} < 0,$$

$$\theta = \frac{1.5\beta t(\alpha - w - t) + 2\beta^2 k\phi}{(2\alpha - 2w - 0.5t)} - s[\beta + \delta z(\alpha - w - 0.5t)].$$

Because the two regions are symmetric, we assume, without loss of generality, that whenever an agglomeration occurs, it occurs in the East. Using (16) we can prove the following:[9]

Proposition 1. *The equilibrium distribution of the firms is as follows:*

(a) If $s\delta zt = 0$,

$$\lambda^* = \begin{cases} 0.5 & if\ \theta < 0 \\ any\ \lambda \in [0,1] & if\ \theta = 0. \\ 1 & if\ \theta > 0 \end{cases}$$

(b) If $s\delta zt \neq 0$,

[9] The proofs of Proposition 1 and subsequent propositions are given in the appendix. Note that $(1 - \lambda^*)$ would also be an equilibrium because the two regions are symmetric. This equilibrium is ruled out by our assumption that the East is larger whenever the two regions have different population.

$$\lambda^* = \begin{cases} 0.5 & if\ \theta < 0 \\ 0.5 + \sqrt{\frac{\theta}{s\delta zt}} & if\ 0 \leq \theta \leq 0.25 s\delta zt \\ 1 & if\ \theta > 0.25 s\delta zt \end{cases}.$$

Proposition 1 shows that all three types of equilibria – full dispersion, partial agglomeration, and full agglomeration – are possible, depending on parameter values. In addition, which type of equilibrium emerges depends not only agglomeration economies (k and ϕ), transport costs (t), consumer preference for variety (β), and marginal production costs (w) (which have been recognized in previous studies), but also on pollution and its characteristics (level of emissions z; mobility of emissions s; and the convexity of marginal pollution damage δ). More specifically, when $s\delta zt \neq 0$, the value of the ratio $\tilde{\theta} = \frac{\theta}{s\delta zt}$ determines the distribution of the firms between the two regions. As $\tilde{\theta}$ increases from a negative value to over 0.25, the equilibrium distribution of firms varies from a perfect dispersion to a complete agglomeration. This implies that the concentration of firms is non-decreasing in $\tilde{\theta}$, and strictly increasing when $0 \leq \tilde{\theta} \leq 0.25$. Because $\tilde{\theta}$ increases with k, ϕ, and β for $0 \leq \tilde{\theta} \leq 0.25$, we have

(17) $$\frac{\partial \lambda^*}{\partial k} > 0,\ \frac{\partial \lambda^*}{\partial \phi} > 0,\ \frac{\partial \lambda^*}{\partial \beta} > 0$$

for $0 \leq \tilde{\theta} \leq 0.25$. These results suggest the firms will be more concentrated in one region with larger economies of scale, more agglomeration economies, and higher levels of consumer preferences for variety.

To understand Proposition 1 intuitively, recall that there are three possible reasons in our model for entrepreneurs to prefer one region over the other: higher income, lower consumer prices, or lower ambient pollution. Consider first the case where there is no pollution ($z = s = 0$) so location decisions are driven solely by relative income and consumer prices. In the absence of transportation costs and agglomeration economies, i.e., if t and ϕ are all equal to zero as well, then both income and consumer prices are always equal in both regions (see (7) and (9)) and as a result entrepreneurs are indifferent regarding where they live. Thus, the equilibrium distribution of firms is indeterminate. This corresponds to case (a) for $\theta = 0$ in Proposition 1.

Suppose now that there is no pollution, but transportation costs or agglomeration economies are positive. This is an example of case (a) for $\theta > 0$ in Proposition 1. To understand the equilibrium in this case, consider an initial distribution where the firms are divided equally between the two regions. Now suppose some of the firms begin to move from the West to East. The prices of the products of these relocated firms will be lower for consumers located in the East due to the saving of transport costs and higher for consumers in the West. However, because there are more consumers enjoying lower prices in the East than those facing higher prices in the West, the total production level of those relocated firms will be higher. Higher production levels reduce the average costs for firms located in the East because of the fixed input. The larger profit margin will attract more firms to the East, which will reinforce the agglomeration economies. This self-fed dynamics will attract all firms to the East in equilibrium. This outcome is similar to the "black hole" condition in Krugman and Venables (1995) and Fujita et al. (1999): the region with the larger initial share of the manufactured sector will attract the whole sector because of agglomeration economies due to transportation costs and economies of scale. Also, note that proposition 1 reduces to Ottaviano et al. (2002)'s classic result when $s = z = 0$ (i.e., no pollution). In this case, the agglomeration configuration (i.e., $\lambda = 0$ or $\lambda = 1$) is the only stable equilibrium when transport costs are low enough to allow interregional trade (i.e., $t < (\alpha - w)$ and $\theta > 0$).

Now consider the effect of pollution in the presence of transportation costs when emissions cause more damage locally (i.e., $s > 0$). This corresponds to case (b) in Proposition 1. In this case, agglomeration benefits are offset by increased pollution costs because, as more firms move from the West to the East, the environmental quality will deteriorate in the East, but improve in the West. When emissions are relatively low, pollution effects will not completely offset the agglomeration benefits, and some agglomeration may still occur. However, when emissions are sufficiently high, environmental costs will outweigh agglomeration benefits, and firms will be perfectly dispersed between the two regions. In particular, Proposition 1 implies the following:

Corollary 1. *Assume transportation costs are positive and damages are strictly convex.*

a) *If emissions are perfectly mobile (i.e., s=0), all firms agglomerate in one region in equilibrium.*
b) *If pollution is not perfectly mobile (i.e., s>0), then there exists a threshold level of emissions, $\bar{z}$, such that equilibrium distribution of firms can be characterized as:*[10]

$$\lambda^* = \begin{cases} 1 & if\ z < \bar{z} \\ 0.5 + \sqrt{\frac{\theta}{s\delta z t}} & if\ \bar{z} \leq z \leq 2\bar{z}. \\ 0.5 & if\ z > 2\bar{z} \end{cases}$$

Note, however, that the pollution effect arises only when the damage function is strictly convex and emissions are not perfectly dispersed. If damages are linear ($\delta = 0$) or emissions are perfectly dispersed ($s = 0$),we are back to case (a) of Proposition 1.[11] In those situations, the distribution of pollution has no effect on damages. Thus, as in the case where there are no emissions, with linear damages or perfect dispersion, firms simply agglomerate to take advantage of transportation cost savings and economies of scale.

In general, under partial agglomeration, we can show the following:

(18a) $$\frac{\partial \lambda^*}{\partial z} = -\frac{1}{2z\sqrt{\tilde{\theta}}}\left[\tilde{\theta} + \frac{(\alpha - w - 0.5t)}{t}\right] < 0$$

(18b) $$\frac{\partial \lambda^*}{\partial s} = -\frac{1}{2s\sqrt{\tilde{\theta}}}\left[\tilde{\theta} + \frac{\beta + \delta z(\alpha - w - 0.5t)}{\delta z t}\right] < 0$$

(18c) $$\frac{\partial \lambda^*}{\partial \delta} = -\frac{1}{2\delta\sqrt{\tilde{\theta}}}\left[\tilde{\theta} + \frac{(\alpha - w - 0.5t)}{t}\right] < 0.$$

These results suggest that higher emissions, more stationary of emissions, and more convexity of marginal pollution damage all reinforce the role of transportation costs and lead to more concentration of firms in one region.[12]

The above results are closely related to the studies that examine the effect of congestion costs on the spatial distribution of firms (Glaeser, 1998; Fujita and Thisse, 2002; Au and Henderson, 2006; Partridge et al., 2007; Broersma and van Dijk, 2008; Fu and Hong, 2011).

---

[10] Again, we assume the East is larger whenever some agglomeration occurs. Otherwise, the symmetric distribution ($1-\lambda^*$) is also an equilibrium.
[11] It is also possible to be in case (a), but with $t > 0$ this would be true in a very limited set of parameter values.
[12] When damages are strictly convex, starting from an equal distribution of pollution, any shifting of pollution marginal damages from one region to the other will increase damages more in the receiving region than it will decrease damages in the other region.

However, unlike congestion, pollution may be transboundary. While it is true that the impact of congestion, such as reduction in the growth rate of a given region, can propagate to neighboring regions, congestion itself does not "transport" to neighboring regions as pollution does.

The above results are also closely related to the studies that examine the effect of pollution on trade patterns. Copeland and Taylor (1999) develop a two-sector model to examine the effects of pollution-created cross-sectoral production externalities on trade patterns, and find that pollution provides a motive for trade because trade can spatially separate incompatible industries. Unteroberdoerster (2001) introduces transboundary pollution into the Copeland and Taylor model and finds that allowing pollution to be transboundary can change some of Copeland and Taylor's results. Hosoe and Naito (2006) and Lange and Quaas (2007) use a new economic geographic model to explore the effect of pollution on trade patterns and firm concentrations, and find that the extent of agglomeration is not independent of pollution.

Consideration of pollution effects also has implications for the role of transportation costs. The standard results in the literature is that reduction in transport costs leads to greater agglomeration. For example, Krugman (1992) shows that under monopolistic competition industrial agglomeration arises when transport costs are sufficiently low. Okubo et al. (2010), Ottaviano et al. (2002), and Saito et al. (2011) suggest that agglomeration tends to follow transport cost reductions.[13] However, consideration of pollution effects can reverse this classic result. Differentiating λ^* with respect to t over the range $0 < \tilde{\theta} < 0.25$ and assuming $s\delta zt \neq 0$ gives:

$$\frac{\partial \lambda^*}{\partial t} = \frac{[\beta + z\delta(\alpha - w)]}{2z\delta t^2\sqrt{\tilde{\theta}}}\left[1 - \left(\frac{1.5t}{2\alpha - 2w - 0.5t}\right)^2 \frac{(\alpha - w)}{s[\beta + z\delta(\alpha - w)]}\right], \tag{19}$$

which is positive when s is sufficiently large. This suggests that a decrease in transport costs can cause firms to disperse rather than agglomerate, if firms generate a large amount of stationary pollution. When transport costs are low enough, there are both markup and cost-of-living advantages to being located in a larger market. Therefore, without pollution, all firms

[13] Similarly, Obstfeld and Rogoff (2000) argue that all the major puzzles of international macroeconomics hang on trade costs, including transport costs. For example, trade costs play a key role in determining market structure and geographic concentration of firms. Details of transport costs also matter to economic geography (Anderson and van Wincoop, 2004). For example, the home market effect hypothesis (big countries produce more of goods with scale economies) hangs on differentiated goods with scale economies having greater trade costs than homogeneous goods (Davis,1998).

agglomerate in one region with sufficiently low transport costs. However, with pollution, the region with more firms will eventually become a less desirable place to live when transport costs are sufficiently low because both the markup and cost-of-living advantages disappear as $t \to 0$. Thus, as $t \to 0$, firms will divide equally between the two regions if they generate more pollution damage locally.

3.2. Aggregate Ambient Pollution and Pollution Damage

As described above, the location decisions of firms reflect concerns about pollution, and, in turn determine the aggregate level of ambient pollution and pollution damages. Substituting the equilibrium distribution of firms λ^* into (11) and (12), we obtain the aggregate level of ambient pollution and the difference in pollution between the two regions in equilibrium:

(20) $$Z_E^* + Z_W^* = \frac{z}{\beta}[t(\lambda^* - 0.5)^2 + (\alpha - w - 0.5t)],$$

(21) $$Z_E^* - Z_W^* = \frac{s}{\beta}(2\alpha - 2w - 0.5t)(\lambda^* - 0.5),$$

where

$$\lambda^* = \begin{cases} 0.5 & if\ \tilde{\theta} < 0 \\ 0.5 + \sqrt{\tilde{\theta}} & if\ 0 \le \tilde{\theta} < 0.25 \\ 1 & if\ \tilde{\theta} \ge 0.25 \end{cases}.$$

These, in turn, determine total damages from pollution, given by:

(22) $$TD^* = (Z_E^* + Z_W^*) + \frac{\delta}{2}[(Z_E^* + Z_W^*)^2 + (Z_E^* - Z_W^*)^2]$$
$$+(\lambda^* - 0.5)(Z_E^* - Z_W^*)[1 + \delta(Z_E^* + Z_W^*)].$$

From (18a), we can see that $(Z_E^* + Z_W^*)$, $(Z_E^* - Z_W^*)$, and TD^* all increase with k and ϕ, which suggest that an increase in agglomeration economies and economies of scale will lead to a higher level of aggregate ambient pollution, increased concentration of pollution, and larger pollution damage.

As can be seen from the above expressions, when firms are not completely dispersed or agglomerated (i.e., when $0 \le \tilde{\theta} < 0.25$), changes in pollution features can affect the total

amount of pollution, as well as its concentration and corresponding total damages, both directly (through changes in *z, s,* and δ) and indirectly (through changes in λ^*). This yields several results that run counter to conventional wisdom based on non-spatial models.

Proposition 2: *Under partial agglomeration, introduction of a greener production technology that reduces emissions per unit of output at a higher capital cost will increase aggregate ambient pollution.*

This follows directly from differentiating the aggregate level of pollution, $Z_E^* + Z_W^*$, in (20) with respect to *z* over the range where $0 \leq \tilde{\theta} < 0.25$ and noting that a reduction in *z* is accompanied with an increase in *k*, which yields (see the proof in the appendix):

$$(23) \quad \frac{d(Z_E^*+Z_W^*)}{dz} = \frac{\partial(Z_E^*+Z_W^*)}{\partial z}\bigg|_{\lambda^*} + \frac{\partial(Z_E^*+Z_W^*)}{\partial \lambda^*}\bigg|_{z}\left(\frac{\partial \lambda^*}{\partial z}\bigg|_{k} + \frac{\partial \lambda^*}{\partial k}\frac{\partial k}{\partial z}\right)$$

$$= \frac{1}{\beta}\left[t(\lambda^* - 0.5)^2 + (\alpha - w - 0.5t)\right] + \frac{2ztz}{\beta}(\lambda^* - 0.5)\left(\frac{\partial \lambda^*}{\partial z}\bigg|_{k} + \frac{\partial \lambda^*}{\partial k}\frac{\partial k}{\partial z}\right)$$

$$= \frac{2ztz}{\beta}(\lambda^* - 0.5)\frac{\partial \lambda^*}{\partial k}\frac{\partial k}{\partial z} < 0.$$

This result is surprising and initially seems counter-intuitive, since an improvement in the production technology that lowers the pollution level per unit of output is expected to reduce the total level of pollution. This would be true if the distribution of production activities were independent of emissions, i.e., holding λ constant, a reduction in z would reduce total pollution. This is reflected in the positive first term in (23). However, if production causes less pollution as a result of the new technology, it becomes less costly for firms to agglomerate. As a result, the agglomeration economies will increase, and each firm will produce more in the more concentrated region. This is reflected in the second term in (23), which is negative. On net, the second term completely offsets the first term. However, because the new technology requires a larger capital investment, there will be a larger economy of scale in production, which will lead to increased concentration of firms and a higher level of total ambient pollution.

The increased concentration of firms also leads to increased concentration of ambient pollution, as shown in the following proposition.

Proposition 3: *Under partial agglomeration, the introduction of a greener production technology that reduces emissions per unit of output at a higher capital cost will lead to (i) a greater concentration of ambient pollution in one region than the other and (ii) higher total pollution damages.*

To see this, note first that, under partial agglomeration,

(24) $$\frac{d(Z_E^* - Z_W^*)}{dz} = \frac{s}{\beta}(\alpha - w - 0.5t)\left(\frac{\partial \lambda^*}{\partial z} + \frac{\partial \lambda^*}{\partial k}\frac{\partial k}{\partial z}\right) < 0.$$

This follows directly from (21) and (18a). In this case, z has only an indirect effect through λ^*. As noted above, agglomeration benefits increase when the new technology is introduced, and as a result pollution will become more concentrated in one region. This will, in turn, lead to an increase in total damages. From (22), (23), and (24), we then have:

(25) $$\frac{dTD^*}{dz} < 0.$$

Thus, rather than improving environmental quality and reducing the damages from pollution, introduction of a green technology actually increases total pollution damages. Since aggregate pollution is unchanged, again the effect on total damages is entirely through the change in the equilibrium distribution of firms that occurs in response to the new technology, i.e., the fact that the greener technology induces a greater concentration of population and pollution.

The equilibrium aggregate pollution and damages are also affected by the extent to which emissions are dispersed.

Proposition 4: *Under partial agglomeration, an increase in the stationarity of emissions (i.e., an increase in s) will (i) reduce the total aggregate pollution, (ii) reduce the concentration of pollution, and (iii) reduce total pollution damages.*

Part (i) of the proposition follows directly from (20) and (18b), which imply that

(26) $$\frac{d(Z_E^*+Z_W^*)}{ds} = \frac{2zt\sqrt{\bar{\theta}}}{\beta}\frac{\partial\lambda^*}{\partial s} < 0.$$

The effect of *s* on aggregate pollution again comes entirely through its indirect effect on location decisions. Intuitively, when emissions are more stationary, the effect of pollution on location decisions described above is stronger, implying that the benefits of agglomeration are lower. As a result, firms will be more dispersed and aggregate output (and hence aggregate pollution) will be reduced. In contrast, as can be seen from (21), *s* has both a direct and an indirect effect on pollution concentration. The direct effect is positive, i.e., greater stationarity leads to greater concentration. However, the indirect effect through λ^* is negative, since greater stationarity reduces agglomeration benefits. This indirect effect will more than offset the direct effect, i.e., the total effect is negative:

(27) $$\frac{d(Z_E^*-Z_W^*)}{ds} = -\frac{(2\alpha-2w-0.5t)[\beta+\delta z(\alpha-w-0.75t)]}{2\beta\delta zt\sqrt{\theta}} < 0.$$

Thus, counter-intuitively, as stated in part (ii) of proposition 4, after location decisions adjust, pollution will actually become less concentrated when pollution is more stationary. Furthermore, because an increase in *s* reduces both aggregate pollution and pollution concentration, i.e., both $(Z_E^* + Z_W^*)$ and $(Z_E^* - Z_W^*)$, from (22) it can easily be shown that total pollution damages are lower as well, i.e., $\frac{dTD^*}{ds} < 0$.

Likewise, we can show:

(28a) $$\frac{d(Z_E^*+Z_W^*)}{d\delta} = \frac{2zt\sqrt{\bar{\theta}}}{\beta}\frac{\partial\lambda^*}{\partial\delta} < 0,$$

(28b) $$\frac{d(Z_E^*-Z_W^*)}{d\delta} = \frac{s}{\beta}(2\alpha - 2w - 0.5t)\frac{\partial\lambda^*}{\partial\delta} < 0,$$

(28c) $$\frac{dTD^*}{d\delta} < 0.$$

Together, these results imply that under partial agglomeration, an increase in the convexity of marginal pollution damage (i.e., an increase in δ) will (i) reduce the total aggregate pollution, (ii) reduce the concentration of pollution, and (iii) reduce total pollution damages. Again, consideration of firm relocation yields results that run counter to conventional wisdom because one would expect that an increase in the marginal pollution damage leads to more pollution damage based on non-spatial models.

4. Effectiveness of Environmental Policy

We turn next to the impact that environmental policy might have on the equilibrium outcome, given endogenous location decisions. We continue to focus on the case of partial agglomeration, where the spatial equilibrium is given by case (b) for $0 \leq \theta \leq 0.25s\delta zt$ in Proposition 1.

4.1. Effectiveness of Technology or Performance Regulations

Emission standards can take different forms. A technology or design standard specifies the type of equipment or technologies that a firm must use, while a performance standard specifies the maximum quantity or minimum quality of a firm's emission. Technology or performance standards are perhaps the most commonly used instruments for pollution control. For example, both the Clear Air Act and the Clear Water Act in the US heavily rely on technology and performance standards. Under the Clean Air Act, for example, new plants locating in counties out of attainment of the national standards for a criteria pollutant are subject to the "Lowest Achievable Emission Rate" (LAER) and are required to install the "Cleanest Available Technology" (CAT), while in attainment counties, only class A polluters are required to install the "Best Available Control Technology" (BACT), and small polluters are exempted from the regulation in attainment areas.

Imagine a policy scenario in which a regulator imposes a technology standard (i.e., requires all manufacturing firms to adopt a cleaner production technology) or, equivalently in terms of our model, puts a limit z^s on allowable emissions per unit of output. The effect of such a technology or performance standard can be captured by our model through a reduction in *z* and an increase in *k*: a reduction in z^s (corresponding to higher standard) reduces *z* and increase *k*. We have already seen in section 3.2 that under partial agglomeration, a reduction in *z* and an increase in *k* can be counter-productive, i.e., it increases aggregate pollution, pollution concentration, and total damages from pollution:

(29a) $$\frac{d\lambda^*}{dz^s} = \frac{\partial\lambda^*}{\partial z}\frac{\partial z}{\partial z^s} + \frac{\partial\lambda^*}{\partial k}\frac{\partial k}{\partial z^s} < 0.$$

(29b) $$\frac{d(Z_E^*+Z_W^*)}{dz^s} = \frac{2\beta\phi}{s\delta}\frac{dk}{dz^s} < 0.$$

(29c) $$\frac{d(Z_E^*-Z_W^*)}{dz^s} = \frac{s(2\alpha-2w-0.5t)}{\beta}\left(\frac{\partial\lambda^*}{\partial z}\frac{\partial z}{\partial z^s} + \frac{\partial\lambda^*}{\partial k}\frac{\partial k}{\partial z^s}\right) < 0.$$

(29d) $$\frac{dTD^*}{dz^s} < 0.$$

These results imply the following:

Proposition 5: *Under partial agglomeration, a performance or technology standard*

a) *increases the concentration of firms*
b) *increases the aggregate ambient pollution level*
c) *increases the pollution concentration*
d) *increases the total pollution damage.*

It may be tempting to impose differential standards across regions to encourage firms to move to less concentrated regions, but such an approach may work only if the differential standard impose much larger cost for firms in the more agglomerated region. Note that the fundamental reason that a uniform standard leads to greater pollution concentration and larger pollution damage is that the standard reduces pollution and thus the cost of agglomeration. The differential standards that require firms to adopt higher technical or performance standard in more concentrated regions may actually exacerbate the problem if they do not impose higher cost on firms in the larger region because it further reduces the cost of agglomeration. Of course, if the differential standards dramatically increase firms' production cost in the more concentrated region, it will discourage agglomeration and may even cause firms to disperse if the additional cost burden is large enough.

4.2. Interaction Effects of Agglomeration Economies and Environmental Standards

As expected, the effects of agglomeration economies and environmental regulation are not independent. Specifically, we can prove the following:

Proposition 6: *Under partial agglomeration,*

(30a) $$\frac{d^2\lambda^*}{dkdz^s} = -\Phi\phi\left(\varepsilon_{z^s}^{z} + \varepsilon_{z^s}^{\lambda^*}\right) > 0$$

(30b) $$\frac{d^2\lambda^*}{d\phi dz^s} = \Phi k\left(\varepsilon_{z^s}^{k} - \varepsilon_{z^s}^{z} - \varepsilon_{z^s}^{\lambda^*}\right) < 0 \; if \; \phi > -\frac{\Psi}{2k\beta^2} \; and \; \varepsilon_z^k < \bar{\varepsilon}_z^k$$

where

$$\Phi = \frac{\beta^2}{(2\alpha - 2w - 0.5t)s\delta zt(\lambda^* - 0.5)} > 0$$

$$\Psi = 1.5\beta t(\alpha - w - t) - s[\beta + \delta z(\alpha - w - 0.5t)](2\alpha - 2w - 0.5t).$$

Equation (30a) indicates that imposing a performance or technology standard (i.e., lowering z^s) reduces the impact of economies of scale on the concentration of firms. This is counter intuitive because one would expect that the effect of economies of scale is stronger under the regulation because it reduces the cost of agglomeration. This would be true if the regulation would not affect the distribution of firms because, holding λ constant, a reduction in z^s would increase the impact of the economy of scale. This is reflected by the negative first term in (30a) (i.e., $-\Phi\phi\varepsilon_{z^s}^{z} < 0$). However, the standard causes more firms to concentrate in one region, which increases the marginal pollution damage and reduce the benefit of agglomeration, as reflected by the positive second term in (30a). This second effect outweighs the first, and overall imposing a performance or technology standard will reduce the effect of the economy of scale on firm concentration. This interaction effect with the economy of scale moderates the direct effect of the standard on firm concentration, making it less likely to be counterproductive.

However, imposing a performance or technology standard (i.e., lowering z^s) could increase or decrease the impact of agglomeration economies on firm concentration, as reflected by equation (30b). The standard may mandate or require capital investment (i.e., increase *k*), but reduces pollution intensity (i.e., *z*). An increase in *k* enhances the benefit of agglomeration, while a decrease in *z* reduces the cost of agglomeration. These two effects are reflected by the first two terms in (30b) (i.e., $\Phi k\left(\varepsilon_{z^s}^{k} - \varepsilon_{z^s}^{z}\right) < 0$). However, the standard causes more firms to concentrate in one region, which increases the marginal pollution damage and reduces the benefit of agglomeration, as reflected by the positive third term in (30b). Overall, the environmental

regulation will increase the effect of agglomeration economies if the agglomeration economies are sufficiently large and the fixed cost is sufficiently sensitive to the standard. In this case, the interaction effect will reinforce the effect of the regulation on concentration, making it more likely to be counterproductive (see figure 2).

4.3. Effectiveness of Emission Taxes

Economists have long advocated emission taxes for pollution control because of their presumed efficiency property (Hanley et al., 1997, p. 61).[14] Here we show that, as with emissions standards, emissions taxes can actually be counter-productive when firms relocate in response to environmental regulation in an economy with agglomeration economies.[15] In particular, we show that imposing an emissions tax can actually lead to greater concentration of pollution and more aggregate pollution damages, even if the total pollution level may be lower.

To determine the overall impact of an emissions tax, note first that the tax will affect the equilibrium outcomes in two ways: (1) through the effect on z, and (2) through the effect on production cost. Regarding the first effect, if we now view z as a choice variable, the tax will induce firms to choose a greener technology or production process and hence a lower level of z. The exact level of z that would be chosen will depend on the benefits of reducing z (in the form of lower tax payments) and the costs of adoption. We can summarize this choice by writing z as a function of the tax rate, τ, i.e., letting $z = z^*(\tau)$, where $\frac{dz^*}{d\tau} < 0$. Thus, a marginal change in the tax rate τ will induce a marginal reduction in z that is analytically identical to the impact of a marginal change in z under the regulatory standard (with the associated fixed cost differential) described above. However, unlike the standard, the tax policy also requires firms to make tax payments for their emissions. Thus, firms incur not only the cost of reducing z (as under the

[14] In addition, emission taxes are believed to be more consistent with the "Polluter-Pays-Principle," which maintains that society's environmental resources, including clean air and water, belong to the public at large, and those who "use" them must then compensate the "owners" (i.e., the public) by paying a tax.

[15] Environmental economists have examined the effect of environmental taxation in a second-best setting, focusing on the double-dividend hypothesis (Bovenbergand de Mooij, 1994; Goulder, 1992, 1995; Parry, 1995; Fullerton, 1997; Koskela and Schöb, 1999; Fullerton and Metcalf, 2001). Here we examine how firms' reactive relocations affect the efficiency of environmental regulation in the presence of agglomeration economies.

standard) but also an increase in their cost of production due to the tax payments. In other words, imposing an emission tax τ increases each firm's marginal production cost from w to $w + \tau z(\tau)$. It may also induce firms to invest in cleaner technologies (i.e., increase k). To simplify notations, we assume $\phi = 0$ here, which means that changes in k have no effect on firms' location choices and therefore no effect on environmental outcomes. Thus, the total effect of a marginal change in the tax on any given outcome variable $F*$ takes the form:

$$\frac{dF^*}{d\tau} = \left\{\frac{\partial F^*}{\partial z}\bigg|_w \frac{\partial z^*}{\partial \tau}\right\} + \left\{\frac{\partial F^*}{\partial w}\bigg|_z \frac{\partial w}{\partial \tau}\right\} = \left\{\frac{\partial F^*}{\partial z}\bigg|_w \frac{\partial z^*}{\partial \tau}\right\} + \left\{\frac{\partial F^*}{\partial w}\bigg|_{\lambda,z} + \frac{\partial F^*}{\partial \lambda}\bigg|_z \frac{\partial \lambda^*}{\partial w}\right\}\frac{\partial w}{\partial \tau}, \tag{31}$$

where $\frac{\partial w}{\partial \tau} = z^*(\tau) + \tau\frac{\partial z^*}{\partial \tau}$. The first bracketed term on the right-hand side of (31) represents the partial effect of the induced change in z, holding w constant but allowing for endogenous effects on firm location. The signs of these effects for the various outcome variables $(\lambda^*, (Z_E^* + Z_W^*), (Z_E^* - Z_W^*), TD^*)$ were derived in the previous section. Those results, coupled with the fact that $\frac{dz^*}{d\tau} < 0$, imply that the first bracketed term on the right-hand side of equation (31) is positive for all outcome variables, i.e., ignoring the impact of the tax payments, with partial agglomeration introducing or increasing an emissions tax that leads to a reduction in *z,* would lead to an increase in agglomeration, total pollution, pollution concentration, and total damages. The question is then whether the second bracketed term, which represents the effect of the tax payments on marginal production costs, works to offset or reinforce these perverse environmental impacts. Note that, as indicated by the last bracketed term in (30), the tax payment effect is also a combination of the direct effect of a change in marginal production costs for a given level of agglomeration and an indirect effect through changes in the equilibrium level of agglomeration.

To determine how the tax payment effect impacts the overall effect of the tax, we need to examine how the outcome variables of interest vary with changes in w. Consider first the impact on agglomeration. Differentiating λ^* with respect to w gives:

$$\frac{\partial \lambda^*}{\partial w} = \frac{\sqrt{\theta}}{2\theta}\left[sz\delta - \frac{(1.5t)^2\beta}{(2\alpha - 2w - 0.5t)^2}\right],$$

which can be positive or negative over the range of partial agglomeration, i.e., when $0 < \theta < 0.25s\delta zt$. Thus, an increase in marginal production costs can lead to more or less agglomeration, depending on parameter values. To understand the intuition behind this result, recall that over the

range of partial agglomeration, the concentration of firms is strictly increasing in $\tilde{\theta}$. However, the effect of w on $\tilde{\theta}$ is ambiguous when $s\delta z \neq 0$. This reflects the divergent influences on utility differences across the two regions and hence relocation incentives, conditional on the level of agglomeration (see (16)). On the one hand, an increase in marginal production costs increases the average price by a larger percent in the more concentrated region, which in turn decreases the benefit of living in that region. On the other hand, the correspondingly greater decrease in output in the more concentrated region reduces the difference in environmental quality across the two regions, which in turn reduces the benefit of living in the less populated region. When this latter pollution effects is relatively strong, i.e., when *s, z,* or δ is relatively large, the increase in w on net creates an incentive for individuals to move to the more concentrated region, thereby increasing the concentration of firms. However, when the latter effect is relatively small, the price effect will dominate and an increase in w will decrease agglomeration. Thus, the total effect of an increase in marginal production costs as a result of the tax payments is ambiguous. It is easily verified that the overall effect of the tax, which combines the effect on z as well as the effect on the tax payments, is also ambiguous (see Table B1 in Appendix B for simulations verifying ambiguous results). However, if the pollution effect is sufficiently strong, i.e., if the term $sz\delta$ is sufficiently large, then the total effect of the increase in marginal production costs will be positive, which, when coupled with the negative effect through the reduction in *z*, implies that the tax would lead to greater agglomeration overall.

Consider next the effect of the tax on aggregate pollution. The tax payment effect, i.e., the second bracketed term in (30), can be seen by differentiating (20) with respect to *w*, which yields:

(32) $$\left.\frac{\partial(Z_E^*+Z_W^*)}{\partial w}\right|_{\lambda} + \frac{\partial(Z_E^*+Z_W^*)}{\partial \lambda}\frac{\partial \lambda^*}{\partial w} =$$

$$\left[-\frac{z}{\beta}\right] + \left[\frac{z}{\beta} - \frac{(1.5t)^2}{s\delta(2\alpha-2w-0.5t)^2}\right] = -\frac{(1.5t)^2}{s\delta(2\alpha-2w-0.5t)^2} < 0.$$

The result in (33) implies that, regardless of whether the emission tax leads to greater or less concentration of firms, the tax payments always reduce the total amount of pollution, i.e., the direct effect of the increase in marginal production costs outweighs any indirect effects through spatial reallocation, implying that the total effect of the tax payments on aggregate pollution is negative. Intuitively, this reflects the fact that an increase in production costs due to the tax

increases product prices and thereby decreases aggregate output and hence aggregate pollution. It does not, however, imply that overall the tax leads to a decrease in pollution. The overall effect of the tax, which incorporates both the effect of the induced change in z and the expected effect of the tax payments, is ambiguous. In addition, since the tax payment effect serves to offset at least partially the perverse effect of the reduction in z, aggregate pollution will be higher under an emission standard than under a comparable tax, i.e., a tax that induces the same reduction in z as the standard would.

Note, however, that the result in (32) also shows that the effectiveness of the tax payments in reducing pollution depends on the magnitude of transportation costs. In particular, as $t \to 0$, the impact of the associated increase in marginal production costs also approaches zero. Glaeser and Kohlhase (2003) estimate that, over the twentieth century, the costs of moving goods declined by over 90% in real terms, and it is now essentially free to move goods. If so, an emission tax is likely to be counterproductive (due to its effect on z), at least, not as effective as an instrument for pollution control now than when transport costs were much higher.

There are two main mechanisms through which declining transport costs reduce the effectiveness of emission taxes. First, as shown by equation (9), an emission tax (corresponding an increase in *w*) reduces the importance of agglomeration economies (i.e., $(Y_E - Y_W)$ decreases with *w*), but the impact will be minimal when transport costs are close to be zero. Second, an emission tax reduces the production level and thus pollution discharge in both regions. However, as shown by equation (12), it improves the environmental quality more in the East than in the West if the pollution is not totally mobile (i.e., $(Z_E - Z_W)$ decrease with *w* for a given λ), and this impact on environmental quality is not affected by transport costs. Thus, when transport costs are sufficiently low, the loss of agglomeration benefits caused by an emission tax will be smaller than the gain in environmental quality in the East. As a result, some firms will move from the less populated West to the more agglomerated East when an emission tax is imposed. The increased concentration of firms in the East will moderate the effect of the emission tax on total production and consumption, and will reduce the effectiveness of the tax as an instrument for pollution control.

The effect of the tax on pollution damages depends not only on its effect on total pollution but also on the distribution of pollution across the two regions. The effect of tax payments on pollution concentration is given by:

(33) $$\frac{d(Z_E^*-Z_W^*)}{dw} = \frac{\partial(Z_E-Z_W)}{\partial w}\bigg|_{\lambda} + \frac{\partial(Z_E^*-Z_W^*)}{\partial\lambda}\frac{\partial\lambda^*}{\partial w} =$$

$$= \left[-\frac{2s\sqrt{\bar{\theta}}}{\beta}\right] + \left[\frac{s(2\alpha-2w-0.5t)}{\beta}\frac{\partial\lambda^*}{\partial w}\right]$$

$$= \left[-\frac{2s\sqrt{\bar{\theta}}}{\beta}\right] + \frac{s\sqrt{\bar{\theta}}(2\alpha-2w-0.5t)}{2\beta\theta}\left[sz\delta - \frac{(1.5t)^2\beta}{(2\alpha-2w-0.5t)^2}\right]$$

$$= -\frac{2s\sqrt{\bar{\theta}}}{\beta}\left\{1 - \frac{(2\alpha-2w-0.5t)}{4\theta}\left[sz\delta - \frac{(1.5t)^2\beta}{(2\alpha-2w-0.5t)^2}\right]\right\}.$$

As can be seen, the direct effect of an increase in marginal production costs for a given concentration of firms is negative, implying that the tax payments will decrease pollution concentration, since it will decrease pollution more in the East than the West if the spatial allocation of firms does not change. However, as noted above, the change in marginal production costs will induce relocation, and both the magnitude and direction of this impact depend on the strength of the pollution effect, i.e., the magnitude of $sz\delta$. If $sz\delta$ is sufficiently large, the indirect effect through relocation will more than offset the direct effect, and the total effect in (28) will be positive, i.e., the tax payments will lead to greater concentration of pollution. This, coupled with the positive effect through changes in z, implies that overall the tax will lead to greater concentration of pollution. However, when $sz\delta$ is small, the overall impact of the tax on pollution concentration is ambiguous.

Because of the increased concentration of pollution, the aggregate pollution damage can be higher under the emission tax. This is confirmed by simulation results presented in Table B1 in Appendix B.

Table 1 summarizes the above results regarding the effect of an emissions tax. Column 2 indicates that the tax has the same perverse effects as the standard when considering only the impact of reducing emissions per unit of output, i.e., agglomeration, aggregate pollution, pollution concentration, and total damages would increase. In addition, with the exception of

aggregate pollution, these effects are amplified by the impact of the tax payments on marginal production costs if the pollution effects are sufficiently large. Conversely, with sufficiently small pollution effects, the tax payments work to offset the perverse effects of changes in z on pollution concentrations and total damages. With regard to aggregate pollution, the tax payment effects always work in the opposite direction of the impact of changes in z, implying that achieving a given reduction in pollution intensity through a standard will always result in higher aggregate pollution than achieving that same reduction through an emissions tax.

5. Optimal Distribution of Firms

We turn next to the question of social optimality. It is well-known that the spatial equilibria that emerge in agglomeration economies can be socially inefficient (e.g., Krugman, 1991; Ottaviano et al., 2002). In this section, we show that both the spatial distributions of firms and pollution are generally not socially optimal in market equilibrium. We follow Ottaviano et al. (2002) and define the optimal distribution as the one that maximizes the aggregate utility for all residents in the economy where all firms price their products at the marginal costs, including marginal pollution damage and transport costs. Formally, the optimal distribution of firms between the two regions, λ^o, solves:

(35) $$\max_{0\leq\lambda\leq 1} TV(\lambda) = 0.5V_E^L + \lambda V_E^C + 0.5V_W^L + (1-\lambda)V_W^C,$$

where V_r^j is the indirect utility function of type-*j* workers (*j* = *L* or *C*) in region *r* (*r* = *E, W*). To simplify notation, we present results for constant marginal pollution damage (i.e., $\delta = 0$) and constant fixed cost (i.e., $\phi = 0$), although the general insights below hold for $\delta \neq 0$. In this case, the maximization problem (36) can be simplified to (see the proof of proposition 7 in the appendix):

(36) $$\max_{0\leq\lambda\leq 1} TV(\lambda) = E + \varphi(\lambda - 0.5)^2,$$

where

$$E = \frac{1}{\beta}\{2\alpha^2 - \alpha(\alpha + w) + \frac{1}{8}[(\alpha + w)^2 + (\alpha + w + t)^2] + \beta + z(\alpha - w - z - 0.5t)\},$$

$$\varphi = \frac{1}{\beta}\,[0.5t(\alpha - w - 0.5t - 4z) + 4s(2\alpha - 2w - 2z - t - s)].$$

This suggests that the optimal distribution of firms depends on the value of φ. If $\varphi > 0$, the agglomerated configuration is the optimal distribution; if $\varphi < 0$, the dispersed configuration is the optimal distribution; and if $\varphi = 0$, any configuration is optimal. Note that $\varphi \lesseqgtr 0$ if and only if

(37) $$z \gtreqless z^* \equiv \frac{0.5t(\alpha+w-0.5t)-8s(\alpha-w-0.5t-0.5s)}{2t+8s}.$$

This suggests the following result:

Proposition 7. *The optimal distribution of firms (λ^o) is as follows:*

$$\lambda^o = \begin{cases} 0.5 & if\ z > z^* \\ any\ \lambda \in [0,1] & if\ z = z^*. \\ 1 & if\ z < z^* \end{cases}$$

A comparison of proposition 7 with corollary 1 reveals that the equilibrium distribution of firms is generally not socially optimal. In particularly, with mobile pollution (s = 0), all firms agglomerate in one region in market equilibrium, which is optimal only if $z < 0.25(\alpha + w - 0.5t)$. When $z > 0.25(\alpha + w - 0.5t)$, the optimal distribution is the perfect dispersion, but the market equilibrium is the complete agglomeration. In an analysis of the effect of pollution and productivity on the spatial distribution of firms, Wu and Reimer (2016) also found that the equilibrium distribution of polluting firms differs from the social optimum when they generate a large amount of stationary pollution and have much higher or lower productivity than clean firms.

In addition to overconcentration of firms, firms' production and pollution level are also above the socially optimal level when the pollution level is sufficiently high. To see this, note that the optimal prices $p^o_{rr} = w + z$ and $p^o_{rr'} = w + z + t$ for $r \neq r'$ are above the equilibrium prices $p^*_{rr} = 0.5(\alpha + w)$ and $p^*_{rr'} = 0.5(\alpha + w + t)$ when $z > 0.25(\alpha - w)$. Because of lower prices, production and consumption levels are too high under the market equilibrium. In this case, both the aggregate level of pollution and the total pollution damage are above the socially optimal level because firms are not only overly concentrated, but also produce too much.

The discrepancy between the optimal and equilibrium distributions arises because firms do not consider the pollution externalities they impose on local residents when making location decisions. As a firm moves from a smaller to larger market, it will produce more and causes more pollution. The private incentive to move to a larger market is larger than the social value of moving when the pollution level is high, causing too many people to be exposed to pollution. Conversely, even if it is socially desirable for a firm to move to a smaller market, the firm may not do so because it cannot capture the social benefit of moving (less pollution damage).

On the other hand, when the pollution level is relatively low (i.e., $z < 0.25(\alpha - w - 0.5t)$), the market may lead to optimal distribution of firms ($\lambda^* = \lambda^o = 1$), but the equilibrium prices tend to be too high (i.e., $p_{rr}^o < p_{rr}^*$ and $p_{rr'}^o < p_{rr'}^*$). As a result, the consumption and production levels are two low, leading to welfare losses. In this case, firms' desires to set higher prices, associated with market power, dominate the tendency to underprice, due to the failure to consider pollution externality.

6. Optimal Policy

The analysis above suggests that optimal environmental regulation must overcome two main challenges when firms are mobile. First, it must achieve both the optimal spatial distribution of firms and the optimal production level of individual firms. Second, it must induce firms to set right prices for both locally consumed goods and exported goods. Because of these challenges, it is difficult to use a single instrument to achieve social optimality. For example, a tax rate that induces firms to set the right price for locally consumed goods may be different from the tax rate that induces firms to set the right price for exported goods.

However, a combination of a differential tax and a lump-sum transfer can be used to achieve the optimal distribution of firms and optimal production and pollution level of individual firms, as well as optimal pricing of local and exported goods. To illustrate, consider an industry that generates completely mobile emissions, with emission intensity $z > z^*$. From proposition 7, the optimal distribution of firms is $\lambda^o = 0.5$, and the optimal prices are

(38a) $$p_{rr} = w + z, \quad p_{ro} = w + z + t.$$

With an emission tax τ, the firms' marginal cost increases to $w + \tau z(\tau)$, and the prices the firms would set under the tax equal

(38b) $$p_{rr} = \frac{\alpha+w+\tau z(\tau)}{2}, \quad p_{ro} = \frac{\alpha+w+t+\tau z(\tau)}{2} \ .$$

A comparison of (38a) and (38b) suggests that the following differential tax can be used to induce firms to price their products at the optimal level:

(39) $$\tau^o = \begin{cases} \frac{w-\alpha+2z^o}{z^o} & for\ locally\ consumed\ goods \\ \frac{w+t-\alpha+2z^o}{z^o} & for\ exported\ goods \end{cases},$$

where z^o denotes a firm's pollution level under the tax when firms are perfectly dispersed between the two regions. This tax will exactly offset firms' incentives to under-price their products.

Under the tax, the utility difference for entrepreneurs located in the two regions equals:

$$V_E - V_W = \frac{6t(\alpha-w-z^o-0.5t)+\beta A(z_0-z^o)}{\beta}(\lambda - 0.5).$$

Thus, without other incentives, all firms agglomerate in one region.

An agglomeration impact fee or a lump-sum transfer can be made to induce optimal distribution of firms. For example, the following agglomeration impact fee imposed on firms in the larger region (East) will make $\lambda = 0.5$ a stable equilibrium:

(40) $$T^* = \left[\frac{6t(\alpha-w-z^o-0.5t)+\beta A(z_0-z^o)}{\beta} + M\right](\lambda - 0.5),$$

where M is any positive number. Under the impact fee, the utility difference for entrepreneurs located in the two regions equals:

$$V_E - V_W = -M(\lambda - 0.5),$$

which implies that $\lambda = 0.5$ is a stable equilibrium under the impact fee.

7. Conclusions

Environmental and regional economists often study the same reality, but focus on the different aspects of the reality. For example, both environmental and regional economists study firms' location decisions, but the former focus on the effect of environmental regulation, while the later focus on the effect of agglomeration economies, without recognizing that the effects of environmental regulation and agglomeration economies are not independent. This disconnection is surprising because the economic activity that is the focus of regional/urban economics often causes the environmental problems that resource and environmental economists care about, and environmental change has feedback effects on economic growth. The disconnection is also costly for both research and policy making because the key dynamics and feedbacks are lost in both fields, and real-world problems may go unaddressed or be addressed in a suboptimal manner. For example, although urban and regional economist have long recognized that agglomeration economies represent some self-reinforcing dynamics, environmental economists have not explored how reactive relocations and agglomeration economies affect the effectiveness of environmental regulation.

In this paper, we take an initial step to analyze the benefits from incorporating insights from urban and regional economics into environmental policy analysis. We find that consideration of firms' reactive relocations and agglomeration economies can change some of the classic results from traditional Pigouvian marginal analysis. Perhaps the most striking result is that a performance or technology standard can lead to higher pollution concentration and larger pollution damage when firm relocations and agglomeration economies are considered. Adoption of cleaner technology makes it less costly for firms to agglomerate, which can lead to increased concentration of firms and higher output per firm in the more concentrated region. An emission tax can also lead to increased concentration of pollution and more pollution damage when firms' reactive relocations are considered, even if the total ambient pollution level may be lower under the tax.

Our results have interesting policy implications. Carbon tax is widely used or proposed as an efficient instrument for reducing greenhouse gas (GHG) emissions and for decelerating climate change and its negative effects on the environment and human health. In 2015, about 40 nations and over 20 cities, states, and regions—representing almost a quarter of GHG emissions—were putting a price on carbon (World Bank, 2015). The increasing adoption of carbon taxes around the world, to a large extent, can be attributed to the belief that carbon taxes

are efficient for controlling greenhouse gas emissions. This paper highlights that the consequences of a carbon tax are far more complex than simply internalizing externalities associated with carbon emissions. For example, carbon taxes can have undesirable leakage effects in the long run - firms and people might relocate in response to a carbon tax, which may affect both the efficiency and effectiveness of a carbon tax. Industries generating GHG emissions are highly diverse: Some are heavy polluters, others generate only a small amount of GHG emissions; some are highly mobile, others are not; some are concentrated in one region, others are more spread out; some have constant return to scale production technologies, others demonstrate strong economies of scale; and some produce goods that are costly to transport, while others face low transport costs. GHG emissions are highly mobile; if the ultimate objective is to reduce global climate change, the location of GHG reduction really does not matter, but matters for social costs. Further research is needed for the design of a carbon policy that takes into account firms' reactive relocations and the diversity of the industries that generate GHG emissions.

References

Artz, Georgeanne, Younjun Kim, and Peter Orazem. 2016. "Does Agglomeration Matter Everywhere? New Firm Location Decisions in Rural and Urban Markets," *Journal of Regional Science*, 56(1), 72-95.

Atkinson, Scott E., and T. H. Tietenberg. 1982. "The Empirical Properties of Two Classes of Designs for Transferable Discharge Permit Markets," *Journal of Environmental Economics and Management*, 9(2), 101-21.

Au, Chun-Chung and Vernon Henderson. 2006. "Are Chinese cities too small?" *Review of Economic Studies*, 73, 549–576.

Baumol, William J., and Wallace E. Oates. 1988. *The Theory of Environmental Policy*. 2nd edition. Cambridge: Cambridge University Press.

Bovenberg, A. L. and R. A. de Mooij. 1994. "Environmental Levies and Distortionary Taxation," *American Economic Review* 84, 1085-1089.

Broersma, Lourens and Jouke van Dijk. 2008. "The effect of congestion and agglomeration on multifactor productivity growth in Dutch regions," *Journal of Economic Geography*, 8, 181–209.

Brunnermeier, S. B., and A. Levinson, "Examining the Evidence on Environmental Regulations and Industry Location," *Journal of Environment and Development* 13(1)(2004), 6–41.

Burkitt, Laurie and Brian Spegele. 2013. "Why Leave Job in Beijing? To Breathe," *The Wall Street Journal*, April 14.

Copeland, Brian R. and M. Scott Taylor. 1999. "Trade, Spatial Separation, and the Environment," *Journal of International Economics*, 47, 137–168.

Duranton, Gilles and Diego Puga. 2014. "Chapter 5–the Growth of Cities," A. Philippe and N. D. Steven (eds.), *Handbook of Economic Growth*. Amsterdam: Elsevier, 781–853.

Ellison, Glenn, Edward L. Glaeser, and William R. Kerr. 2010. "What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns," *The American Economic Review*, 100(3), 1195–1213.

Farrow, R. Scott, Martin T. Schultz, Pinar Celikkol, and George L. Van Houtven. 2005. "Pollution Trading in Water Quality Limited Areas: Use of Benefits Assessment and Cost-Effective Trading Ratios," *Land Economics* 81(2), 191-205.

Feldman, Maryann P. and David B. Audretsch. 1999. "Innovation in Cities: Science-Based Diversity, Specialization and Localized Competition," *European Economic Review*, 43(2), 409–429.

Fu, Shihe and Junjie Hong. 2011. "Testing urbanization economies in manufacturing industries: Urban diversity or urban size?" *Journal or Regional Science*, 51, 585–603.

Fullerton, D. 1997. "Environmental Levies and Distortionary Taxation: Comment," *American Economic Review* 87, 245-251.

Fullerton, D. and G. E. Metcalf. 2001. "Environmental Controls, Scarcity Rents, and Pre-existing Distortions," *Journal of Public Economics* 80, 249-267.

Fujita, Masahisa, Paul Krugman, and Anthony J. Venables. 1999. "The Spatial Economy: Cities, Regions, and International Trade," Cambridge: MIT Press.

Fujita, Masahisa and Jacques-Francois Thisse. 2002. *Economics of Agglomeration: Cities, Industrial Location, and Regional Growth*. New York: Cambridge University Press.

Glaeser, Edward. 1998. "Are cities dying?" *Journal of Economic Perspectives*, 12, 139–160.

Glaeser, Edward L. and Joshua D. Gottlieb. 2008. "The Economics of Place-Making Policies," *Brookings Papers on Economic Activity, Spring* 2008, 155–253.

Glaeser, Edward, Hedi Kallal, José Scheinkman and Andrei Shleifer. 1992. "Growth in cities," *Journal of Political Economy*,100, 1126–1152.

Glaeser, Edward, José Scheinkman and Andrei Shleifer. 1995. "Economic growth in a cross–section of cities," *Journal of Monetary Economics*, 36, 117–143.

Goulder, L. H, ed. 2002. *Environmental Policy Making in Economies with Prior Tax Distortions*. Cheltenham: Elgar.

Goulder, L. H. 1995. "Environmental Taxation and the 'Double Dividend': A Reader's Guide," *International Tax and Public Finance*, 2, 157-184.

Henderson, J. Vernon. 2003. "Marshall's Scale Economies," *Journal of Urban Economics,* 53, 1–28.

Hosoe, Moriki and Tohru Naito. 2006. "Transboundary Pollution Transmission and Regional Agglomeration Effects," *Papers in Regional Science*, 85(1), 99–120.

Jeppesen, T., J. A. List, and H. Folmer, "Environmental regulations and new plant location decisions: evidence from a meta-analysis," *Journal of Regional Science 42*(1)(2002), 19–49.

Koskela, E. and R. Schöb. 1999. "Alleviating Unemployment: The Case for Green Tax Reforms," *European Economic Review* 43, 1723-1746.

Krugman, Paul. 1991. "Increasing Returns and Economic Geography," *Journal of Political Economy*, 99, 483–499.

Krugman, Paul and Anthony J. Venables. 1995. "Globalization and the Inequality of Nations," *Quarterly Journal of Economics*, 110, 857–80.

Krupnick, Alan J., Wallace E. Oates, and Eric Van De Verg. 1983. "On Marketable Air-Pollution Permits: The Case for a System of Pollution Offsets," *Journal of Environmental Economics and Management* 10(3), 233-47.

Lange, Andreas and Martin F. Quaas. 2007. "Economic Geography and the Effect of Environmental Pollution on Agglomeration," *The B.E. Journal of Economic Analysis & Policy*, 7(1), 1935-1682.

Levinson, Arik and M Scott Taylor. 2008. "Unmasking the Pollution Haven Effect," *International Economic Review*, 49, 223–254.

Lucas, Robert E. 1988. "On the Mechanics of Economic Development," *Journal of Monetary Economics*, 22, 3–39.

Markusen, James, Edward R. Morey, and Nancy Olewiler. 1993. "Environmental Policy when Market Structure and Plant Locations are Endogenous," *Journal of Environmental Economics Management*, 24, 69–86.

Marshall, Alfred. 1920. *Principles of Economics*. London: MacMillan.

Montgomery, W. David. 1972. "Markets in Licenses and Efficient Pollution Control Programs," *Journal of Economic Theory* 5(3), 395-418.

Moretti, Enrico. 2004. "Workers' Education, Spillovers, and Productivity: Evidence from Plant-Level Production Functions," *The American Economic Review*, 94(3), 656–690.

Muller, Nicholas Z., and Robert Mendelsohn. 2009. "Efficient Pollution Regulation: Getting the Prices Right," *The American Economic Review* 99(5), 1714-1739.

Okubo, Toshihiro, and Pierre M. Picard, and Jacques-Francois Thisse. 2010. "The Spatial Selection of Heterogeneous Firms," *Journal of International Economics*, 82(2), 230–237.

Ottaviano, Gianmarco I.P., Takatoshi Tabuchi, and Jacques-Francois Thisse. 2002. "Agglomeration and Trade Revisited," *International Economic Review*, 43, 409–435.

Partridge, Mark, Rat Bollman, Rose Olfert and Alessandro Alasia. 2007. "Riding the wave of urban growth in the countryside: Spread, backwash, or stagnation?" *Land Economics*, 83, 128–152.

Parry, I. W. H. 1995. "Pollution Taxes and Revenue Recycling," *Journal of Environmental Economics and Management* 29, S-64-S77.

Porter, Michael E. 1990. *The Competitive Advantage of nations*. New York: The Free Press.

Puga, Diego. 2010. "The Magnitude and Causes of Agglomeration Economies," *Journal of Regional Science*, 50(1), 203–219.

Rosenthal, Stuart S.. 2003. "Geography, Industrial Organization, and Agglomeration," *Review of Economics and Statistics*, 85(2), 377–393.

Rosenthal, Stuart S.. 2004. "Evidence on the Nature and Sources of Agglomeration Economies," in J. V. Henderson and T. Jacques-François, *Handbook of Regional and Urban Economics*. Amsterdam: Elsevier, 2119–2171.

Saito, Hisamitsu, Munisamy Gopinath, and JunJie Wu. 2011. "Heterogeneous Firms, Trade Liberalization and Agglomeration," *Canadian Journal of Economics,* 44(2), 541–560.

Shapiro, Jesse M. 2006. "Smart Cities: Quality of Life, Productivity, and the Growth Effects of Human Capital," *The Review of Economics and Statistics*, 88(2), 324–335.

Tabuchi, Takatoshi and Jacques-Francois Thisse 2006. "Regional Specialization, Urban Heirarchy, and Commuting Costs," *International Economic Review*, 47(4), 1468–2354.

The World Bank. 2015. The State and Trends of Carbon Pricing. Washington DC: The Work Bank Group.

Tietenberg, Thomas H. 1974. "Derived Demand Rules for Pollution Control in a General Equilibrium Space Economy," *Journal of Environmental Economics and Management*, 1, 3–16.

Tietenberg, Thomas H. 1980. "Transferable Discharge Permits and the Control of Stationary Source Air Pollution: A Survey and Synthesis," *Land Economics* 56(4), 391-416.

Unteroberdoerster, Olaf. 2001,. "Trade and Transboundary Pollution: Spatial Separation Reconsidered," *Journal of Environmental Economics and Management*, 41, 269–285.

Wu, JunJie, and J. Reimer. 2016. Pollution Mobility, Productivity Disparity, and the Spatial Distribution of Polluting and Nonpolluting Firms. *Journal of Regional Science* 56: 615–634.

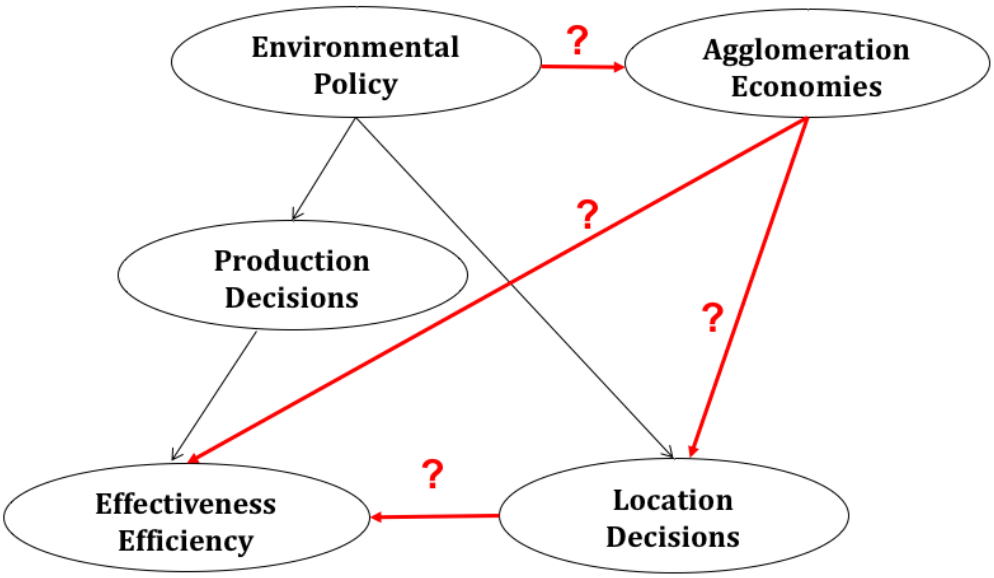

Figure 1. Interaction effects between agglomeration economies and environmental regulation

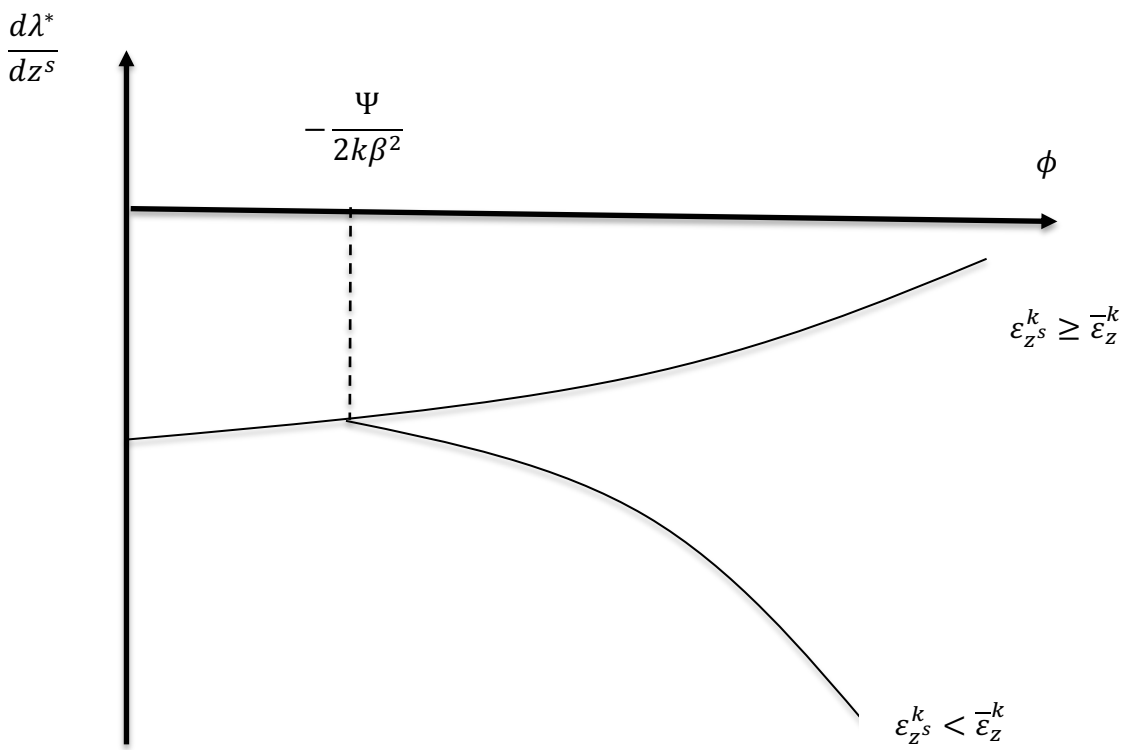

Figure 2. Agglomeration economies can increase or decrease the effect of a regulatory environmental standard on firm concentration

Table 1: Effect of an emissions tax under partial agglomeration

F^*	Effect of reduction in z, for a given marginal production cost $\left\{\frac{\partial F^*}{\partial z}\Big\vert_w \frac{\partial z^*}{\partial \tau}\right\}$ (1)	Direct effect of increase in marginal production cost for a given z and concentration of firms $\left\{\frac{\partial F^*}{\partial w}\Big\vert_{z,\lambda} \frac{\partial w}{\partial \tau}\right\}$ (2)	Indirect effect of increase in marginal production cost through change in concentration of firms, for a given z $\left\{\frac{\partial F^*}{\partial \lambda}\Big\vert_z \frac{\partial \lambda^*}{\partial w}\Big\vert_z \frac{\partial w}{\partial \tau}\right\}$ (3)	Total effect of increase in marginal production cost, for a given z $\left\{\frac{\partial F^*}{\partial w}\Big\vert_z \frac{\partial w}{\partial \tau}\right\}$ (4)=(2)+(3)	Overall effect of an emission tax increase (5)=(1)+(4)
Agglomeration λ^*	(+)	0	+ if $sz\delta$ is sufficiently large	+ if $sz\delta$ is sufficiently large	+if $sz\delta$ is sufficiently large
Aggregate Pollution $(z_E^* + z_W^*)$	(+)	(-)	+ if $sz\delta$ is sufficiently large	(-)	Ambiguous
Pollution Concentration $(z_E^* - z_W^*)$	(+)	(-)	+ if $sz\delta$ is sufficiently large	+ if $sz\delta$ is sufficiently large	+ if $sz\delta$ is sufficiently large
Total Damages TD^*	(+)	(-)	+ if $sz\delta$ is sufficiently large	+ if $sz\delta$ is sufficiently large	+ if $sz\delta$ is sufficiently large

APPENDIX A

Proof of Proposition 1

First, we derive the utility difference in (16). From the indirect utility function (5), the utility difference equals:

(A1) $V_E - V_W = (V_{E0} - V_{W0}) + (Y_E - Y_W) - \big(D(Z_E) - \mathrm{D}(Z_W)\big),$

where V_{r0} is the utility from consumption of the goods in region *r* and is defined by

(A2a) $V_{E0} - = \frac{\alpha^2}{\beta} - \frac{\alpha}{\beta}[\lambda p_{EE} + (1-\lambda)p_{WE}] + \frac{1}{2\beta}[\lambda p_{EE}^2 + (1-\lambda)p_{WE}^2],$

(A2b) $V_{W0} - = \frac{\alpha^2}{\beta} - \frac{\alpha}{\beta}[\lambda p_{EW} + (1-\lambda)p_{WW}] + \frac{1}{2\beta}[\lambda p_{EW}^2 + (1-\lambda)p_{WW}^2],$

(A2c) $$\begin{aligned} V_{E0} - V_{W0} &= -\frac{\alpha}{\beta}[\lambda(p_{EE} - p_{EW}) + (1-\lambda)(p_{WE} - p_{WW})] \\ &\quad + \frac{1}{2\beta}[\lambda(p_{EE}^2 - p_{EW}^2) + (1-\lambda)(p_{WE}^2 - p_{WW}^2)] \\ &= -\frac{\alpha}{\beta}(\lambda - 0.5)(p_{EE} - p_{EW}) + \frac{1}{\beta}(\lambda - 0.5)(p_{EE}^2 - p_{EW}^2) \\ &= \frac{\alpha}{\beta}(\lambda - 0.5)t - \frac{1}{2\beta}(\lambda - 0.5)t(\alpha + w + 0.5t) \\ &= \frac{t}{2\beta}(\lambda - 0.5)(\alpha - w - 0.5t), \end{aligned}$$

where $Y_E = Y_W = 1$ for labor. Entrepreneurs' income in the two regions and are defined by:

(A3a) $Y_E = (p_{EE} - w)q_{EE}M_E + (p_{EW} - w - t)q_{EW}M_W - K_E.$

(A3b) $Y_W = (p_{WE} - w - t)q_{WE}M_E + (p_{WW} - w)q_{WW}M_W - K_W.$

Note that $(p_{EE} - w)q_{EE} = (p_{WE} - w - t)q_{WE}$ and $(p_{EW} - w - t)q_{EW}M_W = (p_{WE} - w - t)q_{WE}M_E$. Therefore, the income difference for entrepreneurs is

(A3c) $$\begin{aligned} (Y_E - Y_W) &= (p_{EE} - w)q_{EE}(M_E - M_W) + (p_{EW} - w - t)q_{EW}(M_W - M_E) - (K_E - K_W) \\ &= [(p_{EE} - w)q_{EE} - (p_{EW} - w - t)q_{EW}](M_E - M_W) - (K_E - K_W) \end{aligned}$$

$$=\left[\left[\frac{(\alpha-w)^2}{4\beta}-\frac{(\alpha-w-t)^2}{4\beta}\right]\right]2(\lambda-0.5)+2k\phi(\lambda-0.5)$$

$$=\left[\frac{t(\alpha-w-0.5t)}{\beta}+2k\phi\right](\lambda-0.5).$$

From equations (10) and (11),

(A4b) $$Z_E-Z_W=s\lambda(q_{EE}M_E+q_{EW}M_W)-s(1-\lambda)(q_{WE}M_E+q_{WW}M_W),$$

$$=s\{(\lambda-0.5)[(q_{EE}+q_{WE})M_E+(q_{EW}+q_{WW})M_W)]$$

$$+0.5[(q_{EE}-q_{WE})M_E+(q_{EW}-q_{WW})M_W)]\}$$

$$=s\{(\lambda-0.5)2(q_{EE}+q_{WE})+(\lambda-0.5)(q_{EE}-q_{WE})\}$$

$$=s[3q_{EE}+q_{WE}](\lambda-0.5)$$

$$=\frac{s}{\beta}(2\alpha-2w-0.5t)(\lambda-0.5).$$

(A4b) $$Z_E+Z_W=z\lambda(q_{EE}M_E+q_{EW}M_W)+z(1-\lambda)(q_{WE}M_E+q_{WW}M_W),$$

$$=z\{(\lambda-0.5)[(q_{EE}-q_{WE})M_E+(q_{EW}-q_{WW})M_W)]$$

$$+0.5[(q_{EE}+q_{WE})M_E+(q_{EW}+q_{WW})M_W)]\}$$

$$=z\{(\lambda-0.5)(M_E-M_W)(q_{EE}-q_{WE})+0.5(M_E+M_W)(q_{EE}+q_{WE})\}$$

$$=z[2(\lambda-0.5)^2(q_{EE}-q_{WE})+(q_{EE}+q_{WE})]$$

$$=z[(\lambda-0.5)^2\frac{t}{\beta}+\frac{1}{\beta}(2\alpha-(\alpha+w+0.5t)]$$

$$=\frac{z}{\beta}[t(\lambda-0.5)^2+(\alpha-w-0.5t)].$$

(A5a) $$D(Z_E)-\mathrm{D}(Z_W)=(Z_E+\delta Z_E^2)-(Z_W+\delta Z_W^2)$$

$$=(Z_E-Z_W)+\delta(Z_E^2-Z_W^2)$$

$$=(Z_E-Z_W)+\delta(Z_E-Z_W)(Z_E+Z_W)$$

$$=(Z_E-Z_W)[1+\delta(Z_E+Z_W)]$$

$$=\frac{s}{\beta^2}(2\alpha-2w-0.5t)[\beta+\delta z[t(\lambda-0.5)^2+(\alpha-w-0.5t)]](\lambda-0.5).$$

(A5b) $$TD = (Z_E + \delta Z_E^2)M_E + (Z_W + \delta Z_W^2)M_W$$

$$= (Z_E + \delta Z_E^2)(0.5 + \lambda) + (Z_W + \delta Z_W^2)[0.5 + (1 - \lambda^*)]$$

$$= (Z_E + \delta Z_E^2)[1 + (\lambda - 0.5)] + (Z_W + \delta Z_W^2)[1 - (\lambda - 0.5)]$$

$$= (Z_E + Z_W) + \frac{\delta}{2}[Z_E + Z_W)^2 + (Z_E - Z_W)^2]$$

$$+(\lambda - 0.5)(Z_E - Z_W)[1 + \delta(Z_E + Z_W)]$$

$$= (Z_E + Z_W) + \frac{\delta}{2}[(Z_E + Z_W)^2 + (Z_E - Z_W)^2]$$

$$+(\lambda - 0.5)^2 \frac{s}{\beta}(2\alpha - 2w - 0.5t)[1 + \delta(Z_E + Z_W)].$$

Substituting (A2c), (A3c) and (A5) into (A1) gives the utility difference for entrepreneurs:

(A6) $$V_E - V_W = (V_{E0} - V_{W0}) + (Y_E - Y_W) - \big(D(Z_E) - \mathrm{D}(Z_W)\big)$$

$$= \frac{t}{2\beta}(\lambda - 0.5)(\alpha - w - 0.5t)$$

$$+ \left[\frac{t(\alpha - w - 0.5t)}{\beta} + 2k\phi\right](\lambda - 0.5)$$

$$-\frac{s}{\beta^2}(2\alpha - 2w - 0.5t)[\beta + \delta z[t(\lambda - 0.5)^2 + (\alpha - w - 0.5t)]](\lambda - 0.5)$$

$$= \left\{\frac{\frac{3t}{2\beta}(\alpha - w - 0.5t) + 2k\phi}{\frac{s}{\beta^2}(2\alpha - 2w - 0.5t)} - \beta - \delta z(\alpha - w - 0.5t) - \delta zt(\lambda - 0.5)^2\right\}$$

$$\times \frac{s}{\beta^2}(2\alpha - 2w - 0.5t)\ (\lambda - 0.5)$$

$$= \left\{\frac{1.5t\beta(\alpha - w - 0.5t) + 2\beta^2 k\phi}{s(2\alpha - 2w - 0.5t)} - \beta - \delta z(\alpha - w - 0.5t) - \delta zt(\lambda - 0.5)^2\right\}$$

$$\times \frac{s}{\beta^2}(2\alpha - 2w - 0.5t)\ (\lambda - 0.5)$$

$$= -\frac{1}{\beta^2}(2\alpha - 2w - 0.5t)\begin{Bmatrix} s\delta zt(\lambda - 0.5)^2 - \frac{1.5t\beta(\alpha - w - 0.5t) + 2\beta^2 k\phi}{(2\alpha - 2w - 0.5t)} \\ +\beta s + s\delta z(\alpha - w - 0.5t) \end{Bmatrix}(\lambda - 0.5)$$

$$= C[s\delta zt(\lambda - 0.5)^2 - \theta](\lambda - 0.5),$$

where

$$C = -\frac{2\alpha-2w-0.5t}{\beta^2} < 0,$$

$$\theta = \frac{1.5\beta t(\alpha-w-t)+2\beta^2 k\phi}{(2\alpha-2w-0.5t)} - s[\beta + \delta z(\alpha - w - 0.5t)].$$

This derives equation (16).

We now use (A6) to prove proposition 1.

(a) If $s\delta zt = 0$ and $\theta = 0$, then $V_E - V_W = 0$ for any λ. Therefore, any $\lambda \in [0,1]$ is a stable equilibrium.

(b) If $s\delta zt = 0$ and $\theta \neq 0$, then $V_E - V_W = -C\theta(\lambda - 0.5)$. Thus, $V_E - V_W = 0$ only when $\lambda = 0.5$. $\partial(V_E - V_W)/\partial\lambda = -C\theta < 0$ if and only if $\theta < 0$. Therefore, $\lambda^* = 0.5$ is a stable only $\theta < 0$. Also, when $\theta > 0$, $V_E - V_W > 0$ at $\lambda = 1$ and $V_E - V_W < 0$ at $\lambda = 0$. Thus $\lambda^* = 0$ and $\lambda^* = 1$ are the only stable equilibria when when $\theta > 0$.

(c) Finally, we consider the case of $s\delta zt \neq 0$.

i) When $\theta \leq 0$, $V_E - V_W = 0$ has only one solution at $\lambda = 0.5$. Because $\frac{\partial(V_E - V_W)}{\partial\lambda} < 0$ at $\lambda = 0.5$, $\lambda = 0.5$ is a stable equilibrium. In addition, $V_E < V_W$ at $\lambda = 1$ and $V_E > V_W$ at $\lambda = 0$. Complete agglomeration is not a stable equilibrium when $\theta \leq 0$.

ii) When $0 \leq \tilde{\theta} < 0.25$, $V_E - V_W = 0$ has three solutions within [0, 1]: $\lambda = 0.5\ \left(0.5 - \sqrt{\tilde{\theta}}\right), \left(0.5 + \sqrt{\tilde{\theta}}\right)$. However, $\frac{\partial(V_E - V_W)}{\partial\lambda} > 0$ at $\lambda = 0.5$ and $\frac{\partial(V_E - V_W)}{\partial\lambda} < 0$ at $\lambda = \left(0.5 - \sqrt{\tilde{\theta}}\right), \left(0.5 + \sqrt{\tilde{\theta}}\right)$. In addition, $V_E < V_W$ at $\lambda = 1$ and $V_E > V_W$ at $\lambda = 0$. Therefore, $\lambda = \left(0.5 - \sqrt{\tilde{\theta}}\right), \left(0.5 + \sqrt{\tilde{\theta}}\right)$ are only stable equilibrium when $0 \leq \tilde{\theta} < 0.25$.

iii) When $\tilde{\theta} \geq 0.25$, $\lambda = 0.5$ is the only solution of $V_E - V_W = 0$ within [0, 1]. However, $\lambda = 0.5$ is not a stable equilibrium because $\frac{\partial(V_E - V_W)}{\partial\lambda} > 0$ at $\lambda = 0.5$. In this case, $\lambda = 0, 1$ are the only stable equilibria because $V_E > V_W$ at $\lambda = 1$ and $V_E < V_W$ at $\lambda = 0$.

Proof of Corollary 1

a) When $s = 0$, the utility difference for entrepreneurs in the two regions equals

(A7) $$V_E^k - V_W^k = \; = \frac{1.5t(\alpha - w - 0.5t)}{\beta}(\lambda - 0.5).$$

$\lambda = 0.5$ is the only solution of $V_E - V_W = 0$ within [0, 1]. However, $\lambda = 0.5$ is not a stable equilibrium because $\frac{\partial(V_E - V_W)}{\partial\lambda} > 0$ at $\lambda = 0.5$. In this case, $\lambda = 0, 1$ are the only stable equilibria because $V_E > V_W$ at $\lambda = 1$ and $V_E < V_W$ at $\lambda = 0$.

b) The results follow directly from proposition 1 because when $\theta < 0$ when $z > 2\overline{z}$; $0 \leq \theta \leq \frac{s\delta zt}{4}$ when $\overline{z} \leq z \leq 2\overline{z}$, and $\theta > \frac{s\delta zt}{4}$ when $z < \overline{z}$, where

$$\overline{z} \equiv \frac{\beta}{2s\delta(\alpha - w - 0.25t)^2}[1.5t(\alpha - w - 0.5t) - 2s(\alpha - w - 0.25t)].$$

Proof of Proposition 2

Totally differentiating (12)' gives:

(A8) $$\frac{d(Z_E^* + Z_W^*)}{dz} = \left.\frac{\partial(Z_E^* + Z_W^*)}{\partial z}\right|_{\lambda^*} + \left.\frac{\partial(Z_E^* + Z_W^*)}{\partial\lambda^*}\right|_z \left(\left.\frac{\partial\lambda^*}{\partial z}\right|_k + \frac{\partial\lambda^*}{\partial k}\frac{\partial k}{\partial z}\right)$$

$$= \frac{1}{\beta}[t(\lambda^* - 0.5)^2 + (\alpha - w - 0.5t)] + \frac{2ztz}{\beta}(\lambda^* - 0.5)\left(\left.\frac{\partial\lambda^*}{\partial z}\right|_k + \frac{\partial\lambda^*}{\partial k}\frac{\partial k}{\partial z}\right).$$

When $0 \leq \tilde{\theta} \leq 0.25$, $\lambda^* = 0.5 + \sqrt{\theta/s\delta zt}$. Differentiating it with respect to z gives

(A9) $$\frac{\partial \lambda^*}{\partial z} = -\frac{1}{2z\sqrt{\tilde{\theta}}}\left[\tilde{\theta} + \frac{(\alpha - w - 0.5t)}{t}\right] < 0.$$

Substituting (A9) and $\lambda^* = 0.5 + \sqrt{\tilde{\theta}}$ into (A8), we obtain

$$\frac{d(Z_E^* + Z_W^*)}{dz} = \frac{2ztz}{\beta}(\lambda^* - 0.5)\frac{\partial \lambda^*}{\partial k}\frac{\partial k}{\partial z} < 0.$$

Proof of Propositions 3, 4, and 5

Omitted because it is given in the text.

Proof of Proposition 6

Differentiating $\lambda^* = 0.5 + \sqrt{\tilde{\theta}}$ with respect to *k*, we obtain:

(A10) $$\frac{\partial \lambda^*}{\partial k} = \frac{1}{2\sqrt{\tilde{\theta}}}\frac{1}{s\delta zt}\frac{2\beta^2\phi}{(2\alpha - 2w - 0.5t)} = \frac{\beta^2\phi}{(2\alpha - 2w - 0.5t)s\delta zt(\lambda^* - 0.5)}.$$

Differentiating (A10) with respect to z^s gives

(A11) $$\frac{d^2\lambda^*}{dkdz^s} = -\Phi\phi\left(\varepsilon_{z^s}^{z} + \varepsilon_{z^s}^{\lambda^*}\right)$$

where $\Phi = \frac{\beta^2}{(2\alpha - 2w - 0.5t)s\delta zt(\lambda^* - 0.5)}$. To show that $\left(\varepsilon_{z^s}^{z} + \varepsilon_{z^s}^{\lambda^*}\right) > 0$, note that

(A12) $$\varepsilon_{z^s}^{z} + \varepsilon_{z^s}^{\lambda^*} = \frac{\partial logz}{\partial logz^s} + \frac{\partial log(\lambda^* - 0.5)}{\partial logz^s}$$

$$= \frac{\partial logz}{\partial logz^s} + \frac{\partial log(\lambda^* - 0.5)}{\partial logz}\frac{\partial logz}{\partial logz^s} + \frac{\partial log(\lambda^* - 0.5)}{\partial logk}\frac{\partial logk}{\partial logz^s}$$

$$= \frac{\partial logz}{\partial logz^s}\left(1 + \frac{\partial log(\lambda^* - 0.5)}{\partial logz}\right) + \frac{\partial log(\lambda^* - 0.5)}{\partial logk}\frac{\partial logk}{\partial logz^s}.$$

Note that

$$\frac{\partial log(\lambda^* - 0.5)}{\partial logz} = \frac{0.5\partial(log\theta - logz)}{\partial logz} = \frac{-s\delta z(\alpha - w - 0.5t)}{2\theta} - 0.5.$$

We have

(A13) $$\varepsilon_{z^s}^{z} + \varepsilon_{z^s}^{\lambda^*} = \frac{\partial logz}{\partial logz^s}\left[0.5 - \frac{s\delta z(\alpha - w - 0.5t)}{2\theta}\right] + \frac{\partial log(\lambda^* - 0.5)}{\partial logk}\frac{\partial logk}{\partial logz^s}$$

$$= \frac{\partial logz}{\partial logz^s}\frac{1}{2\tilde{\theta}}\left[\tilde{\theta} - \frac{(\alpha - w - 0.5t)}{t}\right] + \frac{\partial log(\lambda^* - 0.5)}{\partial logk}\frac{\partial logk}{\partial logz^s}.$$

When firms partially agglomerate,

$$\tilde{\theta} \leq 0.25 < \frac{(\alpha - w - 0.5t)}{t}$$

where the last inequality comes from the assumption that $\alpha - w - 0.75t > 0$. Because both terms in (A13) are negative, we have

$$\varepsilon_{z^s}^{z} + \varepsilon_{z^s}^{\lambda^*} < 0.$$

To derive (31b), differentiating $\lambda^* = 0.5 + \sqrt{\tilde{\theta}}$ with respect to ϕ, we obtain:

(A14) $$\frac{\partial \lambda^*}{\partial \phi} = \frac{1}{2\sqrt{\tilde{\theta}}}\frac{1}{s\delta zt}\frac{2\beta^2\phi}{(2\alpha - 2w - 0.5t)} = \frac{\beta^2 k}{(2\alpha - 2w - 0.5t)s\delta zt(\lambda^* - 0.5)}.$$

Differentiating (A14) with respect to z^s give

(A15) $$\frac{d^2\lambda^*}{d\phi dz^s} = \Phi k\left(\varepsilon_{z^s}^{k} - \varepsilon_{z^s}^{z} - \varepsilon_{z^s}^{\lambda^*}\right).$$

Using (A14),

(A16) $$\varepsilon_{z^s}^{k} - \varepsilon_{z^s}^{z} - \varepsilon_{z^s}^{\lambda^*} = \frac{\partial logk}{\partial logz^s} - \frac{\partial logz}{\partial logz^s} - \frac{\partial \log(\lambda^* - 0.5)}{\partial logz}\frac{\partial logz}{\partial logz^s} - \frac{\partial log(\lambda^* - 0.5)}{\partial logk}\frac{\partial logk}{\partial logz^s}$$

$$= \left(1 - \frac{\partial log(\lambda^* - 0.5)}{\partial logk}\right)\frac{\partial logk}{\partial logz^s} - \left(1 + \frac{\partial log(\lambda^* - 0.5)}{\partial logz}\right)\frac{\partial logz}{\partial logz^s}$$

$$= \left[1 - \frac{\partial log(\lambda^* - 0.5)}{\partial logk}\right]\frac{\partial logk}{\partial logz^s} - \frac{1}{2\tilde{\theta}}\left[\tilde{\theta} - \frac{(\alpha - w - 0.5t)}{t}\right]\frac{\partial logz}{\partial logz^s}.$$

Note that

$$\frac{\partial log(\lambda^* - 0.5)}{\partial logk} = \frac{k}{2(k + \frac{\Psi}{2\beta^2\phi})} \geq 1 \text{ iff } k\phi \leq -\frac{\Psi}{2\beta^2}$$

where $\Psi = 1.5\beta t(\alpha - w - t) - s[\beta + \delta z(\alpha - w - 0.5t)](2\alpha - 2w - 0.5t)$. Thus, if $k\phi \leq -\frac{\Psi}{2\beta^2}$, both terms in (18) are non-negative, and $\frac{d^2\lambda^*}{d\phi dz^s} \geq 0$. If $k\phi > -\frac{\Psi}{2\beta^2}$, from (A16) $\frac{d^2\lambda^*}{d\phi dz^s} \leq 0$ if and only if

$$\frac{\partial logk}{\partial logz^s} \geq \frac{\frac{1}{2\widetilde{\theta}}\left[\widetilde{\theta}-\frac{(\alpha-w-0.5t)}{t}\right]}{\left[1-\frac{\partial \log(\lambda^*-0.5)}{\partial logk}\right]}\frac{\partial logz}{\partial logz^s} = \frac{\widetilde{\theta}-\frac{(\alpha-w-0.5t)}{t}}{2\widetilde{\theta}-\frac{2\beta^2 k}{(2\alpha-2w-0.5t)}}\frac{\partial logz}{\partial logz^s}$$

or

(A17) $$\frac{\partial logk}{\partial logz} \geq \frac{[t\widetilde{\theta}-(\alpha-w-0.5t)](2\alpha-2w-0.5t)}{2t[2\widetilde{\theta}(2\alpha-2w-0.5t)-\beta^2 k]} = \bar{\varepsilon}_{z^s}^{k}.$$

Proof of Proposition 7

We first derive the aggregate utility when firms price their products at the marginal costs. From (5)

(A18) $$V_E^j = \frac{\alpha^2}{2\beta} - \frac{\alpha}{\beta}[\lambda p_{EE} + (1-\lambda)p_{WE}] + \frac{1}{2\beta}[\lambda p_{EE}^2 + (1-\lambda)p_{WE}^2] - Z_E + Y_0 + Y_E^k.$$

When firms price their products at the marginal costs, we have

(A19) $$p_{EE} = w + \frac{(z+s)}{2}(0.5+\lambda) + \frac{(z-s)}{2}(1.5-\lambda) = w + z + s(\lambda - 0.5),$$

where the first term w is the marginal production cost, and the second and third terms following the first equality sign are the marginal pollution damage in the East and the West, respectively. The price of goods produced in the East and sold in the West equals $p_{EW} = (p_{EE} + t) = w + z + t + s(\lambda - 0.5)$. By symmetry, $p_{WW} = w + z - s(\lambda - 0.5)$, and $p_{WE} = w + z + t - s(\lambda - 0.5)$.

When firms price their products at the marginal social costs, the markup over variable and fixed costs for a firm in the East equals (see the proof of proposition 2):

(A20) $$Y_E = \frac{1}{\beta}[z + s(\lambda - 0.5)][2(\alpha - w - z - 0.5t) - (2s - t)(\lambda - 0.5)] - K.$$

The markup for a firm located in the West is found simply by replacing λ with $(1-\lambda)$ in (A20).[16]

The corresponding level of aggregate pollution and the difference in pollution concentrations are given by:

[16] Alternatively, entrepreneurs could also be compensated, through a lump-sum transfers, for the value of their human capital, which can be measured by the aggregate surplus from their products. Because the marginal utility of income is constant across consumers, transfers among consumers do not affect the aggregate utility. Therefore, the level of compensation to entrepreneurs does not affect the optimal firm distribution.

(A21) $$Z_E + Z_W = \frac{2z}{\beta}[(t-2s)(\lambda - 0.5)^2 + (\alpha - w - z - t)],$$

(A22) $$Z_E - Z_W = \frac{s}{\beta}(4\alpha + w + z + t - 2s)(\lambda - 0.5).$$

Substituting (A18)-(A20) into the aggregate utility function gives:

(A23) $$TV(\lambda) = 0.5V_E^L + \lambda V_E^C + 0.5V_W^L + (1-\lambda)V_W^C$$

$$= [(0.5+\lambda)V_{E0} + (1.5-\lambda)V_{W0}] + [1 + \lambda Y_E^C + (1-\lambda)Y_W^C] - TD(\lambda).$$

The terms in the first bracket represent utility from consumption, which equals:

(A24) $$[(0.5+\lambda)V_{E0} + (1.5-\lambda)V_{W0}] = [(V_{E0} + V_{W0}) + (\lambda - 0.5)[(V_{E0} - V_{W0}),$$

(A25) $$(V_{E0} + V_{W0}) = \frac{\alpha^2}{\beta} - \frac{\alpha}{\beta}[\lambda p_{EE} + (1-\lambda)p_{WE}] + \frac{1}{2\beta}[\lambda p_{EE}^2 + (1-\lambda)p_{WE}^2]$$

$$+\frac{\alpha^2}{\beta} - \frac{\alpha}{\beta}[(1-\lambda)p_{WW} + \lambda p_{EW}] + \frac{1}{2\beta}[(1-\lambda)p_{WW}^2 + \lambda p_{EW}^2]$$

$$= \frac{2\alpha^2}{\beta} - \frac{\alpha}{\beta}[2(\alpha + w) + 4s(\lambda - 0.5)^2 + t]$$

$$+\frac{1}{2\beta}[(\alpha + w + t)^2 + (\alpha + w)^2 + 4s(w + z - t + 0.5s)(\lambda - 0.5)^2]$$

$$= \frac{2\alpha^2}{\beta} - \frac{2\alpha}{\beta}[\alpha + w] + \frac{1}{2\beta}[(\alpha + w)^2 + (\alpha + w + t)^2]$$

$$+\frac{2s}{\beta}(w + z - t - 2\alpha + 0.5s)(\lambda - 0.5)^2.$$

(A26) $$(V_{E0} - V_{W0}) = \frac{\alpha^2}{\beta} - \frac{\alpha}{\beta}[\lambda p_{EE} + (1-\lambda)p_{WE}] + \frac{1}{2\beta}[\lambda p_{EE}^2 + (1-\lambda)p_{WE}^2]$$

$$-\left\{\frac{\alpha^2}{\beta} - \frac{\alpha}{\beta}[(1-\lambda)p_{WW} + \lambda p_{EW}] + \frac{1}{2\beta}[(1-\lambda)p_{WW}^2 + \lambda p_{EW}^2]\right\}$$

$$= -\frac{\alpha}{\beta}[-\lambda t + (1-\lambda)t] + \frac{1}{2\beta}[\lambda(p_{EE}^2 - p_{EW}^2) + (1-\lambda)(p_{WE}^2 - p_{WW}^2)]$$

$$= \frac{2\alpha t}{\beta}(\lambda - 0.5) - \frac{t}{\beta}(2\alpha + 2z + t + s)(\lambda - 0.5)$$

$$= -\frac{t}{\beta}(2\alpha + 2z + t + s - 2\alpha)(\lambda - 0.5).$$

Substituting (A25) and (A26) into (A24) gives

(A27) $[(0.5+\lambda)V_{E0}+(1.5-\lambda)V_{W0}] = [(V_{E0}+V_{W0})+(\lambda-0.5)[(V_{E0}-V_{W0})$

$$= \frac{2\alpha^2}{\beta} - \frac{2\alpha}{\beta}[\alpha+w] + \frac{1}{2\beta}[(\alpha+w)^2+(\alpha+w+t)^2]$$

$$+\frac{2s}{\beta}(w+z-t-2\alpha+0.5s)(\lambda-0.5)^2$$

$$-\frac{t}{\beta}(2\alpha+2z+t+s-2\alpha)(\lambda-0.5)^2.$$

The income for entrepreneurs in the East is

$$Y_E^C = (p_{EE}^o - w)q_{EE}^o(0.5+\lambda) + (p_{EW}^o - w - t)q_{EW}^o(1.5-\lambda),$$

where

$$p_{EE}^o = w + z + s(\lambda - 0.5),$$

$$p_{EW}^o = w + z + t + s(\lambda - 0.5),$$

$$(p_{EE}^o - w) = (p_{EW}^o - w - t) = z + s(\lambda - 0.5).$$

(A28) $Y_E^C = [z+s(\lambda-0.5)][(q_{EE}^o+q_{EW}^o)+(q_{EE}^o-q_{EW}^o)(\lambda-0.5)]$

$$= \frac{2}{\beta}[z+s(\lambda-0.5)][(\alpha-w-z-0.5t)-(s-0.5t)(\lambda-0.5)].$$

By symmetry,

(A29) $Y_W^C = \frac{2}{\beta}[z-s(\lambda-0.5)][(\alpha-w-z-0.5t)+(s-0.5t)(\lambda-0.5)].$

Note that

(A30) $(\mathrm{a}+\mathrm{b})(\mathrm{c}-\mathrm{d}) + (\mathrm{a}-\mathrm{b})(\mathrm{c}+\mathrm{d}) = 2(\mathrm{ac}-\mathrm{bd}),$

(A31) $(\mathrm{a}+\mathrm{b})(\mathrm{c}-\mathrm{d}) - (\mathrm{a}-\mathrm{b})(\mathrm{c}+\mathrm{d}) = 2(\mathrm{bc}-\mathrm{ad}).$

By (A30) and (A31),

(A32) $(Y_E^C + Y_W^C) = \frac{4}{\beta}[\mathrm{z}(\alpha-w-z-0.5t) - s(s-0.5t)(\lambda-0.5)^2],$

(A33) $(Y_E^C - Y_W^C) = \frac{4}{\beta}[s(\alpha-w-z-0.5t) - z(s-0.5t)](\lambda-0.5).$

Uing (A32) and (A33), we get:

(A34) $$[\lambda Y_E^C + (1-\lambda)Y_W^C] = 0.5[Y_E^C + Y_W^C] + [Y_E^C - Y_W^C](\lambda - 0.5)$$

$$= \frac{2}{\beta}[\mathrm{z}(\alpha - w - z - 0.5t) - s(s - 0.5t)(\lambda - 0.5)^2]$$

$$+\frac{4}{\beta}[s(\alpha - w - z - 0.5t) - z(s - 0.5t)](\lambda - 0.5)^2$$

$$= \frac{2}{\beta}\mathrm{z}(\alpha - w - z - 0.5t) + \frac{4}{\beta}[s(\alpha - w - z - 0.5t) - (z + 0.5s)(s - 0.5t)](\lambda - 0.5)^2$$

$$= \frac{2}{\beta}\mathrm{z}(\alpha - w - z - 0.5t) + \frac{4}{\beta}[s(\alpha - w - z - 0.75t - 0.5s) - z(s - 0.5t)](\lambda - 0.5)^2.$$

(A35) $$TD(\lambda) \equiv (0.5 + \lambda)Z_E + (1.5 - \lambda)Z_W$$

$$= 0.5(Z_E + Z_W) + (\lambda - 0.5)(Z_E - Z_W),$$

where

(A36) $$Z_E - Z_W = s\lambda(q_{EE}M_E + q_{EW}M_W) - s(1-\lambda)(q_{WE}M_E + q_{WW}M_W),$$

$$= s\{(\lambda - 0.5)[(q_{EE} + q_{WE})M_E + (q_{EW} + q_{WW})M_W)]$$

$$+0.5[(q_{EE} - q_{WE})M_E + (q_{EW} - q_{WW})M_W)]\}$$

$$= s\{(\lambda - 0.5)2(q_{EE} + q_{WE})$$

$$+0.5[(q_{EE} - q_{WE})M_E + (q_{EW} - q_{WW})M_W)]\}$$

$$= s\{(\lambda - 0.5)\frac{2}{\beta}(2\alpha - 2w - 2z - t)$$

$$+\frac{1}{2\beta}\{-[2s(\lambda - 0.5) - t]M_E - [2s(\lambda - 0.5) + t]M_W\}$$

$$= \frac{2s}{\beta}\{(\lambda - 0.5)(2\alpha - 2w - 2z - t)$$

$$-\frac{1}{2\beta}\{4s(\lambda - 0.5) - 2t(\lambda - 0.5)\}$$

$$= \frac{2s}{\beta}\{(2\alpha - 2w - 2z - t) - (s - 0.5t)\}(\lambda - 0.5)\}$$

$$= \frac{2s}{\beta}(2\alpha - 2w - 2z - 0.5t - s)(\lambda - 0.5).$$

(A37) $$Z_E + Z_W = z\lambda(q_{EE}M_E + q_{EW}M_W) + z(1-\lambda)(q_{WE}M_E + q_{WW}M_W),$$

$$= z\{(\lambda - 0.5)[(q_{EE} - q_{WE})M_E + (q_{EW} - q_{WW})M_W)]$$

$$+0.5[(q_{EE} + q_{WE})M_E + (q_{EW} + q_{WW})M_W)]\}$$

$$= z\{(\lambda - 0.5)\left[\left(-\tfrac{1}{\beta}[2s(\lambda - 0.5) - t]M_E - \tfrac{1}{\beta}[2s(\lambda - 0.5) + t]M_W\right)\right]$$

$$+\frac{1}{\beta}(2\alpha - 2w - 2z - t)\}$$

$$= \tfrac{z}{\beta}\{(\lambda - 0.5)[-4s(\lambda - 0.5) + 2t(\lambda - 0.5)] + (2\alpha - 2w - 2z - t)\}$$

$$= \tfrac{z}{\beta}\{2(t - 2s)(\lambda - 0.5)^2 + (2\alpha - 2w - 2z - t)\}.$$

Substituting (A36) and (A37) into (A35) gives

(A38) $$TD(\lambda) \equiv (0.5 + \lambda)Z_E + (1.5 - \lambda)Z_W$$

$$= 0.5(Z_E + Z_W) + (\lambda - 0.5)(Z_E - Z_W)$$

$$= \tfrac{z}{\beta}\{2(t - 2s)(\lambda - 0.5)^2 + (2\alpha - 2w - 2z - t)\}$$

$$+\tfrac{2s}{\beta}(2\alpha - 2w - 2z - 0.5t - s)(\lambda - 0.5)^2$$

$$= \tfrac{z}{\beta}(\alpha - w - z - 0.5t)$$

$$+\tfrac{2}{\beta}\{z(t - 2s) - s(2\alpha - 2w - 2z - 0.5t - s)\}\,(\lambda - 0.5)^2.$$

Substituting (A27), (A34) and (A30) into (A23), we obtain the aggregate utility function:

(A39) $$TV(\lambda) = [(0.5 + \lambda)V_{E0} + (1.5 - \lambda)V_{W0}] + [1 + \lambda Y_E^C + (1 - \lambda)Y_W^C] - TD(\lambda)$$

$$= \frac{2\alpha^2}{\beta} - \frac{2\alpha}{\beta}[\alpha + w] + \frac{1}{2\beta}[(\alpha + w)^2 + (\alpha + w + t)^2]$$

$$+\frac{2s}{\beta}(w + z - t - 2\alpha + 0.5s)(\lambda - 0.5)^2 - \frac{t}{\beta}(2\alpha + 2z + t + s - 2\alpha)(\lambda - 0.5)^2$$

$$+1 + \tfrac{2}{\beta}z(\alpha - w - z - 0.5t) + \tfrac{4}{\beta}[s(\alpha - w - z - 0.75t - 0.5s) - z(s - 0.5t)](\lambda - 0.5)^2$$

$$-\frac{z}{\beta}(\alpha - w - z - 0.5t) - \frac{2}{\beta}\{z(t-2s) - s(2\alpha - 2w - 2z - 0.5t - s)\}\,(\lambda - 0.5)^2$$

$$= \frac{2\alpha^2}{\beta} - \frac{2\alpha}{\beta}[\alpha + w] + \frac{1}{2\beta}[(\alpha + w)^2 + (\alpha + w + t)^2] + 1$$

$$+\frac{2z}{\beta}(\alpha - w - z - 0.5t) - \frac{z}{\beta}(\alpha - w - z - 0.5t)$$

$$+\frac{2}{\beta}\,\{s(w + z - t - 2\alpha + 0.5s) - t(\alpha + z + 0.5t + 0.5s - \alpha)$$

$$+[2s(\alpha - w - z - 0.75t - 0.5s) - 2z(s - 0.5t)]$$

$$-\{z(t-2s) - s(2\alpha - 2w - 2z - 0.5t - s)\}\}(\lambda - 0.5)^2$$

$$= \frac{1}{\beta}\{2\alpha^2 - \alpha(\alpha + w) + 0.5[(\alpha + w)^2 + (\alpha + w + t)^2] + \beta + z(\alpha - w - z - 0.5t)$$

$$+\frac{2}{\beta}\,[s(-w - z - 1.5t - 0.5s) - t(\alpha + z + 0.5t + 0.5s - \alpha)](\lambda - 0.5)^2.$$

Thus, the aggregate utility function can be written as

(A40) $$TV(\lambda) = E + \varphi(\lambda - 0.5)^2,$$

where

$$E = \frac{1}{\beta}\{2\alpha^2 - \alpha(\alpha + w) + \frac{1}{8}[(\alpha + w)^2 + (\alpha + w + t)^2] + \beta + z(\alpha - w - z - 0.5t),$$

$$\varphi = \frac{1}{\beta}\,[0.5t(\alpha - w - 0.5t - 4z) + 2s(4\alpha - 4w - 4z - 2t - 2s)].$$

Because

(A41) $$\varphi \lesseqgtr 0 \text{ iff } z \gtreqless \bar{z} \equiv \frac{0.5t(\alpha + w - 0.5t) - 8s(\alpha - w - 0.5t - 0.5s)}{2t + 8s},$$

we have

(A42) $$\lambda^o = \begin{cases} 1 & if z < \bar{z} \\ 0.5 & if z > \bar{z} \\ any\ \lambda \in [0.1] & if z < \bar{z} \end{cases}.$$

APPENDIX B

Table B1: The effect of an emissions tax on firm and pollution concentration and total pollution damage: simulation results

F^*	Effect of Reduction in z, for a given marginal production cost	Total effect of increase in marginal production cost, for a given z		Overall effect of tax increase	
	$\left.\frac{\partial F^*}{\partial z}\right\|_w \frac{\partial z^*}{\partial \tau}$	$\left.\frac{\partial F^*}{\partial w}\right\|_z \frac{\partial w}{\partial \tau}$		$\frac{dF^*}{d\tau} = \left\{\left.\frac{\partial F^*}{\partial z}\right\|_w \frac{\partial z^*}{\partial \tau}\right\} + \left\{\left.\frac{\partial F^*}{\partial w}\right\|_z \frac{\partial w}{\partial \tau}\right\}$	
		$z = 1, \delta = 10$	$z = 0.01, \delta = 1000$	$z = 1, \delta = 10$	$z = 0.01, \delta = 1000$
Agglomeration λ^*	(+)	(+)	(+)	(+)	(+)
Aggregate Pollution $(z_E^* + z_W^*)$	(+)	(-)	(-)	(-) if $\frac{\partial z^*}{\partial \tau} = 0$	(-) if $\frac{\partial z^*}{\partial \tau} = 0$
Pollution Concentration $(z_E^* - z_W^*)$	(+)	(+)	(+)	(+)	(+)
Total Damages TD^*	(+)	(-)	(+)	(-) if $\frac{\partial z^*}{\partial \tau} = 0$	(+)

Notes: Other parameter values assumed in the numerical example are $\alpha = 1, \beta = 1, t = 0.25$, and s=0.01.