Taxing Pollution: Agglomeration and Welfare Consequences

Marcus Berliant, Washington University in St. Louis

Shin-Kun Peng, Academia Sinica and National Taiwan University

Ping Wang, Washington University in St. Louis and NBER

March 2013

Abstract: This paper demonstrates that a pollution tax with a fixed cost component may lead, by itself, to stratification between clean and dirty firms without heterogeneous preferences or increasing returns. We construct a simple model with two locations and two industries (clean and dirty) where pollution is a by-product of dirty good manufacturing. Under proper assumptions, a completely stratified configuration with all dirty firms clustering in one city emerges as the only equilibrium outcome when there is a fixed cost component of the pollution tax. Moreover, a stratified Pareto optimum can never be supported by a competitive spatial equilibrium with a linear pollution tax that encompasses Pigouvian taxation as a special case. To support such a stratified Pareto optimum, however, an effective but unconventional policy prescription is to redistribute the pollution tax revenue from the dirty to the clean city residents.

JEL Classification: D62, H23, R13.

Keywords: Pollution Tax, Agglomeration of Polluting Producers, Endogenous Stratification.

Acknowledgments: We are grateful for suggestions from an anonymous referee, Sören Blomquist, John Conley, Katherine Cuff, Hideo Konishi, Ailsa Röell, Steve Slutsky, John Weymark, and Jay Wilson, as well as participants at the Public Economic Theory Conference in Taipei, the Regional Science Association International Meetings in Miami, and the Taxation Theory Conference at Vanderbilt University. Financial support from Academia Sinica and the National Science Council of Taiwan (grant no. NSC99-2410-H-001-013-MY3) that enabled this international collaboration by the Three Mouseketeers is gratefully acknowledged. Needless to say, the usual disclaimer applies.

Correspondence: Marcus Berliant, Department of Economics, Washington University in St. Louis, Campus Box 1208, 1 Brookings Drive, St. Louis, MO 63130-4899, U.S.A.; 314-935-8486 (phone); 314-935-4156 (fax); berliant@artsci.wustl.edu (e-mail).

"If you visit American city,
You will find it very pretty.
Just two things of which you must beware:
Don't drink the water and don't breathe the air."
(Tom Lehrer, *Pollution*)

1 Introduction

It is evident that the development of many local economies has featured adjacent but separate clean and dirty cities. Examples of such pairs include Seattle/Tacoma and San Francisco/Oakland with larger clean cities, Ann Arbor/Detroit and Aurora/Denver with smaller clean cities, as well as Washington, D.C./Baltimore with comparably sized clean and dirty cities. A natural question arises: Why are dirty firms clustered in one location and why is such an outcome sustainable over time? Certainly one might address the question with heterogeneity in preferences or increasing returns in production, for example internal increasing returns of the type used in the New Economic Geography literature; see the recent paper by Picard and Tabuchi (2010) and papers cited therein.[1] Our paper proposes an alternative: a pollution tax with a fixed cost tax component may, by itself, lead to stratification between clean and dirty firms without heterogeneous preferences or increasing returns.

Since 1972, the OECD has adopted the polluter pays principle, trying to internalize environmental costs based on the idea first advanced by Pigou (1920). More recently, the OECD (1994) categorized three types of pollution taxes: (i) a proportional tax on the actual pollution output, for example according to the amount of emission; (ii) a proportional tax on a proxy for pollution output, for example according to water consumption, electricity usage or each unit of product when the production process harms the environment; and (iii) a fixed cost tax levied on each company or each household. In this paper, we consider all three types. Whereas a fixed cost tax levied on each firm is considered, the proportional Pigouvian tax is generalized to a linear tax that includes a fixed cost tax component as proposed by Carlton and Loury (1980).[2]

[1] See Porter (1990) for a comprehensive discussion of industrial clustering from a business strategy viewpoint. Our paper is also related to the locational stratification literature, where stratification is caused by human capital (cf. Benabou 1996a,b, Chen, Peng and Wang 2009), local public goods (cf. Nachyba 1997 and Peng and Wang 2005), and the environment (cf. Chen, Huang and Wang 2012).

[2] See also Baumol (1972) and Buchanan and Tullock (1975) on direct control versus taxation, and Chipman and Tian (2012) on markets for rights to pollute in an aspatial context. While there is an existing literature on the welfare consequences of pollution taxation (see citations in Section 7 below), none explores the implication of pollution

In brief, the purpose of this paper is as follows. We construct a simple model with two locations and two industries (clean and dirty) where pollution is a by-product of dirty good manufacturing, and dirty good manufacturing is subject to agglomeration externalities with decreasing private returns to scale. We could obtain our results without agglomeration externalities, but they help simplify the analysis and calculations, as we shall explain below. Next, we establish conditions under which a completely stratified configuration with all dirty firms clustering in one city emerges as the only equilibrium outcome when there is a fixed cost component of the pollution tax. Finally, we show that a stratified Pareto optimum can never be supported by a competitive spatial equilibrium under a linear pollution tax without redistributing the pollution tax revenue from the dirty to the clean city residents.

Regarding our examples of pairs of clean and dirty cities in the US, it is important to point out from where, in our view, the fixed component of a pollution tax arises. As discussed by Karp (2005, pp. 229-230), if firms pay a unit tax based on the aggregate level of pollution, in technical terms an "ambient tax," but if they think that they are so small that they have no effect on the aggregate level of pollution, then they view the tax as a fixed cost.[3] Of course, such an ambient tax may vary by location, as aggregate pollution is generally location-specific. Such a tax is essentially an aggregate amount of required revenue (the tax rate multiplied by total pollution) divided up among firms in the local polluting industry, which is exactly the way we model it.[4]

Our main result establishes that taxing pollution with a fixed component independent of dirty good output can cause firm agglomeration (a variable tax component on top of the fixed component is permitted). The key argument is as follows. At a symmetric, integrated equilibrium, wages equalize both across sectors and locations. Then, in the presence of a fixed total pollution damage payment in each polluted region, dirty factories may not have sufficient profitability to pay the tax and thus no integrated equilibrium exists. Now if dirty firms cluster, as they do in a stratified equilibrium, then they share the fixed pollution tax in the one region where they cluster, implying higher net of tax profit. Moreover, wage equalization between the two locations is no longer required in equilibrium because clean and dirty firms are in two different locations, so there is no wage equalization even

taxation for production agglomeration, in particular when pollution is local (so location is relevant) and agents are mobile. The point of Carlton and Loury (1980) is different, in that they are concerned with firm entry and exit. We do not consider that in our model.

[3]Karp (2005) goes on to consider the case where firms are large so each has an effect on the aggregate level of pollution.

[4]A natural alternative model, that does not capture this idea, is to use a fixed lump-sum tax for each firm entering the local market. This would clearly not be: an ambient tax with a large number of firms.

across sectors. All we need is utility equalization, which only requires that pollution disutility balance with the wage differential. This is consistent with firm profitability under stratification.

The important part of the argument is as follows. Both the fixed component of the pollution tax and decreasing private returns are needed for this result, as indicated above. The fixed component of the pollution tax rules out integrated equilibrium. With decreasing private returns that permit positive rent for dirty firms, agglomeration ensures enough profitability of dirty firms when fully clustered to allow existence of stratified equilibrium with the tax. There is a positive feedback loop: With a variable tax component, more pollution in a region implies more tax revenue that attracts more worker/consumers (who receive the revenue), thus depressing the wage. A potential offsetting factor is that the wage must be higher in the region to compensate for disutility due to more pollution. In the end, the equilibrium configuration is a function of the parameters.

The key difference between this work and the classical literature on Pigouvian taxation is: We assume that there is a local government in each region that must balance its own budget. We take the tax system of each local government to be *exogenous* and uniform, with no tax competition.[5] *The taxes could be set by a higher level of government.* But the revenues stay local. For example, the tax revenues could pass through a higher level of government and be returned to the local government in some form such as funding for a local public good. In other words, we are making an important distinction between the authority that sets the tax on pollution, and the recipients of the revenue.

There are 3 related potential distortions in our framework: a negative pollution externality from dirty firm production imposed on consumers, a positive local agglomeration externality for polluting firms, and a migration incentive for consumers induced by the tax and redistribution schedules in the two regions. Regarding the last distortion, local tax revenue and local profits are distributed back to the residents of that location only. The setting would be classical if there were only one national government with the power to tax differentially and redistribute to consumers independent of region of residence. In that case, the standard welfare theorems would go through under Pigouvian taxes, since correction for the pollution externality and migration incentives (the first and third distortions) can be made in the usual way, whereas the fixed cost component of the firm tax/transfer system can account for the agglomeration externality. However, with independent regional government taxation as in our setting, equilibrium allocations might not be Pareto optimal unless transfers between the regional governments are made so that the regional governments can

[5] The reader is referred to Markusen, Morey, and Oleviler (1995) for modeling fiscal competition in pollution taxes with firms choosing the location of plants.

mimic a national government.

Turning next to a detailed description of our model, to illustrate the possibility that a pollution tax causes agglomeration of dirty firms, we construct a simple model featuring two industries, clean and dirty. Both industries use homogeneous labor as inputs. Whereas the clean service production is Ricardian (constant returns), dirty manufactured good production is socially constant-returns-to-scale and privately diminishing returns with positive spillovers of the Romer type. Pollution is a by-product of dirty good manufacturing. To eliminate unnecessary complications associated with a wealth effect, utility is assumed to be quasi-linear, linear in clean good consumption and pollution but strictly increasing and strictly concave in dirty good consumption. The pollution tax schedule features a fixed cost tax component that is independent of pollution (or dirty good output) and may also contain a marginal tax component that is proportional to dirty good output.

We establish that under proper assumptions, a completely stratified equilibrium with all dirty firms clustered in one city is supported and such a stratified equilibrium cannot emerge in the absence of the fixed payment pollution tax. In some circumstances, an integrated equilibrium is impossible, but a stratified equilibrium exists. Under suitable conditions, we show that *the presence of pollution and a pollution tax with a fixed cost tax component, rather than the Romer-type positive spillovers, are necessary for agglomeration of dirty firms.* Our main findings are robust, and remain valid even when: (i) quasi-linearity of utility is abandoned, or (ii) allowing producers to choose between clean and dirty good technologies. We have assumptions on the model's reduced form that will generate either integrated or stratified equilibrium. We do not push them back to primitives, as there are many exogenous parameters and thus many combinations that will work for each type of configuration. However, in Remarks 6 and 7 below, we fix all but 2 parameters and provide a description of parameter ranges where the respective equilibrium configurations arise.

Next we turn to the examination of Pareto optima. Depending on exogenous parameter values, both integrated and stratified configurations can arise as optima. Whereas an integrated Pareto optimum can be supported by a competitive spatial equilibrium with a linear pollution tax, a stratified Pareto optimum cannot. Specifically, regardless of the linear pollution tax schedule, a stratified equilibrium is always over-polluted compared to the optimum. To support the stratified Pareto optimum, one must redistribute pollution tax revenues from the dirty to the clean city residents. This suggests a new instrument to rectify competitive equilibrium inefficiency when there is pollution generated by dirty good production.[6]

[6] We wish to emphasize that in this paper, we consider only equilibrium or optimal configurations that are completely stratified in terms of production or that are completely integrated in that production is symmetric across locations.

The remainder of the paper is organized as follows. Section 2 contains the notation and basic model. Section 3 provides first order necessary conditions for equilibrium. Section 4 analyzes the two types of equilibria we consider here, namely integrated and stratified. Section 5 analyzes the conditions on parameters that generate each of these types of equilibria. Section 6 gives further results, particularly about stability of equilibrium, that can be derived with specific functional forms, namely an example. Section 7 discusses Pareto optima and the welfare theorems, whereas section 8 concludes.

2 The Model

Consider a local economy consisting of two regions/cities ($i = A, B$) and two sectors (a clean/service good X and a dirty/manufactured good Y). Each region has an abundant supply of land of density one in a featureless landscape. Land is omitted from the benchmark model for tractability reasons, so the model looks more like one of coalition formation than of an urban economy. Goods are freely mobile and there is no cost to transport any commodity between regions. Throughout the paper, the clean good is taken as the numéraire.

This local economy is populated with three groups of active agents: (i) a continuum of households of a fixed mass one, who are all both consumers and workers; (ii) a continuum of clean (non-polluting) firms of mass one, and (iii) a continuum of dirty (polluting) firms of mass $M > 0$. All households are *freely mobile* between the two regions, but once a household has chosen a residential location, it *cannot commute* between the two regions. This latter assumption is equivalent to assuming that the commuting cost between two regions is sufficiently high. Such an assumption is justifiable when the two regions are sufficiently far apart: for many clean-dirty city pairs in the real world such as Ann Arbor-Detroit and Seattle-Tacoma, the fraction of people commuting between cities is essentially negligible. We will discuss in Section 5.1 (see Remark 4) what happens if workers are allowed to commute between the two regions.

In addition to the three groups of active agents, there is a local government ruling each region, whose only activity is to collect pollution taxes/fees for redistribution to consumers. To close the economy, we shall assume that dirty firms in a particular region are owned by consumers in the same region.

We relegate the discussion of other possible configurations to the concluding section.

2.1 Firms

The clean good is produced with labor input under a Ricardian technology,

$$x^i(j) = \psi \cdot n_x^i(j), i = A, B, j \in [0, k^i] \tag{1}$$

where $x^i(j)$ denotes the output of clean firm j in location i, $\psi > 0$ is the inverse of the unit labor requirement for clean good production, $n_x^i(j)$ represents clean firm j's demand for labor, and $k^i \in [0, 1]$ denotes the mass of clean firms in region i. The total local supply of the clean good in region i is given by $X^i = \int_0^{k^i} x^i(j)dj$ and the total local clean industry employees in the region i can be specified as:

$$N_x^i = \int_0^{k^i} n_x^i(j)dj \quad i \in A, B \tag{2}$$

Under *ex post* symmetry of firms in a region, imposed throughout, we have $N_x^i = k^i n_x^i$.

Denote by m^i the mass of dirty firms in region i, by $n_y^i(j)$ the labor demand by a dirty firm j in region i, and by N_y^i the total local dirty industry employees in region i, where:

$$N_y^i = \int_0^{m^i} n_y^i(j)dj \quad i \in A, B \tag{3}$$

Each dirty good firm employs labor as the sole private input under a privately decreasing-returns-to-scale and socially constant-returns-to-scale production technology $\widetilde{f}$:

$$y^i(j) = \widetilde{f}\left(n_y^i(j), N_y^i\right) = N_y^i f\left(\frac{n_y^i(j)}{N_y^i}\right), i \in A, B, j \in [0, m^i] \tag{4}$$

where $y^i(j)$ is the output of dirty firm j in region i. We assume that $\widetilde{f}$ is strictly increasing and strictly concave in each argument, satisfying the boundary condition $\widetilde{f}\left(0, N_y^i\right) = 0$ and the Inada conditions $\lim_{n_y^i(j)\to 0} \frac{\partial \widetilde{f}\left(n_y^i(j), N_y^i\right)}{\partial n_y^i(j)} = \infty$ and $\lim_{n_y^i(j)\to \infty} \frac{\partial \widetilde{f}\left(n_y^i(j), N_y^i\right)}{\partial n_y^i(j)} = 0$. Under social constant returns, we can divide firm output by the total number of local dirty industry employees to obtain f, where the properties of $\widetilde{f}$ imply that f is strictly increasing and strictly concave in the fraction of firm employees in the local dirty industry. The incorporation of N_y^i into a dirty firm's production function captures positive spillovers of the Romer (1986) type, where N_y^i is a positive measure of small firms, and where each firm is of measure zero. Under an *ex post* symmetric equilibrium, $N_y^i = m^i n_y^i$. The presence of uncompensated positive externalities provides an agglomeration force for dirty firms. Nonetheless, we will show in Sections 5.1 and 6.1 that the presence of pollution and a pollution tax with a fixed cost tax component, rather than the Romer-type positive spillovers, are necessary for agglomeration of dirty firms.

Both goods (clean and dirty) are traded and freely mobile. Let p denote the global relative price of the dirty good. Further denote the wage rate prevailing in region i as w^i. Let the region-specific pollution tax in region i be τ^i (to be specified later), where $\frac{\tau^i}{p}$ represents a typical *ad valorem* tax.[7] Each dirty firm in region i chooses labor demand to maximize its profit; its optimization problem is then given by:

$$\pi^i(j) = \max_{n_y^i} p\left[N_y^i f\left(\frac{n_y^i}{N_y^i}\right) - \tau^i\right] - w^i n_y^i(j) \tag{5}$$

The aggregate output of the dirty good in region i is $Y^i = \int_0^{m^i} y^i(j)dj$.

2.2 Households

Each household values the consumption of the clean good and the dirty good but suffers disutility from pollution. Each household is endowed with one unit of labor. Since a household does not value leisure, the entire one unit of labor is supplied inelastically. Let Q^i measure the level of pollution in region i. Following conventional wisdom, we assume that pollution is a by-product of the production of dirty goods, taking a simple linear form:

$$Q^i = \theta Y^i = \theta \int_0^{m^i} y^i(j)dj \tag{6}$$

where $\theta > 0$. The utility of a household residing in region i takes a quasi-linear form:

$$U^i = c_x^i - \gamma \cdot Q^i + u(c_y^i) \tag{7}$$

This utility function is quasi-linear in the spirit of Bergstrom and Cornes (1983): linear in clean good consumption c_x and total pollution Q, but nonlinear in c_y, as $u(c_y)$ is the utility obtained from consuming the dirty good. It is strictly increasing and strictly concave, satisfying the boundary condition $u(0) = 0$ and the Inada conditions $\lim_{c_y^i \to 0} u'(c_y^i) = \infty$ and $\lim_{c_y^i \to \infty} u'(c_y^i) = 0$.

The household's budget constraint in region i is simply specified as follows:

$$c_x^i + pc_y^i = w^i + z^i \tag{8}$$

where z^i represents the sum of government rebates (of pollution tax collection) and firm profit redistribution in region i:

$$z^i = \frac{1}{N^i}\int_0^{m^i} \left[\pi^i(j) + p\tau^i\right] dj, \quad i \in A, B \tag{9}$$

[7] The pollution tax schedule is written in this form for analytical convenience.

Quasi-linear preferences imply that, by substituting in the budget constraint (8), household's utility can be rewritten as:

$$U^i = \left(w^i + z^i\right) - \gamma \cdot Q^i + \left[u(c_y^i) - pc_y^i\right]$$

which is income net of pollution disutility plus the consumer surplus derived from consuming the dirty good. Thus, household's optimization reduces to one variable: maximization of the consumer surplus from dirty good consumption, which simplifies the analysis greatly. We will discuss in Section 5.1 (see Remark 3) what happens if the utility of the clean good is strictly concave.

2.3 The Local Government

The pollution tax levies on the dirty firm are given as follows:

$$\tau^i = \begin{cases} 0, & \text{if } y^i(j) = 0, \forall \ \ j \\ g^i(y^i(j), Y^i), & \text{otherwise} \end{cases}$$

When pollution is nondegenerate, we shall consider two specific regimes of interest, namely, a fixed pollution tax regime and a linear pollution tax regime:[8]

$$g^i = \begin{cases} F/m^i, \text{ under fixed pollution tax regime} \\ L + ty^i(j), \text{ under linear pollution tax regime} \end{cases}$$

Under the fixed pollution tax regime, a fixed levy $F > 0$ is imposed on region i so that each firm pays an equal share $\frac{F}{m^i}$; under the linear pollution tax regime, in addition to a lump-sum tax $L > 0$, a marginal tax $t > 0$ is imposed on firm output y^i. Whereas the former can best illustrate the role of pollution taxation played in firm agglomeration, the latter is important because it encompasses Pigouvian taxation as a special case and allows practical welfare analysis. For notational convenience, we shall denote generally the marginal tax rate as:

$$\zeta \equiv \frac{\partial \tau^i}{\partial y^i(j)} = \begin{cases} 0, \text{ under fixed pollution tax regime} \\ t, \text{ under linear pollution tax regime} \end{cases}$$

3 Optimization and Equilibrium

We are now prepared to derive individual optimizing conditions and to specify market clearing conditions.

[8] This functional form also covers the first type of tax mentioned in the introduction because pollution is proportional to output via θ.

3.1 Optimization

The first-order condition for profit maximization of each clean and dirty firm is, respectively, given by:

$$\psi = w^i \tag{10}$$

$$VMP_L \equiv p(1-\zeta)MPL_y^i = p(1-\zeta)f'\left(\frac{n_y^i}{N_y^i}\right) = w^i \tag{11}$$

where VMP_L denotes the value of the marginal product of labor (or marginal revenue product) and MPL denotes the marginal product of labor. Denote the dirty firm's surplus accrued from uncompensated spillovers as:

$$\sigma(e) \equiv f(e) - (1-\zeta)ef'(e)$$

where $e \equiv \frac{n_y^i}{N_y^i}$. It is convenient to denote the dirty firm's surplus excluding pollution tax as $\widetilde{\sigma}(e) = f(e) - ef'(e)$. Given our assumptions on the production function for dirty firms, both $\sigma(e)$ and $\widetilde{\sigma}(e)$ are strictly increasing in e. Substituting the *ex post* symmetry condition, $N_y^i = m^i n_y^i$ as well as (11) and (3) into (5) yields the profit for every firm j in region i:

$$\pi^i(j) = \pi^i = p\left[n_y^i m^i \sigma\left(1/m^i\right) - \tau^i\right] \tag{12}$$

The lump-sum distribution to each household follows immediately:

$$z^i = \frac{(m^i)^2}{N^i} \cdot \sigma\left(1/m^i\right) \cdot p n_y^i, \; \forall\; \tau^i \tag{13}$$

The household's optimization problem can be written more simply in two steps, solving backward. In the second step, households choose their best consumption bundle subject to their budget in each region. In the first step, they choose their region of residence.

Beginning with the second step, each household residing in region i maximizes their utility subject to the budget constraint by choosing c_y^i:

$$\max_{c_y^i} \; w^i + z^i - pc_y^i - \gamma Q^i + u\left(c_y^i\right) \tag{14}$$

The first-order condition of (14) with respect to c_y^i is given by:

$$u'\left(c_y^i\right) = p \tag{15}$$

It is immediate that, since the relative price of the dirty good across the two regions is one, the consumption of the dirty good in the two regions must be identical too. From the budget constraint (8) and (15), we then solve the clean good consumption as:

$$c_x^i = w^i + z^i - pc_y^i = w^i + z^i - c_y^i u'\left(c_y^i\right) \tag{16}$$

Substituting (11), (13), and (15) into (16), we have the consumption of the clean good in region i as:

$$c_x^i = u'\left(c_y^i\right)\left[(1-\zeta)f'\left(1/m^i\right) + \frac{(m^i)^2}{N^i}\sigma\left(1/m^i\right)n_y^i - c_y^i\right] \tag{17}$$

In the first step, the household's residential location can be determined by:

$$i = \arg\max_i U^i \tag{18}$$

3.2 Market Clearance

Denote region i's labor supply as N^i and recall that total labor supply is normalized to one (the total measure of consumers). The regional and overall labor market clearing conditions are thus:

$$N_x^i + N_y^i = N^i \tag{19}$$

$$N^A + N^B = 1 \tag{20}$$

Moreover, goods market clearing conditions are:

$$\sum_{i=A,B} N^i c_x^i = \sum_{i=A,B} \int_0^{k^i} x^i(j)dj = X^A + X^B \tag{21}$$

$$\sum_{i=A,B} N^i c_y^i = \sum_{i=A,B} \int_0^{m^i} y^i(j)dj = Y^A + Y^B \tag{22}$$

By symmetry, we have:

$$\sum_{i=A,B} N^i c_x^i = \sum_{i=A,B} k^i x^i = X^A + X^B \tag{23}$$

$$\sum_{i=A,B} N^i c_y^i = \sum_{i=A,B} m^i y^i = Y^A + Y^B \tag{24}$$

where $m^A + m^B = M$.

Finally, if both locations are occupied, locational equilibrium requires:

$$U^A = U^B \tag{25}$$

4 Equilibrium Configuration

A *competitive spatial equilibrium* is a tuple of quantities, $\{n_x^i(j), n_y^i(j), N_x^i, N_y^i, N^i, k^i, m^i, c_x^i, c_y^i, x^i(j), y^i(j), Q^i\}$, and prices, $\{w^i, p\}$, such that: (i) all households and firms optimize; (ii) labor markets clear; (iii) goods markets clear; (iv) the population identity holds; and (v) the locational equilibrium condition is met.[9] Among all possible equilibrium configurations, we are particularly interested in

[9] The equilibrium concept is based on the multi-class equilibrium concept constructed by Hartwick, Schweizer and Varaiya (1976).

two equilibria: The first type is an *integrated equilibrium* where all clean and dirty firms are spread symmetrically over the two regions so that both types of firms are completely integrated locationally. The second type is a *stratified equilibrium* where all dirty manufacturing firms agglomerate in one region (without loss of generality, let it be region A) and all clean service firms are located in region B (where workers face better environmental conditions). In order to compare the endogenous variables obtained under the two types of equilibria, we shall use arguments I and S to denote integrated and stratified patterns, respectively.

4.1 Case I: Integrated Equilibrium

In an integrated equilibrium, both firms and households are symmetrically distributed across the two regions. Thus, we have:

$$N_x^A = N_x^B, N_y^A = N_y^B, N^A = N^B = \frac{1}{2}$$
$$k^i = k = \frac{1}{2}, m^i = m = \frac{M}{2},$$
$$n_x^i(j) = n_x = \frac{N_x}{k}, n_y^i(j) = n_y = \frac{N_y}{m}$$
$$n_x + Mn_y = 1$$

Moreover, wages must be equalized between the clean and the dirty sectors in each region. From (10) and (11), we can thus depict in Figure 1 the labor allocation between clean and dirty sectors under the integrated equilibrium.

[Insert Figure 1 here]

Figure 1 illustrates that dirty firms' labor demand, which is a downward-sloping function of n_y^i/N_y^i, is determined by wage equalization between the clean and the dirty sectors (see point $\mathrm{E^I}$), namely where:

$$p(1-\zeta)MPL_y = p(1-\zeta)f'\left(2/M\right) = w = \psi \tag{26}$$

which determines the relative price of the dirty good as a decreasing function of the mass of dirty firms. The Inada conditions assumed are sufficient for the existence of an interior level of dirty industry employment and production.

Under symmetry, a dirty firm's output is now given by, $y = f\left(1/m^i\right) m^i n_y^i = \frac{M}{2} f\left(2/M\right) n_y$. From the dirty good market clearing condition, $c_y^i = c_y = My$, so we have:

$$c_y = M \cdot y = \frac{M^2}{2} f\left(2/M\right) n_y \tag{27}$$

This dirty good market clearing condition enables us to express the dirty good demand as a linear, upward-sloping function of the induced demand for labor starting from the origin, which is referred to as the *dirty good market-clearing* (DM) locus (see Figure 2). Moreover, we can combine (26) and (15), yielding the *dirty good optimization* (DO) locus:

$$u'(c_y) = \frac{\psi}{(1-\zeta)f'(2/M)} \tag{28}$$

Thus, the demand for the dirty good is independent of the induced demand for labor.

[Insert Figure 2 here]

As depicted in Figure 2, one can see that the integrated equilibrium quantity of the dirty good and employment are jointly determined at point E^{I}.

Clean good market clearance implies:

$$c_x = x = \psi n_x(I) = \psi[1 - M \cdot n_y(I)]$$

One may easily check that one of (8), (27) and the above equation are redundant, i.e., Walras' law is verified. Substituting the equilibrium $n_y(I)$ and (28) into (12), we have:

$$\pi(I) = \frac{M}{2}\frac{\psi}{(1-\zeta)f'(2/M)}\left[\sigma\left(2/M\right)n_y(I) - \left(2/M\right)\tau^i\right] \tag{29}$$

Finally, locational equilibrium (25) in this case is trivial. See Table 1 for a summary of the values of the endogenous variables at equilibrium.

4.2 Case II: Stratified Equilibrium

Now, we move to examine stratified equilibrium. At a stratified equilibrium, assume that the dirty firms agglomerate in region A, and the clean firms agglomerate in region B. Then stratified equilibrium is as shown in Table 1, and we have:

$$k^A = 0, k^B = 1, m^A = M, m^B = 0, \pi^B = z^B = 0$$

$$N^A = Mn_y, N^B = n_x, n_x + Mn_y = 1$$

Thus, we obtain the dirty good production for each dirty firm in region A as: $y = f\left(\frac{1}{m^A}\right)m^A n_y^A = Mf\left(1/M\right)n_y$. In this case, wages need not be equalized between the two regions: those residing in the dirty region receive a higher wage but suffer from pollution. The utility levels of workers in the two regions are equal. The wages in the regions A and B are $w^A = p(1-\zeta)f'(1/M)$ and $w^B = \psi$, respectively. The dirty good market clearing condition implies:

$$c_y = My = M^2 f(1/M)\, n_y \tag{30}$$

which can be combined with (15) to yield:

$$p = u'\left[M^2 f(1/M)\, n_y\right] \tag{31}$$

By diminishing marginal utility, the above expression entails a negative relationship between dirty good price and employment (see the bottom panel of Figure 3). From the clean good market clearing condition, one obtains: $N^A c_x^A + N^B c_x^B = x = \psi n_x$, or, using (8) and Table 1,

$$c_x(S) = x = \psi n_x = \psi(1 - Mn_y)$$

which can again be used with (8) and (30) to verify Walras' law.

Next, we can rewrite (12) under stratified equilibrium as:

$$\pi(S) = Mu'\left[f(1/M)\, M^2 n_y(S)\right]\left[\sigma(1/M)\, n_y(S) - (1/M)\,\tau^i\right] \tag{32}$$

The equilibrium level of pollution in region A is given by: $Q^A = \theta Y^A = \theta My = \theta M^2 f(1/M)\, n_y(S)$. We can derive the utility level attained by households residing in region A as:

$$U^A = Mf(1/M)\left\{u'\left[f(1/M)\, M^2 n_y(S)\right]\left[1 - Mn_y(S)\right] - \gamma\theta Mn_y(S)\right\} + u\left[f(1/M)\, M^2 n_y(S)\right]$$

Since there are no dirty firms and thus no pollution in region B, in equilibrium there is no pollution tax revenue nor redistribution of dirty firm profits in region B. The utility level attained by a household residing in region B is:

$$U^B = \psi - f(1/M)\, M^2 n_y(S) u'\left[f(1/M)\, M^2 n_y(S)\right] + u\left[f(1/M)\, M^2 n_y(S)\right]$$

We can then compute the utility difference between regions A and B as:

$$\Delta U \equiv U^A - U^B = Mf(1/M)\left\{u'\left[f(1/M)\, M^2 n_y(S)\right] - \gamma\theta Mn_y(S)\right\} - \psi \tag{33}$$

By employing (31) and (33), we determine the stratified equilibrium relative price p and dirty firm labor demand $n_y(S)$ as shown in Figure 3. Specifically, from the top panel of Figure 3, utility equalization pins down the equilibrium level of dirty industry employment under stratification, which can be plugged into the bottom panel to obtain the relative price of the dirty good.

[Insert Figure 3 here]

5 Characterization of Equilibrium

Before turning to each of the two specific pollution tax regimes, one may compare dirty sector employment per firm, $n_y(I)$ and $n_y(S)$, under integrated and stratified equilibrium, respectively.

In an integrated equilibrium, we can use the dirty good market clearing condition and the dirty good demand, (27) and (28), to derive:

$$u'\left[\frac{M^2}{2} f\left(2/M\right) n_y\right] = \frac{\psi}{(1-\zeta) f'(2/M)} \tag{34}$$

In a stratified equilibrium, we can apply the location equilibrium condition in (33) to obtain:

$$\delta(n_y) \equiv u'\left[M^2 f\left(1/M\right) n_y\right] - \gamma\theta M n_y = \frac{\psi}{M f\left(1/M\right)} \tag{35}$$

where $\delta(n_y)$ measures the household's net surplus from consuming the dirty good.

These equilibrium relationships can be referred to as the *dirty good market equilibrium* loci, DE(I) and DE(S), respectively, under integrated and stratified configurations (see Figure 4). Whereas the DE(I) locus yields the equilibrium $n_y(I)$ as shown in the top panel of Figure 4, the DE(S) locus pins down the equilibrium $n_y(S)$ as depicted in the bottom panel of Figure 4. In the top panel of Figure 4, the LHS of DE(I) yields a downward sloping locus as a result of diminishing marginal utility, whereas the RHS is simply a constant that is decreasing in the exogenous mass of dirty firms. Thus, the integrated equilibrium is pinned down at point E^I. In the bottom panel, the LHS of DE(S), $\delta(n_y)$, is also a downward sloping locus and the RHS a constant depending negatively on the exogenous mass of dirty firms. These loci determine the stratified equilibrium at point E^S. To establish nice sufficient conditions for stratification in the next two subsections, we shall restrict our attention to a plausible scenario with $n_y(I) < n_y(S)$, i.e., dirty industry employment under integration is lower than that under stratification. It is clear from the definition of $\delta(n_y)$ that the above scenario is more likely to arise the smaller $\gamma\theta$ is. *In other words, for all of the results below, we shall assume that $\gamma\theta$ is small, a condition sufficient to ensure that dirty industry employment under integration is smaller than under stratification.*

[Insert Figure 4 here]

5.1 Fixed Pollution Tax Regime

We examine under what conditions the stratified equilibrium emerges under the fixed pollution tax regime but the integrated equilibrium does not, where the pollution tax levied by the local government under the two different configurations is given by:

$$\tau^i = \begin{cases} 2F/M, & \text{for Case } I \\ F/M, & \text{for Case } S \end{cases}$$

For purposes of comparison, in the stratified case only one local government raises pollution tax revenue, whereas in the integrated case each local government raises the same revenue as the dirty city in the stratified case. One interpretation of this assumption is that the simple presence of pollution in a city is enough to trigger a tax.

We impose a regularity condition on the dirty firm's surplus from uncompensated spillovers:

Condition R-1: (Regularity Condition on a Dirty Firm's Surplus)

$$\frac{1}{4}\widetilde{\sigma}\left(2/M\right) < \widetilde{\sigma}\left(1/M\right)$$

Under Condition R-1, we then consider the following:

Condition S-1: (Sufficient Condition for Stratification Under a Fixed Tax)

$$\frac{1}{4}\widetilde{\sigma}\left(2/M\right) < \frac{F}{M^2 n_y(S)} < \widetilde{\sigma}\left(1/M\right)$$

We can then establish:

Theorem 1: (Stratified Equilibrium) *Consider a local economy in which pollution production is not too severe and pollution disutility is not too high, in other words $\gamma\theta$ is sufficiently small. Under Condition R-1, we suppose that the fixed pollution tax is moderate so that the inequalities in Condition S-1 are met. Then the stratified configuration arises as an equilibrium outcome, but the integrated configuration does not.*

Proof. The proofs of all the theorems and propositions are relegated to the Appendix. ■

Thus, under Condition R-1, Condition S-1 is sufficient to ensure that the stratified configuration is an equilibrium outcome, but the integrated configuration is not. Intuitively, the first inequality of Condition S-1 implies negative profit received by dirty firms under integration, whereas the second inequality guarantees positive profit obtained by dirty firms under stratification. The main tipping point here is the gains from clustering under the fixed pollution tax regime.

Remark 1: (Impossibility of Integrated Equilibrium) It is not difficult to show that when F is large enough to satisfy $F > \frac{M}{4}\sigma\left(2/M\right)$, then dirty firms always incur negative profit, implying that an integrated configuration can never arise in equilibrium.

Remark 2: (On the Role of Agglomerative Externalities) It is important to note that despite the agglomeration force from uncompensated spillovers, the key driving force for all dirty firms to cluster in one region (A) is the presence of a fixed pollution tax that is independent of an individual firm's output. Specifically, with $F = 0$, it is clear that $\pi(I) > 0$, implying that the integrated configuration always arises in equilibrium. Moreover, we can compute the profits under the two configurations as follows:

$$\begin{aligned}\pi(I) &= \frac{M}{2}\frac{\psi}{f'(2/M)}\widetilde{\sigma}\,(2/M)\,n_y(I)\\ \pi(S) &= Mu'\left[M^2 f\,(1/M)\,n_y(S)\right]\widetilde{\sigma}\,(1/M)\,n_y(S)\end{aligned}$$

Further, assume that $\frac{1}{2}\widetilde{\sigma}\,(2/M) > \widetilde{\sigma}\,(1/M)$. Then, dirty firms will incur higher profit under integrated equilibrium compared to stratified equilibrium when the following inequality is met:

$$\frac{\frac{1}{2}\widetilde{\sigma}\,(2/M)}{\widetilde{\sigma}\,(1/M)} > \frac{u'\left[M^2 f\,(1/M)\,n_y(S)\right]n_y(S)}{\frac{\psi}{f'(2/M)}n_y(I)}$$

Refer to the top panel of Figure 4. The ratio on the right-hand side of the above inequality is measured by the ratio of the lightly shaded area covering E^O to the shaded area covering E^I. As long as this ratio is less than $\frac{\frac{1}{2}\widetilde{\sigma}(2/M)}{\widetilde{\sigma}(1/M)}$ (which is greater than one under the additional condition stated above), dirty firms will earn higher profits under an integrated equilibrium compared to a stratified equilibrium, and thus is viable whenever the stratified equilibrium is viable.

An important, related point is that we could accomplish our goal without any agglomeration externalities at all. Suppose that we simply used a decreasing returns technology for dirty firms, so that they make positive profits in any equilibrium without taxes. With the tax as specified, for F low both integrated and stratified equilibria will exist, with profits higher under stratified equilibrium. For higher F, only stratified equilibrium exists, as dirty firm profits are negative at integrated equilibrium. But this argument neglects an important issue. In comparing the integrated and stratified equilibria, there is movement along the supply curve for the dirty good due to wage differences (for the compensating differential from pollution), resulting in price changes and thus demand changes for the consumption goods as well. So the comparison is not that easy. Allowing for agglomeration externalities with socially constant returns actually simplifies the analysis because the dirty good production function (inclusive of the agglomeration externality) is linear in equilibrium. Nonetheless, once we have specific functional forms for the dirty good production technology, we will be able to return to this issue and provide a more concrete discussion (see Remark 7 in Section 6.1 below).

Remark 3: (On Strictly Concave Utility of the Clean Good) Suppose the utility of the clean good is strictly concave but the clean good (say, food) is more of a necessity than the dirty good in the sense that the income elasticity of demand for the clean good is lower than that of the dirty good. (In the current specification, the income elasticity of the demand for clean good is, by construction, one.) Then, for a richer jurisdiction, the willingness to pay for the pollution-generating good is higher, making integration more likely to survive the equilibrium profitability test. That is, consideration of a more general utility function specification with the clean good being more of a necessity than the dirty good reduces the likelihood of dirty firms clustering. Thus, the presence of income effects *per se* is *not* as important as the *relative income elasticity* of demand for the two consumption commodities.

Remark 4: (On Interregional Commuting of Workers) Recall that, in our benchmark model, at any stratified equilibrium, utility levels, but not wages, are equated between cities. What happens if commuting between cities is allowed? Notice that in our model all households have identical utility functions. Suppose we go to another extreme, setting commuting cost to zero. Free commuting implies that wage equalization also holds even under stratification. This wage equalization condition restricts the dirty firm's profitability, making stratification less likely to emerge as an equilibrium outcome. In conclusion, *sufficiently high intercity commuting cost* is necessary for dirty firm clustering to arise in equilibrium.

5.2 Linear Pollution Tax Regime

Under the linear pollution tax regime, $\zeta = t$ and

$$g = \begin{cases} L + \frac{tM}{2} f(\frac{2}{M}) n_y(I), \text{ in integrated equilibrium} \\ L + tMf(\frac{1}{M}) n_y(S), \text{ for stratified equilibrium} \end{cases}$$

We impose a stronger regularity condition on the dirty firm's surplus from uncompensated spillovers:

Condition R-2: (Regularity Condition on a Dirty Firm's Surplus)

$$\frac{1}{2}\widetilde{\sigma}(2/M) < \widetilde{\sigma}(1/M)$$

Under Condition R-2, we further consider the following condition:

Condition S-2: (Sufficient Condition for Stratification Under Linear Tax)

$$\frac{1}{2}\widetilde{\sigma}(2/M) < \frac{L}{(1-t)Mn_y(S)} < \widetilde{\sigma}(1/M)$$

This ensures:

Theorem 2: (Stratified Equilibrium) *Consider a local economy in which pollution production is not too severe and pollution disutility is not too high, in other words $\gamma\theta$ is sufficiently small. Under Condition R-2, we suppose that the lump-sum component of the linear pollution tax is moderate and the marginal tax rate is not too high so that the inequalities in Condition S-2 are met. Then the stratified configuration arises as an equilibrium outcome but the integrated configuration does not.*

In Section 6 below, we shall verify that both the presence of pollution and the presence of a fixed tax are crucial for a stable stratified equilibrium to arise.

6 The Case with Specific Functional Forms

Under the fixed pollution tax regime, we are left to check whether the stratified equilibrium is stable. Due to the difficulty of examining stability in the general setting, we shall conduct our analysis under specific functional forms for the dirty good production technology and the subutility for the dirty good. Specifically, we assume that $\widetilde{f}$ and u both take simple Cobb-Douglas forms:

$$\begin{aligned} \widetilde{f}\left(n_y^i(j), N_y^i\right) &= \phi[n_y^i(j)]^{\beta}[N_y^i]^{1-\beta}, \phi > 0 \text{ and } \beta \in (0,1) \\ u\left(c_y\right) &= \eta\left(c_y\right)^{\alpha}, \eta > 0 \text{ and } \alpha \in (0,1) \end{aligned}$$

Before deriving the stability condition, it is useful to provide explicit conditions in this special case under which the stratified configuration is an equilibrium outcome but the integrated configuration is not.

6.1 Fixed Pollution Tax Regime

Under the fixed pollution tax regime with the specific functional forms, we can derive a sufficient condition to ensure existence of a stratified equilibrium as follows:

Condition S-1′: (Stratified Equilibrium)

$$\beta\left[1 + \frac{\gamma\theta F}{(1-\beta)\phi\psi M^{1-\beta}}\right] < \frac{\alpha\beta\eta\phi}{\psi} M^{1-\beta}\left(\frac{1-\beta}{F}\right)^{1-\alpha} < 2^{2-\alpha-\beta}$$

We can establish:

Proposition 1: (Stratified Equilibrium under Fixed Pollution Tax) *Consider a local economy in which pollution production is not too severe and pollution disutility is not too high, in other words*

$\gamma\theta$ is sufficiently small, and Condition R-1 is met. Then, under a fixed pollution tax regime with Condition S-1$'$, a stratified competitive spatial equilibrium emerges.

We are now ready to check whether the stratified equilibrium is stable. Informally, stability is defined using small perturbations of firms from one region to the other, checking to see whether or not they would return to their equilibrium region.

Consider,

Condition I: (Instability without Pollution Tax)

$$\psi + \gamma\theta \left[\phi(\frac{\alpha\eta\beta}{\psi})^{\beta} M^{1-\beta}\right]^{\frac{1}{1-\alpha\beta}} > \left\{\psi^{\beta(1-\alpha)}(\alpha\eta\beta\phi^{\alpha})^{1-\beta} M^{\alpha[2-\beta(2-\beta)]}\right\}^{\frac{1}{1-\alpha\beta}}$$

We can then obtain:

Proposition 2: (Instability of Stratified Equilibrium) *Consider a local economy in which pollution production is not too severe and pollution disutility is not too high, in other words $\gamma\theta$ is sufficiently small, and Condition R-1 is met. Then, under Condition I, a stratified competitive spatial equilibrium is unstable in the absence of the pollution tax.*

Remark 5: (On Pollution vs. Corporate Tax) One may inquire whether our analysis applies to general corporate taxation. First, thinking of τ as a corporate tax in an economy without pollution concerns is not economically sensible, since it is not a tax on profits. Second, even if we ignore economic considerations, should $\gamma\theta = 0$, Conditions S and I would contradict each other if

$$M^{\frac{2\alpha-(1-\alpha)\beta^2}{1-\beta}} > \left[\frac{(1-\beta)\phi}{F}\right]^{1-\alpha}$$

That is, should the above inequality be met, pollution concerns are crucial for supporting the stratified equilibrium as a stable equilibrium configuration. Third, Condition S-1$'$ (particularly the second inequality) cannot hold when there is no fixed pollution tax ($F = 0$). In summary, we have shown that pollution and a fixed tax are crucial for a stable stratified equilibrium to arise.[10]

Remark 6: (Equilibrium Classification and Bifurcation Diagram) It is possible to delineate numerically a diagram in (M, F), namely the exogenous measure of dirty firms and the exogenous fixed cost tax revenue, that shows how changes in the values of (M, F) result in different types

[10]If there is no tax and both industries have (different) CRS production functions (with no Romer externality) but there is still pollution, then an integrated equilibrium will arise with both wage and utility equalized but with a corner solution in consumption (only the dirty good is consumed).

of equilibria, i.e. integrated versus stratified. Specifically, we set $\alpha = \beta = \gamma = 0.5$, $\phi = \eta = 1$, $\psi = 0.1$ and $\theta = 0.02$. We can then vary the values of each of M and F from 0 to 30. As shown in Figure 5(a), stratification is more profitable than integration for lower values of M and higher values of F: the indifference boundary between the two configurations is given by $\widetilde{BEC}$. Of course, a configuration can be supported only under positive profit, which is met for the area under $\widetilde{AES}$ in the case of stratification and for the area under $\widetilde{OEI}$ in the case of integration. Thus, a stratified equilibrium arises in the area of $OAEB$ (shaded with horizontal lines) whereas an integrated equilibrium emerges in the area of $BEID$ (shaded with vertical lines). We now set $F = 5$ and vary the measure of dirty firms, M. As long as $M > 1.60$, a nondegenerate equilibrium exists where firms earn sufficient profits to pay for the pollution tax. Over the range $M \in (1.60, 12.42)$, the equilibrium configuration is stratified and the fraction of dirty firms in city A is one. As firms continue to enter, the equilibrium configuration becomes integrated and the fraction of dirty firms in city A drops to $\frac{1}{2}$. This is depicted in the bifurcation diagram, Figure 5(b).[11]

[Insert Figures 5(a,b) here]

The intuition behind the equilibrium configuration of firms under various parameter values is as follows. At moderate levels of fixed tax cost F, if there are few dirty firms M, they must cluster together to be able to pay the tax. However, as the number of dirty firms increases, enough profit is generated to allow them to separate into halves and afford to pay the tax out of profits. Although they could conceivably generate even more profit if they stratified, the missing items are the level of pollution and the wage. With many dirty firms, the level of pollution in a stratified configuration is high, so the wage rate must also be high to attract workers, and this makes such a configuration impossible. For other parameter values that we have not discussed, for example if the number of dirty firms is low but the fixed tax cost is moderate or high, the dirty firms will all shut down even with the Inada condition on utility. If they all produce just a little, profits are insufficient to pay the tax.

Remark 7: (Reexamining the Role of Agglomerative Externalities) In our benchmark setup, by incorporating a regional-specific agglomeration externality following the Romer (1986) convention, dirty goods production exhibits private decreasing-returns-to-scale and social constant-returns-to-scale. Social constant returns simplify the analysis greatly, enabling a clean analysis of the equilibrium configuration, namely integration versus stratification. Nonetheless, the parameter β on the one hand captures the degree of externality and on the other hand measures the magnitude of

[11] If one allows the location of dirty city to be in either A or B, then a standard fork bifurcation diagram is obtained.

producer rent. Let us examine how the magnitude of β would affect the result by looking at the two inequalities in Condition S-1$'$ separately:

$$\begin{aligned}\left[1+\frac{\gamma\theta F}{(1-\beta)\phi\psi M^{1-\beta}}\right] &< \frac{\alpha\eta\phi}{\psi}M^{1-\beta}\left(\frac{1-\beta}{F}\right)^{1-\alpha} \\ M^{1-\beta}\left(\frac{1-\beta}{F}\right)^{1-\alpha} &< \frac{\psi 2^{2-\alpha-\beta}}{\alpha\beta\eta\phi}\end{aligned}$$

It is clear that higher β (lower producer rent and lower magnitude of externality) implies that the first inequality of the stratification condition is less likely to hold. Both the left hand side and the right hand side of the second inequality are, however, lower and the net effect on the likelihood of stratification is thus ambiguous. For sufficiently high β (for example, taking the extreme case as $\beta \to 1$), the first inequality always fails to hold while the second always holds. Intuitively, as the region-specific agglomeration externality diminishes, dirty firms will have less incentive to cluster. However, when β becomes too high, dirty firms can never generate enough rent to cover fixed costs regardless of the underlying configuration (integration versus stratification). This is because we cannot separate the role of the agglomerative externality from producer rent.[12] We can further resort to our numerical example, presented in Remark 6 above, to discuss how equilibrium classification changes in response to β.

[Insert Figure 6(a,b) here]

From Figure 6(a), we can see that, under the benchmark parametrization, when $\beta < 0.6273$ and F takes intermediate values falling between the long dashed curve and the solid curve, a stratified configuration arises in equilibrium; when β is higher than the critical value and F takes low values falling below both the the solid curve and the dashed curve, an integrated configuration emerges. As β increases, the range of F that can support a nondegenerate competitive spatial equilibrium (namely one with positive dirty good production and finite relative prices) narrows. We can also reproduce the bifurcation diagram 5(b) in Figure 6(b): it indicates that, given the benchmark values $F = 5$ and $M = 12.571$, the equilibrium configuration is stratified over the range $\beta \in (0, 0.5000)$; the equilibrium configuration turns integrated when $\beta \in (0.5000, 0.7220)$, and only a degenerate competitive spatial equilibrium (with no dirty good production and infinite relative price) exists when β continues to rise, exceeding 0.7220.

[12] In order to separate these two channels (producer rent and magnitude of externality), one must give up social constant returns, assuming instead social decreasing returns: $\widetilde{f}\left(n_y^i(j), N_y^i\right) = \phi[n_y^i(j)]^{\beta_1}[N_y^i]^{\beta_2}$, $\phi > 0$, $\beta_1, \beta_2 \in (0,1)$ and $\beta_1 + \beta_2 < 1$. We have tried this but lost analytic tractability.

Remark 8: (On Endogenous Choice of Production Technologies) In the benchmark economy, we have followed an Arrow-Debreu convention by assuming that producers are endowed with specific technologies for producing particular goods. One may inquire what happens if producers are allowed to make an endogenous choice of production technologies (clean versus dirty). Consider the model modified so that a potential producer can choose between clean and dirty good production, followed by location choice and then production. Then the value of being a clean good producer (zero profit) must be equal to the value of being a dirty good producer, implying: $\max\{\pi(I), \pi(S)\} = 0$, where

$$\begin{aligned}\pi(I) &= \frac{M}{2}\frac{\psi}{(1-\zeta)f'(2/M)}\left[\sigma\left(2/M(I)\right)n_y(I) - \left(2/M(I)\right)\tau^i\right] \\ \pi(S) &= Mu'\left[f\left(1/M\right)M^2 n_y(S)\right]\left[\sigma\left(1/M(S)\right)n_y(S) - \left(1/M(S)\right)\tau^i\right]\end{aligned}$$

Let the equilibrium mass of dirty firms as a result of free entry be denoted as $M^*(I)$ and $M^*(S)$, respectively, for the cases of integration and stratification. Thus,

$$\begin{aligned}\frac{\sigma\left(2/M^*(I)\right)}{2/M^*(I)} &= \frac{\tau^i}{n_y(I)} \\ \frac{\sigma\left(1/M^*(S)\right)}{1/M^*(S)} &= \frac{\tau^i}{n_y(S)}\end{aligned}$$

Recall from our discussion of the bifurcation diagram that as M rises, $\frac{\sigma(e)}{e}$ would increase in e and hence $\frac{\sigma(2/M))}{2/M}$ would become more likely to dominate $\frac{\sigma(1/M)}{1/M}$. Using this property and manipulating (see the Appendix), we arrive at:

$$\begin{aligned}M^*(I) &= 2\left[\frac{\psi}{\alpha\beta\eta\phi}\left(\frac{2F}{1-\beta}\right)^{1-\alpha}\right]^{\frac{1}{1-\beta}} \\ M^*(S) &= \left\{\frac{[\psi(1-\beta)+\gamma\theta F]}{\alpha(1-\beta)^{2-\alpha}\eta\phi}F^{1-\alpha}\right\}^{\frac{1}{1-\beta}}\end{aligned}$$

This is familiar from the monopolistic competition literature with endogenous entry of firms, as in the work by Melitz (2003). While one may impose constraints in various ways to determine the equilibrium outcome, a land requirement is natural in our economy (as proposed by Helpman, 1998). Suppose there is a fixed land requirement for all households and firms, each at an inelastic unit normalized to one (which can be justified as an equilibrium outcome as in Berliant, Peng and Wang, 2002). Consider the public land ownership structure delineated in Fujita (1989, pp. 60-61)[13] with a total supply of $L > 2$ in the entire local economy. Then the total land demand is: $2 + M$ (clean firms of mass one + dirty firms of mass M + households of mass one). It is straightforward

[13] The public land ownership model features land rent collections that are refunded equally to all inhabitants of a location.

to see that the necessary and sufficient condition for the stratified configuration to arise as the only equilibrium outcome is:

$$2 + M^*(S) \leq L < 2 + M^*(I)$$

The above inequalities can be manipulated to yield (see the Appendix):

$$1 + \frac{\gamma\theta F}{(1-\beta)\psi} \leq \frac{\alpha\eta\phi(1-\beta)^{1-\alpha}(L-2)^{1-\beta}}{\psi F^{1-\alpha}} < \frac{2^{2-\alpha-\beta}}{\beta} \tag{36}$$

which can hold true only if

$$1 + \frac{\gamma\theta F}{(1-\beta)\psi} < \frac{2^{2-\alpha-\beta}}{\beta} \tag{37}$$

The latter inequality is satisfied when β is sufficiently small (i.e., with a sufficiently strong spillover externality). Given Condition (37), then Condition (36) ensures that the equilibrium is stratified. This condition requires that (i) the fixed payment F is not too large (otherwise, too many firms must cluster, implying that land demand exceeds land supply) and (ii) land supply is not too large (otherwise, even an integrated equilibrium can be supported). Notice that neither the firms' nor the households' optimization problems change with the introduction of the land market. This is because, for producers, the same amount of land rent is added and subtracted from profits, whereas for consumers the same amount of land rent is added to both sides of the budget. In short, adding endogenous technology choice by allowing potential producers choose between clean and dirty good production would not alter our main findings once we introduce a simple land market under public ownership with an appropriate land supply satisfying Condition (36).

6.2 Linear Pollution Tax Regime

We turn next to examining the case of a linear pollution tax. Consider,

Condition S-2′: (Stratified Equilibrium)

$$\beta(1-t) + \frac{\beta\gamma\theta ML}{\phi\psi(1-\beta)} < \frac{\alpha\beta\eta\phi M^{\alpha-\beta}(1-t)^{2-\alpha}}{\psi}\left(\frac{1-\beta}{L}\right)^{1-\alpha} < 2^{1-\beta}$$

We now have:

Proposition 3: (Stratified Equilibrium under Linear Pollution Tax) *Consider a local economy in which pollution production is not too severe and pollution disutility is not too high, in other words $\gamma\theta$ is sufficiently small, and Condition R-2 is met. Then, under a linear pollution tax regime with Condition S-2′, a stratified competitive spatial equilibrium emerges.*

7 Can Pareto Optimum Be Price Supported?

Since individuals are *ex ante* identical, we restrict our attention to within-region-equal-treatment Pareto optimum in the sense that all households within a given region reach an identical indirect utility level and consume the same bundle. Such a Pareto optimum must satisfy the following constraints:

$$\begin{aligned} x^i(j) &= \psi n_x^i(j), \quad i = A, B, j \in [0, k^i] \\ y^i(j) &= N_y^i f\left(\frac{n_y^i(j)}{N_y^i}\right), \quad i \in A, B, j \in [0, m^i] \end{aligned}$$

$$N_x^i = \int_0^{k^i} n_x^i(j)dj, \quad N_y^i = \int_0^{m^i} n_y^i(j)dj, \quad N_x^A + N_x^B + N_y^A + N_y^B = 1$$

$$\sum_{i=A,B} N^i c_x^i = \sum_{i=A,B} \int_0^{k^i} x^i(j)dj, \quad \sum_{i=A,B} N^i c_y^i = \sum_{i=A,B} \int_0^{m^i} y^i(j)dj$$

where the first two equations specify production technologies, the third gives labor material balance and the population identity, and the last represents commodity material balance. Such Pareto optima are found by solving the following optimization problem:

$$\begin{aligned} \max U^A &= c_x^A - \gamma\theta \int_0^{m^A} y^A(j)dj + u(c_y^A) \\ \text{s.t. } U^B &= c_x^B - \gamma\theta \int_0^{m^B} y^B(j)dj + u(c_y^B) = \overline{U} \\ &\quad \text{and the above technology and material balance constraints.} \end{aligned}$$

We consider equilibria with linear taxes in the next two subsections.

Remark 9: (Pareto Optimal Configuration) It is natural to inquire at this point whether, for given parameters, the Pareto optimum features an integrated or stratified configuration. Indeed, although it is not central to our analysis, in general it depends on the comparison of utility from dirty good production and pollution damage in a region with half or all of the dirty firms. In our quasi-linear setting, it amounts to $u(2f(2/M)) - \gamma\theta f(2/M)$ for an integrated configuration compared with $u(f(1/M)) - \gamma\theta f(1/M)$ for a stratified configuration.

7.1 Case I: Integrated Optimum

At an integrated optimum, we have all interior allocations. We can establish:

Theorem 3: (Equilibrium Support of Integrated Configuration) *Consider a local economy in which pollution production is not too severe and pollution disutility is not too high, in other words*

$\gamma\theta$ is sufficiently small, and Condition R-2 is met. Then, the Pareto optimum with an integrated configuration can be supported by a competitive spatial equilibrium under the following marginal tax rate: $\zeta = \{1 + 2\psi/[\gamma\theta f'(2/M)]\}^{-1}$.

Intuitively, the higher the pollution damage (captured by larger $\gamma\theta$) is, the greater the marginal pollution tax will be.

7.2 Case II: Stratified Optimum

At a stratified optimum, we have: $k^A = m^B = 0$. We can establish:

Theorem 4: (Suboptimality of Stratified Equilibrium) *Consider a local economy in which pollution production is not too severe and pollution disutility is not too high, in other words $\gamma\theta$ is sufficiently small, and Condition R-2 is met. Then, a stratified competitive spatial equilibrium is suboptimal with over-employment and over-production in the dirty goods sector relative to the stratified Pareto optimum.*

Thus, a stratified equilibrium can never reach Pareto optimality by means of a linear pollution tax (which encompasses Pigouvian taxation). In fact, the equilibrium employment in the dirty sector under the stratified configuration is always too large, implying that dirty goods and pollution are both over-produced. Such an over-polluting equilibrium outcome can never be corrected by a linear pollution tax.

To understand the result, it is best to refer to Figure 7, where we plot the downward-sloping after-tax MPL locus in the top panel and repeat the locational equilibrium diagram (the top panel of Figure 3) in the bottom panel of Figure 7. A high marginal tax will shift down the after-tax MPL locus without altering any other curves. Thus, the only change is the corresponding reduction in the dirty industry wage, w^A. As long as $w^A > \psi$ still holds after the tax increase, the lower wage will be fully offset by the tax and profit redistribution, keeping consumers in region A as well off as before the tax increase. This is equivalent to saying that although dirty good demand is elastic, dirty good supply is perfectly inelastic. As a result, dirty good employment and production in stratified equilibrium remain at levels higher than the respective optimum quantities, regardless of the linear pollution tax levied.

[Insert Figure 7 here]

Another way of interpreting our welfare results is as follows. The same number of distortions and tax instruments are available no matter the equilibrium configuration of firms, so that one would expect that Pareto optimum would be either supportable or unsupportable with prices and taxes independent of the equilibrium configuration. However, the integrated Pareto optimal configuration has the unique feature that it is symmetric across locations, so the distortion associated with migration is not present. Thus, one fewer instrument is needed to support the integrated configuration, in contrast with the stratified configuration.

In the conventional literature, Pigouvian taxes (a special form of a linear tax without the lump-sum component) need not work in practice due to the difficulty of computing marginal damages at the optimum (Baumol 1972), or when firms have monopoly power so that they can transfer the tax burden (Buchanan and Tullock, 1975), or when oligopolistic firms have dynamic strategic interactions (Benchekroun and Van Long 1998), or when lobbying groups care about the distribution of income in political games (Aidt 1998)). In our paper, assuming away all of these issues, we show that even a generalized Pigouvian tax as proposed by Carlton and Loury (1980) cannot restore first best under a static, competitive environment, when we allow locational choice with endogenous clustering.

Whereas the linear pollution tax cannot correct equilibrium inefficiency, it should be noted that an appropriate redistribution scheme may do the job. In particular, consider a lump-sum redistribution from polluted region A to clean region B. This induces ΔU to shift down and hence equilibrium employment in the dirty industry to fall. Thus, as long as $\gamma\theta$ is not too large, there exists an appropriate level of such a redistribution to support the Pareto optimal level of dirty industry employment as an equilibrium.

8 Concluding Remarks

In this work, we have shown how a fixed charge component of a tax system can cause agglomeration of polluting firms as an equilibrium phenomenon. We have also established that whereas an integrated Pareto optimum can be supported by a competitive spatial equilibrium with a linear pollution tax, a stratified Pareto optimum cannot. Regardless of the linear pollution tax schedule, a stratified equilibrium is always over-polluted compared to the optimum. To support the stratified Pareto optimum, however, an effective (but practically not implementable) policy prescription is to redistribute the pollution tax revenue from the dirty to the clean city residents. Such a policy will induce migration to the clean city, thereby reducing production of the dirty good and thus of pollution.

In this paper, we have considered only equilibrium configurations that are completely stratified in terms of production or that are completely integrated in that production is symmetric across locations. One may inquire whether other configurations may emerge in equilibrium. The answer is positive: it is possible that one city is mixed with both clean and dirty industries present, whereas another has only the clean industry. In this configuration, clean industry workers must have equal utility across locations and all workers must have the same wage in the city with mixed industries. Under the Ricardian technology where clean workers are paid an exogenously fixed wage, the two equalization conditions can be met only in knife-edge cases. It is therefore innocuous to ignore this partially integrated configuration.

Many extensions of the model are possible. For example, global or interregional pollution could be present in addition to the local pollution we have considered. Naturally, agglomeration is a product of local pollution, as global pollution affects everyone in the same way. Aside from the extension to endogenous technology choice delineated in Remark 8, we have refrained from adding land to the model for tractability reasons. Future work should proceed in this direction. Capitalization of pollution damages and lump-sum transfers to localities could change the results. It would also be interesting to examine zoning in this context.

References

[1] Aidt, T.S., 1998. Political internalization of economic externalities and environmental policy. *Journal of Public Economics* 69, 1-16.

[2] Baumol, W. J., 1972. On taxation and the control of externalities. *American Economic Review* 62, 307-322.

[3] Benabou, R., 1996a. Heterogeneity, stratification and growth. *American Economic Review* 86, 584-609.

[4] Benabou, R., 1996b. Equity and efficiency in human capital investment: The local connection. *Review of Economic Studies* 63, 237-264.

[5] Benchekroun, H. and N. Van Long, 1998. Efficiency inducing taxation for polluting oligopolists. *Journal of Public Economics* 70, 325-342.

[6] Bergstrom, T. and R. Cornes, 1983. Independence of allocative efficiency from distribution in the theory of public goods. *Econometrica* 51, 1753-1765.

[7] Berliant, M., S. Peng and P. Wang, 2002. Production externalities and urban configuration. *Journal of Economic Theory* 104, 275-303.

[8] Buchanan, J. M. and G. Tullock, 1975. Polluter's profits and political response: Direct control versus taxes. *American Economic Review* 65, 139-147.

[9] Carlton, D.W. and G.C. Loury, 1980. The limitation of Pigouvian taxes as a long-run remedy for externalities. *Quarterly Journal of Economics* 95, 559-567.

[10] Chen, B., S. Peng and P. Wang, 2009. Intergenerational human capital evolution, local public good preferences, and stratification. *Journal of Economic Dynamics and Control* 33, 745-757.

[11] Chen, B., C. Huang and P. Wang, 2012. Stratification by environment. *Journal of Public Economic Theory* 14, 711-735.

[12] Chipman, J.S. and G. Tian, 2012. Detrimental externalities, pollution rights, and the 'Coase theorem'. *Economic Theory* 49, 309-327.

[13] Fujita, M., 1989. *Urban Economic Theory: Land Use and City Size*. New York, Cambridge University Press.

[14] Hartwick, J., U. Schweizer and P. Varaiya, 1976. Comparative statics of a residential economy with several classes. *Journal of Economic Theory* 13, 396-413.

[15] Helpman, E., 1998. The size of regions, in D. Pines, E. Sadka and Y. Zilcha (eds.), *Topics in Public Economics*. New York, Cambridge University Press, pp. 33-54.

[16] Karp, L., 2005. Nonpoint source pollution taxes and excessive tax burden. *Environmental & Resource Economics* 31, 229-251.

[17] Markusen, J.R., E.R. Morey and N. Oleviler, 1995. Competition in regional environmental policies when plant locations are endogenous. *Journal of Public Economics* 56, 55-77.

[18] Melitz, M.J., 2003. The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica* 71, 1695-1725.

[19] Nechyba, T., 1997. Existence of equilibrium and stratification in local and hierarchical Tiebout economies with property taxes and voting. *Economic Theory* 10, 277-304.

[20] OECD, 1994. *Managing the Environment: The Role of Economic Instruments*. Paris, OECD.

[21] Peng, S. and P. Wang, 2005. Sorting by foot: 'travel-for' local public goods and equilibrium stratification. *Canadian Journal of Economics* 38, 1224-1252.

[22] Picard, P.M. and T. Tabuchi, 2010. Self-organized agglomerations and transport costs. *Economic Theory* 42, 565-589.

[23] Pigou, A. C., 1920. *The Economics of Welfare*. London, Macmillan.

[24] Porter, M.E., 1990. *The Competitive Advantage of Nations*. New York, Free Press.

Appendix

Proof of Theorem 1: A condition sufficient to show that the stratified configuration is an equilibrium but the integrated configuration is not is:

$$\pi(I) < 0 < \pi(S)$$

That is, the dirty firms only operate under stratification. From (29) and (32), in turn, we need the following inequality condition:

$$\frac{1}{4}\widetilde{\sigma}\,(2/M)\,n_y(I) < \frac{F}{M^2} < \widetilde{\sigma}\,(1/M)\,n_y(S)$$

When $n_y(I) < n_y(S)$, the above inequality holds under Condition S-1, which can be met only under Condition R-1. Since $n_y(I)$ and $n_y(S)$ are endogenous, we must further investigate their magnitudes in order to establish precise sufficient conditions on primitives. This can be accomplished utilizing Figure 4, by comparing the positions of point E^I and point E^S. We can see from the top panel of Figure 4 that, as long as $\gamma\theta$ is not too large, we can have $n_y(I) < n_y(S)$ (as shown in the bottom panel of Figure 4). Given this and Condition R-1, we can always choose F to satisfy Condition S-1, which subsequently ensures the existence of a stratified configuration but not the integrated configuration as an equilibrium outcome, as illustrated diagrammatically where $n_y(S)$ is pinned down by Figure 4 with Conditions R-1 and S-1 met as in Figure 7(a). ∎

Proof of Theorem 2: In this case, the profits generated by each dirty firm under integrated and stratified configurations become:

$$\begin{aligned}
\pi(I) &= \frac{M}{2}\frac{\psi}{(1-t)f'(\frac{2}{M})}[(1-t)\widetilde{\sigma}(\frac{2}{M})n_y(I) - \frac{2}{M}L] \\
\pi(S) &= Mu'[f(\frac{1}{M})M^2 n_y(S)][(1-t)\widetilde{\sigma}(\frac{1}{M})n_y(S) - \frac{1}{M}L]
\end{aligned}$$

Similar to the fixed tax case, here is a sufficient condition to ensure that the stratified configuration is an equilibrium but the integrated configuration is not:

$$\pi(I) < 0 < \pi(S)$$

which can be rewritten as the following inequalities:

$$\frac{1}{2}\widetilde{\sigma}\,(2/M)\,n_y(I) < \frac{L}{(1-t)M} < \widetilde{\sigma}\,(1/M)\,n_y(S)$$

Under Condition R-2, as shown in Figure 7(b) and the circumstances delineated by Figure 4, $n_y(I) < n_y(S)$ and thus Condition S-2 is sufficient to ensure the inequalities above. ∎

Proof of Proposition 1: With specific functional forms, dirty firm employment in integrated equilibrium can be solved explicitly:

$$n_y(I) = \frac{1}{2}\left[\frac{\alpha\eta\beta\phi^{\alpha}}{\psi(M/2)^{1-\alpha(2-\beta)}}\right]^{\frac{1}{1-\alpha}}$$

whereas dirty firm employment in stratified equilibrium must satisfy:

$$\frac{\alpha\eta\phi^{\alpha}}{M^{1-\alpha(2-\beta)}[n_y(S)]^{1-\alpha}} = \psi + \gamma\theta\phi M^{2-\beta} n_y(S)$$

Substituting these into Condition S-1 gives the result. ■

Proof of Proposition 2: Suppose the stratified equilibrium is not stable. We have deviation of dirty firms of positive measure ε moving from A to B, receiving joint profit given by:

$$\pi_{\varepsilon} = p\phi(\tilde{n}_{y_{\varepsilon}})^{\beta}\varepsilon^{1-\beta} - \psi\tilde{n}_{y_{\varepsilon}} - p\tilde{\tau}$$

From the clean and dirty firms' first-order conditions for profit optimization, we have:

$$\tilde{n}_{y_{\varepsilon}} = \varepsilon(\frac{\beta\phi p}{\psi})^{\frac{1}{1-\beta}}$$

Combining these expressions, we obtain:

$$\pi_{\varepsilon} = \psi(\frac{1-\beta}{\beta})\varepsilon(\frac{\beta\phi p}{\psi})^{\frac{1}{1-\beta}} - p\tilde{\tau}$$

Thus, the per deviating firm profit can be computed as follows:

$$\tilde{\pi} = \frac{\pi_{\varepsilon}}{\varepsilon} = \psi(\frac{1-\beta}{\beta})(\frac{\beta\phi p}{\psi})^{\frac{1}{1-\beta}} - p\frac{F}{\varepsilon}$$

Recall that the profit of a firm that doesn't deviate is:

$$\pi^A = p\left[(1-\beta)\phi M^{1-\beta} n_y - \frac{F}{M}\right]$$

To ensure stability, we therefore need: $\lim_{\varepsilon\to 0}\tilde{\pi} < \pi^A$, which holds trivially as $\lim_{\varepsilon\to 0}\tilde{\pi} = -\infty$. It remains to check that the Romer positive externality alone cannot lead to stable dirty firm agglomeration. This is equivalent to showing that, with $F = 0$,

$$\tilde{\pi} > \pi^A$$

which requires:

$$n_y < \left[\frac{\alpha\eta\beta\phi^{\alpha}}{\psi M^{\frac{1-\alpha\beta(2-\beta)}{\beta}}}\right]^{\frac{\beta}{1-\alpha\beta}}$$

Using (35), we can rewrite the inequality above in primitives, yielding Condition I. ■

Proof of Remark 8: Suppose that the locational choice and production in stage 2 yields an integrated configuration. Then from

$$\pi(I) = \frac{M(I)}{2} \frac{\psi}{(1-\zeta) f'(\frac{2}{M(I)})} \left[\sigma \left(\frac{2}{M(I)} \right) n_y(I) - \left(\frac{2}{M(I)} \right) \tau^i \right] = 0$$

and the proof of Proposition 1 (in this Appendix), we have

$$n_y(I) = \frac{1}{2} \left[\frac{\alpha\beta\eta\phi^\alpha}{\psi[\frac{M(I)}{2}]^{1-\alpha(2-\beta)}} \right]^{\frac{1}{1-\alpha}}$$

Since $\tau^I = 2F/M(I)$ and under the fixed tax pollution scheme $\zeta = 0$, free entry implies:

$$\sigma \left(\frac{2}{M^*(I)} \right) n_y(I) - \left[\frac{2}{M^*(I)} \right]^2 F = 0$$

Using the definition of σ, the specific functional form for the dirty good production technology and the population identity under the integrated configuration, $N_y^i(I) = \frac{M(I)}{2} n_y^i(I)$, we have:

$$f(\frac{n_y^i(I)}{N_y^i(I)}) = \phi[\frac{n_y^i(I)}{\frac{M(I)}{2} n_y^i(I)}]^\beta = \phi[\frac{2}{M(I)}]^\beta \quad \text{and} \quad f'(\frac{n_y^i(I)}{N_y^i(I)}) = \phi\beta[\frac{2}{M^*(I)}]^{\beta-1}$$

Plugging in $n_y(I)$ into the free entry condition yields:

$$\left[f(\left(\frac{2}{M^*(I)} \right) - \frac{2}{M^*(I)} f'(\left(\frac{2}{M^*(I)} \right) \right] \frac{1}{2} \left\{ \frac{\alpha\beta\eta\phi^\alpha}{\psi \left[\frac{M^*(I)}{2} \right]^{1-\alpha(2-\beta)}} \right\}^{\frac{1}{1-\alpha}} = \left[\frac{2}{M^*(I)} \right]^2 F$$

or

$$\phi(1-\beta) \left[\frac{2}{M^*(I)} \right]^\beta \left\{ \frac{\alpha\beta\eta\phi^\alpha}{\psi \left[\frac{M^*(I)}{2} \right]^{1-\alpha(2-\beta)}} \right\}^{\frac{1}{1-\alpha}} = 2 \left[\frac{2}{M^*(I)} \right]^2 F$$

which can be manipulated to derive the expression for $M^*(I)$ in Remark 8.

Suppose now that the locational choice and production in stage 2 yields a stratified configuration. Then from

$$\pi(S) = Mu' \left[f\left(1/M\right) M^2 n_y(S) \right] \left[\sigma\left(1/M(S)\right) n_y(S) - \left(1/M(S)\right) \tau^S \right] = 0$$

the proof of Proposition 1, and $\tau^S = \frac{F}{M(S)}$, we have:

$$\sigma \left(\frac{1}{M(S)} \right) n_y(S) - \left[\frac{1}{M(S)} \right]^2 F = 0$$

where $n_y(S)$ satisfies:

$$\frac{\alpha\eta\phi^\alpha}{M(S)^{1-\alpha(2-\beta)}[n_y(S)]^{1-\alpha}} = \psi + \gamma\theta\phi M(S)^{2-\beta} n_y(S)$$

Straightforward manipulations of the free entry condition using the definition of σ, the specific functional form for the dirty good production technology and the population identity imply:

$$\left[f(\left(\frac{1}{M^*(S)}\right) - \frac{1}{M^*(S)}f'\left(\frac{1}{M^*(S)}\right)\right] n_y(S) = \left[\frac{1}{M(S)}\right]^2 F$$

or

$$\left[\phi[\frac{1}{M^*(S)}]^{\beta} - \frac{1}{M^*(S)}\phi\beta[\frac{1}{M^*(S)}]^{\beta-1}\right] n_y(S) = [\frac{1}{M^*(S)}]^2 F$$

or

$$n_y(S) = \frac{F}{\phi(1-\beta)}\left[\frac{1}{M^*(S)}\right]^{2-\beta}$$

Substituting the above expression into the $n_y(S)$ equation, we obtain:

$$\frac{\alpha\eta\phi^{\alpha}}{M^*(S)^{1-\alpha(2-\beta)}\left\{\frac{F}{\phi(1-\beta)}\left[\frac{1}{M^*(S)}\right]^{2-\beta}\right\}^{1-\alpha}} = \psi + \frac{\gamma\theta F}{(1-\beta)}$$

which can be manipulated to derive the expression for $M^*(S)$ in Remark 8.

Given a total supply of $L > 2$ in the entire local economy and one unit inelastic demand for land by each household and producer, the necessary and sufficient condition for the stratified configuration to arise as the only equilibrium outcome can be manipulated to yield:

$$M^*(S) \leq L - 2 < M^*(I)$$

or

$$\left\{\frac{[\psi(1-\beta)+\gamma\theta F]}{\alpha(1-\beta)^{2-\alpha}\eta\phi}F^{1-\alpha}\right\}^{\frac{1}{1-\beta}} \leq L-2 < 2\left\{\frac{\psi}{\alpha\beta\eta\phi}\left(\frac{2F}{1-\beta}\right)^{1-\alpha}\right\}^{\frac{1}{1-\beta}}$$

or Condition (36), which can hold true only if Condition (37) is met. ■

Proof of Proposition 3: Under a linear pollution tax, it is easily verified that the sufficient condition S-2 becomes Condition S-2$'$. ■

Proof of Theorem 3: Upon substituting out the production technologies, this problem can be solved by setting up the Lagrangian as follows:

$$\begin{aligned}
\mathcal{L} \;=\;& c_x^A - \gamma\cdot\theta\int_0^{m^A} N_y^A f\left(\frac{n_y^A(j)}{N_y^A}\right)dj + u(c_y^A) \\
&+\lambda_U[c_x^B - \gamma\theta\int_0^{m^B} N_y^B f\left(\frac{n_y^B(j)}{N_y^B}\right)dj + u(c_y^B) - \overline{U}] \\
&+\lambda_N\left[1 - \int_0^{k^A} n_x^A(j)dj - \int_0^{k^B} n_x^B(j)dj - \int_0^{m^A} n_y^A(j)dj - \int_0^{m^B} n_y^B(j)dj\right] \\
&+\lambda_X\sum_{i=A,B}\left[\int_0^{k^i}\psi n_x^i(j)dj - \left(\int_0^{k^i} n_x^i(j)dj + \int_0^{m^i} n_y^i(j)dj\right)c_x^i\right] \\
&+\lambda_Y\sum_{i=A,B}\left[\int_0^{m^i} N_y^i f\left(\frac{n_y^i(j)}{N_y^i}\right)dj - \left(\int_0^{k^i} n_x^i(j)dj + \int_0^{m^i} n_y^i(j)dj\right)c_y^i\right]
\end{aligned}$$

where λ_U, λ_N, λ_X and λ_Y are Lagrange multipliers associated with the utility constraint and labor and goods material balance constraints, respectively. Since k^A and m^B are zero under a stratified configuration, we must derive the Pareto optimum under each configuration separately.

The first-order conditions with respect to the 4 consumption and the 4 labor variables are given by:

$$\frac{\partial \mathcal{L}}{\partial c_x^A} = 1 - \lambda_X \left(\int_0^{k^A} n_x^A(j)dj + \int_0^{m^A} n_y^A(j)dj \right) = 0$$

$$\frac{\partial \mathcal{L}}{\partial c_x^B} = \lambda_U - \lambda_X \left(\int_0^{k^B} n_x^B(j)dj + \int_0^{m^B} n_y^B(j)dj \right) = 0$$

$$\frac{\partial \mathcal{L}}{\partial c_y^A} = u'(c_y^A) - \lambda_Y \left(\int_0^{k^A} n_x^A(j)dj + \int_0^{m^A} n_y^A(j)dj \right) = 0$$

$$\frac{\partial \mathcal{L}}{\partial c_y^B} = \lambda_U u'(c_y^B) - \lambda_Y \left(\int_0^{k^B} n_x^B(j)dj + \int_0^{mB} n_y^B(j)dj \right) = 0$$

$$\frac{\partial \mathcal{L}}{\partial n_x^A(i)} = -\lambda_N + \lambda_X(\psi - c_x^A) - \lambda_Y c_y^A = 0$$

$$\frac{\partial \mathcal{L}}{\partial n_x^B(i)} = -\lambda_N + \lambda_X(\psi - c_x^B) - \lambda_Y c_y^B = 0$$

$$\frac{\partial \mathcal{L}}{\partial n_y^A(i)} = -\gamma\theta f'\left(\frac{n_y^A(j)}{N_y^A}\right) - \lambda_N - \lambda_X c_x^A + \lambda_Y \left[f'\left(\frac{n_y^A(j)}{N_y^A}\right) - c_y^A \right] = 0$$

$$\frac{\partial \mathcal{L}}{\partial n_y^B(i)} = -\lambda_U\gamma\theta f'\left(\frac{n_y^B(j)}{N_y^B}\right) - \lambda_N - \lambda_X c_x^B + \lambda_Y \left[f'\left(\frac{n_y^B(j)}{N_y^B}\right) - c_y^B \right] = 0$$

Straightforward manipulation and simplification yields:

$$u'(c_y^A) = u'(c_y^B) = \frac{\lambda_Y}{\lambda_X} \tag{A1}$$

$$c_x^B - c_x^A = u'(c_y^A)c_y^A - u'(c_y^B)c_y^B \tag{A2}$$

$$\left\{ [u'(c_y^A) - \gamma\theta N^A] f'\left(\frac{n_y^A}{N_y^A}\right) - \psi \right\} = n_y^B \left\{ [u'(c_y^B) - \gamma\theta N^B] f'\left(\frac{n_y^B}{N_y^B}\right) - \psi \right\} = 0 \tag{A3}$$

$$1 = k^A n_x^A + k^B n_x^B + m^A n_y^A + m^B n_y^B \tag{A4}$$

$$\psi(k^A n_x^A + k^B n_x^B) = (k^A n_x^A + m^A n_y^A)c_x^A + (k^B n_x^B + m^B n_y^B)c_x^B \tag{A5}$$

$$m^A N_y^A f\left(\frac{n_y^A}{N_y^A}\right) + m^B N_y^B f\left(\frac{n_y^B}{N_y^B}\right) = (k^A n_x^A + m^A n_y^A)c_y^A + (k^B n_x^B + m^B n_y^B)c_y^B \tag{A6}$$

Under an integrated configuration, we have: $N^A = N^B = \frac{1}{2}$, $n_y^A = n_y^B = n_y$ and $n_x = 1 - Mn_y$. From (A1) and (A2), we must have the same consumption bundles across the two locations. Using (A4) and (A5) then yields:

$$c_x = \psi(1 - Mn_y) \tag{A7}$$

Combining (A4) and (A6), we obtain:

$$c_y = 2(\frac{M}{2})^2 f(\frac{2}{M}) n_y \tag{A8}$$

Both consumptions are identical to the equilibrium ones. Substituting (A8) into (A3) implies:

$$\left[u'\left(2(\frac{M}{2})^2 f(\frac{2}{M}) n_y\right) - \frac{1}{2}\gamma\theta\right] = \frac{\psi}{f'(\frac{2}{M})} \tag{A9}$$

By setting n_y in the equilibrium captured by (34) and in the Pareto optimum captured by (A9) equal to one another, one obtains:

$$\frac{\psi}{f'(\frac{2}{M})} + \frac{1}{2}\gamma\theta = \frac{\psi}{(1-\zeta)f'(\frac{2}{M})}$$

which can be manipulated to derive the marginal tax rate given in the statement of the theorem. ■

Proof of Theorem 4: At a stratified optimum, we have: $k^A = m^B = 0$, together with the following 6 first-order conditions:

$$\frac{\partial \mathcal{L}}{\partial c_x^A} = 1 - \lambda_X \left(\int_0^{k^A} n_x^A(j)dj + \int_0^{m^A} n_y^A(j)dj \right) = 0$$

$$\frac{\partial \mathcal{L}}{\partial c_x^B} = \lambda_U - \lambda_X \left(\int_0^{k^B} n_x^B(j)dj + \int_0^{m^B} n_y^B(j)dj \right) = 0$$

$$\frac{\partial \mathcal{L}}{\partial c_y^A} = u'(c_y^A) - \lambda_Y \left(\int_0^{k^A} n_x^A(j)dj + \int_0^{m^A} n_y^A(j)dj \right) = 0$$

$$\frac{\partial \mathcal{L}}{\partial c_y^B} = \lambda_U u'(c_y^B) - \lambda_Y \left(\int_0^{k^B} n_x^B(j)dj + \int_0^{mB} n_y^B(j)dj \right) = 0$$

$$\frac{\partial \mathcal{L}}{\partial n_x^B(i)} = -\lambda_N + \lambda_X(\psi - c_x^B) - \lambda_Y c_y^B = 0$$

$$\frac{\partial \mathcal{L}}{\partial n_y^A(i)} = -\gamma\theta f'\left(\frac{n_y^A(j)}{N_y^A}\right) - \lambda_N - \lambda_X c_x^A + \lambda_Y \left[f'\left(\frac{n_y^A(j)}{N_y^A}\right) - c_y^A\right] = 0$$

in conjunction with two corner labor allocations: $n_x^A(i) = n_y^B(i) = 0$. Manipulations similar to those in the proof of Theorem 3 above give (A1) and (A2) – so consumption bundles must still be the same across the two locations – together with:

$$[u'(c_y^A) - \gamma\theta N^A] f'\left(\frac{n_y^A}{N_y^A}\right) - \psi = 0 \tag{A10}$$

$$1 = k^B n_x^B + m^A n_y^A \tag{A11}$$

$$
\begin{aligned}
\psi k^B n_x^B &= m^A n_y^A c_x^A + k^B n_x^B c_x^B && \text{(A12)} \\
m^A N_y^A f\left(\frac{n_y^A}{N_y^A}\right) &= m^A n_y^A c_y^A + k^B n_x^B c_y^B && \text{(A13)}
\end{aligned}
$$

Under a stratified configuration, we have: $n_x^A = n_y^B = 0$, $n_y^A = n_y$, $N^A = Mn_y^A$, $N^B = 1 - Mn_y^A$, and $n_x^B = n_x = 1 - Mn_y$. Whereas clean good consumption still takes the same form as in (A7), (A4) and (A6) together yield:

$$c_y = M^2 f(\frac{1}{M})n_y \tag{A14}$$

implying again that both (location-specific) Pareto optimal consumption bundles are identical to the equilibrium ones. From (A10) and (A14), we have:

$$\left[u'\left(M^2 f(\frac{1}{M})n_y\right) - \gamma\theta M n_y\right] = \frac{\psi}{f'(\frac{1}{M})} \tag{A15}$$

Since n_y in a stratified equilibrium is determined by (35), and since $\widetilde{\sigma}(\frac{1}{M}) = f(\frac{1}{M}) - \frac{1}{M}f'(\frac{1}{M}) > 0$, we can see that:

$$\frac{\psi}{Mf(\frac{1}{M})} < \frac{\psi}{f'(\frac{1}{M})}$$

This implies that n_y in a stratified equilibrium exceeds the Pareto optimal level, which completes the proof. ■

Table 1. Population Accounting

Integrated	City A		City B		Aggregate	
	Firms	Workers	Firms	Workers	Firms	Workers
Clean Sector	$k^A = \frac{1}{2}$	$N_x^A = \frac{1}{2} n_x$	$k^B = \frac{1}{2}$	$N_x^B = \frac{1}{2} n_x$	1	n_x
Dirty Sector	$m^A = \frac{M}{2}$	$N_y^A = \frac{M}{2} n_y$	$m^B = \frac{M}{2}$	$N_y^B = \frac{M}{2} n_y$	M	$N_y = Mn_y$
Total	$\frac{1}{2}(1+M)$	$\frac{1}{2}$	$\frac{1}{2}(1+M)$	$\frac{1}{2}$	$1+M$	1

Stratified	City A (Dirty)		City B (Clean)		Aggregate	
	Firms	Workers	Firms	Workers	Firms	Workers
Clean Sector	0	0	1	$N_x = n_x$	1	$N_x = n_x$
Dirty Sector	M	$N_y = Mn_y$	0	0	M	$N_y = Mn_y$
Total	M	Mn_y	1	$N_x = n_x$	$1+M$	1

Figure 1. Labor Allocation Under Integrated Equilibrium

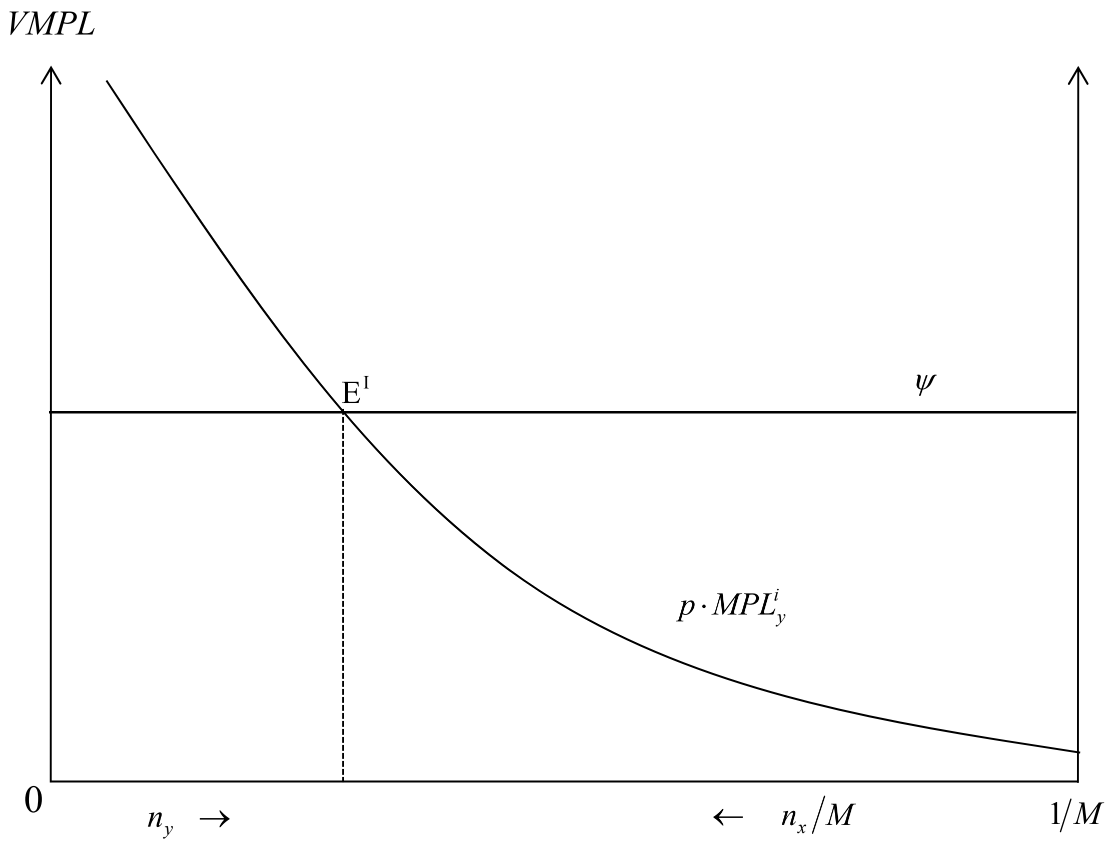


Figure 2. Dirty Good Equilibrium under Integrated Configuration

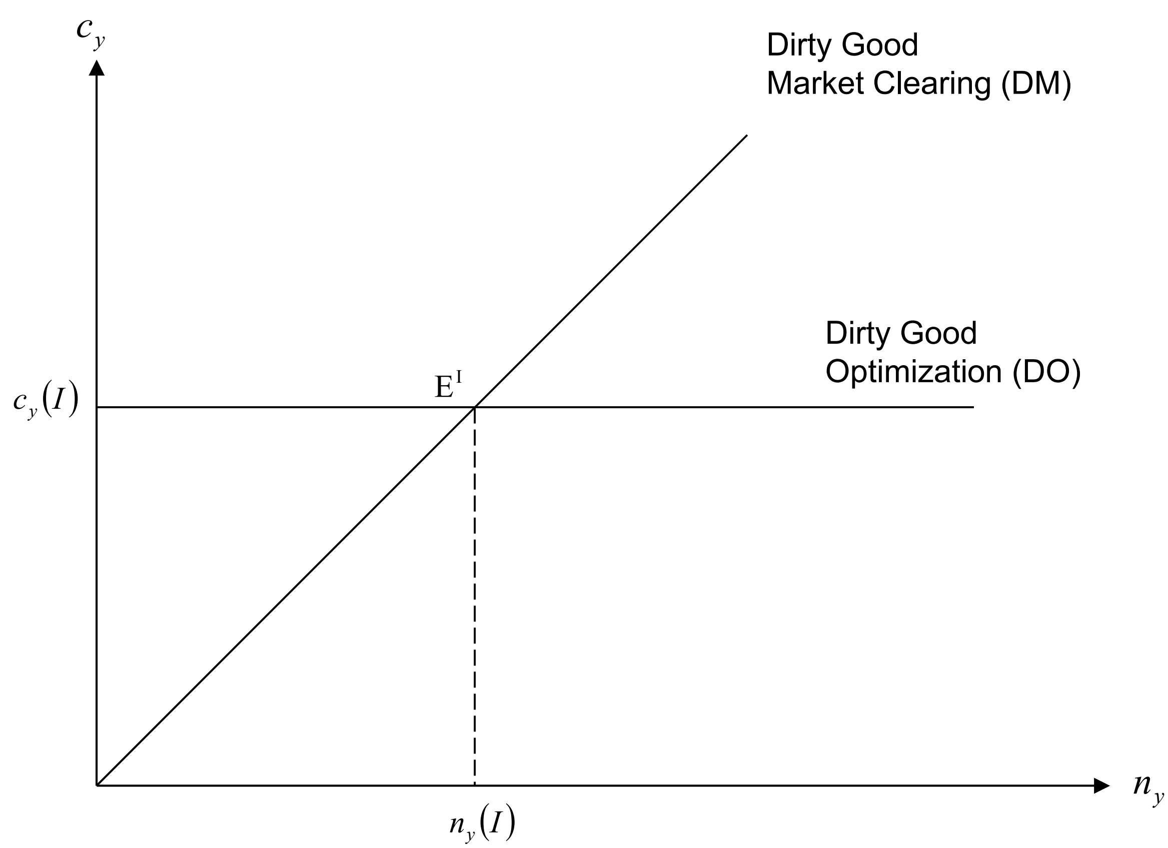


Figure 3. Locational Equilibrium Under Stratified Configuration

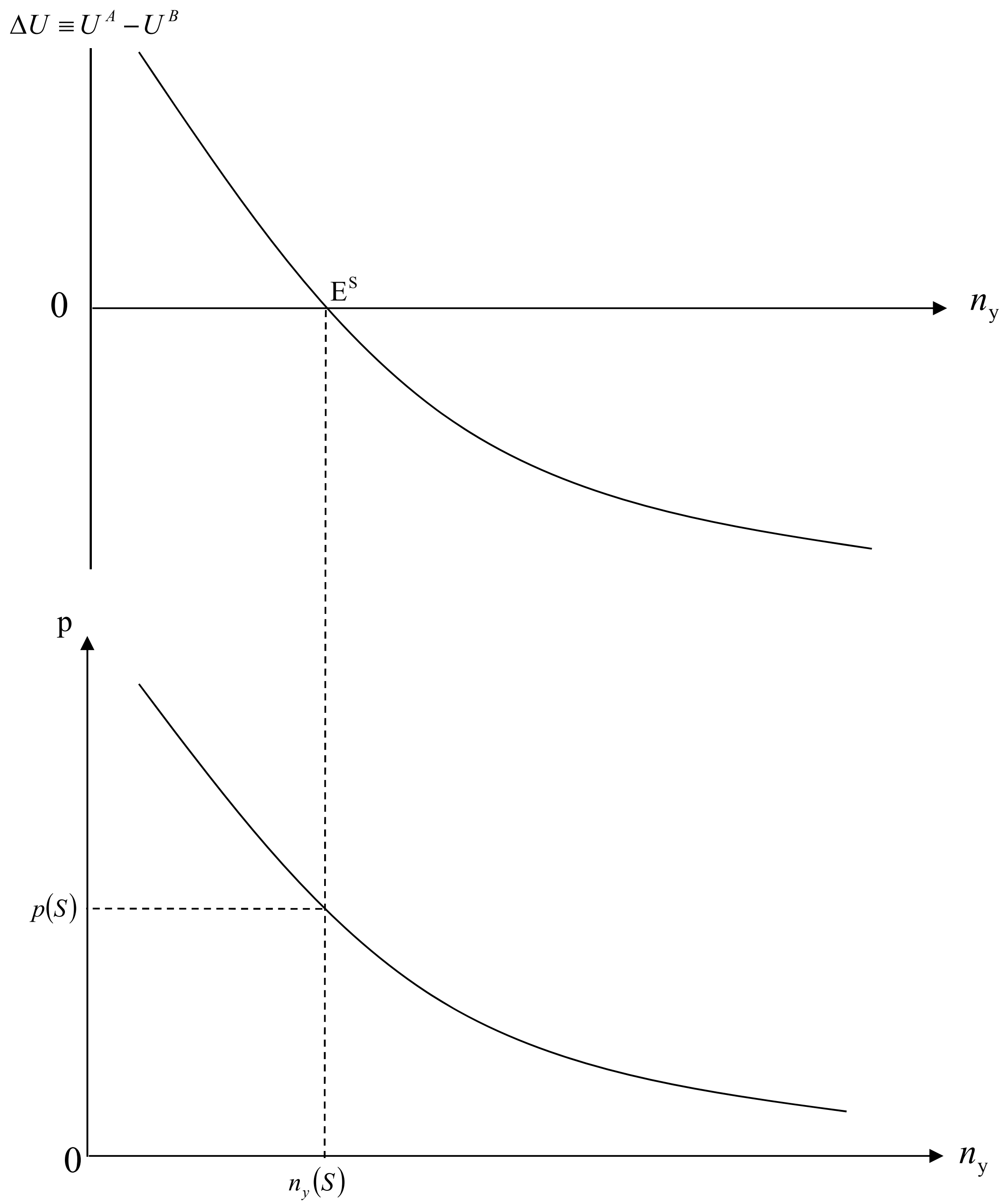


Figure 4. Integrated vs. Stratified Equilibrium

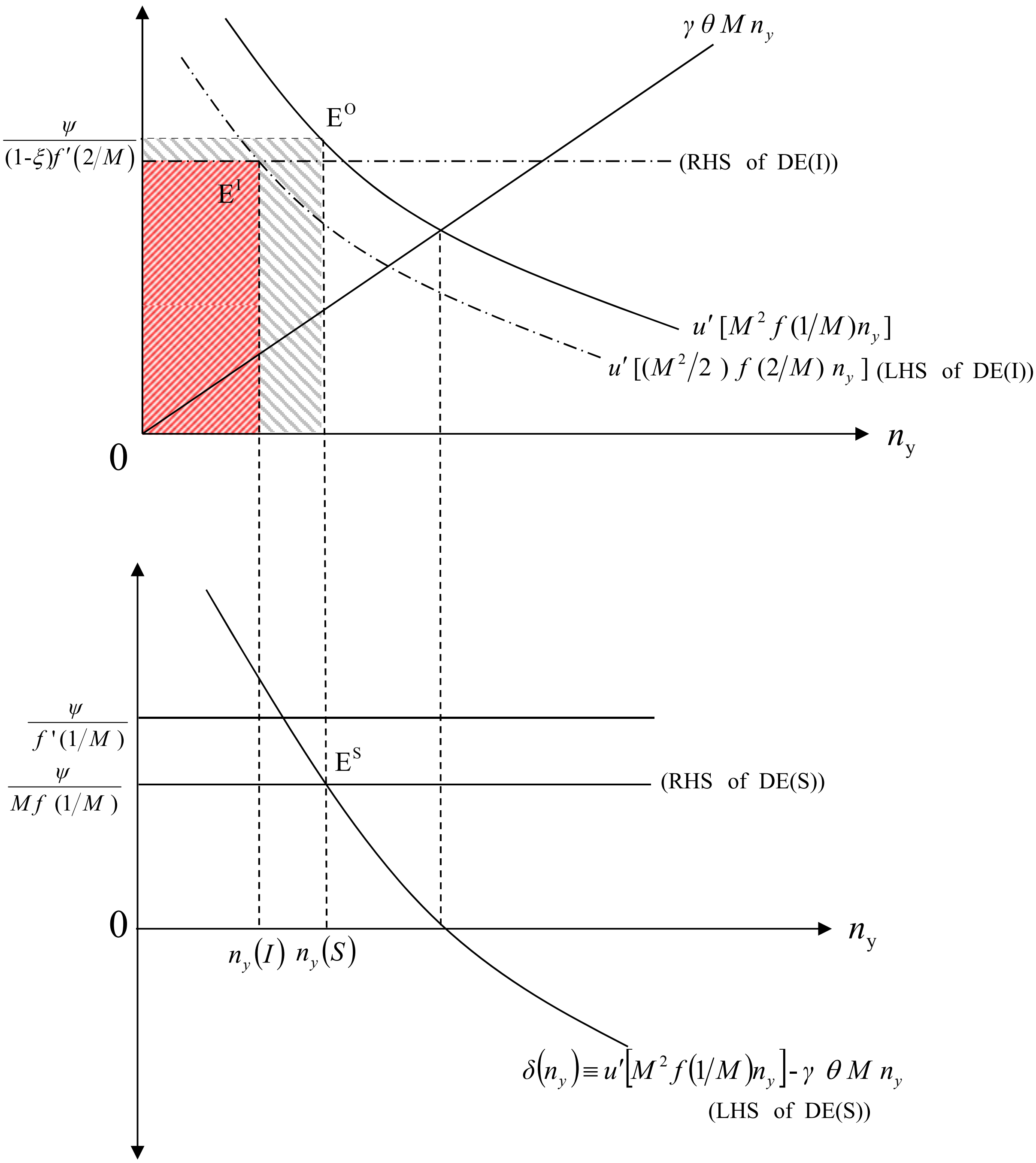
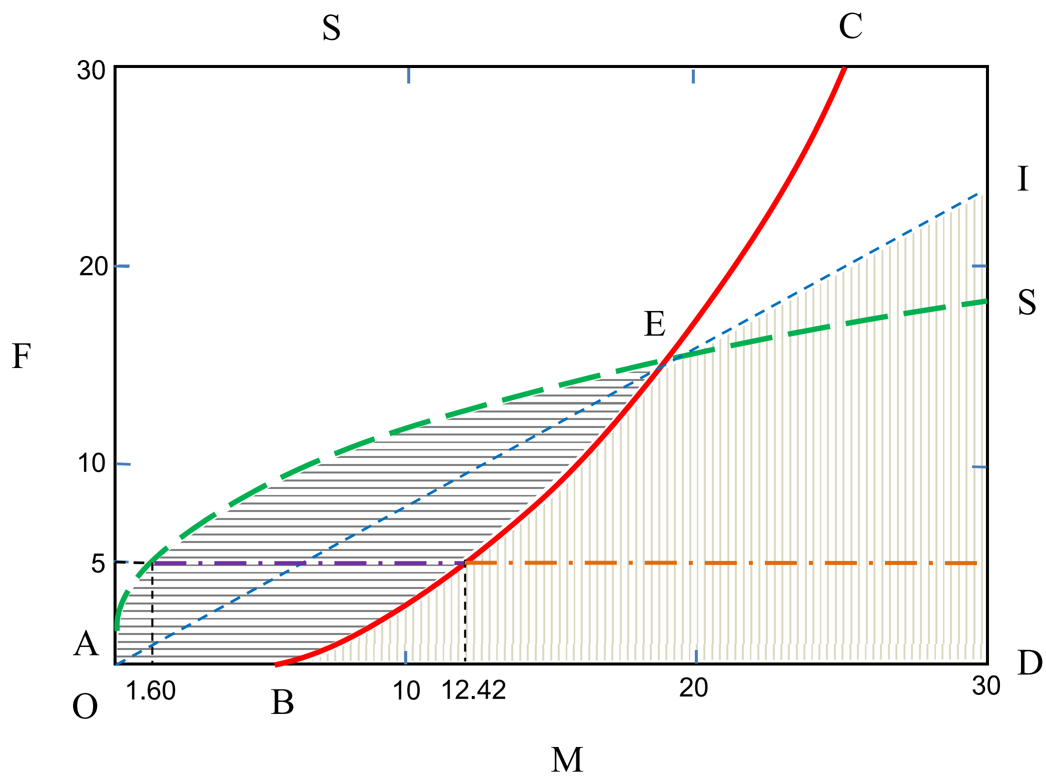


Figure 5. Equilibrium Configuration

(a) Equilibrium Classification



(b) Bifurcation Diagram

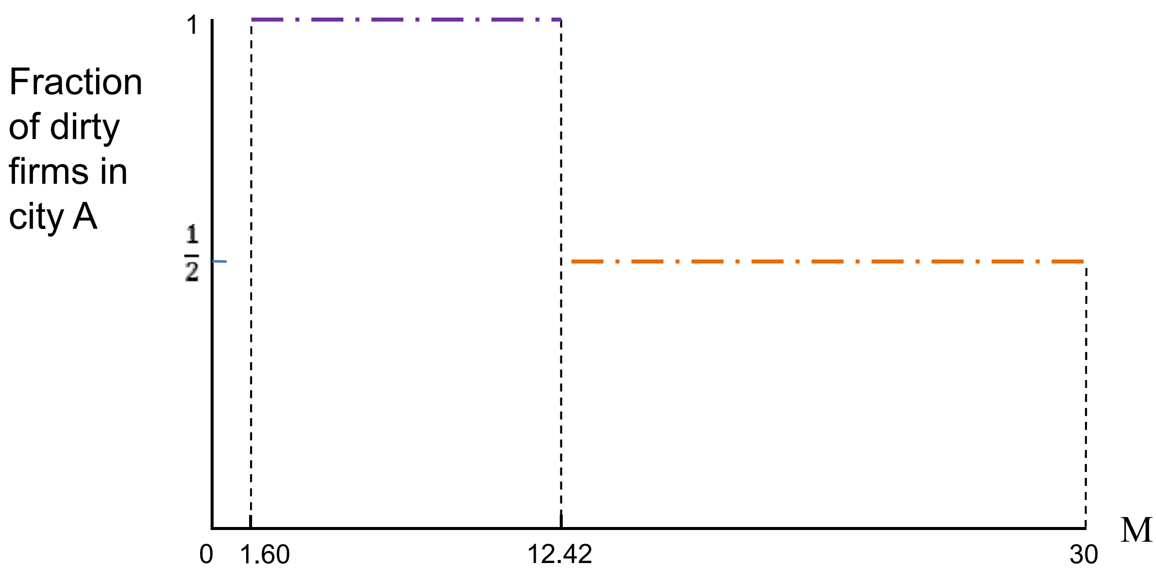
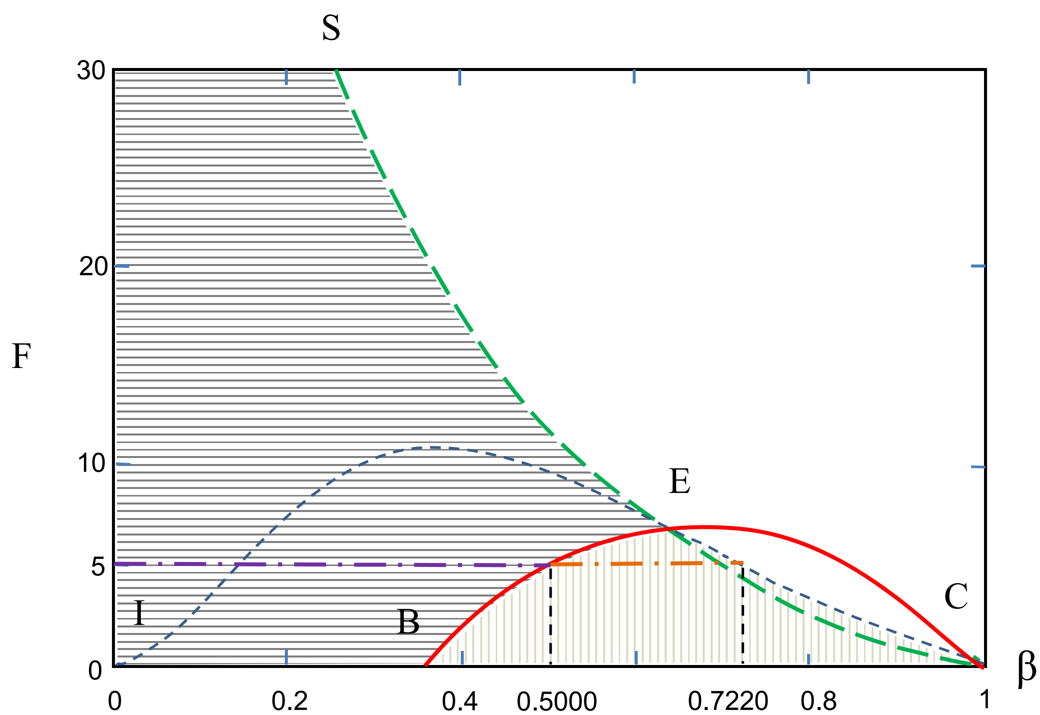


Figure 6. Equilibrium Configuration with β

(a) Equilibrium Classification



(b) Bifurcation Diagram

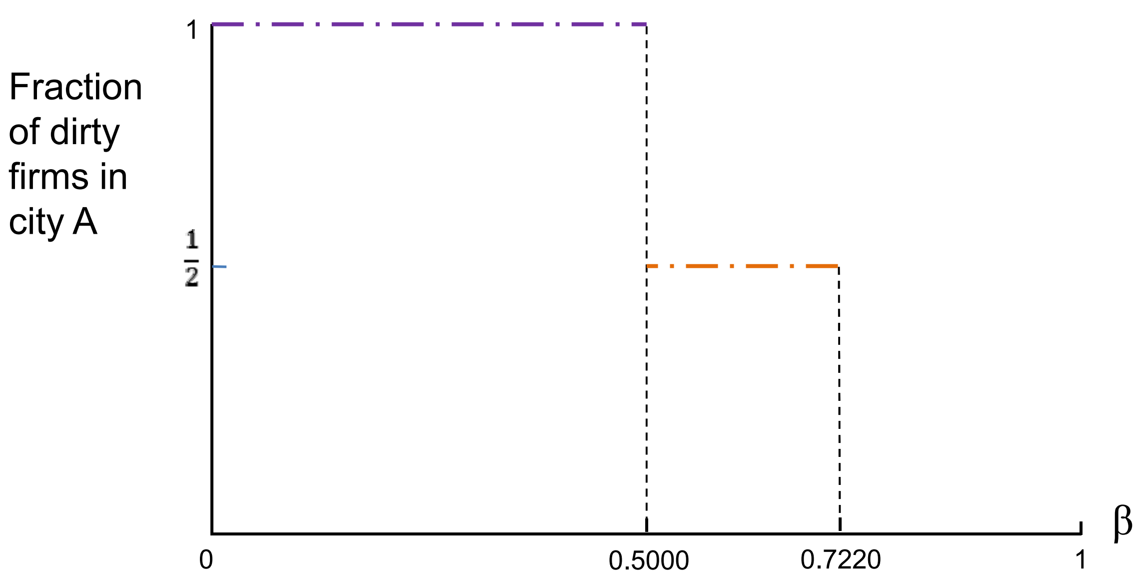


Figure 7. Tax Effects in Stratified Equilibrium

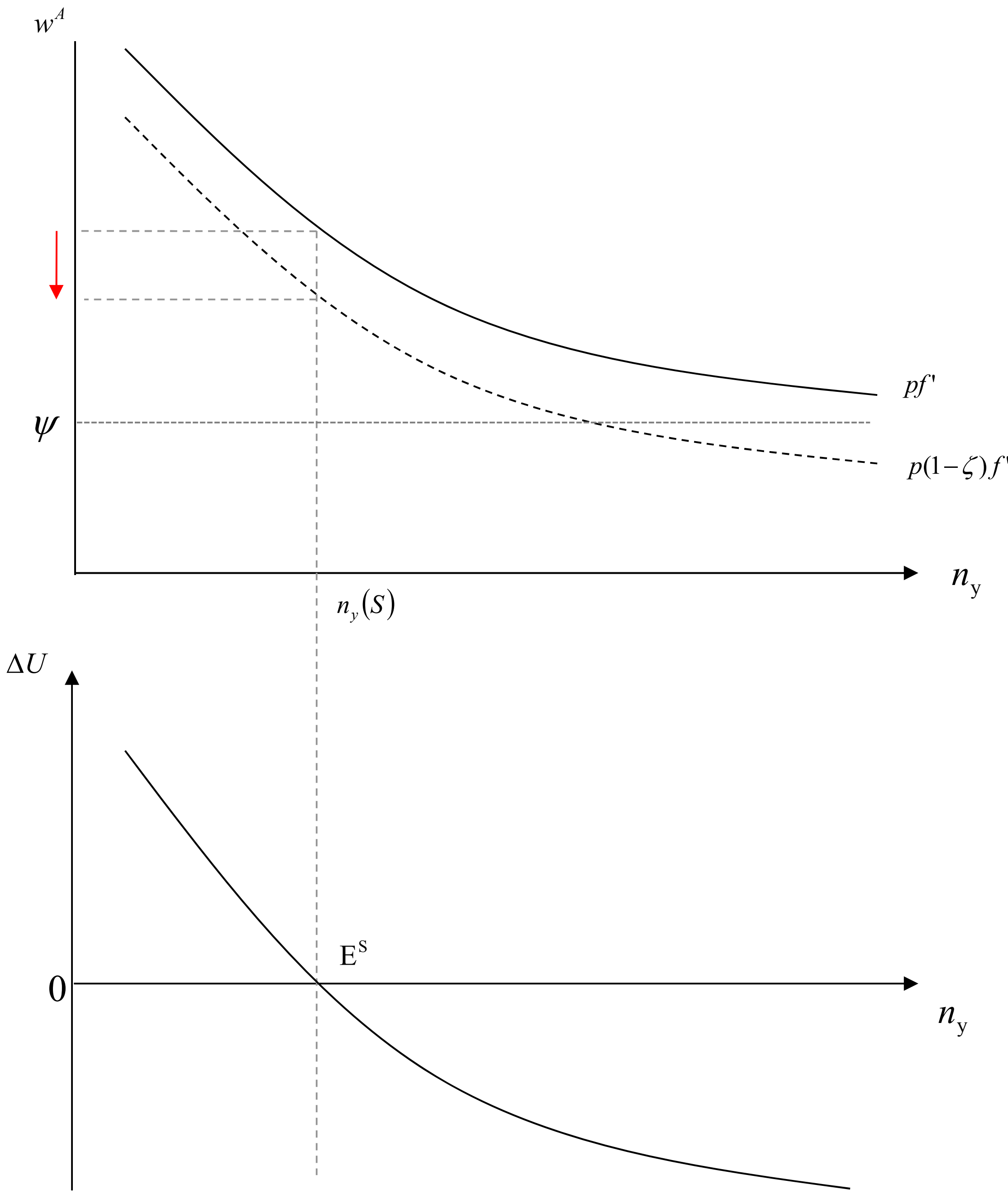
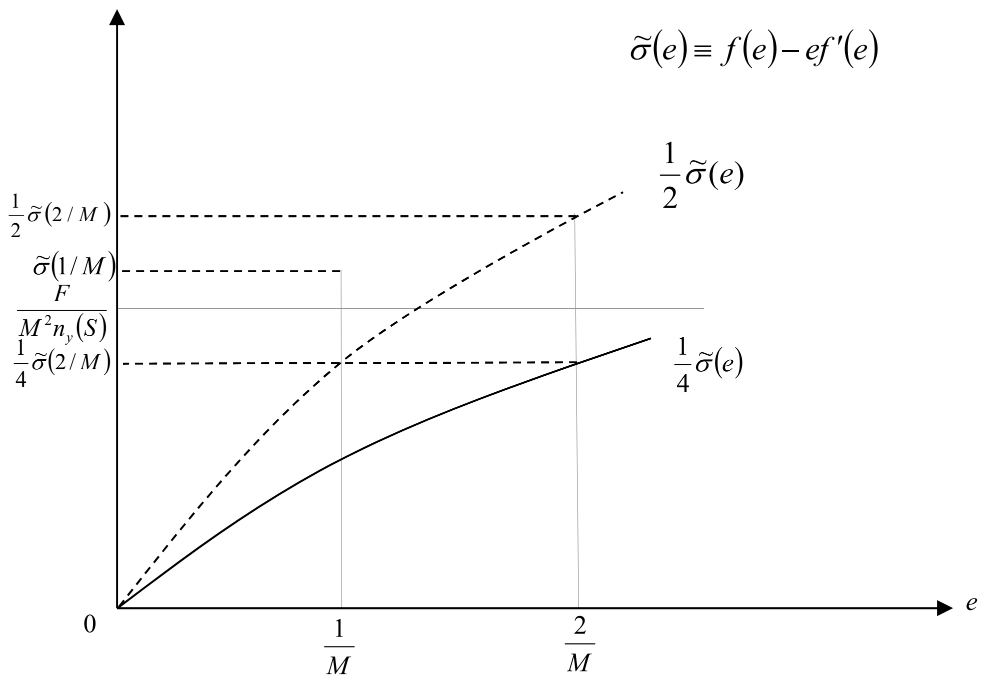


Figure 8. Surplus from Uncompensated Spillovers

a. Fixed Pollution Tax Regime: Conditions S-1 and R-1



b. Linear Pollution Tax Regime: Conditions S-2 and R-2