7-6-2016

Cap and trade under transactions costs and factor irreversibility

Rajesh Singh
Iowa State University, rsingh@iastate.edu

Quinn Weninger
Iowa State University, weninger@iastate.edu

Cap and trade under transactions costs and factor irreversibility

Abstract
We study production capacity utilization and emission permit utilization in a model where firms jointly produce a valued good and an environmental bad, pollution. Firms are ex ante identical but experience random productivity shocks after factor employment. A regulator imposes a cap-and-trade policy to control pollution emissions. Trade in emission permits entails transactions costs which follow two specifications: constant per unit trading costs or fixed trading costs. Under constant per unit trading costs, the equilibrium outcome depends only on the total unit trading costs; the incidence of costs borne by buyers and sellers does not matter. Under fixed costs, both buyers' and sellers' costs matter. Under proportional costs permit trade always occurs, with either full or partial market clearing, as long as the total trading costs are below the permit trade surplus. With fixed costs, trade is either partial or non-existent. The implication is that firms fully utilize their production capacity for a range of proportional trading costs; capacity is never fully utilized under fixed costs. Under proportional costs, trade is impeded most, even with small costs, when the emission cap is either relatively high or low. There exists a non-monotonic relationship between the aggregate emissions cap and a lower bound for trading costs that obstruct or preclude trade. Under fixed costs, a similar relationship between emission cap and the cost threshold that precludes trade holds only if the output variance is exogenously fixed. Otherwise, the higher the emission cap the higher is this cost threshold. In contrast to proportional costs where capacity utilization decreases with productivity variance, the result is the opposite under fixed costs.

Keywords
cap-and-trade regulation, transactions costs, permit trade, capacity and permit utilization

Disciplines
Economics

Cap-and-Trade under Transactions Costs and Factor Irreversibility

Rajesh Singh*and Quinn Weninger†

Abstract

We study production capacity utilization and emission permit utilization in a model where firms jointly produce a valued good and an environmental bad, pollution. Firms are ex ante identical but experience random productivity shocks after factor employment. A regulator imposes a cap-and-trade policy to control pollution emissions. Trade in emission permits entails transactions costs which follow two specifications: constant per unit trading costs or fixed trading costs. Under constant per unit trading costs, the equilibrium outcome depends only on the total unit trading costs; the incidence of costs borne by buyers and sellers does not matter. Under fixed costs, both buyers' and sellers' costs matter. Under proportional costs permit trade always occurs, with either full or partial market clearing, as long as the total trading costs are below the permit trade surplus. With fixed costs, trade is either partial or non-existent. The implication is that firms fully utilize their production capacity for a range of proportional trading costs; capacity is never fully utilized under fixed costs. Under proportional costs, trade is impeded most, even with small costs, when the emission cap is either relatively high or low. There exists a non-monotonic relationship between the aggregate emissions cap and a lower bound for trading costs that obstruct or preclude trade. Under fixed costs, a similar relationship between emission cap and the cost threshold that precludes trade holds only if the output variance is exogenously fixed. Otherwise, the higher the emission cap the higher is this cost threshold. In contrast to proportional costs where capacity utilization decreases with productivity variance, the result is the opposite under fixed costs.

Key words: Cap-and-trade regulation, transactions costs, permit trade, capacity and permit utilization.

JEL code: Q5, L5

* *Singh (Corresponding Author)*: Associate Professor, Department of Economics, 260 Heady Hall, Iowa State University, Ames IA 50011-1070; Phone 515-294-6203; Fax: 515- 294-0221; Email: `rsingh@iastate.edu`

† *Weninger:* Associate Professor, Department of Economics, 260 Heady Hall, Iowa State University, Ames IA 50011-1070; Phone 515-294-5886; Fax: 515-294-0221; Email: `weninger@iastate.edu`

1 Introduction

We study production of a consumption good and jointly produced pollution emissions in a stochastic environment. In our model, emissions are a by-product of a valued-good production process and increase with the quantity of the good that is produced (Murty et. al., 2012). A regulator sets an emissions cap that is initially allocated/auctioned to ex ante identical firms. The regulation limits the quantity of emissions and indirectly, the quantity of consumption good produced per firm and in aggregate. The role of markets for trading emissions permits arises naturally due to the uncertainty firms face in production. In our model, firms experience idiosyncratic productivity shocks after factors are employed. Firms that experience high productivity want to produce a relatively higher amount of goods and to do so they need to purchase more permits than they hold. On the other side of the market, there are firms with relatively lower productivity who hold permits that exceed their emissions.[1] The purpose of our study is to understand the efficiency implications of transactions costs in the markets for these emissions permits.

Our model is designed to allow a version of the first welfare theorem to hold: a perfectly competitive market for costless trade in emission permits produces a *constrained efficient* equilibrium (Montgomery, 1972). Under decentralization, firms employ factor inputs, trade in permits, and produce output that is identical to the allocations of a planner who is constrained by a total emissions target. Efficiency requires that firms ex ante choose a production scale such that in the aggregate full capacity emissions obey the cap. *Ex post*, i.e., after realizing their productivities, firms with excess emission permits trade with those with shortfalls and the permit market clears. An efficient production-emission plan is thus implemented by a joint determination of optimal factor employment and permit market prices. For example, if the cap is relatively low emission permit prices are relatively high that by raising the opportunity cost of production ensure a suitably low factor employment. The reverse is the case when the cap is relatively high.

In the absence of permit trade, the equilibrium is inefficient. While some firms are unable to fully utilize their production capacity due to emission constraints, others let their permits go unutilized. Ex ante, firms optimally choose a higher production capacity relative to the capacity chosen under trade when the emission cap is sufficiently low. Conversely, when the cap is sufficiently high, the capacity chosen by firms is relatively lower. In either case, there is a *latent* supply-demand mismatch: under high emissions cap, for example, the aggregate supply of permits exceeds the emissions that could potentially be generated under full capacity production. The opposite is the case when the emission cap is relatively low.[2] This mismatch continues to exist and drive the equilibrium outcomes when trade occurs with transactions costs.

While the costless trade and no-trade environments build a strong case for a cap-and-trade programs, trade in permits is neither costless nor prohibited in practice. We study intermediate cases where trade entails finite transaction costs, e.g., the costs of collecting information, bargaining, monitoring and enforcing sales agreements, and perhaps, costs of fulfilling regulatory paperwork or other administrative requirements. Empirical evidence suggests these costs can be of fixed form, independent of trading volume, and/or proportional to quantity traded.[3]

---

[1] Alternatively, the model can be modified to allow firms to differ in their emission efficiencies ex post. Qualitatively, the results are similar.

[2] If trade were to occur ex post the latent mismatch would transform into an excess supply of permits when emission caps are relatively high and excess demand for permits when emission caps are low.

[3] Heindl (2012) finds transactions costs in the European Union emissions trading scheme to be non-linear and declining with the size of emitting firm. Generally, empirical estimates vary depending on the age of the permit trading program among other factors. Gangadharan (2000) estimates transactions costs reduced the likelihood of a permit trade in the Los Angeles Regional Clean Air Incentives Market by 32% in 1995 and 12% in 1996. Kerr and Maré (1998) estimate transactions costs in the US Lead Phasedown program account for between 10-20% of potential trade surplus. Similar findings are reported by Jaraitė and Kužukauskas, 2012. Estimate of transactions costs associated with the Clean Development Mechanism are considered to be large (Michaelowa et al., 2003).

We contrast production and emission outcomes under both fixed and proportional transactions costs. We derive equilibrium factor employment, potential and actual production of the good output, the emissions, and permit prices under the two forms for trading costs. We introduce two metrics, *capacity utilization* and *permit utilization,* for evaluating inefficiencies generated by trading costs. The former is defined as the ratio of actual output produced to full capacity output, and the latter is the ratio of permits utilized relative to the cap issued by the regulator. The key results are summarized by two propositions.

Under proportional costs, firms' input employment and aggregate potential output (capacity) depend on the sum of a buyer's and a seller's per permit trading costs (henceforth, *total* costs) and not on the buyer's and the seller's share of these costs. As long as the total cost falls below a threshold, the equilibrium is efficient: production capacity as well as emission permits are fully utilized as under costless trade. The explanation is the following. For the production-emission outcome under trading costs to match that under costless trade, two conditions have to be met: (1) an ex ante optimal factor employment condition that roughly requires expected marginal permit cost to match the permit price under costless trade, and (2) an ex post participation constraint that requires that both buyers and sellers participate in the permit market. Condition (1) roughly demands that the mean of the buyers' and sellers' opportunity cost of emissions, i.e., the buyers' per permit expenditure and the sellers' per permit sales income, equal the permit price under costless trade. Condition (2) places an upper bound on a buyer's expenditure to ensure non-negative surplus from a permit purchase, and a floor of zero on a seller's income from a permit sale. Finally, by placing a wedge between what a buyer pays and what a seller receives for a permit, trading costs engender a third condition: (3) For each permit trade, the buyers' expenditure and the sellers' income differ precisely by the total trading costs - an accounting identity that must hold.

If one of the participation constraints in (2) binds, (1) and (3) can not hold simultaneously for all values of trading costs. To see this, suppose (1) holds and the equilibrium is efficient. Then the participation constraint in (2) along with condition (1) determine per permit buyers' expenditure and sellers' income, which in turn uniquely determine the *total trading cost threshold* that satisfies (3). This cost threshold is non-monotonic in emissions cap: it first rises with the cap and then falls. To understand this, consider first a sufficiently small emission cap such that the permit price under costless trade is near the upper bound of the expenditure a permit buyer is willing to incur. With trading costs, a relatively small cap induces firms to ex ante overemploy factors as in the absence of permit trade. Ex post, there is an excess demand for permits raising its price such that the buyers' expenditure hits its upper bound, i.e, the participation constraint binds. Now (1) can only be satisfied if a buyer's per permit expenditure and a seller's per permit income, both, are sufficiently close to the permit price under costless trade. That is, their difference (i.e., total trading cost) is sufficiently small. Conversely, consider a sufficiently large cap such that the permit price under costless trade is close to zero. In equilibrium, permits are in excess supply and sellers' participation constraints bind. Now satisfying condition (1) requires that buyers' expenditure and sellers' income, per permit, be close to zero, thus setting the total cost threshold also close to zero. Finally, with an intermediate emission cap the permit price under costless trade is sufficiently far from the participation bounds in (2), which gives enough room for a buyer's per permit expenditure and a seller's per permit income to differ substantially before one of the participation constraints in (2) triggers. As a result, the cost thresholds are relatively higher.

For any emission cap if the total cost lies below the threshold, either buyers' per permit expenditure has to decline or sellers' per permit income has to rise in a manner that the binding participation constraint in (2) becomes slack. Now (1) and (3) hold simultaneously and the equilibrium is efficient. On the other hand, if the total cost is above the threshold, the participation constraint continues to bind, but (1) can no longer hold. Some trade, however, always occurs as long as the total trading cost remains below the market value of goods produced due to the marginal permit. Consider a relatively small cap with total cost above the threshold but below the market surplus due to an additional permit use. There is an excess demand for permits and the buyers' participation constraints bind. All emission permits are utilized but some firms end up with an idle production capacity. Conversely, with large emission caps there is an excess supply of permits and the sellers' participation constraints bind. Now all firms utilize their full production capacity though some permits remain unutilized.

Fixed trading costs preclude full utilization of either production capacity or permits because some emission-constrained firms with idle production capacity and some firms with unutilized permits optimally choose not to enter the permit market. The second proposition in the paper formally establishes a joint (two-dimensional) upper bound on buyers' and sellers' trading costs that separate cap-dependent production-emission equilibria with and without trade. Here, both buyers' and sellers' costs matter for trade. If the emission cap is relatively small, the amount of unutilized permits are low and a small fixed trading cost discourages sellers from entering the permit market, even if buyers are willing to forfeit their entire surplus to buy a permit. The opposite is the case when the cap is relatively large: few firms have ex post excess production capacity and even with small trading costs they prefer not to enter the market even when the permits are offered for free. Under the assumption that sellers' and buyers' costs are equal, the two-dimensional upper bound on fixed costs converges to a unidimensional upper bound, as under proportional costs, and a non-monotonic relationship between emission caps and the common trading cost reappears: When the emission caps are set either relatively low or high even small costs preclude trade, whereas for the intermediate values of caps, the cost bounds are higher.

Under both types of trading costs, there is an excess demand for permits and some production capacity remains idle for relatively low emission caps. In such cases, capacity utilization is increasing in the cap. Conversely, with relatively high emission caps, there is an excess supply of permits and some permits are wasted. The higher the cap, the lower is the permit utilization. For any cap, the permit demand-supply gap widens with an increase in productivity shocks. However, the capacity/permit utilization response to the variance of shocks is opposite under the two cost specifications. Under proportional costs all firms participate in the permit market and an increase in productivity variance aggravates the utilization problem by widening the demand-supply gap. Under fixed costs, a rise in the output variance increases the mass of firms with an excess demand and/or supply who now find it optimal to enter the permit market. A higher output uncertainty now alleviates the utilization problem.

Earlier literature has examined the impact of transactions costs on the performance of cap-and-trade environmental regulations.[4] In a model of decentralized cap-and-trade emissions regulation, Montgomery (1972) showed that with frictionless permit trading an aggregate abatement target can be achieved at minimum cost, and that the abatement outcome is independent of the initial distribution of pollution rights. Stavins (1995) studied the cost efficiency of transactions costs in a setting where firms employ a convex abatement technology and are initially heterogeneous in terms of emissions permits holdings and/or abatement demand. Stavins (1995) showed, in a partial equilibrium framework, that cost-efficient abatement outcomes do not emerge in the presence of non-linear transactions costs.[5]

These authors and others in the cap-and-trade literature have studied pollution abatement and emission reduction by employing an increasing and convex cost-of-abatement function. In this framework, the input and output choices of firms are assumed independent of emissions abatement decisions (see Färe et. al., 2013 for an exception). Most papers adopting this approach also assume a social objective of minimizing the cost of meeting a given emissions target (e.g., Montgomery, 1972; Stavins, 1995; Zhao, 2003). Reigning wisdom holds that transaction costs, generally, act as a wedge between marginal abatement costs and benefits and thereby preclude efficient abatement.

Our work departs from earlier literature in two important ways. We derive permit prices, factor allocations, industrial output and emissions under permit market transactions costs in a general equilibrium model. We track the transactions cost wedge through buyer and seller interactions, equilibrium permit prices, and its effect on the ex ante factor allocation decisions of firms. We show that reigning

[4] See Krutilla and Krause (2010) for a review.

[5] Montero (1997) extends Stavins' (1995) analysis by introducing uncertainty in approval of permit trades by the regulator. This essentially adds another layer of transactions costs. The analytical results of this paper however rely on an exogenous permit price.

wisdom does not always hold, particularly when transactions costs are proportional and not too large. Second, our model features joint production of a good and an environmentally bad output.[6] We focus explicitly on the link between emission regulations and firms' production and investment plans and, as a consequence, derive industry-specific metrics for evaluating permit program performance in various trading environments.

Our assumption of irreversibility of factor inputs is akin to Zhao (2003), who shows uncertainty in abatement costs discourage investment under fixed emission charges relatively more than under cap and trade. In our model, the irreversible factor proxies for a firm's installed capital and/or long-term labor contracts. A setting in which our model fits well is energy intensive industries operating under cap-and-trade CO_2 emissions allowance regulation, although the model applies broadly across polluting industries.[7]

The rest of the paper is organized as follows. We present our basic production and emission model in the next section, which also lays out the costless trade and no-trade equilibrium as the two limiting benchmarks. The following two sections analyze the model's equilibrium with proportional and fixed trading costs, sequentially. Concluding remarks are offered in Section 5. An appendix, Section 7, provides all algebraic proofs while extensions of the benchmark model are discussed in Section 8

2 Model

There is a consumer good with price p, produced by a unit mass of ex ante identical firms indexed by $i \in [0,1]$. A common production technology denoted by $y(n,\varepsilon)$ is used by all firms, where n denotes the quantity employed of a composite factor with its purchase price w, and ε is a firm-specific idiosyncratic productivity shock. Shocks are independently and identically distributed, with $g(\varepsilon)$ and $G(\varepsilon)$ respectively denoting the pdf and cdf of $\varepsilon \in \left[\varepsilon^{\min}, \varepsilon^{\max}\right]$. We assume that $y \geq 0,\ y_n \geq 0,\ y_\varepsilon > 0$, $y_{nn} \leq 0$, and $y_{n\varepsilon} > 0$; $y(0,\varepsilon) = 0\ \forall\varepsilon$, $\lim_{n\to 0} y_n = \infty$, and $\lim_{n\to\infty} y_n = 0$, where subscripts denote differentiation with respect to the subscripted argument. Below we assume that n is chosen before firm-specific ε is realized and that n is irreversible. In equilibrium, all firms choose a common n. Let $y^{\min}(n) \equiv y\left(n,\varepsilon^{\min}\right)$ and $y^{\max}(n) \equiv y\left(n,\varepsilon^{\max}\right)$. Let $y^f(n) \equiv E\{y(n,\varepsilon)\} \equiv \int_\varepsilon y(n,\varepsilon)\ g(\varepsilon)\ d\varepsilon$. Note that y^f denotes the expected *full-capacity* output of a firm if it did not expect to face any further production constraints. Since there is no aggregate uncertainty in the model, y^f also stands for the aggregate output under full capacity utilization. Whenever obvious, we use y^f, $y^{\min}$, and $y^{\max}$ below without n as an explicit argument.

It is assumed that producing y units of goods generates $z(y)$ units of a pollutant as a by-product.[8] Regulations require a firm i to cap its emission up to its holding of emission permits. The emissions are monitored and perfectly enforced. Let e^R denote the total amount of permits (the emission cap) issued by the regulator. If firm i holds e^i units of emission permits at the time of production, its emissions must follow $z^i \equiv z\left(y^i\right) \leq e^i$. We assume that $z_y > 0$. Empirical evidence that can guide the sign of z_{yy} is not available, and henceforth for analytical convenience we assume that $z(y) = \gamma y$. A firm i with factor employment n and shock ε^i can produce at its full capacity up to $y\left(n,\varepsilon^i\right)$. Actual output y^i

[6] Our approach to modeling pollution technology as joint production of a consumer good and an environmentally bad output follows Fare, Grosskopf, Noh, and Weber (2005), Murty et al. (2012) and others. We assume weak output disposability in emissions, while maintaining strong output disposability in production of the consumption good.

[7] Cap-and-trade regulatory programs for CO_2 emissions now exist in Australia, Japan, New Zealand, Switzerland and the United States. Canada, China and South Korea are in the process of developing their own programs. The European Union Emissions Trading System covers 45% of total European Union emissions. Firms for which our model is particularly applicable include power and heat generating firms, and firms in the energy-intensive sector (oil refineries, firms producing steel and iron and other metals, ceramics, pulp, paper and chemicals).

[8] An alternative assumption could be that firms have a deterministic technology $y(n)$ for consumer goods and shock ε^i instead is specific to emissions: y units of output produced by a firm i generates $z\left(\varepsilon^i, y\right)$ units of emissions. Appendix 7.3 shows that the two specifications are equivalent.

however may be constrained by its emission permits:[9]

$$y^i \equiv \min\left\{y\left(n,\varepsilon^i\right), e^i/\gamma\right\}, \tag{1}$$

A production plan consists of two-stage ex ante and ex post decisions, before and after firms' productivity shock is realized. In the first stage, that we think of as the planning stage, firms trade and acquire emission permits e^{i0} in a pre-production market at per unit price ρ, and employ n to maximize expected profits. The productivity shock is realized at the production stage: after permits are purchased and n is employed. At this stage, firms choose $\left\{\left(e^i, y^i\right) : \gamma y^i \leq e^i\right\}$ depending on the permit market environment. For example, with permit trade, firm i with $\gamma y\left(n,\varepsilon^i\right) > e^{i0}$ can choose to purchase additional emission permits from firm j with $\gamma y\left(n,\varepsilon^j\right) < e^{j0}$. Thus, e^i denotes the production-stage permit holding of firm i. To differentiate between planning and production stage permit prices, we denote the latter by r.

Our assumption of irreversible inputs refers to a situation where firms build production capacity through irreversible investments in capital and/or sign binding labor contracts. The associated investment and labor costs are sunk prior to the realization of productivity shocks.[10] We assume throughout that permit trading at the planning stage is frictionless. A few words of justification for this assumption are warranted. First, planning stage permit adjustments are not motivated by the need to balance unanticipated emissions with permit holdings. It is only after firms realize their potential productivity shock that they adjust permit holdings to balance their actual emissions to their permitted level. The latter motive for trade is the focus of our analysis. Second, since firms in our model are ex ante identical, there is no planning stage permit adjustment beyond the initial allocation of cap/permits from the regulator.[11]

We assume that buyers and sellers of permits at the production stage must incur a cost of κ^b and κ^s, respectively, that can be either a fixed cost of entering the permit market, or proportional to the quantity of permits traded. Under fixed costs, buyers must pay κ^b and sellers must pay κ^s to engage in permit trading, irrespective of the quantity traded. Under proportional costs, if the market permit price is r, a seller receives $r - \kappa^s$ per unit traded, while a buyer pays $r + \kappa^b$. In both cases, if $\kappa^b = \kappa^s = 0$, trade is costless. If either of κ^b or κ^s is prohibitively large, no trade will occur, and firm i will produce and market $y^i = \min\left\{y\left(n,\varepsilon^i\right), \frac{e^{i0}}{\gamma}\right\}$.

We are now ready to study equilibria under alternative trading environments. For concreteness, we will also employ parametric examples and compute various quantities of interest, such as full capacity and realized outputs, permits resold and unsold in the market along with the equilibrium permit prices in the planning and the production stage. A unit mass assumption for firms, their ex ante similarity, and an absence of aggregate uncertainty implies that firms' expected outcomes at the planning stage coincide with ex post aggregate outcomes. Let $y^a \equiv \int y^i di$ denote the aggregate output produced by firms, where y^i follows (38). Then, aggregate emission $z^a \equiv \int z^i di = \gamma y^a$. Thus, $e^a \equiv z^a$ denotes the aggregate permits utilized. Two metrics are used to compare equilibria under various trading environments: industry-wide *capacity utilization* which we express as ratio y^a/y^f and *permit utilization,* expressed as e^a/e^R. We say an equilibrium is *efficient* when $y^a/y^f = e^a/e^R = 1$.

It is useful to characterize three benchmark equilibria: (i) no regulation; (ii) regulation with costless permit trade and; (iii) regulation without permit trade. A tilde ($\tilde{\ }$) is placed over endogenous variables n, ρ, y^f, y^a, z^a, and e^a to denote outcomes in the no regulation case. An asterisk (*) and hat symbol ($\hat{\ }$) are used under the costless permit and no permit trade benchmarks, respectively.

[9] Our technology can be taken as a special case of more general specification offered by Färe et al. (2005).

[10] An alternative characterization would allow a portion of the input n to be reversible, perhaps at a cost. The case of input reversibility is discussed in Appendix 8.2.

[11] Alternatively, we can assume that firms have an identical initial allocation of emission permits at the planning stage and simply abstract from permit trade at this stage. None of the results in the paper will change. The shadow price of these permits then equals the market price under frictionless trade. Having a planning stage frictionless permit trade essentially allows us to quantify the value of these permits for the regulator.

Unregulated emissions Here, a firm's expected profit is

$$E\{\pi(n)\} = p\ y^f(n) - w\ n;$$

which obtains $\tilde{n}$ as a solution to

$$y_n^f(\tilde{n}) = \frac{w}{p}. \tag{2}$$

Thus, aggregate output $\tilde{y}^a = \tilde{y}^f$. Define $\tilde{z}^{\max} \equiv \gamma y^{\max}(\tilde{n})$ as the maximum amount of emission caused by any firm when emissions are unregulated. Then, if $e^R > \tilde{z}^{\max}$, no firm is emission constrained. Since in practice regulations are restrictive, we assume $e^R < \tilde{z}^{\max}$ in what follows. It is useful to note for future that $\tilde{z}^a = \gamma\tilde{y}^a = \gamma\tilde{y}^f$ is the equilibrium emissions in the absence of a regulation cap.

Costless permit trade Here, firms' optimal input choice equalizes its marginal cost with its marginal expected revenue (See Appendix 7.1):

$$w = (p - \gamma\rho^*) \int_{\varepsilon^{\min}}^{\varepsilon^{\max}} y_n(n^*, \varepsilon)\ g(\varepsilon) d\varepsilon, \tag{3}$$

and in equilibrium

$$y^a(n^*) = y^f(n^*) = \frac{e^R}{\gamma}. \tag{4}$$

The above two determine $\rho^*(e^R)$ and $n^*(e^R)$. A comparison of (3) with (2) implies that if the regulator sets $e^R = \tilde{z}^a$, $\rho^* = 0$ and $n^* = \tilde{n}$. The aggregate outcome is the same as that in the absence of any regulatory cap.[12] However, if $e^R < \tilde{z}^a$, $n^* < \tilde{n}$ and $\rho^* > 0$. The assumption $y_{nn} < 0$ implies that $\frac{d\rho^*}{de^R} < 0$ and the assumption $\lim_{n\to 0} y_n = \infty$ further implies that as $e^R \to 0$, $\rho^* \to \frac{p}{\gamma}$; ρ^* monotonically maps $(0, \tilde{z}^a)$ to $\left(\frac{p}{\gamma}, 0\right)$.

In the equilibrium given by (3) and (4), all firms begin with identical amount of permits, i.e., $e^{i0} = e^R$ for all i, and employ $n^* : y^{f*} = \frac{e^R}{\gamma}$. After productivity shocks are realized those with $y(n^*, \varepsilon^i) > y^{f*}$ purchase more emission permits to fully utilize their production capacity, while those with $y(n^*, \varepsilon^i) < y^{f*}$ sell, and markets clear: $e^{a*} = \int_0^1 z^i di = e^R$. All excess permits are sold since $\rho^* > 0$; also, $\rho^* < \frac{p}{\gamma}$ ensures all unused permits are purchased in equilibrium. As a result, full utilization occurs. That is, $\frac{y^{\alpha}}{y^f} = \frac{e^a}{e^R} = 1$ holds.

Consider a central planner who aims to produce $y^f = \frac{e^R}{\gamma}$. It is optimal for the planner to command each firm to employ an equal input amount n, such that $y^f(n) = \frac{e^R}{\gamma}$, and then produce at full capacity after firms' productivity is realized. The decentralized outcome under perfectly competitive and costless permit trade replicates the planner's goal, and thus a version of the first welfare theorem holds.[13] This notion of efficiency breaks down in the absence of permit trade, which we discuss next.

Without permit trade The optimal input employment and planning stage permit price (in equilibrium) is given by (see Appendix 7.2)

$$\hat{\rho} = \frac{p}{\gamma}\left(1 - G(\varepsilon^R)\right); \tag{5a}$$

$$w = p \int_{\varepsilon^{\min}}^{\varepsilon^R} y_n(\hat{n}, \varepsilon)\ g(\varepsilon) d\varepsilon, \tag{5b}$$

[12] This is precisely due to the exchange (trade) of permits between firms whose potential emissions exceed their permit holdings with those whose emissions fall below. In the absence of this exchange the aggregate output equals y^f only when $e^R \geq \tilde{z}^{\max}$ as discussed earlier.

[13] See Proposition 5.F.1 in Mas-Colell, Whinston, and Green (1995).

where ε^R is defined by $y\left(\hat{n},\varepsilon^R\right) \equiv \frac{e^R}{\gamma}$. The above two equations determine $\hat{n}\left(e^R\right)$ and $\hat{\rho}\left(e^R\right)$. Equation (5a) implies that the marginal emission permit is chosen to equate its up-front cost with its expected benefit. A unit of emission permit allows the firm to market $\frac{1}{\gamma}$ units of output with a sale value of $\frac{p}{\gamma}$. But the permit is used only if the firm's full capacity emissions exceeds its permitted level, which occurs with probability $1 - G(\varepsilon^R)$. The condition for the optimal input choice (5b) is standard; n is chosen to equate the input price with its expected marginal value product.

Note that in general $\hat{y}^f \neq \frac{e^R}{\gamma}$. Under costless trade, firms utilize their full capacity at the production stage. In the absence of trade, a firm i can fully utilize its production capacity only if $e^R \geq z\left(\hat{n},\varepsilon^i\right)$; otherwise its output is constrained by $y^i = \frac{e^R}{\gamma}$. The aggregate production is

$$\hat{y}^a = \int_{\varepsilon^{\min}}^{\varepsilon^R} y\left(\hat{n},\varepsilon\right)\; g(\varepsilon)d\varepsilon + \frac{e^R}{\gamma}\left(1 - G\left(\varepsilon^R\right)\right) < \hat{y}^f,$$

The output lost due to unutilized capacity and the quantity of unutilized permits can be expressed as

$$\begin{aligned} \hat{y}^f - \hat{y}^a &= \int_{\varepsilon^R}^{\varepsilon^{\max}} \left(y\left(\hat{n},\varepsilon\right) - \frac{e^R}{\gamma}\right)\; g(\varepsilon)d\varepsilon; \\ e^R - \hat{e}^a &= \int_{\varepsilon^{\min}}^{\varepsilon^R} \left(e^R - \gamma y\left(\hat{n},\varepsilon\right)\right)\; g(\varepsilon)d\varepsilon. \end{aligned}$$

For analytical convenience, henceforth, we assume that y is separable in n and ε.

Let $f\left(n\,,\varepsilon\right) = f\left(n\right) + \varepsilon$ or $f\left(n,\,\varepsilon\right) = \varepsilon f\left(n\right)$

Under this assumption, a result that is useful in the following is stated next.

Lemma 1 *There exists $e^X < \tilde{z}^{\max}$ such that*

$$e^R \lesseqqgtr e^X \iff \hat{y}^f \gtreqqless \frac{e^R}{\gamma}.$$

Proof. See Appendix 7.5. ■

With no permit trade, each firm faces a cap e^R at the production stage and can produce no more than $\frac{e^R}{\gamma}$. While the marginal cost of n is constant at w, on the benefit side a marginal increase in n has two mutually opposite effects. Output rises for all ex post productivity realizations, which, in the standard manner, would raise firms' revenue. However, a marginal increase in n also triggers the emission constraint to bind at a lower productivity, i.e., ε^R falls. This reduces the probability that the marginal n will be utilized in production. The two marginal revenue effects of n balance its cost w when $y^f\left(\hat{n}\right) = \frac{e^X}{\gamma}$. When $e^R < e^X$, at a relatively lower scale of production, n is relatively small: its marginal product effect dominates the reduced probability of utilization effect, inducing a factor *over*-employment such that $\hat{y}^f > \frac{e^R}{\gamma}$. The situation reverses for $e^R > e^X$.

The optimal choice of n is increasing in e^R, when $e^R \to \tilde{z}^{\max}$, $\hat{y}^f \to \tilde{y}^f < \frac{e^R}{\gamma}$, but for e^R sufficiently close to zero $\hat{y}^f > \frac{e^R}{\gamma}$ since $\lim_{n\to 0} y_n = \infty$. By continuity there exists e^X, which is unique when $y\left(n,\varepsilon\right) = y\left(n\right) + \varepsilon$ or $= \varepsilon y\left(n\right)$. Irrespective of whether $e^R \lesseqqgtr e^X$, for some firms operating at full capacity will violate their permissible emissions, while for others full capacity emissions are below the permits they hold. As a result, both capacity and emission permit utilizations are below 100%. Suppose, as a counterfactual, permit markets open unexpectedly at the operations stage when firms have already employed $\hat{n}$. In this case, there would be an *excess demand* for permits when $e^R < e^X$ and an *excess supply* of permits for $e^R > e^X$. We show below that this market mismatch exists in equilibrium even when the permit market is open at the operations stage but trade incurs transactions costs.

At this point, It is worth noting that in all trading environments ρ, r, n, y^f, $e^a = \gamma y^a$ are endogenously determined as functions of model parameters p, w, technological specification $y(.)$, ε, and the regulator's cap e^R. Our objective is to understand how transactions costs, in its alternate forms and magnitudes impact permit trade, production plans, and marketable output over the entire range of binding emission caps $e^R < \tilde{z}^{\max}$, where $\tilde{z}^{\max}$ follows from (2). Any alteration in $\{p, w\}$ will only re-scale the range of e^R for our analysis and henceforth we keep w and p fixed in the rest of the paper.[14] All endogenous quantities and prices are implicitly treated as functions of e^R. We use e^R as an argument explicitly only when clarity demands.

3 Proportional trading costs

Under proportional trading costs, a firm in need of additional emission permits pays $r + \kappa^b$ per unit, while a firm with surplus permits receives $r - \kappa^s$ per unit. Let $\kappa^T \equiv \kappa^b + \kappa^s$ denote the total per unit trading costs. Notice that for permits to be bought and sold in equilibrium, we must have $r + \kappa^b \leq \frac{p}{\gamma}$ and $r - \kappa^s \geq 0$, respectively. This implies $\kappa^b \leq \frac{p}{\gamma} - r \leq \frac{p}{\gamma} - \kappa^s$. Thus for trade to occur $\kappa^T \leq \frac{p}{\gamma}$: no trade occurs when the goods market revenue from an extra permit is below total trading costs.

The expected profit of a firm employing n and initially holding e^0 permits can be written as

$$\begin{aligned} E\left\{\pi\left(n, e^0\right)\right\} \quad = \quad & p \int_{\varepsilon^{\min}}^{\varepsilon^{\max}} y(n, \varepsilon)\ g(\varepsilon) d\varepsilon + (r - \kappa^s) \int_{\varepsilon^{\min}}^{\varepsilon^0} \left(e^0 - \gamma y(n, \varepsilon)\right) g(\varepsilon) d\varepsilon \\ & - \left(r + \kappa^b\right) \int_{\varepsilon^0}^{\varepsilon^{\max}} \left(\gamma y(n, \varepsilon) - e^0\right)\ g(\varepsilon) d\varepsilon - \rho e^0 - wn, \end{aligned} \tag{6}$$

where $\varepsilon^0\left(n, e^0\right)$ is determined from $y\left(n, \varepsilon^0\right) = e^0$. The second and the third term above denote expected revenue from permits sold and expected cost of permits purchased, respectively. In equilibrium with $\varepsilon^0 = \varepsilon^R$, the optimal choice for n equates its marginal cost with its marginal benefit:

$$w = p y_n^f - \underbrace{\left((r - \kappa^s) \int_{\varepsilon^{\min}}^{\varepsilon^R} \gamma y_n(n, \varepsilon)\ g(\varepsilon) d\varepsilon + \left(r + \kappa^b\right) \int_{\varepsilon^R}^{\varepsilon^{\max}} \gamma y_n(n, \varepsilon)\ g(\varepsilon) d\varepsilon\right)}_{\equiv EMEC(n)}. \tag{7}$$

The first term on the RHS is the expected marginal revenue from goods' sales. The terms within brackets represent the *expected marginal emission cost* of n (henceforth, EMEC(n)) due to pollution regulation. The first term relates to ε realizations for which the firm has a permit surplus, i.e., it is a seller. A marginal increase in n reduces revenues from permit sales at the rate of $(r - \kappa^s)\,\gamma y_n(n, \varepsilon)$. The second term captures ε realizations for which the firm has a permit shortfall, i.e., it is a buyer. A marginal increase in n now increases firms' permit costs at the rate of $\left(r + \kappa^b\right) \gamma y_n(n, \varepsilon)$. Thus, the first and the second components of EMEC(n) reflect expected loss of revenue from permit sales and expected increase in costs, respectively, due to the marginal increase in n.

The profit maximizing equilibrium choice for e^0 dictates:

$$\rho = \left(r + \kappa^b\right)\left(1 - G(\varepsilon^R)\right) + (r - \kappa^s)\, G\left(\varepsilon^R\right), \tag{8}$$

which equates the permit's planning stage market value with its expected worth at the production stage: A marginal permit is used against the firm's emissions with probability $1 - G(\varepsilon^R)$, in which case it is worth $r + \kappa^b$; with probability $G\left(\varepsilon^R\right)$ the firm does not need it against its emissions, and sells it in the market at $r - \kappa^s$.

[14] One of them is irrelevant as it can easily be set as the numeraire.

As of now, we have two optimality conditions (7) and (8) but three unknowns: ρ, r and n. The equilibrium is characterized by (7) and (8) holding at $e^0 = e^R$, and the market clearing conditions:

$$\begin{aligned} e^R &= e^a \text{ and } y^a < y^f \text{ if and only if } r + \kappa^b = \frac{p}{\gamma}; \\ e^R &> e^a \text{ and } y^a = y^f \text{ if and only if } r = \kappa^s; \\ e^R &= e^a \text{ and } y^a = y^f \text{ if and only if } \kappa^s < r < \frac{p}{\gamma} - \kappa^b. \end{aligned} \tag{9}$$

The first condition relates to cases when firms' full capacity emissions exceed permits held in the aggregate. Firms with relatively high productivity would want to buy additional permits as long as $r + \kappa^b \leq \frac{p}{\gamma}$. There is an excess demand for permits in the aggregate at $r + \kappa^b < \frac{p}{\gamma}$ that drives its equilibrium market price all the way to $r = \frac{p}{\gamma} - \kappa^b$. The second relates to cases when outstanding permits exceed full capacity emissions creating an excess supply at the production stage. In equilibrium permit prices are driven down to just cover the sellers' trading cost. It is only in the third case that both sellers and buyers obtain a positive surplus from trade and markets clear: Every firm produces at full capacity and no permit is wasted.

Note that with $\kappa^b = \kappa^s = 0 = \kappa^T$, we have $r = \rho = \rho^*$ from (8) and (7), and the latter reduces to (3). On the other hand, when $\kappa^b + \kappa^s = \kappa^T > \frac{p}{\gamma}$, permit trade ceases and instead of (6) a firm's expected profit becomes

$$E\left\{\pi\left(n, e^0\right)\right\} = p\int_{\varepsilon^{\min}}^{\varepsilon^0} y\left(n,\varepsilon\right)\; g(\varepsilon)d\varepsilon + \frac{pe^0}{\gamma}\left(1 - G\left(\varepsilon^0\right)\right) - \rho e^0 - wn,$$

which yields the pre-production permit price and the optimal input choice as given by (5b) and (5a). These two conditions also hold in the limit when $\kappa^T = \frac{p}{\gamma}$ and trade occurs with the only possibility that $r = \kappa^s$ and $r + \kappa^b = \frac{p}{\gamma}$ as can be verified by substituting in (7) and (8); these two equations along with $e^0 = e^R$ also determine the equilibrium for $\kappa^T \geq \frac{p}{\gamma}$. The following proposition characterizes equilibria for all $\kappa^T \in \left[0, \frac{p}{\gamma}\right]$.

Proposition 2 *There exists* $\bar{\kappa}^T\left(e^R\right) \in \left[0, \frac{p}{\gamma}\right]$, *with* $\bar{\kappa}^T(0) = 0; \bar{\kappa}^T\left(e^X\right) = \frac{p}{\gamma}; \bar{\kappa}^T\left(e^R\right) = 0$ *for all* $e^R \geq \tilde{z}^a$*; and for all* $e^R \leq \tilde{z}^a$

$$\frac{\partial \bar{\kappa}^T}{\partial e^R} \gtreqless 0 \text{ for } e^R \lesseqgtr e^X,$$

such that (i) if $\kappa^T \leq \bar{\kappa}^T\left(e^R\right)$, $n = n^*$, $y^a = y^{f*} = \frac{e^R}{\gamma}$ *for all* e^R*; (ii) if* $\kappa^T \in (\bar{\kappa}^T, \frac{p}{\gamma}]$,

$$e^R \lesseqqgtr e^X \iff \begin{cases} y^a = \frac{e^R}{\gamma} < y^f, \\ y^a = y^f = \frac{e^R}{\gamma}, \\ y^a = y^f < \frac{e^R}{\gamma}, \end{cases}$$

and where n is determined from

$$w = \gamma\kappa^T \int_{\varepsilon^{\min}}^{\varepsilon^R} y_n\left(n,\varepsilon\right)\; g(\varepsilon)d\varepsilon \quad \text{for } e^R \leq e^X; \tag{10a}$$

$$w = p\; y^f\left(n\right) - \gamma\kappa^T \int_{\varepsilon^R}^{\varepsilon^{\max}} y_n\left(n,\varepsilon\right)\; g(\varepsilon)d\varepsilon \quad \text{for } e^R \geq e^X. \tag{10b}$$

Proof. *See Appendix 7.6.* ■

Part (i) of the proposition states that the equilibrium is efficient as long as the total per unit trading cost falls below an upper bound $\bar{\kappa}^T$ which varies with e^R non-monotonically. This bound is sufficiently small for both large and small values of e^R. To understand this, note that for the equilibrium with trading costs to be efficient, r must induce the input choice n^* despite trading costs. Comparing the RHS of (3) with the RHS of (7) requires that the EMEC(n) with trading costs must equal that under costless trade for $n = n^*$. This can be rewritten as

$$\rho^* = (r - \kappa^s)\,\frac{E\{y_n(n^*|s)\}}{E\{y_n(n^*)\}} + \left(r + \kappa^b\right)\frac{E\{y_n(n^*|b)\}}{E\{y_n(n^*)\}}, \tag{11}$$

where $\frac{E\{y_n(n^*|s)\}}{E\{y_n(n^*)\}}$ and $\frac{E\{y_n(n^*|b)\}}{E\{y_n(n^*)\}}$ denote the fraction of expected marginal emission due to n, conditional on the firm being a seller and a buyer, respectively. Thus, efficient factor employment demands that the average of the sellers' and buyers' permit prices, weighted by their respective contributions to the expected marginal emissions due to n, equal ρ^*. Clearly, r is decreasing in κ^b and increasing in κ^s. Intuitively, a higher κ^b reduces permit demand and lowers its market price; a higher κ^s hinders sales and thereby requires the equilibrium permit price to rise. Finally, ex post, permit markets must clear for the all emissions permits to be utilized, requiring $r \in \left(\kappa^s, \frac{p}{\gamma} - \kappa^b\right)$.

Consider a value of e^R that is slightly less than emissions under no regulatory constraints, $\tilde{z}^a$. Under costless trade, the permit price ρ^* is then close to zero. Sellers' ex post *participation constraint* $r \geq \kappa^s$ and firms' ex ante efficient factor employment condition, (11), are satisfied only if both κ^s and κ^b are sufficiently small. Similarly, consider the case when $e^R \to 0$ and thus $\rho^* \to \frac{p}{\gamma}$. Once again, buyers' ex post participation constraint $r + \kappa^b \leq \frac{p}{\gamma}$ and (11) are satisfied only if both κ^s and κ^b are sufficiently small. Thus, full efficiency occurs under high and low emission caps if and only if $\bar{\kappa}^T$ is sufficiently small. Now, consider an intermediate value of e^R, e.g. such that $\rho^* = \frac{p}{2\gamma}$. Such a cap allows a range of $r \in \left(0, \frac{p}{\gamma}\right)$ and a corresponding $\kappa^T \in [0, \frac{p}{\gamma})$ such that both buyers' and sellers' participation constraints are satisfied and (11) holds.

When $\kappa^T > \frac{p}{\gamma}$, there is no surplus from trade and the permit market ceases to exist. For example, if buyers' gains are zero, i.e., $r + \kappa^b = \frac{p}{\gamma}$, the sellers' revenue $r - \kappa^s = r + \kappa^b - \kappa^T < 0$, and sellers remain inactive; if sellers' surplus $r - \kappa^s = 0$, the buyers' gain $\frac{p}{\gamma} - \left(r + \kappa^b\right) = \frac{p}{\gamma} - \kappa^T < 0$. The equilibrium outcomes are the same as in the no trade benchmark derived earlier. Following Lemma 1, either production capacity or permits remain underutilized depending on whether $e^R \gtreqless e^X$. We next consider both cases to explain how the aggregate output-permit mismatch prevails under $\kappa^T \in \left(\bar{\kappa}^T, \frac{p}{\gamma}\right)$.

3.0.1 $e^R < e^X$

Equation (11) can be rewritten as

$$\rho^* = \underbrace{r + \kappa^b}_{\leq \frac{p}{\gamma}} - \kappa^T \frac{E\{y_n(n^*|s)\}}{E\{y_n(n^*)\}}, \tag{12}$$

which, for a given $\kappa^b < \frac{p}{\gamma}$, places an upper bound on κ^T for ensuring that buyers participate: $r + \kappa^b \leq \frac{p}{\gamma}$.When $\kappa^T = \bar{\kappa}^T$, $r + \kappa^b = \frac{p}{\gamma}$, and (12) continues to hold. For $\kappa^T > \bar{\kappa}^T$, $r = \frac{p}{\gamma} - \kappa^b$; (12) ceases to hold and $n \neq n^*$ is determined by part (ii) of Proposition 2. Recalling from Lemma 1 that for $\kappa^T = \frac{p}{\gamma}$,

$y^f = \hat{y}^f > \frac{e^R}{\gamma} = y^{f*}$, we have $y^f \in \left(y^{f*}, \hat{y}^f\right)$ for $\kappa^T \in \left(\bar{\kappa}^T, \frac{p}{\gamma}\right)$. There is an excess demand for permits and the participation condition for the buyers binds, i.e., $r + \kappa^b = \frac{p}{\gamma}$. Any increase in κ^T essentially lowers the sellers' opportunity cost as $r - \kappa^s = \frac{p}{\gamma} - \kappa^T$. Thus EMEC($n$) decreases and equilibrium n and y^f are higher than n^* and y^{f*}.

The reason $y^f < \hat{y}^f$ under costly trade is that a marginal unit of output obtains $p - \gamma\left(r + \kappa^b\right) = 0$ for a firm i with $y\left(n, \varepsilon^i\right) > \frac{e^R}{\gamma}$ as in the absence of trade. However, if the firm realizes $y\left(n, \varepsilon^i\right) < \frac{e^R}{\gamma}$, it can sell its excess permits at $r - \kappa^s = \frac{p}{\gamma} - \kappa^T > 0$ unlike in the absence of trade. This ex post asymmetry – no gains if excess capacity but positive earnings if excess permits – induces firms to choose a relatively lower output capacity under costly trade.

3.0.2 $e^R > e^X$

Now, (11) can be rewritten as

$$\rho^* = \underbrace{r - \kappa^s}_{\geq 0} + \kappa^T \frac{E\left\{y_n\left(n^*|b\right)\right\}}{E\left\{y_n\left(n^*\right)\right\}}, \tag{13}$$

which places an upper bound of $\bar{\kappa}^T$ on κ^T for sellers to participate in the permit market, i.e., $r - \kappa^s \geq 0$. At $\kappa^T = \bar{\kappa}^T$, $r = \kappa^s$, and (13) continues to hold. For $\kappa^T > \bar{\kappa}^T$, $r = \kappa^s$, and (13) ceases to hold and $n \neq n^*$ is determined by part (ii) of Proposition 2. Now (from Lemma 1) with $\kappa^T = \frac{p}{\gamma}$, $y^f = \hat{y}^f < \frac{e^R}{\gamma} = y^{f*}$. Thus, for all $\kappa^T \in \left(\bar{\kappa}^T, \frac{p}{\gamma}\right)$, $y^f \in \left(\hat{y}^f, y^{f*}\right)$. In equilibrium, there is an excess supply of permits and the sellers' participation condition binds, i.e., $r = \kappa^s$. Any increase in κ^T essentially increases the buyers' opportunity cost $r + \kappa^b = \kappa^T$. Thus EMEC(n) increases and the equilibrium n and y^f are lower than n^* and y^{f*}, causing an excess supply of permits by $\frac{y^{f*} - y^f}{y^{f*}}$. As for $y^f > \hat{y}^f$, a firm i gets $r - \kappa^s = 0$ if $y\left(n, \varepsilon^i\right) < \frac{e^R}{\gamma}$, whereas if $y\left(n, \varepsilon^j\right) > \frac{e^R}{\gamma}$ it can buy additional permits and get a net value of $p - \gamma\left(r + \kappa^b\right) > 0$. This encourages a larger factor employment relative to $\hat{n}$.

To sum up, efficiency demands that ex post permit markets clear. Since all sellers of unused permits receive the same marginal revenue, and all buyers receive an identical marginal benefit, they are all active in the permit market as long as these marginal values are positive. In the event of an excess permit supply or demand, the short side of the market captures all the surplus from trade. Finally, in case of an excess demand, i.e., when $r + \kappa^b = \frac{p}{\gamma}$, some permit-constrained firms are unable to fully utilize their production capacity.

For a concrete illustration, we parameterize the model and derive the thresholds stated in Proposition 1.

3.1 Parametric examples

Below, both specifications, $y\left(n, \varepsilon\right) = y\left(n\right) + \varepsilon$ and $y\left(n, \varepsilon\right) = \varepsilon y\left(n\right)$, are discussed sequentially. Under the first specification we assume that ε is uniformly distributed. This allows simple closed form solutions of endogenous quantities and prices as functions of targeted emissions and trading costs (both proportional and fixed). However, the results may have a limited scope because (i) the output variance is exogenously fixed and is invariant to the production scale, (ii) the shock distribution is restricted to be finite. To address these limitations, we assume ε to be log-normally distributed for the second specification. We show that with proportional trading costs the results are qualitatively similar under both specifications, whereas with fixed trading costs the results diverge (Section 4).

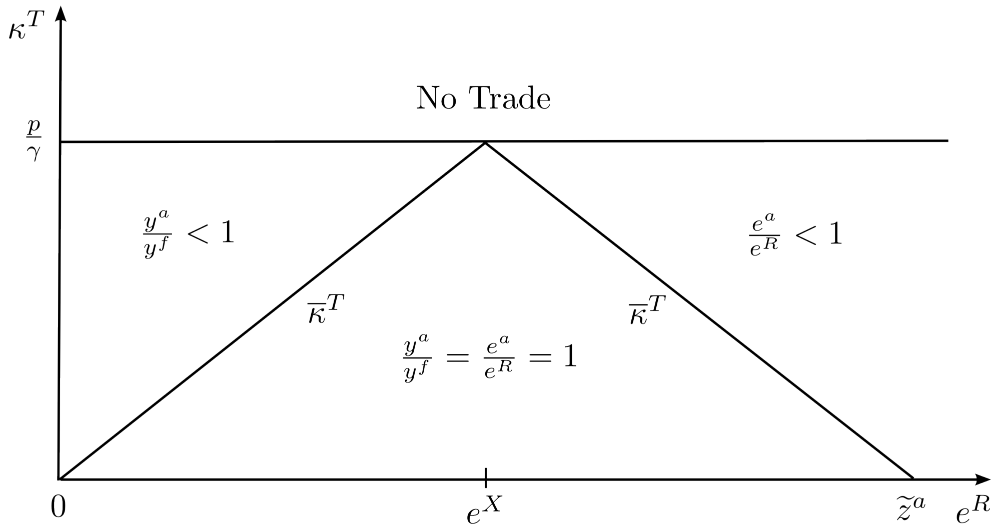


Figure 1: Full, partial and no trade outcomes: trading costs and emissions caps.

3.1.1 Additive, uniformly distributed shock

Let $y\left(n\,,\varepsilon\right) = \sqrt{n} + \varepsilon$, with $\varepsilon \sim U\ [-c, c]$.[15] The cost thresholds stated in Proposition 2 are (see Appendix 7.7 for the details)

$$\bar{\kappa}^T\left(e^R\right) \equiv \begin{cases} \frac{4e^R w}{\gamma^2} & \text{for} \quad e^R < e^X = \frac{\gamma\ p}{4w} \\ \frac{2p}{\gamma} - \frac{4e^R w}{\gamma^2} & \text{for} \quad e^R \in \left[\frac{\gamma\ p}{4w}, \frac{\gamma\ p}{2w}\right] \\ 0 & \text{for} \quad e^R > \frac{\gamma\ p}{2w} = \tilde{z}^a \end{cases} \tag{14}$$

Figure 1 below exhibits $\bar{\kappa}^T\left(e^R\right)$ in the two dimensional space of e^R and κ^T. No trade occurs if $\kappa^T > \frac{p}{\gamma}$ for any value of e^R. For $\kappa^T \in \left(\bar{\kappa}^T, \frac{p}{\gamma}\right)$, permit trade occurs with excess demand when $e^R < e^X$, and with excess supply when $e^R > e^X$. In the former, $r = \frac{p}{\gamma} - \kappa^b$. In the latter, $r = \kappa^s$. Within the boundaries of the triangle, the production-emission outcome is replicates that under costless trade.

As stated in Proposition 2, it is easily checked for all values of e^R that both ex ante efficiency and ex post market participation conditions hold when $\kappa^T \leq \bar{\kappa}^T$, and $y^a = y^{f*} = \frac{e^R}{\gamma}$ (see Appendix 7.7). For $\kappa^T > \bar{\kappa}^T$, there are two cases to consider.

3.1.2 $e^R < e^X = \frac{\gamma\ p}{4w}; \kappa^T > \bar{\kappa}^T = \frac{4e^R w}{\gamma^2}$

As stated in Proposition 2 $\frac{e^a}{e^R} = 1$. On the other hand, the production capacity utilization can be expressed as

$$\frac{y^a}{y^f} = 1 - \frac{\gamma c}{e^R + \gamma c}\left(1 - \frac{\bar{\kappa}^T}{\kappa^T}\right),$$

As intuition suggests, capacity utilization is decreasing in c and κ^T. As e^R increases, the demand-supply mismatch in the permit market narrows, and capacity utilization improves. In the limit, $e^R \to e^X$, we have 100% capacity utilization ($y^a = y^f$).

[15] To ensure that $y\left(n\,,\varepsilon_{\min}\right) > 0$ requires restrictions on c. See Appendix 7.7.

3.1.3 $e^R > e^X = \frac{\gamma\ p}{4w}; \kappa^T > \bar{\kappa}^T = \frac{2p}{\gamma} - \frac{4e^R w}{\gamma^2}$

Now, following Proposition 2 $\frac{y^a}{y^f} = 1$. However, the permit utilization ratio can be expressed as

$$\frac{e^a}{e^R} = 1 - \frac{\gamma c}{e^R} \frac{1 - \frac{\bar{\kappa}^T}{\kappa^T}}{1 + \frac{4wc}{\gamma \kappa^T}}.$$

Permit utilization is decreasing in κ^T and c. It is also decreasing in e^R. In this case, the demand-supply gap narrows as e^R approaches e^X; as $e^R \longrightarrow e^X$, permit utilization is 100%.

3.1.4 Multiplicative, log-normally distributed shock

We let $y(n, \varepsilon) = \varepsilon\ n^{0.5}$ where $\ln \varepsilon\ \tilde{}\ N(0, \sigma)$. This implies $\tilde{y}^a = \tilde{y}^f = y^f(\tilde{n}) = \sqrt{\tilde{n}} \exp\left(\frac{1}{2}\sigma^2\right) = \frac{p}{2w} \exp\left(\sigma^2\right) = \frac{\tilde{z}^a}{\gamma}$. Appendix 7.8 shows that

$$e^X = \frac{\gamma\ p}{2w} \exp\left(\sigma^2\right) \Phi\left[-\frac{1}{2}\sigma\right],$$

where Φ denotes the cdf of a standard normal distribution. Applying proposition 2 equations (10a) and (10b), we have (see Appendix 7.8)

$$\bar{\kappa}^T\left(e^R\right) \equiv \begin{cases} \frac{2w \exp\left(-\sigma^2\right)}{\gamma \Phi\left[-\frac{1}{2}\sigma\right]} \frac{e^R}{\gamma} \text{ for } e^R \le e^X \\ \frac{p - 2w \exp\left(-\sigma^2\right) \frac{e^R}{\gamma}}{\gamma \Phi\left[\frac{1}{2}\sigma\right]} \text{ for } e^R \in \left[e^X, \tilde{z}^a\right] \\ 0 \text{ for } e^R \ge \tilde{z}^a \end{cases}$$

Figure 1, as in the previous example, exhibits these thresholds. Once again, it is easily checked for all values of e^R (see Appendix 7.8) that both ex ante efficiency and ex post participation conditions hold when $\kappa^T \le \bar{\kappa}^T$, and $y^a = y^f = \frac{e^R}{\gamma}$. For $\kappa^T > \bar{\kappa}^T$, the two cases are

$e^R < e^X$ Although a closed form expression for capacity utilization can not be obtained, $\frac{y^a}{y^f}$ follows from

$$\frac{y^a}{y^f} \Phi\left[\frac{1}{\sigma} \log\left[\frac{y^a}{y^f}\right] - \frac{1}{2}\sigma\right] = \frac{2e^R}{w\gamma^2 \kappa^T} \exp\left[-\sigma^2\right].$$

Thus, capacity utilization is decreasing in σ and κ^T. As e^R increases, the demand-supply mismatch in the permit market narrows, and capacity utilization improves. In the limit, $e^R \to e^X$, we have $y^a = y^f$.

$e^R > e^X$ Permit utilization $\frac{e^a}{e^R}$ follows from

$$\frac{\frac{e^a}{e^R}}{1 - \frac{\gamma \kappa^T}{p}\left(1 - \Phi\left[-\frac{1}{\sigma} \log\left[\frac{e^a}{e^R}\right] - \frac{1}{2}\sigma\right]\right)} = \frac{\gamma p}{2we^R} \exp\left[\sigma^2\right].$$

The LHS is increasing in $\frac{e^a}{e^R}$ and κ^T; the RHS is decreasing in e^R. Thus, permit utilization $\frac{e^a}{e^R}$ decreases with κ^T as well as e^R. Once again, as $e^R \to e^X$, $e^a \to e^X$. The variation of $\frac{e^a}{e^R}$ with σ can not be analytically signed. When checked numerically, it is decreasing in σ.

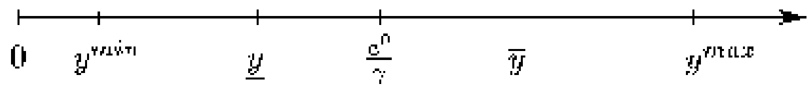


Figure 2: Productivity shock thresholds and trade activity under fixed costs.

4 Fixed trading costs

This section assumes trading in emission permits entails a fixed (quantity invariant) market entry cost of κ^b and κ^s for buyers and sellers, respectively. The expected profit of a firm that acquires permits e^0 and employs input n is

$$E\left\{\pi\left(n, e^0\right)\right\} = \int_{\varepsilon^{\min}}^{\underline{\varepsilon}} \left\{py\left(n,\varepsilon\right) + r\left(e^0 - \gamma y\left(n,\varepsilon\right)\right)\right\}\ g(\varepsilon)d\varepsilon + p\int_{\underline{\varepsilon}}^{\varepsilon^0} y\left(n,\varepsilon\right)g(\varepsilon)d\varepsilon \tag{15}$$

$$+p\frac{e^0}{\gamma}\int_{\varepsilon^0}^{\bar{\varepsilon}} g(\varepsilon)d\varepsilon + \int_{\bar{\varepsilon}}^{\varepsilon^{\max}} \left\{py\left(n,\varepsilon\right) + r\left(\gamma y\left(n,\varepsilon\right) - e^0\right)\right\}\ g(\varepsilon)d\varepsilon$$

$$-\kappa^s G\left(\underline{\varepsilon}\right) + \kappa^b\left(1 - G\left(\bar{\varepsilon}\right)\right) - \rho e^0 - wn, \tag{16}$$

where $\underline{\varepsilon}$ and $\bar{\varepsilon}$ are obtained from

$$\gamma r\left(\frac{e^0}{\gamma} - \underbrace{y\left(n,\underline{\varepsilon}\right)}_{\equiv \underline{y}}\right) = \kappa^s; \tag{17a}$$

$$(p - \gamma r)\left(\underbrace{y\left(n,\bar{\varepsilon}\right)}_{\equiv \bar{y}} - \frac{e^0}{\gamma}\right) = \kappa^b. \tag{17b}$$

Thresholds, $\underline{\varepsilon}$ and $\bar{\varepsilon}$ depend on the firm's permit holding e^0 and factor allocation n, choices that are examined shortly.

Equation (16) is easily understood with the help of Figure 2. The four integrals decompose expected profits into potential output realizations in the intervals, $\left[y^{\min}, \underline{y}\right]$, $\left[\underline{y}, \frac{e^0}{\gamma}\right]$, $\left[\frac{e^0}{\gamma}, \bar{y}\right]$, and $[\bar{y}, y^{\max}]$. The first two capture the cases of a firms' potential emissions falling below its permit holdings, $z\left(y\left(n,\varepsilon\right)\right) < e^0$. The first integral in (16) is the net expected return when productivity is very low and $y \in \left[y^{\min}, \underline{y}\right]$ is realized. In this case, the firm will enter the permit market and sell excess permits $e^0 - z$. The second integral states that firms with emission between $\gamma \underline{y}$ and e^0 will not sell excess permits because the sale earns less than the market entry cost. The third and the fourth integrals represent high potential output realizations. A firm realizing a modestly high output between $\frac{e^0}{\gamma}$ and $\bar{y}$ will not buy additional permits because the gains from trade are less than the market entry cost. Only if y exceeds $\bar{y}$, will the net revenues from purchasing additional permits offset the entry cost. The first two terms in the third line of (16) represent expected trading costs. Put simply, permit trades occur only if the gains are sufficiently large to offset fixed costs. For future reference, we term firms with $y \in \left[y^{\min}, \underline{y}\right] \cup [\bar{y}, y^{\max}]$ as *active* and those with $y \in \left[\underline{y}, \bar{y}\right]$ as *inactive* in the permit market.

The necessary and sufficient condition for the optimal choice of n in equilibrium with $e^0 = e^R$ is given by

$$w = (p - \gamma r)\int_{\varepsilon \notin [\underline{\varepsilon}, \bar{\varepsilon}]} y_n\left(n,\varepsilon\right)\ g(\varepsilon)d\varepsilon + p\int_{\underline{\varepsilon}}^{\varepsilon^R} y_n\left(n,\varepsilon\right)g(\varepsilon)d\varepsilon. \tag{18}$$

The first term on the RHS represents the expected revenue from marginal n, conditional on the firm being active in the permit market. For shocks in this range, the market value of a unit of output is

$p-\gamma r$. The second RHS term represents expected marginal revenue conditional on emissions lying below permit holdings but above the threshold that warrants entry to the permit market. Here, producing goods has a market value of p but no opportunity costs.[16]

The necessary condition for the optimal choice of e^0, in equilibrium $\left(e^0 = e^R\right)$, gives

$$\rho = \frac{p}{\gamma}\left[G(\bar{\varepsilon}) - G(\varepsilon^R)\right] + r\left[1 - G(\bar{\varepsilon}) + G(\underline{\varepsilon})\right]. \tag{19}$$

The probability that the emission cap binds and permits are utilized at the margin is $\left[G(\bar{\varepsilon}) - G(\varepsilon^R)\right]$; the probability that the firm is active in the permit market and either buys or sells permits at r is $1 - G(\bar{\varepsilon}) + G(\underline{\varepsilon})$. Thus, the RHS represents expected value from a marginal permit.

Thus far, we have two equations, (18) relating optimal n to input and output prices, and (19) relating ρ to r. One remaining is the permit market clearing condition:

$$\int_{\varepsilon^{\min}}^{\underline{\varepsilon}} \left(e^R - \gamma y\left(n, \varepsilon\right)\right)\ g(\varepsilon)\ d\varepsilon = \int_{\bar{\varepsilon}}^{\varepsilon^{\max}} \left(\ \gamma y\left(n, \varepsilon\right) - e^R\right)\ g(\varepsilon)\ d\varepsilon. \tag{20}$$

Thus, ρ, r and n are jointly determined by (18) - (20). It is easily seen that as κ^b (κ^s) increases, the mass of firms with potentially high (low) emissions that buy (sell) in the permit market shrinks. If either of these costs become sufficiently large, the permit market ceases to function.[17] In this case, $\varepsilon \notin [\underline{\varepsilon}, \bar{\varepsilon}] = \varnothing$, the first term in (18) vanishes and $\underline{\varepsilon} = \varepsilon^{\min}$ in the second term. The necessary conditions expressed by (18) and (19) then reduce to (5b) and (5a). On the contrary, if $\kappa^s = \kappa^b = 0$, $\rho = r$ and equation (18) reduces to (3).

Once again, with the two cost bounds described above, we now present our general results.

Proposition 3 *For all $e^R \leq \tilde{z}^{\max}$, define*

$$\begin{aligned} \bar{\kappa}^b\left(e^R\right) &\equiv p\left(y\left(\hat{n}\left(e^R\right), \varepsilon^{\max}\right) - \frac{e^R}{\gamma}\right); \\ \bar{\kappa}^s\left(e^R\right) &\equiv p\left(\frac{e^R}{\gamma} - \gamma y\left(\hat{n}\left(e^R\right), \varepsilon^{\min}\right)\right). \end{aligned}$$

Then for each $e^R < \tilde{z}^{\max}$ a permit market with the equilibrium price $r \in \left(0, \frac{p}{\gamma}\right)$ exists if and only if

$$\frac{\kappa^b}{\bar{\kappa}^b\left(e^R\right)} + \frac{\kappa^s}{\bar{\kappa}^s\left(e^R\right)} < 1. \tag{21}$$

Proof. See Appendix 7.9. ■

When $\kappa^b > \bar{\kappa}^b$, no firm enters the permit market as a buyer even when sellers face no trading costs and are willing to sell their surplus permits for free. The reverse is the case when $\kappa^s > \bar{\kappa}^s$. Now even when buyers face no fixed costs and are willing to pay $\frac{p}{\gamma}$ per permit, the firm with the highest unused (excess) permits does not break even by selling them. Condition (21) ensures that there is some surplus from trade for some sellers and/or buyers, by requiring that when either sellers' or buyers' costs are high their counterparts' costs be offsettingly low. If this condition holds, the equilibrium price adjusts, $r \in \left(0, \frac{p}{\gamma}\right)$, and market clears. In this case, permits sold by the firms with $y \in \left[y^{\min}, \frac{1}{\gamma}\left(e^R - \frac{\kappa^s}{r}\right)\right]$ equal those bought by firms with $y \in \left[\frac{e^R}{\gamma} + \frac{\kappa^b}{p-\gamma r}, y^{\max}\right]$ and (20) holds.

[16] Recall that when potential output lies between $y\left(n, \varepsilon^R\right)$ and $y\left(n, \bar{\varepsilon}\right)$, the firm does not break even in purchasing additional emission permits. A marginal unit of potential output is simply lost.

[17] The argument is clearly valid for a finite support of the productivity distribution. If instead the support is infinite, a sufficiently large κ_s will drive out all permit sellers from the market.

Under fixed costs, permit price equilibrates the demand from active sellers with supply from active buyers by altering the *extensive* margin. When $r \to \frac{\kappa^s/\gamma}{\frac{e^R}{\gamma} - y^{\min}} > 0$, the mass of active sellers vanishes, while as $r \to \frac{p}{\gamma} - \frac{\kappa^b/\gamma}{y^{\max} - \frac{e^R}{\gamma}} > 0$, the mass of active buyers vanishes. As a result, r adjusts to equilibrate the two. Yet, firms with $y \in \left[\underline{y}, \frac{e^R}{\gamma}\right]$ have excess permits and firms with $y \in \left[\frac{e^R}{\gamma}, \bar{y}\right]$ have unutilized production capacity. As a result, neither the production capacity nor the emission permits is fully utilized:

$$y^f - y^a = \int_{\varepsilon^R}^{\bar{\varepsilon}} \left(y(n, \varepsilon) - \frac{e^R}{\gamma} \right) g(\varepsilon) d\varepsilon > 0; \tag{22a}$$

$$e^R - e^a = \int_{\underline{\varepsilon}}^{\varepsilon^R} \left(e^R - \gamma y(n, \varepsilon) \right) g(\varepsilon) d\varepsilon > 0. \tag{22b}$$

It is instructive to contrast this result with that under proportional costs. Recall that when proportional trading costs are sufficiently small, all firms strictly benefit by trading in the permit market and both permit and production capacity is fully utilized. When κ^T is sufficiently large (but $< \frac{p}{\gamma}$), trade occurs with either excess permit demand or supply. Trade ceases only when $\kappa^T > \frac{p}{\gamma}$. Also, capacity utilization is less than 100% only when $e^R < e^X$, and κ^T is sufficiently high. In contrast under fixed costs, even when permits are traded, both capacity and permit utilizations are less than 100%.

It is easily checked that $\lim_{e^R \to 0} \bar{\kappa}^s \left(e^R\right) = 0$ and $\lim_{e^R \to \tilde{z}^{\max}} \bar{\kappa}^s \left(e^R\right) > 0$; $\lim_{e^R \to 0} \bar{\kappa}^b \left(e^R\right) > 0$ and $\lim_{e^R \to \tilde{z}^{\max}} \bar{\kappa}^b \left(e^R\right) = 0$. Since y satisfies Inada conditions, $\frac{e^R}{\gamma}$ lies in the lower tail of the equilibrium output distribution when e^R is sufficiently low. There are many potential buyers but few sellers and even a small $\bar{\kappa}^s$ drives sellers out of the market. Conversely, a sufficiently high e^R may lie in the upper tail of output distribution in equilibrium. Then, there are many potential sellers and few buyers and even a small $\bar{\kappa}^b$ drives out buyers. This leads one to conjecture that $\bar{\kappa}^b$ is decreasing whereas $\bar{\kappa}^s$ is increasing in e^R.

The empirical transactions costs literature does not inform us to take a stand on whether $\kappa^b \gtrless \kappa^s$. For analytical convenience in what follows we therefore often assume $\kappa^b = \kappa^s = \kappa$. Then, κ proxies for total *per trade* transactions costs and allows us to contrast cost thresholds under fixed *vis à vis* proportional costs.[18] Under proportional costs, emission permits are either in excess demand or supply for all $\kappa^T \in \left(\bar{\kappa}^T, \frac{p}{\gamma}\right)$, and whether $\frac{e^R}{\gamma} \lesseqgtr y^f$ depends uniquely on whether $e^R \lesseqgtr e^X$ (see Proposition 2). Under fixed costs, however, this result only holds with $\kappa^b = \kappa^s$ and when ε is symmetric around its mean.

We illustrate the above results more concretely by extending our parametric examples studied in section 3.

4.1 Parametric examples

We continue with the specifications studied under proportional costs in Section 3.1.1.

4.1.1 Additive, uniformly distributed shock

The fixed cost bounds defined in Proposition 3 are (see Appendix 7.10):

$$\bar{\kappa}^b = 2pc \frac{\tilde{z}^{\max} - e^R}{\tilde{z}^{\max} + \gamma c}; \; \bar{\kappa}^s = 2pc \frac{e^R + \gamma c}{\tilde{z}^{\max} + c}.$$

[18] With uniform distribution the mass of sellers equals that of buyers. Thus, each transaction can be thought of entailing a total cost of 2κ.

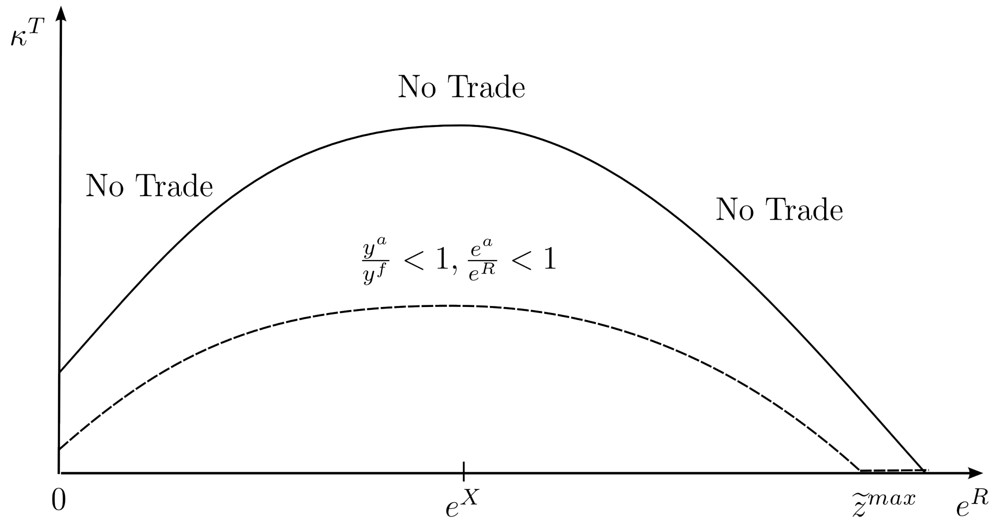


Figure 3: Trade outcomes under fixed trading costs.

When e^R is higher, the proportion of firms that have excess capacity is smaller and, therefore, a lower threshold discourages sellers from entering the permit market. The opposite is the case for sellers: A higher e^R increases the number of excess permits held by sellers, and thus allows them to break even at a higher trading cost. It is not possible to obtain a closed form solution for n and r in terms of model parameters. They are jointly determined from (18), (19), and (20):

$$\sqrt{n} = \frac{p - \gamma r}{2w} - \frac{\kappa^b - \kappa^s}{2c} = \frac{e^R}{\gamma} + \frac{1}{4c}\left[\left(\frac{\kappa^b}{p - \gamma r}\right)^2 - \left(\frac{\kappa^s}{\gamma r}\right)^2\right] \tag{23}$$

Evidently, unlike under proportional costs, where the equilibrium potential output depends only on κ^T, the equilibrium now depends on both κ^b and κ^s. For $\kappa^b \neq k^s$, $\frac{e^R}{\gamma} \lesseqgtr y^f \not\Rightarrow e^R \lesseqgtr e^X = \frac{\gamma p}{4w}$ in general (unlike Proposition 2 under proportional costs). For example, let $\kappa^b = 2\kappa^s$. Let $e^R = e^X$; then $\bar{\kappa}^b = \bar{\kappa}^s = pc$. Appendix 7.11 shows that $y^f \neq \frac{e^X}{\gamma}$ for $e^R = e^X$ for any $\kappa^s < \frac{pc}{3}$. On the other hand, if $\kappa^s \geq \frac{pc}{3}$, condition (21) stated in Proposition 3 does not hold and there is no permit trade in equilibrium.

Henceforth, we assume $\kappa^b = \kappa^s = \kappa$. Now (21) in Proposition 3 can be expressed as

$$\kappa < \bar{\kappa} \equiv \frac{\bar{\kappa}^b \bar{\kappa}^s}{\bar{\kappa}^b + \bar{\kappa}^s} = \frac{pc}{2}\left[1 - \left(\frac{e^R - e^X}{e^X + \gamma c}\right)^2\right]. \tag{24}$$

To facilitate a contrast with a similar result under proportional costs (see Figure 1), the relationship expressed in (24) is exhibited below in Figure 3.

Following (24), $\bar{\kappa}$ reaches its maximum of $\frac{pc}{2}$ at $e^R = e^X$ and monotonically declines on either side. However, in contrast with Figure 1, trade is only partial in the region below the cost threshold since some potential buyers and sellers prefer to remain inactive. Also, due to fixed costs, the threshold intersects with the x axis at $\tilde{z}^{\max} = \gamma\left(\frac{p}{2w} + c\right)$ instead of $\tilde{z}^a = \frac{\gamma p}{2w}$. The upshot is that even small transactions costs are detrimental to trade when the emission cap is either small or large. Once again, when the cap is small, there are few potential sellers, and a small trading cost drives them out of the market. When the cap is large, the same logic holds for a small number of buyers.

As is standard, n and r exhibit an inverse relationship at the planning stage as expressed by the first equation in (23)). However, in the permit market at the production stage, a higher n induces a higher demand for permits. Thus, as evident from the second part of (23), r is increasing in n. Thus, both $y^f(n)$ and r are uniquely determined as functions of e^R. It is easily checked that $\frac{\partial y^f}{\partial e^R} > 0$ and $\frac{\partial r}{\partial e^R} < 0$. In general, y^f and r not only depend on e^R, but also on κ and c. However, when $e^R = e^X$, $\gamma r = \frac{p}{\gamma}$ and $y^f = \frac{e^R}{\gamma}$ is independent of κ and c. Once again[19]

$$y^f \gtreqqless \frac{e^R}{\gamma} \Leftrightarrow e^R \lesseqqgtr e^X.$$

This result is identical to those in the absence of trade (see Lemma 1) and under proportional costs (see Proposition 1). But, as shown earlier, this does not generally hold with fixed trading costs.[20] Although a common underlying economic argument – that the optimal choice of n equates its upfront marginal cost with its expected marginal benefits – underlies the result as well, it is yet instructive to understand this result since it involves a novel extensive margin, i.e., the mass of active and inactive traders, that has not been encountered earlier. First, it is easy to check that $e^R = e^X = \frac{\gamma p}{4w}$, i.e., $\gamma r = \frac{p}{2}$ and $y^f = \sqrt{n} = \frac{p}{4w}$ satisfies (23). Here, a marginal increase in n increases potential output by $y_n^f = \frac{1}{2\sqrt{n}} = \frac{2w}{p}$. The expected marginal revenue however depends on the distribution of active/inactive firms in the permit market as shown in Figure 4(a). With a probability $\frac{1}{2} - \frac{\kappa}{pc}$, a firm has excess permits and is active in the market: its marginal revenue is $(p - \gamma r)\, y_n = w$; with probability $\frac{\kappa}{pc}$ the firm is inactive and the marginal revenue is $p\ y_n = 2\ w$. With probability $\frac{\kappa}{pc}$ the firm has an excess capacity but lacks matching emission permits: the marginal revenue is 0; with probability $\frac{1}{2} - \frac{\kappa}{pc}$ the firm fully utilizes its capacity by purchasing additional permits: the marginal revenue is $(p - \gamma r)\, y_n = w$. Thus, the expected marginal revenue is $2\left(\frac{1}{2} - \frac{\kappa}{pc}\right) w + \frac{\kappa}{pc} 2w = w$, thus precisely matching the input cost. Hence, $\sqrt{n} = \frac{p}{4w}$ is indeed optimal when $e^R = \frac{\gamma\ p}{4w}$. Continuing this argument, it can be checked that the expected marginal revenue exceeds w for a firm that sets $\sqrt{n} \le \frac{e^R}{\gamma}$ when $e^R < \frac{\gamma p}{4w}$ or sets $\sqrt{n} \ge \frac{e^R}{\gamma}$ when $e^R > \frac{\gamma p}{4w}$.

Equations (23) imply that equilibrium depends on $\frac{\kappa^2}{4c}$ jointly. When trading costs are proportional, every permit buyer and seller faces the same marginal returns from trade, whereas fixed costs affect the extensive margin, i.e., the mass of active traders. This mass depends on the trading costs relative to firms' heterogeneity (productivity spread). Define $\varkappa \equiv \frac{\kappa^2}{4c}$ as the *effective* trading cost. Then,

$$e^R \lesseqgtr e^X \Rightarrow \frac{\partial y^f}{\partial \varkappa} \gtreqless 0$$

That y^f may be increasing in trading costs is counterintuitive. However, this happens when $e^R < e^X$ and therefore $r > \frac{p}{2\gamma}$. The potential output distribution of firms active/inactive in the permit market is shown above in Figure 4(b). Evidently, higher trading costs drive out buyers relatively more than the sellers. The permit price must decrease in equilibrium. This induces firms to increase their ex ante input employment. The opposite is the case when $e^R > e^X$.

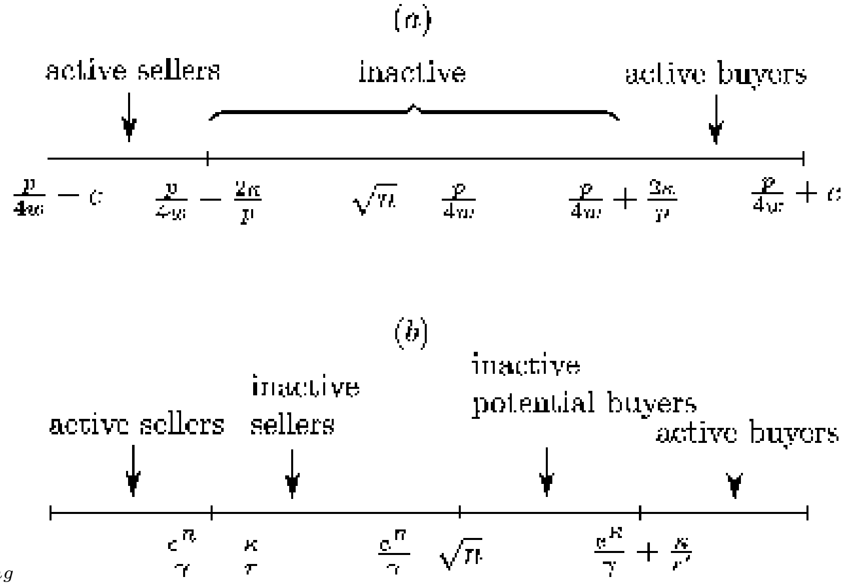
Finally, capacity and permit utilizations can be expressed as

$$\frac{y^a}{y^f} = 1 - \frac{\kappa^2}{2c} \frac{w}{(p - \gamma r)^3} < 1; \ \frac{e^a}{e^R} = 1 - \frac{\kappa^2}{4c} \frac{\gamma}{e^R} \left(\frac{1}{\gamma r}\right)^2 < 1,$$

[19] Recall that under proportional costs the following holds more generally, irrespective of (i) whether shocks are additive or multiplicative, (ii) the distribution of productivity shocks, and (iii) the relative magnitudes of buyers' and sellers' trading costs.

[20] It only holds for the special case of $\kappa^b = \kappa^s$ and a uniform distribution for ε. The two assumptions togther by offering a symmetry around $y^f = \frac{e^R}{\gamma} \Leftrightarrow e^R = e^X$ make this result possible.

4



4.png

Figure 4: Trading activity under fixed trading costs.

where the first inequality utilizes $r \in \left(0, \frac{p}{\gamma}\right)$. Neither production capacity nor permits are fully utilized in contrast with proportional costs where at least one of them is fully utilized. When $e^R = e^X$, the mass of inactive firms with excess capacity and with excess permits are identical and the above two expressions reduce to

$$\frac{y^a}{y^f} = \frac{e^a}{e^R} = 1 - \frac{4w}{p^3}\frac{\kappa^2}{c}$$

Both utilizations reduce with the effective trading cost $\varkappa$.[21] Finally,

$$e^R \lesseqgtr e^X \Rightarrow \frac{y^a}{y^f} \lesseqgtr \frac{e^a}{e^R}.$$

This follows from the fact that at sufficiently low emission caps firms over-employ n relative to permitted emissions, whereas firms under-employ n when the cap is sufficiently high.

4.1.2 Multiplicative, log-normally distributed shocks

From Proposition 2, $\bar{\kappa}^b = \infty$ for all $e^R \in (0, \tilde{z}^{\max})$ and $\bar{\kappa}^s = \frac{p}{\gamma} e^R$. For all finite κ^b, (21) reduces to

$$\kappa^s < \frac{p}{\gamma} e^R.$$

Under this specification, there are always some firms willing to purchase permits for any finite trading costs. It is the mass of active sellers that may vanish as κ^s gets larger. Thus, some trade always occurs as long as the preceding inequality is satisfied. Note also that $\bar{\kappa}^s$ is linearly increasing in e^R.[22]

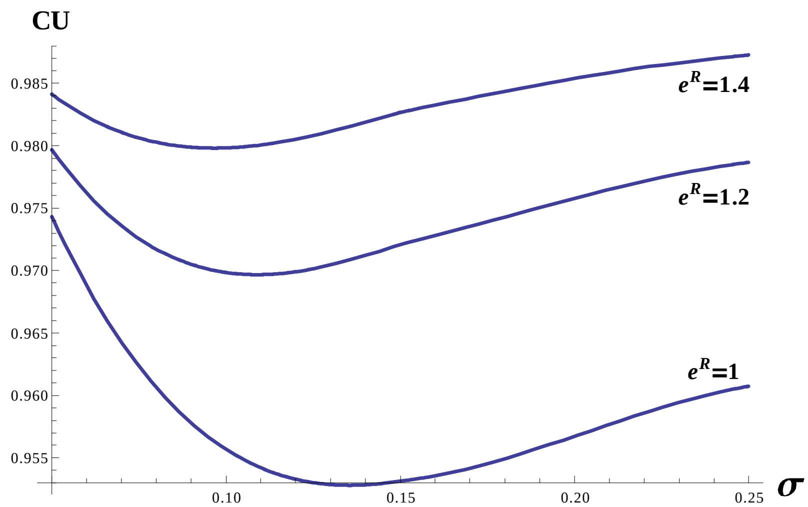
Even when $\kappa^b = \kappa^s = \kappa$, $y^f \neq \frac{e^R}{\gamma}$ for $e^R = e^X$ (see Appendix 7.12) unlike with additive and uniform ε. Under proportional costs, when $e^R = e^X$, the equilibrium quantity of permits demanded

[21] The result holds even when $e^R \neq e^X$.

[22] Both $\bar{\kappa}^b$ and $\bar{\kappa}^s$ will be finite if the support for ε is finite. For example, if $\varepsilon \in U[1-\epsilon, 1+\epsilon]$, with $\epsilon < 1$, it is easily checked that $\bar{\kappa}^b$ is finite and decreasing in e^R.

equals the quantity supplied, exactly as in costless trade, as long as $\kappa^T < \frac{p}{\gamma}$. The shape of firms' productivity distribution is irrelevant. However, since fixed costs affect the extensive margin, the shape of the distribution along with κ determines the prices and quantity traded in the permit market. Since $\log\varepsilon \sim N\left[0,\sigma^2\right]$ is asymmetric around the mean, $e^R = e^X$ does not lead to an ex ante choice of $y^f(n) = \frac{e^R}{\gamma}$.[23]

In the alternative specification with $y(n,\varepsilon) = y(n) + \varepsilon$, the equilibrium and the utilization rates depend solely on $\frac{\kappa^2}{\sigma}$.[24] The results however focus on the cases where $\kappa < \bar{\kappa}$. Clearly, for a given κ, an increase in σ improves utilization ratios. However, when $\kappa > \bar{\kappa}$, it is worth noting that in the absence of trade, an increase in σ would instead worsen utilization ratios. With $y(n,\varepsilon) = \varepsilon y(n)$ and $\log\varepsilon\tilde{}N\left[0,\sigma^2\right]$, a sharply peaked distribution of full capacity outputs may imply that very little trade occurs for a sufficiently large κ even when $\kappa < \bar{\kappa}^s = \frac{p}{\gamma}e^R$. In this case, an increase in σ may lead to a decrease in utilization ratios. Figure 5 exhibits capacity utilization for $e^R = 1$, 1.2, and 1.4, with $\kappa = 0.5$ and $\gamma = 1$, $p = 5$, and $w = 1$. The capacity utilization first declines (since it is practically under a no trade regime) and then improves as an increasing σ diminishes the trade-impeding effect of κ.



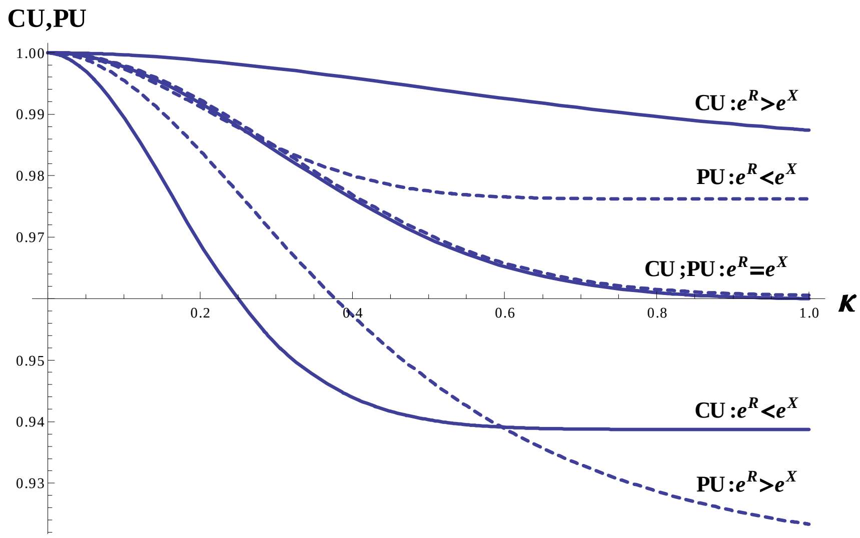
Capacity utilization (CU) vis--vis the variance of output shocks.

Notice that capacity utilization improves with a higher emission cap. However, as under additive uniform shocks, permit utilization decreases with e^R. Figure 6 shows how permit utilization flips around $e^R = e^X$. Here we assume $\sigma = 0.1$ along with the parameter values mentioned earlier. When $e^R = e^X$ then as under additive and uniform ε, capacity and permit utilizations, although not exactly equal, are

[23] It can be verified that when ε is symmetric (which requires its support to be finite as well), the symmetry of results around $e^R = e^X$ are similar as under additive and uniform shocks. When ε is asymmetric, then for all $\kappa < \bar{\kappa}$ there exists another $e^{X'} \neq e^X$ such that $y^f(n) = \frac{e^R}{\gamma} = \frac{e^{X'}}{\gamma}$. However, now $e^{X'}$ depends also on κ.

[24] Note that $\sigma = \frac{c}{\sqrt{3}}$ when $f(n,\varepsilon) = n + \varepsilon$ with $\varepsilon \sim U[-c,c]$.

almost identical. This is because the distribution of ε for $\sigma = 0.1$ is fairly symmetric around the mean.



As expected, both permit and capacity utilizations decrease with κ. Both ratios converge to their no trade levels as κ becomes sufficiently large.

5 Conclusions

We study equilibrium production and capacity utilization in a model with random production and emissions, and costly trade in emission permits. The equilibria critically depend on whether the marginal trading costs are constant (proportional) or whether firms incur a fixed cost of entering the permit market. Trading costs impede permit trade relatively more when the emission cap is either sufficiently low or sufficiently high. At these extremes, small fixed costs may totally shut trade off by driving either buyers or sellers out of the permit market.

Under proportional costs, trade always occurs as long as total per unit trading costs are below the total surplus, which is derived from added output revenue generated when firms utilize available production capacity. There exists an emission cap-dependent threshold such that if the trading cost lies below the threshold, emission caps and all firm output capacities are fully utilized. This bound is non-monotonic, it first increases with the emission cap and then decreases as the cap approaches a point where it no longer binds. If trading cost lies above the threshold, emission permit trade is partial, i.e., there remains either an excess supply or an excess demand for permits. A similar non-monotonic cost bound exists under fixed trading costs if the output randomness has a finite support. If trading costs lie above the bound there is no permit trade. If they lie below, trade occurs but is only partial. There are always firms who are unable to sell their excess permits and others who end up with unutilized production capacity.

Under both types of trading costs, firms' capacity utilization is increasing in the emission cap. With proportional costs, capacity utilization is less than 100% only when the emission cap is sufficiently low and trading costs are sufficiently high. On the other hand, under fixed costs, some production capacity remains idle for all levels of emission caps and trading costs. The bottomline is that proportional trading costs cause permit excesses or shortages only if these costs are sufficiently high, whereas there are always some firms with excess permits and some with permit shortage under fixed costs.

The implications of trading costs – either fixed or proportional - for firms production plans and capacity utilization have not been addressed in the literature. This is the main novelty of our work.

To highlight the role that transactions costs play in impeding trade and efficient equilibria, we have presented a simple model of choice under uncertainty. Sans uncertainty, there is no role for trade or trading frictions. The model, however, abstracts from aggregate uncertainty. We conjecture that introducing aggregate uncertainty in the model will not change the qualitative nature of our results. We also conjecture that corrective policies can be used that induce full capacity utilization and regain efficient production equilibria even in the absence of permit trade. This is left for future research.

6 References

Rolf Färe, Shawna Grosskopf, Dong-Woon Noh, and William Weber, Characteristics of a Polluting Technology: Theory and Practice, *Journal of Econometrics*, Vol. 126 (2005): 469 - 492.

Rolf Färe, Shawna Grosskopf, and Carl A. Pasurka Jr., Tradable permits and unrealized gains from trade, *Energy Economics*, 40 (2013): 416-424.

Gangadharan, Lata, Transactions costs in pollution markets: An empirical study. *Land Economics*, Vol. 76 (2000): 601-614.

Heindl, Peter, Transactions Costs and Tradable Permits: Empirical Evidence from the EU Emissions Trading Scheme, Discussion paper No. 12-021, Center for European Economic Research, March 2012. Available at http://ftp.zew.de/pub/zew-docs/dp/dp12021.pdf.

Jaraitė, Jūrate and Andrius Kužukauskas, Firm Trading Behavior and Transactions Costs in the European Union's Emission Trading System: An Empirical Assessment, CERE Working Paper, 2012:9, November 11, 2012.

Kerr, Suzi and David C. Maré, Transactions costs and tradable permit markets: The United States lead phasedown. Available at SSRN http:/ssrn.com/abstract=1082596 (June 30, 1998).

Krutilla, Kerry and Rachel Krause, Transactions Costs and Environmental Policy: An Assessment Framework and Literature Review. *International Review of Environmental and Resource Economics*, Vol. 4 (2010): 261-354.

Mas-Colell, Andreu, Michael D. Whinston and Jerry R. Green, Microeconomic Theory, Oxford university Press, 1995.

Michaelowa, Axel, Marcus Stronzik, Frauke Eckermann and Alistair Hunt, Transactions costs and the Kyoto Protocol, *Climate Policy*, Vol. 3 (2003): 261-278.

Montero, Juan-Pablo, Marketable pollution permits with uncertainty and transactions costs, *Resource and Energy Economics*, Vol. 20 (1997): 27-50.

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

Murty, Sushama., R. Robert Russell and Steven B. Levkoff, On modeling pollution-generating technologies, *Journal of Environmental Economics and Management*, 64 (2012): 117-135.

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

Zhao, Jinhua, Irreversible abatement investment under cost uncertainties: tradable emission permits and emissions charges, *Journal of Public Economics*, Vol. 87 (2003): 2765-2789.

7 Appendix

7.1 Equilibrium with costless permit trade

While acquiring input n and permits e^0 at the planning stage, firms expect their profits to be:

$$\int_{\varepsilon^{\min}}^{\varepsilon^{\max}} \left(p\ y\left(n,\varepsilon\right) - r\left(\gamma y\left(n,\varepsilon\right) - e^0\right)\right)\ g(\varepsilon)d\varepsilon = \left(p-\gamma r\right) y^f\left(n\right) + r\ e^0. \tag{25}$$

The firm's choice at the production stage is trivial: as long as $p > \gamma r$, a firm i with $\gamma y\left(n,\varepsilon^i\right) > e^0$ will buy the difference as additional permits. On the other hand, every firm with $\gamma y\left(n,\varepsilon^i\right) < e^0$ will sell excess permits as long as $r > 0$. We will require the price in the permit market at the production stage to be consistent with its forecasted value at the planning stage. At the planning stage, the firm's profit maximization problem is to choose e^0 and n to maximize expected profit:

$$E\left\{\pi\left(n,e^0\right)\right\} \equiv \left(p-\gamma r\right) y^f\left(n\right) + re^0 - wn - re^0.$$

An equilibrium requires that $r = \rho$. Suppose $r > \rho$. Then a firm can buy all the permits at ρ, and sell it later at production stage with per unit profit of $r - \rho$. If $r < \rho$, it is better for everyone to sell permits in the first stage and then buy back later, which in equilibrium is impossible since all firms are identical ex ante. To implement $r = \rho$ we assume that a forward market, potentially through intermediaries or a centralized exchange, exists that credibly promises to buy and sell permits at the production stage at the pre-production price.

By assumption $y^f\left(n\right)$ is strictly concave in n. Therefore, a necessary and sufficient condition for the optimal choice of n is

$$y_n^f\left(n\right) \equiv \frac{dy^f\left(n\right)}{dn} = \frac{w}{p-\gamma\rho},$$

where we have used $r = \rho$. This is stated as (3) in the main text. It follows from the strict concavity of $y^f\left(.\right)$ that n is decreasing in ρ. Since $\rho < 0$ is ruled out, n can, at most, equal $\tilde{n}$. Further, since $n \geq 0$, $\gamma\rho \leq p$.

We have shown above that $\rho \in \left[0, \frac{p}{\gamma}\right]$. In particular if n is finite, $\rho > 0$ and firms utilize all permits acquired at the planning stage, either for its own use or for resale; the latter, in turn, is in demand from firms whose emission permits fall short of their potential output level of emissions. Further, $n > 0$ implies $\gamma\rho < p$ and no firm wastes its production capacity. A firm i with productivity realization ε^i readjusts its emission permits to $e^i = \gamma y\left(n,\varepsilon^i\right)$, and therefore $e^i - e^{i0}$ denotes the amount of permits it buys (sells, if negative) in the permit market. In the aggregate $\int e^i di = e^R$, which obtains (4) in the main text.

7.2 Equilibrium without permit trade

When permit trade is prohibitively costly, the choice problem at the production stage is trivial. The firm can not exceed its permitted emissions: $y^i \leq y\left(n,\varepsilon^0\right) = \frac{e^0}{\gamma}$ must hold. The firm's first stage choice problem is

$$\max_{\{n,e^0\}}\left\{\int_{\varepsilon^{\min}}^{\varepsilon^0} py\left(n,\varepsilon\right)\ g\left(\varepsilon\right)d\varepsilon+\int_{\varepsilon^0}^{\varepsilon^{\max}} p\frac{e^0}{\gamma}\ g\left(\varepsilon\right)d\varepsilon-\rho e^0-wn\right\},$$

where the first two terms summarize firm's expected sales revenue. The optimal choices of n and e^0 respectively obtain (5b) and (5a) in the main text.

7.3 Equilibrium with uncertain emission efficiency

Here, we assume that a firm i employing n can produce up to $y\left(n\right)$ consumer goods; but a unit of output produced generates $\varepsilon^i\gamma$ units of emissions. Thus,

$$y^i\left(n\right)=\min\left\{y\left(n\right),\frac{e^i}{\varepsilon^i\gamma}\right\},$$

where e^i is its permit holding. Then, firms with $\gamma\varepsilon^i y\left(n\right)<e^0$ will have excess emission permits while those with $\gamma\varepsilon^i y\left(n\right)>e^0$ would like to acquire additional permits. A costless permit trade will yield an equilibrium n^* and ρ^* given by

$$\begin{aligned}
y_n\left(n^*\right) &= \frac{w}{p-\gamma E\left\{\varepsilon\right\}\rho^*};\\
y^{a*} &= y^{f*}=\frac{e^R}{E\left\{\varepsilon\right\}\gamma}.
\end{aligned}$$

The equilibrium $\{\rho^*,n^*\}$ is identical to that stated in (3) and (4) for $y\left(n,\varepsilon\right)=\varepsilon y\left(n\right)$ and $E\left\{\varepsilon\right\}=1$. In the absence of permit trade, firms' planning problem is

$$\max_{\{n,e^0\}}\left\{\int_{\varepsilon^{\min}}^{\varepsilon^0} py\left(n\right)\ g\left(\varepsilon\right)d\varepsilon+\frac{p}{\gamma}\int_{\varepsilon^0}^{\varepsilon^{\max}} \frac{e^0}{\varepsilon}\ g\left(\varepsilon\right)d\varepsilon-\rho e^0-wn\right\},$$

where $\varepsilon^0=\frac{e^0}{\gamma y(n)}$. In equilibrium $\hat{\rho}$ and $\hat{n}$ will be given by

$$\begin{aligned}
\hat{\rho} &= \frac{p}{\gamma}\int_{\varepsilon^R}^{\varepsilon^{\max}} \varepsilon^{-1}\ g\left(\varepsilon\right)d\varepsilon;\\
w &= py_n\left(\hat{n}\right)G\left(\varepsilon^R\right),
\end{aligned}$$

Clearly, by re-scaling γ and y, $\{\hat{\rho},\hat{n}\}$ can be made identical to that given by (5a) and (5b). Total production of consumer goods is

$$\hat{y}^a=y\left(n\right)G\left(\varepsilon^R\right)+\frac{e^R}{\gamma}\int_{\varepsilon^R}^{\varepsilon^{\max}}\frac{1}{\varepsilon}g\left(\varepsilon\right)d\varepsilon<y^f=y\left(n\right)$$

and the output lost due to partial capacity utilization is

$$\hat{y}^f-\hat{y}^a=\int_{\varepsilon^R}^{\varepsilon^{\max}}\left(y\left(\hat{n}\right)-\frac{e^R}{\gamma}\frac{1}{\varepsilon}\right)g\left(\varepsilon\right)d\varepsilon$$

Thus, the results under this alternative specification are equivalent to that in the main text. It can be verified that qualitatively the results will be similar to that in the main text under partial trade with either proportional or fixed trading costs.

7.4 Equilibrium with reversible inputs

With reversible inputs, firms trade and acquire emission permits e^0 in a pre-production market at per unit price ρ. After the permits are obtained, they draw their productivity. Without any emission cap, their optimal input employment follows

$$py_n\left(n,\varepsilon^i\right)=w.$$

With emission caps, there are two possibilities. If permits can be traded, the optimal input employment is given by

$$y_n\left(n,\varepsilon^i\right)=\frac{w}{p-\rho\gamma}, \tag{26}$$

and in equilibrium:

$$y^a=\int y\left(n,\varepsilon\right)g\left(\varepsilon\right)d\varepsilon=\frac{e^R}{\gamma} \tag{27}$$

In the absence of trade, and assuming that the emission cap binds for all firms, input employment follows

$$y\left(n^i,\varepsilon^i\right)=\frac{e^R}{\gamma}$$

and once again (53) holds. Allocations implied by (52) and (53) replicate the allocations of a planner who only faces aggregate constraint $\gamma y^a=e^R$, whereas no-trade outcome implies a constraint on each firm. Thus, the two outcomes are different – producing the same output would utilize more inputs in the absence of trade.

7.5 Proof of Lemma 1

We first determine e^X and show that it is unique. Let $\hat{n}^X\equiv\hat{n}\left(e^X\right)$. For $y\left(n\,,\varepsilon\right)=y\left(n\right)+\varepsilon$

$$y_n\left(\hat{n}^X\right)G\left[\frac{e^X}{\gamma}-y\left(\hat{n}^X\right)\right]=\frac{w}{p}\text{ and }\frac{e^X}{\gamma}=y\left(\hat{n}^X\right)+E\left\{\varepsilon\right\}$$

Substituting the second into the first obtains $\hat{n}^X$ uniquely from

$$y_n\left(\hat{n}^X\right)=\frac{w}{pG\left[E\left\{\varepsilon\right\}\right]}$$

which in turn determines e^X uniquely. For $y\left(n,\,\varepsilon\right)=\varepsilon y\left(n\right)$

$$y_n\left(\hat{n}^X\right)E\left[\varepsilon|\varepsilon\leq\frac{e^X}{\gamma y\left(\hat{n}^X\right)}\right]=\frac{w}{p}.\text{ and }\frac{e^X}{\gamma}=E\left\{\varepsilon\right\}y\left(\hat{n}^X\right)$$

Here, $\hat{n}^X$ is uniquely determined from

$$y_n\left(\hat{n}^X\right)=\frac{w}{pE\left[\varepsilon|\varepsilon\leq E\left\{\varepsilon\right\}\right]}$$

Note that $\hat{n}\left(\gamma y\left(\tilde{n},\varepsilon^{\max}\right)\right)=\tilde{n}$ and $\hat{y}^f\left(\tilde{n}\right)=\int_{\varepsilon^{\min}}^{\varepsilon^{\max}}y\left(\tilde{n},\varepsilon\right)g\left(\varepsilon\right)d\varepsilon<y\left(\tilde{n},\varepsilon^{\max}\right)$. Thus, as $e^R\to\gamma y\left(\tilde{n},\varepsilon^{\max}\right)>e^X$, $\hat{y}^f\to\tilde{y}^f<\frac{e^R}{\gamma}$. Since $\frac{e^X}{\gamma}$ is unique, for all $e^R<e^X$, $\hat{y}^f>\frac{e^R}{\gamma}$ must hold.

7.6 Proof of Proposition 1

Consider the $y^f = \frac{e^R}{\gamma}$ equilibrium first. Here $r > \kappa^s$ and $r < \frac{p}{\gamma} - \kappa^b$. Begin with $\kappa^b = \kappa^s = 0$. Here $y^{a*} = y^f(n^*) = \frac{e^R}{\gamma}$ determines n^*, (7) and (8) imply $\rho = r = \rho^*$ as determined from (3).

We verify below that as long as κ^b, κ^s, and $\kappa^b + \kappa^s = \kappa^T$ are sufficiently small, n^* continues to be the optimal input choice and ρ and r are determined by (7) and (8). Below, we consider the two cases with $e^R \lesseqgtr e^X$.

7.6.1 $e^R < e^X$

Fix $e^R < e^X$. Note that $\kappa^T = \frac{p}{\gamma}$ satisfies equation (10a), i.e.,

$$w = \gamma\kappa^T \int_{\varepsilon^{\min}}^{\varepsilon^R} y_n(n,\varepsilon)\ g(\varepsilon)d\varepsilon,$$

with $n = \hat{n}$ and then from Lemma 1 $\hat{y}^f > \frac{e^R}{\gamma}$. It is easily shown from the above that $\frac{\partial n}{\partial \kappa^T} > 0$, so if $y^f > \frac{e^R}{\gamma}$ for $\kappa^{T'} < \frac{p}{\gamma}$, $y^f > \frac{e^R}{\gamma}$ for all $\kappa^T \in \left[\kappa^{T'}, \frac{p}{\gamma}\right]$. Now define $\bar{\kappa}^T$ as

$$w = \gamma\bar{\kappa}^T \int_{\varepsilon^{\min}}^{\varepsilon^R} y_n\left(n^*\left(e^R\right),\varepsilon\right)\ g(\varepsilon)d\varepsilon = (p - \gamma\rho^*) \int_{\varepsilon^{\min}}^{\varepsilon^{\max}} y_n\left(n^*\left(e^R\right),\varepsilon\right)\ g(\varepsilon)d\varepsilon$$

Clearly, there exists such a $\bar{\kappa}^T \in \left(0, \frac{p}{\gamma}\right)$ for all $\rho^* \in \left(0, \frac{p}{\gamma}\right)$. Thus for all $\kappa^T \in (\bar{\kappa}^T, \frac{p}{\gamma}]$, $y^f > \frac{e^R}{\gamma}$, and for $\kappa^T = \bar{\kappa}^T$, $y^f = \frac{e^R}{\gamma}$ by definition. Continuity then implies that

$$y^f > \frac{e^R}{\gamma} \text{ for all } e^R < e^X \text{ and } \kappa^T \in (\bar{\kappa}^T, \frac{p}{\gamma}]$$

When $\kappa^T < \bar{\kappa}^T$, equation (10a) no longer holds because it would then imply $y^f < e^R$. Also, since $\bar{\kappa}^T < \frac{p}{\gamma}$, equation (10b) can not hold simultaneously with (10a) when $\kappa^T = \bar{\kappa}^T$. Therefore, $y^f = e^R$, $n = n^*$, and $\rho = r = \rho^*$ are given by (3).

Finally, it is easily checked that $\frac{\partial \bar{\kappa}^T}{\partial e^R} > 0$.

7.6.2 $e^R > e^X$

Now, fix $e^R > e^X$. Note that $\kappa^T = \frac{p}{\gamma}$ satisfies

$$w = p \int_{\varepsilon^{\min}}^{\varepsilon^{\max}} y_n(n,\varepsilon)\ g(\varepsilon)d\varepsilon - \gamma\kappa^T \int_{\varepsilon_q}^{\varepsilon^{\max}} y_n(n,\varepsilon)\ g(\varepsilon)d\varepsilon$$

with $n = \hat{n}$ and then from Lemma 1, $y^f < e^R$. It is easily shown from the above that $\frac{\partial n}{\partial \kappa^T} < 0$, so $y^f < \frac{e^R}{\gamma}$ for $\kappa^{T'} < \frac{p}{\gamma}$, $y^f < \frac{e^R}{\gamma}$ for all $\kappa^T \in \left[\kappa^{T'}, \frac{p}{\gamma}\right]$. Now define $\bar{\kappa}^T$ as

$$\begin{aligned} w &= p \int_{\varepsilon^{\min}}^{\varepsilon^{\max}} y_n\left(n^*\left(e^R\right),\varepsilon\right)\ g(\varepsilon)d\varepsilon - \gamma\bar{\kappa}^T \int_{\varepsilon^R}^{\varepsilon^{\max}} y_n\left(n^*\left(e^R\right),\varepsilon\right)\ g(\varepsilon)d\varepsilon \\ &= (p - \gamma\rho^*) \int_{\varepsilon^{\min}}^{\varepsilon^{\max}} y_n\left(n^*\left(e^R\right),\varepsilon\right)\ g(\varepsilon)d\varepsilon \end{aligned}$$

Clearly, there exists such a $\bar{\kappa}^T \in \left(0, \frac{p}{\gamma}\right)$ for all $\rho^* \in \left(0, \frac{p}{\gamma}\right)$. Thus for all $\kappa^T \in (\bar{\kappa}^T, p]$, $y^f < \frac{e^R}{\gamma}$, and for $\kappa^T = \bar{\kappa}^T$, $y^f = \frac{e^R}{\gamma}$ by definition. Continuity then implies that

$$y^f < \frac{e^R}{\gamma} \text{ for all } e^R > e^X \text{ and } \kappa^T \in (\bar{\kappa}^T, \frac{p}{\gamma}]$$

When $\kappa^T < \bar{\kappa}^T$, equation (10b) no longer holds because it would then imply $y^f > \frac{e^R}{\gamma}$. Also, since $\bar{\kappa}^T < \frac{p}{\gamma}$, equation (10a) can not hold simultaneously with (10b) when $\kappa^T = \bar{\kappa}^T$. Therefore, $y^f = \frac{e^R}{\gamma}$, $n = n^*$, and $\rho = r = \rho^*$ are given by (3).

Finally, it is easily checked that $\frac{\partial \bar{\kappa}^T}{\partial e^R} < 0$.

7.6.3 $e^R = e^X$

Here, $y^f = \frac{e^X}{\gamma}$ holds for $\kappa^T = \frac{p}{\gamma}$ simultaneously in both (10b) and (10a) with $n = \hat{n}$. However, any other value of $\kappa^T < \frac{p}{\gamma}$ is not consistent with either $y^f\left(n\left(e^X\right)\right) > \frac{e^X}{\gamma}$ (as governed by (10a)) or $y^f\left(n\left(e^X\right)\right) < \frac{e^X}{\gamma}$ (as governed by (10b)) because in either case the potential output versus emission cap mismatch continues to hold for higher values of κ^T which will be violated as $\kappa^T \to \frac{p}{\gamma}$. Hence $\bar{\kappa}^T = \frac{p}{\gamma}$ for $e^R = e^X$.

7.7 Derivations with uniformly distributed shocks and proportional costs

With $y(n, \varepsilon) = \sqrt{n} + \varepsilon$, with $\varepsilon \sim U[-c, c] \Rightarrow y^f(n) = \sqrt{n}$; $\varepsilon^R = \frac{e^R}{\gamma} - \sqrt{n}$; $\tilde{y}^a = y^f(\tilde{n}) = \sqrt{\tilde{n}} = \frac{p}{2w}$; and $\tilde{z}^{\max} = \gamma \tilde{y}^{\max} = \gamma\left(\frac{p}{2w} + c\right)$. To ensure that $y\left(n, \varepsilon^{\min}\right) > 0$, i.e., $\sqrt{n} > c$, we assume that $c < \frac{1}{2}\sqrt{\frac{pe^R}{\gamma w}}$ for all $e^R < \gamma\left(\frac{p}{2w} + c\right)$, and $c < \frac{p}{2w}$, for all $e^R \geq \gamma\left(\frac{p}{2w} + c\right)$. $G\left(\varepsilon^R\right) = \frac{\frac{e^R}{\gamma} - \sqrt{n} + c}{2c}$. Evaluating (5b) with $\frac{e^X}{\gamma} = y^f(\hat{n}) \Rightarrow \varepsilon^X = 0$ obtains $e^X = \frac{\gamma\, p}{4w}$.

Applying proposition 2 equations (10a) and (10b), we have (61) in the main text. Below we verify that as long as $\kappa^T < \bar{\kappa}^T\left(e^R\right)$

$$n = n^* = \left(\frac{e^R}{\gamma}\right)^2,$$

If the above holds, (3), (8), and (11):

$$\rho = r + \frac{\kappa^b - \kappa^s}{2} = \rho^* = \frac{1}{\gamma}\left(p - 2\frac{e^R}{\gamma}w\right)$$

Thus,

$$r = \frac{1}{\gamma}\left(p - 2\frac{e^R}{\gamma}w\right) - \frac{\kappa^b - \kappa^s}{2} = \rho^* - \frac{\kappa^b - \kappa^s}{2} \tag{28}$$

We now identify the set of e^R and $\kappa^T \in \left[0, \frac{p}{\gamma}\right]$ that jointly characterize the three possible permit market equilibria as stated by (9).

7.7.1 Case I: $e^R < e^X = \frac{\gamma p}{4w}$;

Suppose $\kappa^T < \bar{\kappa}^T = 4\frac{e^R}{\gamma^2}w$ and $\kappa^s = 0$, i.e., $\kappa^b < 4w\frac{e^R}{\gamma^2}$. Then from (28) $r + \kappa^b < \frac{p}{\gamma}$. If instead $\kappa^b = 0$ and $\kappa^s < 4w\frac{e^R}{\gamma^2}$, $r > \kappa^s$ because $e^R < \gamma\frac{p}{4w}$. Thus, the third condition in (9) holds.

Now if $\kappa^T > \bar{\kappa}^T = 4\frac{e^R}{\gamma^2}w$ proposition 2 equation (10a) yields

$$w = \gamma \frac{\kappa^T}{2c} \int_{-c}^{\frac{e^R}{\gamma}-\sqrt{n}} \frac{1}{2\sqrt{n}} d\varepsilon \Rightarrow \sqrt{n} = \frac{\frac{e^R}{\gamma}+c}{1+4wc/\gamma\kappa^T} \tag{29}$$

Since $\kappa^T > 4\frac{e^R}{\gamma^2}w$, $\sqrt{n} > \frac{e^R}{\gamma}$ and $r = \frac{p}{\gamma} - \kappa^b$. To obtain ρ, we use (8):

$$\rho = \frac{p}{\gamma} - \kappa^T \frac{e^R/\gamma - \sqrt{n} + c}{2c}.$$

which as $\kappa^T \to \frac{p}{\gamma}$ reduces to

$$\hat{\rho} = \frac{p}{\gamma} \frac{\sqrt{\hat{n}} + c - e^R/\gamma}{2c}.$$

The last expression can also be obtained by using (5b) and (5a). With $\kappa^T < \frac{p}{\gamma}$, (29) implies $\sqrt{n} = \frac{\frac{e^R}{\gamma}+c}{1+4w\frac{c}{\gamma\kappa^T}} < \sqrt{\hat{n}} = \frac{e^R/\gamma+c}{1+4w\frac{c}{p}}$. Thus, $y^f \in \left(\frac{e^R}{\gamma}, \hat{y}^f\right)$. Finally,

$$y^f - y^a = \sqrt{n} - \frac{e^R}{\gamma} = c\frac{1-\frac{\bar{\kappa}^T}{\kappa^T}}{1+4w\frac{c}{\gamma\kappa^T}} > 0; \frac{y^f - y^a}{y^f} = \frac{\gamma c}{e^R + \gamma c}\left(1 - \frac{\bar{\kappa}^T}{\kappa^T}\right)$$

The output loss increases with κ^T and reaches its upper bound as $\kappa^T \to \frac{p}{\gamma}$.[25]

7.7.2 Case II: $e^R > e^X = \frac{\gamma}{4w}\frac{p}{}$;

Let $\kappa^T < \bar{\kappa}^T = \frac{2p-4we^R/\gamma}{\gamma}$ and $\kappa^s = 0$, i.e., $\kappa^b < \frac{2p-4we^R/\gamma}{\gamma}$. Then from (28) $r + \kappa^b < \frac{2p-4we^R/\gamma}{\gamma} = \frac{p}{\gamma} + \frac{4w}{\gamma}\left(\frac{p}{4w} - \frac{e^R}{\gamma}\right) < p$ and $r > \kappa^s = 0$. If $\kappa^b = 0$ and $\kappa^s < \frac{2p-4we^R/\gamma}{\gamma}$, $r = \frac{1}{\gamma}\left(p - 2w\frac{e^R}{\gamma}\right) + \frac{\kappa^s}{2} > \kappa^s$ and $r < \frac{2p-4we^R/\gamma}{\gamma} < \frac{p}{\gamma}$.

Now if $\kappa^T > \frac{2p-4we^R/\gamma}{\gamma}$ proposition 2 equation (10b) yields

$$w = \frac{p}{2c}\int_{-c}^{c} \frac{1}{2\sqrt{n}} d\varepsilon - \frac{\gamma\kappa^T}{2c}\int_{\frac{e^R}{\gamma}-\sqrt{n}}^{c} \frac{1}{2\sqrt{n}} d\varepsilon \Rightarrow \sqrt{n} = \frac{\frac{e^R}{\gamma} + c\left(\frac{2p}{\gamma\kappa^T} - 1\right)}{1+4w\frac{c}{\gamma\kappa^T}} \tag{30}$$

Obviously, since $\kappa^T > \frac{2p-4we^R/\gamma}{\gamma}$, $\sqrt{n} = y^f(n) < \frac{e^R}{\gamma}$, $r = \kappa^s$, we use equation (8) to get

$$\rho = \kappa^T \frac{\sqrt{n} + c - \frac{e^R}{\gamma}}{2c}$$

[25] When $\kappa^T = \frac{p}{\gamma}$, there is no surplus from trade because $r + \kappa_b = \frac{p}{\gamma}$ and $r = \kappa_s$. If we assume that firms continue to trade, the loss is given by the above expression with κ^T replaced by $\frac{p}{\gamma}$. On the other hand, if we assume that no trade occurs when $\kappa^T = \frac{p}{\gamma}$, all firms with $y(n, \varepsilon) > \frac{e^R}{\gamma}$ are unable to utilize full capacity. Then it can be shown that

$$y^f - y^a = c\left(\frac{\frac{p}{2w} + c - \frac{e^R}{\gamma}}{\frac{p}{2w} + 2c}\right)^2 > 0 \text{ for all } \frac{e^R}{\gamma} < \frac{p}{2w} + c.$$

Thus, the lost output due to capacity underutilization exhibits a discontinuity at $\kappa^T = \frac{p}{\gamma}$.

which, again, in the limit as $\kappa^T \to \frac{p}{\gamma}$ reduces to $\hat{\rho}$ obtained earlier. However, since $\kappa^T < \frac{p}{\gamma}$, input choice given by (30) $\sqrt{n} = \frac{\frac{e^R}{\gamma}+c\left(\frac{2p}{\gamma\kappa^T}-1\right)}{1+4w\frac{c}{\gamma\kappa^T}} > \sqrt{\hat{n}} = \frac{\frac{e^R}{\gamma}+c}{1+4w\frac{c}{p}}$. Now, $y^f \in \left(\hat{y}^f, \frac{e^R}{\gamma}\right)$. Finally,

$$e^R - e^a = e^R - \frac{e^R + \gamma c\left(\frac{2p}{\gamma\kappa^T} - 1\right)}{1 + 4w\frac{c}{\gamma\kappa^T}}; \frac{e^R - e^a}{e^R} = \frac{\gamma c}{e^R} \frac{1 - \frac{\bar{\kappa}^T}{\kappa^T}}{1 + \frac{c}{e^X}\frac{p}{\kappa^T}}$$

7.7.3 Case III: $e^R = e^X = \frac{\gamma p}{4w}$;

Let $\kappa^T < \bar{\kappa}^T = \frac{p}{\gamma}$ and $\kappa^s = 0$, i.e., $\kappa^b < \frac{p}{\gamma}$. Then from (28) $r + \kappa^b < \frac{p - 2w\frac{e^R}{\gamma}}{\gamma} + \frac{p}{2\gamma} = p$ and $r > \kappa^s = 0$. If $\kappa^b = 0$ and $\kappa^s < \frac{p}{\gamma}$, $r = \frac{1}{\gamma}\left(p - 2w\frac{e^R}{\gamma}\right) + \frac{\kappa^s}{2} = \frac{\frac{p}{\gamma}+\kappa^s}{2} > \kappa^s$ and $r < \frac{p}{\gamma}$. Thus, for all $\kappa^T \leq \frac{p}{\gamma}$, $y^f = y^{f*} = \hat{y}^f = \frac{e^X}{\gamma}$.

7.8 Derivations for lognormal shocks and proportional trading costs

For future reference if $\ln\varepsilon \sim N\left[0, \sigma^2\right]$, then the following definite integral can be expressed as

$$\int_{\varepsilon^l}^{\varepsilon^h} \varepsilon\ g(\varepsilon)\ dy$$
$$= \exp\left(\frac{\sigma^2}{2}\right)\left[\Phi\left(\frac{\ln\varepsilon^h - \sigma^2}{\sigma}\right) - \Phi\left(\frac{\ln\varepsilon^l - \sigma^2}{\sigma}\right)\right]$$

Also note that $\sqrt{n^*}\exp\left[\frac{1}{2}\sigma^2\right] = \frac{e^R}{\gamma}$. Then, $\varepsilon^R = \exp\left[\frac{1}{2}\sigma^2\right]$

$$G\left[\varepsilon^R|n^*\right] = \Pr\left[\ln\varepsilon \leq \frac{1}{2}\sigma^2\right] = \Phi\left[\frac{\frac{1}{2}\sigma^2 - 0}{\sigma}\right] = \Phi\left[\frac{1}{2}\sigma\right]$$

For obtaining e^X, first obtain $\hat{n}$:

$$w = \frac{p}{2\sqrt{\hat{n}}}\int_{\varepsilon^{\min}}^{\varepsilon^R} \varepsilon g(\varepsilon)\, dy = \frac{p}{2\sqrt{\hat{n}}}\exp\left(\frac{\sigma^2}{2}\right)\Phi\left(\frac{\ln\left(\frac{e^R/\gamma}{\sqrt{\hat{n}}}\right) - \sigma^2}{\sigma}\right)$$

When $e^R = e^X$

$$y^f(\hat{n}) = \sqrt{\hat{n}}\exp\left(\frac{1}{2}\sigma^2\right) = \frac{e^X}{\gamma},$$

Combining the above two gets:

$$e^X = \gamma\frac{p}{2w}\exp\left(\sigma^2\right)\Phi\left(-\frac{1}{2}\sigma\right)$$

With $n = n^*$,

$$y^{a*} = y^f(n^*) = \sqrt{n^*}\exp\left(\frac{1}{2}\sigma^2\right) = \frac{e^R}{\gamma}$$

Following (3):

$$w = \frac{(p - \gamma\rho^*)}{2\sqrt{n^*}}\exp\left(\frac{1}{2}\sigma^2\right).$$

Combining the above two gets:

$$\rho^* = \frac{p - 2w\frac{e^R}{\gamma}\exp\left[-\sigma^2\right]}{\gamma}$$

For $e^R < e^X$, to get $\bar{\kappa}^T$ use (10a) to get (with $e^R = \gamma\sqrt{n^*}\exp\left(\frac{1}{2}\sigma^2\right)$)

$$\begin{aligned} w &= \frac{\gamma\bar{\kappa}^T}{2\sqrt{n}}\int_0^{\varepsilon^R}\varepsilon g\left(\varepsilon\right)d\varepsilon = \frac{\gamma\bar{\kappa}^T}{2\sqrt{n}}\exp\left(\frac{\sigma^2}{2}\right)\Phi\left(\frac{\ln\varepsilon^R - \sigma^2}{\sigma}\right) \\ &= \frac{\gamma\bar{\kappa}^T}{2\sqrt{n}}\exp\left(\frac{\sigma^2}{2}\right)\Phi\left(-\frac{1}{2}\sigma\right); \end{aligned}$$

Thus,

$$\bar{\kappa}^T = \frac{2we^R\exp\left[-\sigma^2\right]}{\gamma^2\Phi\left[-\frac{1}{2}\sigma\right]}$$

Likewise, for $e^R > e^X$, use (10b) to get

$$\bar{\kappa}^T = \frac{1}{\gamma}\frac{p - 2w\frac{e^R}{\gamma}\exp\left[-\sigma^2\right]}{1 - \Phi\left[-\frac{1}{2}\sigma\right]}$$

To get r for $\kappa^T < \bar{\kappa}^T$ use (7) to get

$$w = \left(p - \gamma r\right)\frac{1}{2\sqrt{n}}\exp\left[\frac{1}{2}\sigma^2\right] + \frac{1}{2\sqrt{n}}\gamma\kappa^s\exp\left[\frac{1}{2}\sigma^2\right]\Phi\left[-\frac{1}{2}\sigma\right] - \frac{1}{2\sqrt{n}}\gamma\kappa^b\exp\left[\frac{1}{2}\sigma^2\right]\left(1 - \Phi\left[-\frac{1}{2}\sigma\right]\right),$$

which using $e^R = \gamma\sqrt{n^*}\exp\left(\frac{1}{2}\sigma^2\right)$ simplifies to

$$\begin{aligned} r &= \frac{p - 2w\frac{e^R}{\gamma}\exp\left[-\sigma^2\right]}{\gamma} + \kappa^s\Phi\left[-\frac{1}{2}\sigma\right] - \kappa^b\Phi\left[\frac{1}{2}\sigma\right] \\ &= .\rho^* - \kappa^b\Phi\left[\frac{1}{2}\sigma\right] + \kappa^s\Phi\left[-\frac{1}{2}\sigma\right] \end{aligned} \tag{31}$$

Using $G\left[\varepsilon^R|n^*\right] = \Phi\left[\frac{1}{2}\sigma\right]$ in (8) gets

$$\rho = r + .\kappa^b\Phi\left[-\frac{1}{2}\sigma\right] - \kappa^s\Phi\left[\frac{1}{2}\sigma\right],$$

which combined with the previous equation gives

$$\rho = \rho^* - \kappa^T\left(\Phi\left[\frac{1}{2}\sigma\right] - \Phi\left[-\frac{1}{2}\sigma\right]\right). \tag{32}$$

Notice that under uniformly distributed additive random shocks, $\rho = \rho^*$, but here $\rho < \rho^*$; $\rho^* - \rho$ is increasing in both κ^T and σ.

It is worth noting that all of the expressions derived up to now converge to the expressions for additive uniformly distributed shock when $\sigma \to 0$. All the differences otherwise are essentially due to the fact that, unlike a uniform distribution, a log-normal distribution is not symmetric around the mean.

Once again, we identify the set of e^R and $\kappa^T \in \left[0, \frac{p}{\gamma}\right]$ that jointly characterize the three possible equilibria as stated by (9).

7.8.1 Case I: $e^R < e^X = \gamma \frac{p}{2w} \exp\left(\sigma^2\right) \Phi\left[-\frac{1}{2}\sigma\right]$;

Let $\kappa^T < \bar{\kappa}^T = \frac{2w\exp\left(-\sigma^2\right)\frac{e^R}{\gamma}}{\gamma\Phi\left[-\frac{1}{2}\sigma\right]}$ and $\kappa^s = 0$, i.e., $\kappa^b < \bar{\kappa}^T$. Then from (31) $r + \kappa^b < \frac{p}{\gamma}$. If $\kappa^b = 0$ and $\kappa^s < \bar{\kappa}^T$, $r > \kappa^s$ because $e^R < \gamma \frac{p}{2w} \exp\left(\sigma^2\right) \Phi\left[-\frac{1}{2}\sigma\right]$.

Now if $\kappa^T > \bar{\kappa}^T = \frac{2w\exp\left(-\sigma^2\right)\frac{e^R}{\gamma}}{\gamma\Phi\left[-\frac{1}{2}\sigma\right]}$, proposition 2 equation (10a) yields

$$w = \frac{\gamma\kappa^T}{2\sqrt{n}} \exp\left[\frac{1}{2}\sigma^2\right] \Phi\left[\frac{\log\left[\frac{e^R/\gamma}{\sqrt{n}}\right] - \sigma^2}{\sigma}\right] \tag{33}$$

Clearly, n is increasing in κ^T, and therefore for $\kappa^T \in \left(\bar{\kappa}^T, \frac{p}{\gamma}\right)$, $n \in (n^*, \hat{n})$, and $y^f \in \left(\frac{e^R}{\gamma}, \hat{y}^f\right)$. The rationale behind $n < \hat{n}$ is similar to the one offered in the example with uniform additive shocks.

7.8.2 Case II: $e^R \in (e^X, e^R]$

Let $\kappa^T < \bar{\kappa}^T = \frac{p - 2w\exp\left(-\sigma^2\right)\frac{e^R}{\gamma}}{\gamma\Phi\left[\frac{1}{2}\sigma\right]}$ and $\kappa^s = 0$, i.e., $\kappa^b < \frac{p - 2w\exp\left(-\sigma^2\right)\frac{e^R}{\gamma}}{\gamma\Phi\left[\frac{1}{2}\sigma\right]}$. Then, from (31) $r + \kappa^b < \frac{p}{\gamma}$ because

$$\begin{aligned} r + \kappa^b &= \frac{p - 2w\frac{e^R}{\gamma}\exp\left[-\sigma^2\right]}{\gamma} + \kappa^b\Phi\left[-\frac{1}{2}\sigma\right] \\ &< \frac{p - 2w\frac{e^R}{\gamma}\exp\left[-\sigma^2\right]}{\gamma} + \frac{p - 2w\exp\left(-\sigma^2\right)\frac{e^R}{\gamma}}{\gamma\Phi\left[\frac{1}{2}\sigma\right]}\Phi\left[-\frac{1}{2}\sigma\right] \\ &= p + \underbrace{\frac{p\Phi\left[-\frac{1}{2}\sigma\right] - 2w\exp\left(-\sigma^2\right)\frac{e^R}{\gamma}}{\gamma\Phi\left[\frac{1}{2}\sigma\right]}}_{<0 \text{ for } e^R > e^X} \end{aligned}$$

If $\kappa^b = 0$ and $\kappa^s < \frac{p - 2w\exp\left(-\sigma^2\right)\frac{e^R}{\gamma}}{\gamma\Phi\left[\frac{1}{2}\sigma\right]}$, $r > \kappa^s$ because

$$\begin{aligned} r &= \frac{p - 2w\exp\left(-\sigma^2\right)\frac{e^R}{\gamma}}{\gamma\Phi\left[\frac{1}{2}\sigma\right]}\Phi\left[\frac{1}{2}\sigma\right] + \kappa^s\Phi\left[-\frac{1}{2}\sigma\right] \\ &> \kappa^s\Phi\left[\frac{1}{2}\sigma\right] + \kappa^s\Phi\left[-\frac{1}{2}\sigma\right] = \kappa^s \end{aligned}$$

The conditions for permit market clearing are thus satisfied.

However, if $\kappa^T > \frac{p - 2w\exp\left(-\sigma^2\right)\frac{e^R}{\gamma}}{\gamma\Phi\left[\frac{1}{2}\sigma\right]}$, proposition 2 equation (10b) yields

$$w = \frac{p}{2\sqrt{n}} \exp\left[\frac{1}{2}\sigma^2\right]\left[1 - \frac{\gamma\kappa^T}{p}\left(1 - \Phi\left[\frac{\log\left[\frac{e^R/\gamma}{\sqrt{n}}\right] - \sigma^2}{\sigma}\right]\right)\right]$$

Clearly, n is decreasing in κ^T as can be verified from the above equation. Thus, for $\kappa^T \in \left(\bar{\kappa}^T, \frac{p}{\gamma}\right)$, $n \in (\hat{n}, n^*)$, and therefore $y^f \in \left(\hat{y}^f, e^R/\gamma\right)$. There is no output capacity lost because emission permits are in excess supply. Again, the rationale that $n < \hat{n}$ is the same as offered in the previous example.

Case III: $e^R = e^X = \gamma \frac{p}{2w} \exp\left(\sigma^2\right) \Phi\left[-\frac{1}{2}\sigma\right]$; Let $\kappa^T < \frac{p}{\gamma}$ and $\kappa^s = 0$, i.e., $\kappa^b < \frac{p}{\gamma}$. Then from (31) $r + \kappa^b = \frac{p - 2w\frac{e^R}{\gamma}\exp\left[-\sigma^2\right]}{\gamma} + \kappa^b \Phi\left[-\frac{1}{2}\sigma\right] < \frac{p}{\gamma}$ and $r = \frac{p - 2w\frac{e^R}{\gamma}\exp\left[-\sigma^2\right]}{\gamma} - \kappa^b \Phi\left[\frac{1}{2}\sigma\right] > \frac{p - 2w\frac{e^R}{\gamma}\exp\left[-\sigma^2\right]}{\gamma} - \frac{p}{\gamma}\Phi\left[\frac{1}{2}\sigma\right] = 0 = \kappa^s$. If $\kappa^b = 0$ and $\kappa^s < \frac{p}{\gamma}$, $r + \kappa^b = \frac{p - 2w\frac{e^R}{\gamma}\exp\left[-\sigma^2\right]}{\gamma} + \kappa^s \Phi\left[-\frac{1}{2}\sigma\right] < \frac{p}{\gamma}$ and $r - \kappa^s = \frac{p - 2w\frac{e^R}{\gamma}\exp\left[-\sigma^2\right]}{\gamma} - \kappa^s \Phi\left[\frac{1}{2}\sigma\right] > 0$. Thus, for all $\kappa^T \leq \frac{p}{\gamma}$, $y^a = y^f = \hat{y}^f = y^{f*} = \frac{e^X}{\gamma}$.

7.9 Proof of Proposition 3

Suppose trade occurs. Then, it must be the case that

$$\begin{aligned} p - \gamma r &> \frac{\kappa^b}{y\left(\hat{n}, \varepsilon^{\max}\right) - \frac{e^R}{\gamma}} = p\frac{\kappa^b}{\bar{\kappa}^b}\text{; and} \\ r &> \frac{\kappa^s}{e^R - \gamma y\left(\hat{n}, \varepsilon^{\min}\right)} = \frac{p}{\gamma}\frac{\kappa^s}{\bar{\kappa}^s} \end{aligned}$$

combining them together gets

$$\frac{\kappa^b}{\bar{\kappa}^b} + \frac{\kappa^s}{\bar{\kappa}^s} < 1.$$

To show the result in the other direction, suppose the above holds and without loss of generality let $\frac{\kappa^b}{\bar{\kappa}^b} + \frac{\kappa^s}{\bar{\kappa}^s} = \alpha < 1$. Fix $\kappa^b \leq \alpha\, \bar{\kappa}^b$, and define

$$x \equiv p - \frac{\kappa^b}{y\left(\hat{n}, \varepsilon^{\max}\right) - \frac{e^R}{\gamma}} = p\left(1 - \frac{\kappa^b}{\bar{\kappa}^b}\right) > 0.$$

Let $\gamma r = x - \epsilon > 0$ for ϵ infinitesimally small. Some buyers find it rational to enter the market provided there are some sellers willing to sell at r. If $\kappa^b = \alpha\, \bar{\kappa}^b$, then $\kappa^s = 0$ and all sellers are active. If $\kappa^b < \alpha\, \bar{\kappa}^b$, then $\kappa^s = \left(\alpha - \frac{\kappa^b}{\bar{\kappa}^b}\right)\kappa^s = \left(\alpha\, p - \frac{\kappa^b}{y(\hat{n},\varepsilon^{\max}) - \frac{e^R}{\gamma}}\right)\left(\frac{e^R}{\gamma} - y\left(\hat{n}, \varepsilon^{\min}\right)\right) < \left(p - \frac{\kappa^b}{y(\hat{n},\varepsilon^{\max}) - \frac{e^R}{\gamma}}\right)\left(\frac{e^R}{\gamma} - y\left(\hat{n}, \varepsilon^{\min}\right)\right)$ and there are sellers who find it rational to sell at r. Trade occurs. The argument is symmetric for $\kappa^s \leq \alpha\, \bar{\kappa}^s$ and it holds for all $\alpha < 1$.

The proof that $r \in \left(0, \frac{p}{\gamma}\right)$ when trade occurs in equilibrium follows from the market clearing, (20). This condition always holds with equality. The proof is obtained by contradiction. Suppose the LHS (supply) exceeds the RHS (demand). Then, at the margin $r \to 0$ and the LHS shrinks to zero. Likewise, if the RHS exceeds the LHS, at the margin $r \to p$, and the RHS shrinks to zero.

7.10 Derivations for the example with uniformly distributed shocks and fixed costs

Substituting $\sqrt{\hat{n}} = \frac{p}{2w}\frac{\frac{e^R}{\gamma}+c}{\frac{p}{2w}+2c} = \frac{\left(\frac{\tilde{z}^{\max}}{\gamma}-c\right)\left(\frac{e^R}{\gamma}+c\right)}{\frac{\tilde{z}^{\max}}{\gamma}+c}$ (Section 7.6) in $\bar{\kappa}^b\left(e^R\right)$ and $\bar{\kappa}^b\left(e^R\right)$ as defined in Proposition 2

$$\begin{aligned} \bar{\kappa}^b\left(e^R\right) &\equiv p\left(y\left(\hat{n}\left(e^R\right), \varepsilon^{\max}\right) - \frac{e^R}{\gamma}\right) = p\left(\frac{\left(\frac{\tilde{z}^{\max}}{\gamma} - c\right)\left(\frac{e^R}{\gamma} + c\right)}{\frac{\tilde{z}^{\max}}{\gamma} + c} - \left(\frac{e^R}{\gamma} - c\right)\right) \\ \bar{\kappa}^s\left(e^R\right) &\equiv \frac{p}{\gamma}\left(e^R - \gamma y\left(\hat{n}\left(e^R\right), \varepsilon^{\min}\right)\right) = p\left(\frac{e^R}{\gamma} + c - \frac{\left(\frac{\tilde{z}^{\max}}{\gamma} - c\right)\left(\frac{e^R}{\gamma} + c\right)}{\frac{\tilde{z}^{\max}}{\gamma} + c}\right) \end{aligned}$$

obtains the expressions in the main text.

When (21) in Proposition 2 holds, equations (17a) and (17b) imply

$$\underline{\varepsilon} = \frac{e^R}{\gamma} - \frac{\kappa^s}{\gamma r} - \sqrt{n}; \bar{\varepsilon} = \frac{e^R}{\gamma} + \frac{\kappa^b}{p - \gamma r} - \sqrt{n}.$$

Substituting these in (19) yields

$$\rho = \frac{p}{2c\gamma} \frac{\kappa^b}{p - \gamma r} + \frac{r}{2c} \left[2c - \frac{\kappa^s}{\gamma r} - \frac{\kappa^b}{p - \gamma r} \right]$$

The pre-production permit price equals its expected marginal market revenue at the production stage (on the RHS). The probability that a firm with some excess capacity does not enter the permit market as a buyer is $\frac{1}{2c}\frac{\kappa^b}{p-\gamma r}$; here, a marginal permit is worth $\frac{p}{\gamma}$. The first term captures the expected benefit. The expected benefit of a marginal permit for an inactive firm with excess permits is zero.[26] Finally, the probability that a firm is active in the permit market either as a buyer or a seller equals $\frac{1}{2c}\left(2c - \frac{\kappa^s}{\gamma r} - \frac{\kappa^b}{p-\gamma r}\right)$. The above equation simplifies to

$$\rho = r + \frac{\kappa^b - \kappa^s}{2\gamma c}. \tag{35}$$

When $\kappa^b \neq \kappa^s$, the pre-production permit price is higher than its post-production value if the buyers' fixed costs are higher than that of the sellers. Intuitively, selling is easier than buying in the post-production permit market, thus raising its ex ante value. Instead, if selling is harder, $\rho < r$.

Equations (18) and (20) yield, respectively

$$y^f = \sqrt{n} = \frac{p - \gamma\rho}{2w} = \frac{p - \gamma r}{2w} - \frac{\kappa^b - \kappa^s}{2c}; \tag{36}$$

$$y^f = \sqrt{n} = \frac{e^R}{\gamma} + \frac{1}{4c}\left[\left(\frac{\kappa^b}{p - \gamma r} \right)^2 - \left(\frac{\kappa^s}{\gamma r} \right)^2 \right]. \tag{37}$$

where we have utilized (19) in deriving (36). Thus the above two equations along with (35) solve for ρ, r, and n. When $\kappa^b = \kappa^s = \kappa$, the above two reduce to (23) in the main text.

Clearly, when $e^R = e^X = \frac{\gamma p}{4w}$, $r = \frac{p}{2\gamma}$ and $y^f = \sqrt{n} = \frac{p}{4w} = \frac{e^X}{\gamma}$. That $\frac{\partial r}{\partial e^R} < 0$ directly follows from (37) after substituting for n from (36). then (36) implies that $\frac{\partial y^f}{\partial e^R} > 0$. Once $\frac{\partial r}{\partial e^R} < 0$ is shown (37) implies

$$y^f \gtreqqless \frac{e^R}{\gamma} \Leftrightarrow e^R \lesseqgtr e^X = \frac{\gamma\, p}{4w}$$

as stated in the main text.

Also, note that with $\kappa^b = \kappa^s = \kappa$,

$$\underline{\varepsilon} = \frac{e^R}{\gamma} - \frac{\kappa}{\gamma r} - \sqrt{n}; \bar{\varepsilon} = \frac{e^R}{\gamma} + \frac{\kappa}{p - \gamma r} - \sqrt{n};$$

$$\varepsilon^R = \frac{e^R}{\gamma} - \sqrt{n} = -\frac{\kappa^2}{4c}\left[\left(\frac{1}{p - \gamma r} \right)^2 - \left(\frac{1}{\gamma r} \right)^2 \right].$$

Further,

$$\bar{\varepsilon} \to \varepsilon^{\max} \Longleftrightarrow \underline{\varepsilon} \to \varepsilon^{\min}.$$

[26] With a probability $\frac{1}{2c}\frac{\kappa_s}{\gamma r'}$.

To show this, suppose $\bar{\varepsilon} \to \varepsilon^{\max}$, i.e.,

$$\frac{e^R}{\gamma} + \frac{\kappa}{p - \gamma r} - > \sqrt{n} + c.$$

Then, using (37), we get

$$\frac{e^R}{\gamma} - \frac{\kappa}{\gamma r} = \sqrt{n} - c$$

which proves the $\Rightarrow$ part. To show the reverse, begin with $\frac{\kappa}{\gamma r} = \frac{e^R}{\gamma} - \sqrt{n} + c$ to obtain $\frac{\kappa}{p - \gamma r} = \sqrt{n} + c - \frac{e^R}{\gamma}$. Hence, the mass of sellers and buyers in the permit market vanishes simultaneously in the limit. Clearly, when κ satisfies the above equalities, $\frac{p-\gamma r}{\gamma r} = \frac{\frac{e^R}{\gamma} - \sqrt{\hat{n}} + c}{\sqrt{\hat{n}} + c - \frac{e^R}{\gamma}}$, and

$$\gamma r = p \frac{\sqrt{\hat{n}} + c - \frac{e^R}{\gamma}}{2c} = p \left[1 - \frac{\frac{e^R}{\gamma} - \sqrt{\hat{n}} + c}{2c} \right].$$

Also, then (36) implies[27]

$$w = p \frac{\frac{e^R}{\gamma} - \sqrt{\hat{n}} + c}{4c} \frac{1}{\sqrt{\hat{n}}}; \sqrt{\hat{n}} = \frac{p}{4w} \frac{\frac{e^R}{\gamma} + c}{\frac{p}{4w} + c}.$$

To find $\bar{\kappa}\left(e^R\right)$:

$$\begin{aligned} \frac{\bar{\kappa}}{p - \gamma r} &= \frac{\bar{\kappa}}{p \frac{\frac{e^R}{\gamma} - \sqrt{\hat{n}} + c}{2c}} = \sqrt{\hat{n}} + c - \frac{e^R}{\gamma} \Rightarrow \\ \kappa &= \frac{p}{2c} \left[c^2 - \left(\frac{e^R}{\gamma} - \sqrt{\hat{n}} \right)^2 \right] = \frac{pc}{2} \left[1 - \left(\frac{\frac{e^R}{\gamma} - \frac{p}{4w}}{\frac{p}{4w} + c} \right)^2 \right]. \end{aligned}$$

To find r, when $\kappa = \bar{\kappa}$,

$$\begin{aligned} \frac{\bar{\kappa}}{p - \gamma r} + \frac{\bar{\kappa}}{\gamma r} &= 2c; \\ (\gamma r)^2 - p(\gamma r) + \frac{\bar{\kappa} p}{2c} &= 0 \Rightarrow r = \frac{p}{2\gamma} \frac{\frac{p}{2w} + c - \frac{e^R}{\gamma}}{\frac{p}{4w} + c}. \end{aligned}$$

When $e^R = e^X = \frac{p}{4w}$, $\bar{\kappa}\left(e^X\right) = \frac{pc}{2}$, When $e^R = \tilde{z}^{\max} = \gamma\left(\frac{p}{2w} + c\right)$, then $\bar{\kappa} = 0$; $\bar{\kappa}\left(e^R\right)$ has a minimum when $\left(\frac{\frac{e^R}{\gamma} - \frac{p}{4w}}{\frac{p}{4w} + c}\right)^2$ is maximized. It equals 1 when $e^R = e^X$, and it equals $\frac{pc}{2}\left[1 - \left(\frac{\frac{4wc^2}{p} - \frac{p}{4w}}{\frac{p}{4w} + c}\right)^2\right]$ when $\frac{e^R}{\gamma} = \frac{4wc^2}{p}$. Recall that $y^{\min} > 0$ requires $\frac{e^R}{\gamma} > \frac{4wc^2}{p}$. For all values of $\kappa \in (0, \bar{\kappa}\left(e^R\right)$ a permit market with trading costs exists, and some capacity is unutilized in equilibrium. Alternatively, $\bar{\kappa}$ can also be derived as $\bar{\kappa} \equiv \frac{\kappa^b \kappa^s}{\kappa^b + \kappa^s}$, as stated in the main text.

Capacity and permit utilization can be calculated by using (22a) and (22b) as expressed in the main text. Whether $\frac{e^a}{e^R} \lesseqgtr \frac{y}{y^f}$ requires checking whether

$$\frac{\gamma}{e^R} \sqrt{n} \gtreqless \left(\frac{\gamma r}{p - \gamma r} \right)^2$$

Note that as $e^R \to 0$, the RHS $\to \infty$ and when $e^R \to \tilde{z}^{\max}$ the RHS $\to 0$. We already know that LHS = RHS when $e^R = e^X$.

[27] For $y^{\min *} > 0$, i.e., $\sqrt{n^*} > c$, we assume $c < \frac{1}{2}\sqrt{\frac{pe^R}{\gamma w}}$. Since $\frac{e^R}{\gamma} < \sqrt{\hat{n}} + c$, this also implies $c < \frac{p}{2w}$.

7.11 $y^f \neq \frac{e^X}{\gamma}$ for $e^R = e^X$ when $\kappa^b \neq \kappa^s$ in the uniform case

The three equilibrium equations are

$$\begin{aligned}
\rho &= r + \frac{\kappa^b - \kappa^s}{2c}.; \\
y^f &= \sqrt{n} = \frac{p - \gamma\rho}{2w}; \\
y^f &= \sqrt{n} = \frac{e^R}{\gamma} + \frac{1}{4c}\left[\left(\frac{\kappa^b}{p-\gamma r}\right)^2 - \left(\frac{\kappa^s}{\gamma r}\right)^2\right].
\end{aligned}$$

Let $\kappa^b = 2\kappa^s$. Then

$$\begin{aligned}
\rho &= r + \frac{\kappa^s}{2c};\ p - \gamma\rho = p - \gamma r - \frac{\kappa^s}{2c};\ p - \gamma r = 2w\sqrt{n} + \frac{\kappa^s}{2c}, \\
\gamma r &= p - 2w\sqrt{n} - \frac{\kappa^s}{2c}, \\
\sqrt{n} &= \frac{e^R}{\gamma} + \frac{1}{4c}\left[4\left(\frac{\kappa^s}{2w\sqrt{n} + \frac{\kappa^s}{2c}}\right)^2 - \left(\frac{\kappa^s}{p - 2w\sqrt{n} - \frac{\kappa^s}{2c}}\right)^2\right].
\end{aligned}$$

When $\sqrt{n} = \frac{p}{4w}$

$$\frac{p}{4w} = \frac{e^R}{\gamma} + \frac{1}{4c}\left[4\left(\frac{\kappa^s}{\frac{p}{2} + \frac{\kappa^s}{2c}}\right)^2 - \left(\frac{\kappa^s}{\frac{p}{2} - \frac{\kappa^s}{2c}}\right)^2\right].$$

Then $\frac{e^R}{\gamma} = \frac{p}{4w}$ if and only if

$$\begin{aligned}
\frac{p}{2} + \frac{\kappa^s}{2c} &= 2\left(\frac{p}{2} - \frac{\kappa^s}{2c}\right) \\
3\frac{\kappa^s}{2c} &= \frac{p}{2}; \\
\kappa^s &= \frac{pc}{3}; \kappa^b = \frac{2pc}{3}
\end{aligned}$$

But when $\frac{e^R}{\gamma} = \frac{p}{4w}$, $\kappa^b = \kappa^s = pc$ and then

$$\frac{\kappa^s}{\kappa^s} + \frac{\kappa^b}{\kappa^b} = 1$$

7.12 $y^f \neq \frac{e^X}{\gamma}$ for $e^R = e^X$ even when $\kappa^b = \kappa^s$ in the log-normal case

Suppose $e^R = e^X$ and $y^f = \sqrt{n}\exp\left[\frac{1}{2}\sigma^2\right] = e^X$. For the actively trading firms in the permit market, this implies

$$e^X\left(1 - G[\bar{\varepsilon}] + G[\underline{\varepsilon}]\right) = \sqrt{n}\left(\int_0^{\underline{\varepsilon}} \varepsilon\ g(\varepsilon)\ dh + \int_{\underline{\varepsilon}}^{\infty} \varepsilon\ g(\varepsilon) dh\right),$$

which in turn requires

$$1 - \Phi\left[\frac{\ln\bar{\varepsilon} - \sigma^2}{\sigma}\right] + \Phi\left[\frac{\ln\underline{\varepsilon} - \sigma^2}{\sigma}\right] = 1 - \Phi\left[\frac{\ln\bar{\varepsilon}}{\sigma}\right] + \Phi\left[\frac{\ln\underline{\varepsilon}}{\sigma}\right],$$

which can not hold for any $\sigma > 0$ and $\kappa > 0$.

8 Extensions

Below we discuss three variations of our model. The first questions the validity of our results once the assumption of linear relationship between emissions and intended output is dropped. Specifically, we let the emissions be convex in (or concave) in output and revisit our key results.[28] Next, we explore how the relationship between trade-impeding cost thresholds are modified when the assumption of irreversible inputs is dropped. Finally, we argue why our model in Section 2 most parsimoniously fits the framework proposed by Murty et. al. (2012), by contrasting it with a plausible alternative.

8.1 Non-linear emissions

There is no a priori reason to think that emissions vary linearly in the output as we have proposed in Section 2. To allow for emissions as a by-product that is convex or concave in the goods output, let

$$z(y) = \gamma y + \frac{1}{2}\xi y^2.$$

In case of concave z, we restrict ξ sufficiently enough to ensure that $z'(.) > 0$ in the relevant domain for y. With emissions defined as above, the output produced by firms as expressed in (38) is modified to

$$y^i \equiv \min\left\{y\left(n, \varepsilon^i\right), y\left(e^i\right)\right\}, \tag{38}$$

where

$$y\left(e^i\right) = \begin{cases} \sqrt{\frac{2e}{\xi} + \left(\frac{\gamma}{\xi}\right)^2} - \frac{\gamma}{\xi}, \text{ if } \xi > 0 \\ -\sqrt{\frac{2e}{\xi} + \left(\frac{\gamma}{\xi}\right)^2} - \frac{\gamma}{\xi}, \text{ if } \xi < 0 \end{cases}$$

Following the outline in the main text, we first discuss the equilibria under proportional costs and then under fixed costs.

Observe first that under costless trade

$$w = \int_{\varepsilon^{\min}}^{\varepsilon^{\max}} \left(p - \rho \underbrace{\left(\gamma + \xi y\left(n^*, \varepsilon\right)\right)}_{\equiv e_y^*}\right) y_n\left(n^*, \varepsilon\right) \; g(\varepsilon) d\varepsilon, \tag{39}$$

and $e^R = E\left[e\left(y\left(n^*, \varepsilon\right)\right)\right]$.

In the absence of permit trade $w = p\int_{\varepsilon^{\min}}^{\varepsilon^R} y_n\left(\hat{n}, \varepsilon\right) \; g(\varepsilon)d\varepsilon$ with ε^R determined by $e^R = e\left(y\left(n, \varepsilon^R\right)\right)$. It can be checked that Lemma 1 continues to apply with $e^X \equiv E\left[e\left(y\left(\hat{n}\left(e^X\right), \varepsilon\right)\right)\right]$, and can be restated as

$$e^R \lesseqqgtr e^X \iff E\left[e\left(y\left(\hat{n}\left(e^R\right), \varepsilon\right)\right)\right] \gtreqqless e^R. \tag{40}$$

That is, there is a *latent* excess demand for permits for all $e^R < e^X$, and excess supply for all $e^R > e^X$.

8.1.1 Proportional costs

For reasons that become clear below, the two cases of convex and concave emissions are considered sequentially.

[28] Is there a clear evidence on the one versus another?

Convex emissions: $\xi > 0$ Notice that a firm producing y at the margin now requires $\gamma + \xi y$ units of permits. Thus, permit costs are increasing in the output scale. The firms whose full capacity emissions fall below their permits continue to sell at the market price r as long as $r \geq \kappa^s$. As for permit buyers, however, the permit opportunity cost is increasing in the output scale. All firms (whose full capacity emissions exceed their permit holdings) are in the market if and only if $\left(r + \kappa^b\right)\left(\gamma + \xi\ y\left(n, \varepsilon^{\max}\right) y_n\left(n, \varepsilon^{\max}\right)\right) < p$, demanding that the most productive firm derive a positive surplus from the marginal permit purchase.

Thus, for all $\kappa^s < r < \frac{p}{\gamma + \xi\ y(n^*, \varepsilon^{\max})} - \kappa^b$, the input choice continues to be $n^*\left(e^R\right) : e^R = E\left\{e\left(y\left(n^*, \varepsilon\right)\right)\right\}$, while r and ρ^* are now related by (see 11):

$$\rho^* = \left(r - \kappa^s\right) \frac{E\left[e_y^*\ y_n\left(n^*|s\right)\right]}{E\left[e_y^*\ y_n\left(n^*\right)\right]} + \left(r + \kappa^b\right) \frac{E\left[e_y^*\ y_n\left(n^*|b\right)\right]}{E\left[e_y^*\ y_n\left(n^*\right)\right]}.$$

The equilibrium allocations and utilizations continues to replicate costless trade as long as total cost $\kappa^s + \kappa^b < \bar{\kappa}^T\left(e^R\right)$, which is determined below.

The equilibrium broadly follows the one described in Section 3 and Proposition 2. In case of an excess demand, when $r > \kappa^s$, but $r + \kappa^b = \frac{p}{\gamma + \xi\ y(n, \varepsilon^c)}$, for $\varepsilon^c \in \left[\varepsilon^R, \varepsilon^{\max}\right]$, the input employment is determined from

$$\begin{aligned} w = p \int_{\varepsilon^{\min}}^{\varepsilon^c} y_n\left(n, \varepsilon\right) g(\varepsilon) d\varepsilon - \left(\frac{p}{\gamma + \xi\ y\left(n, \varepsilon^c\right)} - \kappa^T\right) \int_{\varepsilon^{\min}}^{\varepsilon^R} \left(\gamma + \xi\ y\left(n, \varepsilon\right)\right)\ y_n\left(n, \varepsilon\right)\ g(\varepsilon) d\varepsilon \\ - \frac{p}{\gamma + \xi\ y\left(n, \varepsilon^c\right)} \int_{\varepsilon^R}^{\varepsilon^c} \left(\gamma + \xi\ y\left(n, \varepsilon\right)\right)\ y_n\left(n, \varepsilon\right)\ g(\varepsilon) d\varepsilon \end{aligned} \tag{41}$$

Since all permits are sold, the permit market equilibrium requires

$$\int_{\varepsilon^{\min}}^{\varepsilon^c} e\left(\sqrt{n} + \varepsilon\right) g\left(\varepsilon\right) d\varepsilon + e\left(\sqrt{n} + \varepsilon^c\right)\left(1 - G\left[\varepsilon^c\right]\right) = e^R \tag{42}$$

The equilibrium input employment is then determined jointly from (41) and (42).

On the other hand, when there is an excess supply, $r = \kappa^s$, and the input choice is determined from

$$w = p \int_{\varepsilon^{\min}}^{\varepsilon^c} y_n\left(n, \varepsilon\right)\ g(\varepsilon) d\varepsilon - \kappa^T \int_{\varepsilon^R}^{\varepsilon^c} \left(\gamma + \xi y\left(n, \varepsilon\right)\right) y_n\left(n, \varepsilon\right)\ g(\varepsilon) d\varepsilon, \tag{43a}$$

where $\varepsilon^c = \varepsilon^{\max}$ if $\kappa^T \leq \frac{p}{\gamma + \xi\ y(n, \varepsilon^{\max})}$. Otherwise, $\varepsilon^c < \varepsilon^{\max}$ is determined from $\kappa^T = \frac{p}{\gamma + \xi\ y(n, \varepsilon^c)}$.

Now, let n satisfy

$$w = p \int_{\varepsilon^{\min}}^{\varepsilon^{\max}} y_n\left(n, \varepsilon\right)\ g(\varepsilon) d\varepsilon - \frac{p}{\gamma + \xi\ y\left(n, \varepsilon^{\max}\right)} \int_{\varepsilon^R}^{\varepsilon^{\max}} \left(\gamma + \xi y\left(n, \varepsilon\right)\right) y_n\left(n, \varepsilon\right)\ g(\varepsilon) d\varepsilon$$

along with $e\left(y\left(n, \varepsilon^R\right)\right) = e^R = E\left[e\left(y\left(n, \varepsilon\right)\right)\right]$. The above satisfies (41) - (43a) with $\kappa^T = \frac{p}{\gamma + \xi\ y(n, \varepsilon^{\max})}$: all firms with excess permits sell, while all those with permit shortages buy. Define this level of permits as e^{X^*} and the corresponding input employment as $n^*\left(e^{X^*}\right)$, which equals that under costless trade.

Using using (41), define

$$\bar{\kappa}^T\left(e^R\right) \equiv \frac{w - p\ \int_{\varepsilon^{\min}}^{\varepsilon^{\max}} \left(1 - \frac{\gamma + \xi\ y(n^*, \varepsilon)}{\gamma + \xi\ y(n^*, \varepsilon^{\max})}\right)\ y_n\left(n^*, \varepsilon\right)\ g(\varepsilon) d\varepsilon}{\int_{\varepsilon^{\min}}^{\varepsilon^R} \left(\gamma + \xi\ y\left(n^*, \varepsilon\right)\right)\ y_n\left(n^*, \varepsilon\right)\ g(\varepsilon) d\varepsilon} \leq \frac{p}{\gamma + \xi\ y\left(n^*, \varepsilon^{\max}\right)}$$

The expression follows the same line as in the proof of Proposition 2. For $\xi = 0$, and with $y\left(n, \varepsilon\right) =$ either $f\left(n\right) + \varepsilon$ or $\varepsilon f\left(n\right)$ .we know that $\bar{\kappa}^T\left(e^R\right)$ is increasing in e^R. Also, for e^R and $n^*\left(e^R\right)$ sufficiently small, even with $\xi > 0$ the result continues to be similar to that under linear emissions, i.e., $\bar{\kappa}^T\left(e^R\right)$ is increasing in e^R. We conjecture (and verify with our two parametric examples) that $\bar{\kappa}^T\left(e^R\right)$ is increasing in all $e^R < e^{X^*}$, with its maximum at $\bar{\kappa}^T\left(e^{X^*}\right) = \frac{p}{\gamma+\xi\ y\left(n^*\left(e^{X^*}\right),\varepsilon^{\max}\right)}$.

Similarly, using (43a) define

$$\bar{\kappa}^T\left(e^R\right) \equiv \frac{p\int_{\varepsilon^{\min}}^{\varepsilon^{\max}} y_n\left(n^*,\varepsilon\right)\ g(\varepsilon)d\varepsilon - w}{\int_{\varepsilon^R}^{\varepsilon^{\max}}\left(\gamma+\xi y\left(n^*,\varepsilon\right)\right)y_n\left(n^*,\varepsilon\right)\ g(\varepsilon)d\varepsilon} \leq \frac{p}{\gamma+\xi\ y\left(n^*,\varepsilon^{\max}\right)},$$

Clearly, $\bar{\kappa}^T\left(e^R\right)$ is decreasing in n^*. It is at its maximum at e^{X^*} with $\bar{\kappa}^T\left(e^{X^*}\right) = \frac{p}{\gamma+\xi\ y\left(n^*\left(e^{X^*}\right),\varepsilon^{\max}\right)}$ and therefore the last inequality holds for all $e^R > e^{X^*}$.

The above discussion leads to the following:

Conclusion 4 *For all $\kappa^T \leq \bar{\kappa}^T\left(e^R\right)$, the input employment and equilibrium utilizations are identical to that under costless trade. $\bar{\kappa}^T\left(e^R\right)$ is increasing in e^R for all $e^R < e^{X^*}$ and decreasing for all $e^R > \varepsilon^{X^*}$. For all $\kappa^T > \bar{\kappa}^T\left(e^R\right)$, input choice is either determined from (41) and (42) or from (43a). In the former case, all permits are utilized while some capacity remains utilized; the utilizations are switched in the latter case.*

It can be checked that the excess supply and the excess demand equilibria captured by (41) and (43a) converge to no trade equilibria as $\kappa^T \to \frac{p}{\gamma+\xi\ y\left(\hat{n}\left(e^R\right),\varepsilon^R\right)}$.

Observe by invoking continuity that there would be a continuum of pairs of $\left\{e^R, \kappa^T\left(e^R\right)\right\}$ for which both (41) and (43a) would hold as they do for $\left\{e^{X^*}, \frac{p}{\gamma+\xi\ y\left(n^*\left(e^{X^*}\right),\varepsilon^{\max}\right)}\right\}$. When (41) holds with $\varepsilon^c < \varepsilon^{\max}$, all excess permits holders sell their permit surplus off to the buyers, but there are still firms that do not produce to their full capacity. Since both (41) and (43a) hold together with $\kappa^T = \frac{p}{\gamma+\xi\ y(n,\varepsilon^c)}$, the threshold that separates excess demand and excess supply equilibria are now jointly determined by (42) and

$$w = p\int_{\varepsilon^{\min}}^{\varepsilon^c} y_n\left(n,\varepsilon\right)\ g(\varepsilon)d\varepsilon - \frac{p}{\gamma+\xi\ y\left(n,\varepsilon^c\right)}\int_{\varepsilon^R}^{\varepsilon^c}\left(\gamma+\xi y\left(n,\varepsilon\right)\right)y_n\left(n,\varepsilon\right)\ g(\varepsilon)d\varepsilon, \tag{44a}$$

Given e^R, (42) and (44a) jointly determine n and ε^c, which in turn determines $\frac{p}{\gamma+\xi\ y(n,\varepsilon^c)}$. Denote these set of $\left\{e^R, \frac{p}{\gamma+\xi\ y\left(n\left(e^R\right),\varepsilon^c\right)}\right\}$ values as $\left\{e^R, \bar{\kappa}^{T^c}\left(e^R\right)\right\}$. Obviously, $\left\{e^{X^*}, \frac{p}{\gamma+\xi\ y\left(n^*\left(e^{X^*}\right),\varepsilon^{\max}\right)}\right\}$ is one such point on this separating threshold. The end point of this line in the limit as $\varepsilon^c \to \varepsilon^R$ lies at $\left\{e^R, \frac{p}{\gamma+\xi\ y\left(\hat{n}\left(e^R\right),\varepsilon^R\right)}\right\} : E\left[e\left(\hat{n},\varepsilon\right)|\varepsilon < \varepsilon^R\right] = e^R$. This happens in the limit when $e^R \to 0$ and $\hat{n} \to 0$. The above discussion leads to the following:

Conclusion 5 *For $e^R > e^{X^*}$ and for all $\kappa^T \in \left[\bar{\kappa}^T\left(e^R\right), \frac{p}{\gamma+\xi\ y\left(\hat{n}\left(e^R\right),\varepsilon^R\right)}\right]$, input employment is determined by (43a) with ε^c determined from $\frac{p}{\gamma+\xi\ y(n,\varepsilon^c)} = \kappa^T$. For all $e^R < e^{X^*}$ and for all $\kappa^T \in \left[\bar{\kappa}^T\left(e^R\right), \bar{\kappa}^{T^c}\left(e^R\right)\right]$, input employment is jointly determined from (41) and (42), whereas for all $\kappa^T \in \left[\bar{\kappa}^{T^c}\left(e^R\right), \frac{p}{\gamma+\xi\ y\left(\hat{n}\left(e^R\right),\varepsilon^R\right)}\right]$, input employment is determined by (43a) with ε^c determined from $\frac{p}{\gamma+\xi\ y(n,\varepsilon^c)} = \kappa^T$.*

Example 6 *Let $y = \sqrt{n} + \varepsilon$, with $\varepsilon \sim U\left[-c, c\right]$. Let $p = w = 1$ and $c = 0.05$. Let $\gamma = \xi = 1$. Figure 1* *below illustrates various types of equilibria detailed above:*

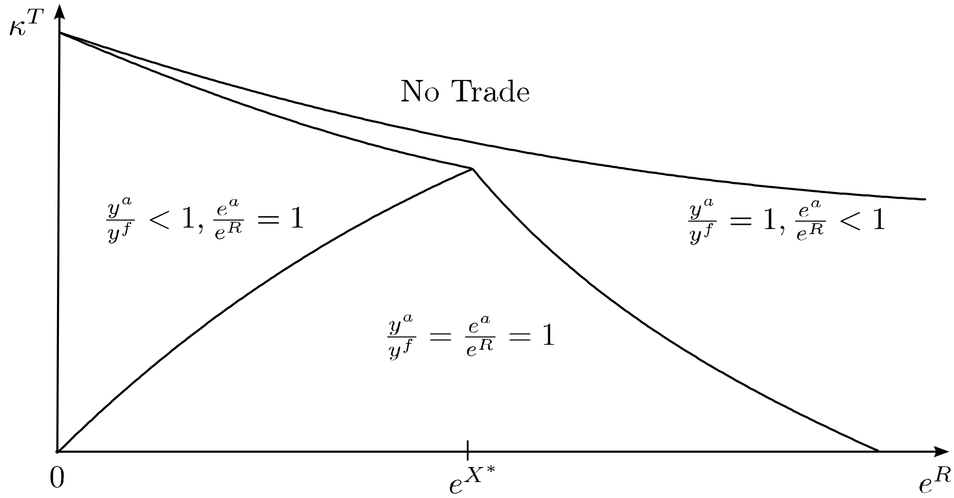


Figure 1

Concave emissions: $\xi < 0$ With firms still requiring $\gamma + \xi y$ units of permits at the margin, permit costs are decreasing in the output scale. The firms whose full capacity emissions fall below their permits continue to sell at the market price r as long as $r \geq \kappa^s$. Firms whose full capacity emissions exceed their permit holdings are in the market if and only if $\left(r + \kappa^b\right)\left(\gamma + \xi\ y\left(n, \varepsilon^R\right)\right) < p$, requiring that the least productive firm with $\varepsilon \geq \varepsilon^R$ derive a positive surplus from the marginal permit purchase. In case of an excess demand, when $r > \kappa^s$, but $r + \kappa^b = \frac{p}{\gamma+\xi\ y(n,\varepsilon^c)}$, for $\varepsilon^c \in \left[\varepsilon^R, \varepsilon^{\max}\right]$, the input employment is determined from

$$w = p\int_{\varepsilon^{\min}}^{\varepsilon^R} y_n(n,\varepsilon)\,g(\varepsilon)d\varepsilon - \left(\frac{p}{\gamma + \xi\ y(n,\varepsilon^c)} - \kappa^T\right)\int_{\varepsilon^{\min}}^{\varepsilon^R} \left(\gamma + \xi\ y(n,\varepsilon)\right)\ y_n(n,\varepsilon)\ g(\varepsilon)d\varepsilon$$
$$- \frac{p}{\gamma + \xi\ y(n,\varepsilon^c)}\int_{\varepsilon^c}^{\varepsilon^{\max}} \left(\gamma + \xi\ y(n,\varepsilon)\right)\ y_n(n,\varepsilon)\ g(\varepsilon)d\varepsilon + p\int_{\varepsilon^c}^{\varepsilon^{\max}} y_n(n,\varepsilon)\,g(\varepsilon)d\varepsilon \tag{45}$$

Since all permits are sold, the permit market equilibrium requires

$$\int_{\varepsilon^{\min}}^{\varepsilon^R} e\left(\sqrt{n}+\varepsilon\right)g(\varepsilon)\,d\varepsilon + e^R\left(G\left[\varepsilon^c\right] - G\left[\varepsilon^R\right]\right) + \int_{\varepsilon^c}^{\varepsilon^{\max}} e\left(\sqrt{n}+\varepsilon\right)g(\varepsilon)\,d\varepsilon = e^R \tag{46}$$

The equilibrium is jointly determined by (45) and the above permit market equilibrium

On the other hand, when there is an excess supply, $r = \kappa^s$, and the input choice is determined from

$$\begin{aligned} w &= p\left(\int_{\varepsilon^{\min}}^{\varepsilon^R} y_n(n,\varepsilon)\ g(\varepsilon)d\varepsilon + \int_{\varepsilon^c}^{\varepsilon^{\max}} y_n(n,\varepsilon)\ g(\varepsilon)d\varepsilon\right) \\ &\quad -\kappa^T\int_{\varepsilon^c}^{\varepsilon^{\max}} \left(\gamma + \xi y(n,\varepsilon)\right) y_n(n,\varepsilon)\ g(\varepsilon)d\varepsilon, \end{aligned} \tag{47a}$$

where $\varepsilon^c = \varepsilon^R$ if $\kappa^T \leq \frac{p}{\gamma+\xi\ y(n,\varepsilon^R)}$. Otherwise, $\varepsilon^c < \varepsilon^{\max}$ is determined from $\kappa^T = \frac{p}{\gamma+\xi\ y(n,\varepsilon^c)}$.

Now, let n satisfy

$$w = p\int_{\varepsilon^{\min}}^{\varepsilon^{\max}} y_n(n,\varepsilon)\ g(\varepsilon)d\varepsilon - \frac{p}{\gamma + \xi\ y(n,\varepsilon^R)}\int_{\varepsilon^R}^{\varepsilon^{\max}} \left(\gamma + \xi y(n,\varepsilon)\right) y_n(n,\varepsilon)\ g(\varepsilon)d\varepsilon$$

along with $e\left(y\left(n,\varepsilon^{R}\right)\right) = e^{R} = E\left[e\left(y\left(n,\varepsilon\right)\right)\right]$. The above satisfies both (45) and (47a) with $\kappa^{T} = \frac{p}{\gamma+\xi\ y(n,\varepsilon^{R})}$: all firms with excess permits sell, while all those with permit shortages buy. Define this level of permits as e^{X^*} and the corresponding input employment as $n^*\left(e^{X^*}\right)$, which equals that under costless trade.

Using using (45), define

$$\bar{\kappa}^{T}\left(e^{R}\right) \equiv \frac{w - p\ \int_{\varepsilon\notin[\varepsilon^{R},\varepsilon^{c}]}^{\varepsilon^{\max}} \left(1 - \frac{\gamma+\xi\ y(n^*,\varepsilon)}{\gamma+\xi\ y(n^*,\varepsilon^{R})}\right)\ y_n\left(n^*,\varepsilon\right)\ g(\varepsilon)d\varepsilon}{\int_{\varepsilon^{\min}}^{\varepsilon^{R}} \left(\gamma+\xi\ y\left(n^*,\varepsilon\right)\right)\ y_n\left(n^*,\varepsilon\right)\ g(\varepsilon)d\varepsilon} \leq \frac{p}{\gamma+\xi\ y\left(n^*,\varepsilon^{R}\right)}$$

Following the argument for convex emissions, $\bar{\kappa}^{T}\left(e^{R}\right)$ is increasing in e^{R} for ξ close to 0. Again, we conjecture that $\bar{\kappa}^{T}\left(e^{R}\right)$ is increasing in all $e^{R} < e^{X^*}$, with its maximum at $\bar{\kappa}^{T}\left(e^{X^*}\right) = \frac{p}{\gamma+\xi\ y\left(n^*\left(e^{X^*}\right),\varepsilon^{R}\right)}$.

Similarly, using (43a) define

$$\bar{\kappa}^{T}\left(e^{R}\right) \equiv \frac{p\int_{\varepsilon^{\min}}^{\varepsilon^{\max}} y_n\left(n^*,\varepsilon\right)\ g(\varepsilon)d\varepsilon - w}{\int_{\varepsilon^{R}}^{\varepsilon^{\max}} \left(\gamma+\xi y\left(n^*,\varepsilon\right)\right) y_n\left(n^*,\varepsilon\right)\ g(\varepsilon)d\varepsilon} \leq \frac{p}{\gamma+\xi\ y\left(n^*,\varepsilon^{R}\right)},$$

Clearly, $\bar{\kappa}^{T}\left(e^{R}\right)$ is decreasing in n^*. It is at its maximum at e^{X^*} with $\bar{\kappa}^{T}\left(e^{X^*}\right) = \frac{p}{\gamma+\xi\ y\left(n^*\left(e^{X^*}\right),\varepsilon^{R}\right)}$ and therefore the last inequality holds for all $e^{R} > e^{X^*}$.

The above discussion leads to the following:

Conclusion 7 *For all $\kappa^{T} \leq \bar{\kappa}^{T}\left(e^{R}\right)$, the input employment and equilibrium utilizations are identical to that under costless trade. $\bar{\kappa}^{T}\left(e^{R}\right)$ is increasing in e^{R} for all $e^{R} < e^{X^*}$ and decreasing for all $e^{R} > \varepsilon^{X^*}$. For all $\kappa^{T} > \bar{\kappa}^{T}\left(e^{R}\right)$, input choice is either determined from (45) and (46) or from (47a). In the former case, all permits are utilized while some capacity remains unutilized; the utilizations are switched in the latter case.*

It can be checked that the excess supply and the excess demand equilibria captured by (45) and (47a) converge to no trade equilibria as $\kappa^{T} \to \frac{p}{\gamma+\xi\ y(\hat{n}(e^{R}),\varepsilon^{\max})}$.

Again, there is a continuum of pairs of $\left\{e^{R},\kappa^{T}\left(e^{R}\right)\right\}$ for which both (45) and (47a) hold as they do for $\left\{e^{X^*},\frac{p}{\gamma+\xi\ y\left(n^*\left(e^{X^*}\right),\varepsilon^{R}\right)}\right\}$. When (45) holds with $\varepsilon^{c} > \varepsilon^{R}$, all excess permits holders sell their permit surplus off to the buyers, but there are still firms that do not produce to their full capacity. Since both (45) and (47a) hold together with $\kappa^{T} = \frac{p}{\gamma+\xi\ y(n,\varepsilon^{c})}$, the threshold that separates excess demand and excess supply equilibria are now jointly determined by (46) and

$$w = p\int_{\varepsilon\notin[\varepsilon^{R},\varepsilon^{c}]} y_n\left(n,\varepsilon\right)\ g(\varepsilon)d\varepsilon - \frac{p}{\gamma+\xi\ y\left(n,\varepsilon^{c}\right)}\int_{\varepsilon^{c}}^{\varepsilon^{\max}} \left(\gamma+\xi y\left(n,\varepsilon\right)\right) y_n\left(n,\varepsilon\right)\ g(\varepsilon)d\varepsilon, \tag{48a}$$

Given e^{R}, (46) and (48a) jointly determine n and ε^{c}, which in turn determines $\frac{p}{\gamma+\xi\ y(n,\varepsilon^{c})}$. Denote these set of $\left\{e^{R},\frac{p}{\gamma+\xi\ y(n(e^{R}),\varepsilon^{c})}\right\}$ values as $\left\{e^{R},\bar{\kappa}^{T^{c}}\left(e^{R}\right)\right\}$. Obviously, $\left\{e^{X^*},\frac{p}{\gamma+\xi\ y\left(n^*\left(e^{X^*}\right),\varepsilon^{R}\right)}\right\}$ is one such point on this separating threshold. The end point of this line in the limit as $\varepsilon^{c} \to \varepsilon^{\max}$ lies at $\left\{e^{R},\frac{p}{\gamma+\xi\ y(\hat{n}(e^{R}),\varepsilon^{\max})}\right\} : E\left[e\left(\hat{n},\varepsilon\right)|\varepsilon < \varepsilon^{R}\right] = e^{R}$. This happens in the limit when $e^{R} \to 0$ and $\hat{n} \to 0$. The above discussion leads to the following:

Conclusion 8 *For $e^{R} > e^{X^*}$ and for all $\kappa^{T} \in \left[\bar{\kappa}^{T}\left(e^{R}\right),\frac{p}{\gamma+\xi\ y(\hat{n}(e^{R}),\varepsilon^{\max})}\right]$, input employment is determined by (47a) with ε^{c} determined from $\frac{p}{\gamma+\xi\ y(n,\varepsilon^{c})} = \kappa^{T}$. For all $e^{R} < e^{X^*}$ and for all $\kappa^{T} \in$*

$\left[\bar{\kappa}^T\left(e^R\right), \bar{\kappa}^{T^c}\left(e^R\right)\right]$, *input employment is jointly determined from (41) and (42), whereas for all* $\kappa^T \in \left[\bar{\kappa}^{T^c}\left(e^R\right), \frac{p}{\gamma+\xi\ y(\hat{n}(e^R),\varepsilon^{\max})}\right]$, *input employment is determined by (43a) with* ε^c *determined from* $\frac{p}{\gamma+\xi\ y(n,\varepsilon^c)} = \kappa^T$.

Let $y = \sqrt{n} + \varepsilon$, with $\varepsilon \sim U\,[-c, c]$. Let $p = w = 1$ and $c = 0.05$. Let $\gamma = 1,\ \xi = -0.25$. Figure 2 below illustrates various types of equilibria detailed above:

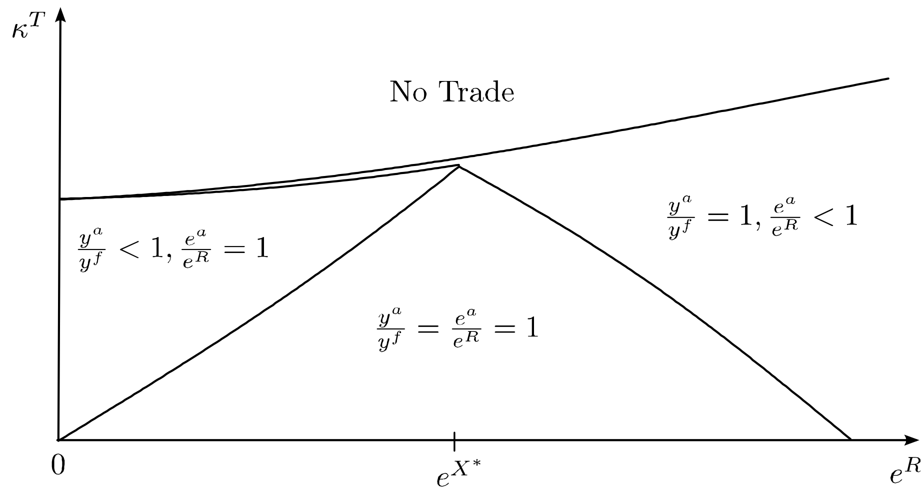


Figure 2

8.1.2 Fixed costs

The expected profit of a firm is now given by

$$E\left[\pi\left(n, e^0\right)\right] = \int_{\varepsilon^{\min}}^{\underline{\varepsilon}} \left\{py\left(n,\varepsilon\right) + r\left(e^0 - e\left(y\left(n,\varepsilon\right)\right)\right)\right\}\ g(\varepsilon)d\varepsilon + p\int_{\underline{\varepsilon}}^{\varepsilon^0}\ y\left(n,\varepsilon\right)g(\varepsilon)d\varepsilon$$
$$+py\left(n,\varepsilon^0\right)\int_{\varepsilon^0}^{\bar{\varepsilon}} g(\varepsilon)d\varepsilon + \left(py\left(n,\varepsilon^c\right) - r\left(e\left(y\left(n,\varepsilon^c\right)\right) - e^0\right)\right)\int_{\varepsilon^c}^{\varepsilon^{\max}} g(\varepsilon)d\varepsilon$$
$$+\int_{\bar{\varepsilon}}^{\varepsilon^c} \left\{py\left(n,\varepsilon\right) - r\left(e\left(y\left(n,\varepsilon\right)\right) - e^0\right)\right\}\ g(\varepsilon)d\varepsilon$$
$$-\kappa^s G\left(\underline{\varepsilon}\right) + \kappa^b\left(1 - G\left(\bar{\varepsilon}\right)\right) - \rho e^0 - wn,$$

where $e\left(y\left(n,\varepsilon^0\right)\right) \equiv e^0$ and $\underline{\varepsilon}$ and $\bar{\varepsilon}$ are obtained from

$$r\left(e^0 - e\left(\underbrace{y\left(n,\underline{\varepsilon}\right)}_{\equiv \underline{y}}\right)\right) = \kappa^s;$$
$$p\left(\underbrace{y\left(n,\bar{\varepsilon}\right)}_{\equiv \bar{y}} - y\left(n,\varepsilon^0\right)\right) - r\left(e\left(\underbrace{y\left(n,\bar{\varepsilon}\right)}_{\equiv \bar{y}}\right) - e^0\right) = \kappa^b.$$

The fourth term in the profit expression captures the fact that under convex emissions high productivity firms cap their emissions at $e\left(y\left(n,\varepsilon^c\right)\right)$, where ε^c is determined in equilibrium from

$$re_y\left(y\left(n,\varepsilon^c\right)\right) = p$$

if $\varepsilon^c < \varepsilon^{\max}$. Otherwise, $\varepsilon^c = \varepsilon^{\max}$ and the fourth term does not exist.

The first order condition for the input employment in equilibrium is given by

$$\begin{aligned} w &= \int_{\varepsilon^{\min}}^{\underline{\varepsilon}} \left(p - re'\left(y\left(n,\varepsilon\right)\right)\right) y_n \; g(\varepsilon) d\varepsilon + p \int_{\underline{\varepsilon}}^{\varepsilon^0} y_n\left(n,\varepsilon\right) g(\varepsilon) d\varepsilon \\ &\quad + p y_n\left(n,\varepsilon^0\right) \int_{\varepsilon^0}^{\bar{\varepsilon}} g(\varepsilon) d\varepsilon + \underbrace{\left(p - re'\left(y\left(n,\varepsilon^c\right)\right)\right)}_{=0 \text{ iff } \varepsilon^c < \varepsilon^{\max}} y_n\left(n,\varepsilon^c\right) \int_{\varepsilon^c}^{\varepsilon^{\max}} g(\varepsilon) d\varepsilon \\ &\quad + \int_{\bar{\varepsilon}}^{\varepsilon^c} \left(p - re'\left(y\left(n,\varepsilon\right)\right)\right) y_n \; g(\varepsilon) d\varepsilon \end{aligned}$$

with $e^0 = e^R$.

Similar to the condition described in Proposition 2 in the main text, below we derive a condition that determines cost thresholds that separate trade and no-trade equilibria for a given value of emission cap. Suppose

$$\begin{aligned} r\left(e^R - e\left(y\left(\hat{n},\varepsilon_{\min}\right)\right)\right) &> \kappa^s; \\ p\left(y\left(\hat{n},\varepsilon^{\max}\right) - y\left(\hat{n},\varepsilon^R\right)\right) - r\left(e\left(y\left(\hat{n},\varepsilon^{\max}\right)\right) - e^R\right) &> \kappa^b. \end{aligned}$$

The first and the second condition respectively drive the sellers and buyers out of the permit market. For the symmetric case, letting $\kappa^b = \kappa^s = \kappa$, combining the two above gets the required condition:

$$\kappa < \bar{\kappa}\left(e^R\right) \equiv p \frac{\left(e^R - e\left(y\left(\hat{n}\left(e^R\right),\varepsilon_{\min}\right)\right)\right)\left(y\left(\hat{n}\left(e^R\right),\varepsilon^{\max}\right) - y\left(\hat{n},\varepsilon^R\right)\right)}{e\left(y\left(\hat{n}\left(e^R\right),\varepsilon^{\max}\right)\right) - e\left(y\left(\hat{n}\left(e^R\right),\varepsilon_{\min}\right)\right)}$$

which for the uniform case as a counterpart to equation (24) in the main text reduces to

$$\bar{\kappa}\left(e^R\right) = \frac{p}{2} \frac{\left(e^R - \gamma\hat{n}\left(e^R\right) + \gamma c - \frac{1}{2}\xi\left(\hat{n}\left(e^R\right) - c\right)^2\right)}{\gamma + \xi\hat{n}\left(e^R\right)} \left(1 - \frac{\varepsilon^R}{c}\right) \tag{51}$$

with

$$\begin{aligned} \hat{n}\left(e^R\right) &= \frac{\frac{p}{4w}}{\frac{p}{4w} + c} \left(\sqrt{\left(\frac{\gamma}{\xi}\right)^2 + 2\frac{e^R}{\xi}} - \frac{\gamma}{\xi} + c\right) \\ \varepsilon^R &= \frac{c}{\frac{p}{4w} + c} \left(\sqrt{\left(\frac{\gamma}{\xi}\right)^2 + 2\frac{e^R}{\xi}} - \frac{\gamma}{\xi} - \frac{p}{4w}\right) \end{aligned}$$

For $p = w = 1$ and $c = 0.05$ and $\gamma = \xi = 1$, Figure 3 below verifies that the cost thresholds are very similar to that with linear emissions.

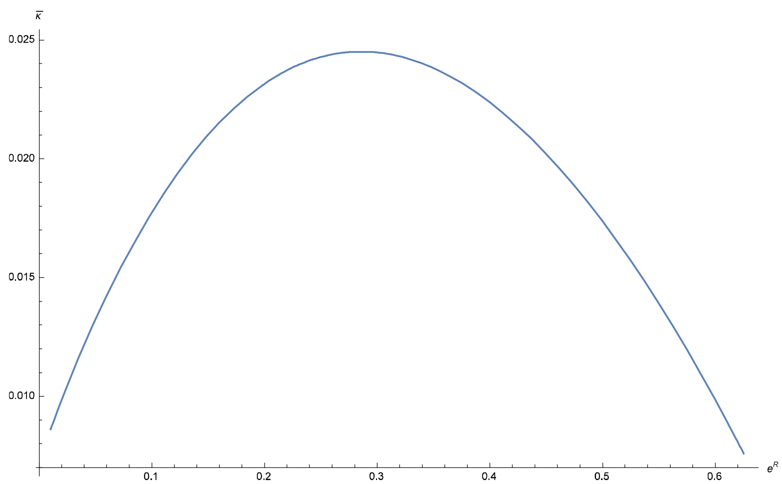


Figure 3

Under concave emissions, the fourth term in the profit function described above does not exist. When a firm with $\bar{\varepsilon}$ breaks even by purchasing at r, all firms with $\varepsilon > \bar{\varepsilon}$, having a lower marginal emission (e_y) at their output scale will also find it optimal to purchase additional permits and produce to full capacity. The cost threshold continues to be defined by (51), but now

$$\hat{n}\left(e^R\right) = \frac{\frac{p}{4w}}{\frac{p}{4w}+c}\left(-\sqrt{\left(\frac{\gamma}{\xi}\right)^2 + 2\frac{e^R}{\xi}} - \frac{\gamma}{\xi} + c\right)$$

$$\varepsilon^R = \frac{c}{\frac{p}{4w}+c}\left(-\sqrt{\left(\frac{\gamma}{\xi}\right)^2 + 2\frac{e^R}{\xi}} - \frac{\gamma}{\xi} - \frac{p}{4w}\right)$$

For $p = w = 1$ and $c = 0.05$ and $\gamma = 1, \xi = -0.25$, Figure 4 below verifies that the cost thresholds are very similar to that with linear emissions.

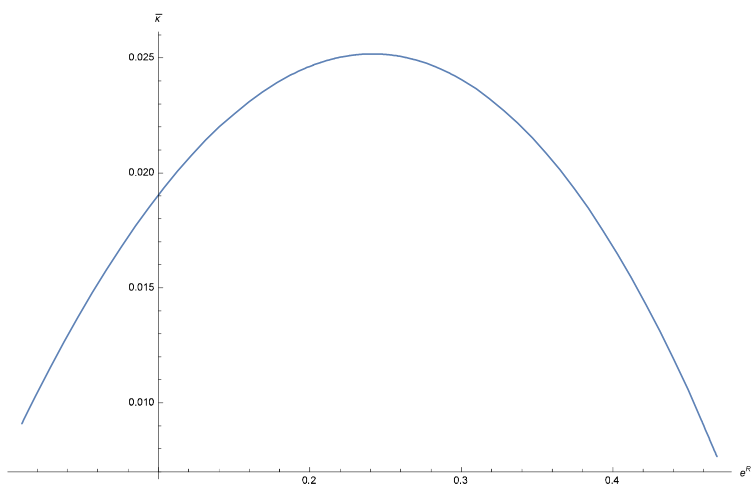


Figure 4

8.2 Equilibrium with reversible inputs

With reversible inputs, firms trade and acquire emission permits $e^0 = e^R$ in a pre-production market at per unit price ρ. After the permits are obtained, they draw their productivity. Without any emission cap, their optimal input employment follows

$$py_n(n,\varepsilon) = w.$$

A firm with higher productivity realization employs a larger amount of input. The optimal $n(\varepsilon; p, w)$ depends on ε and input/output prices. In what follows, to avoid clutter, we continue to use n without its arguments.

With emission caps, there are two possibilities. If permits can be traded, the optimal input employment is given by

$$y_n(n,\varepsilon) = \frac{w}{p-\rho\gamma}, \tag{52}$$

and in equilibrium:

$$y^a = \int y(n,\varepsilon)\, g(\varepsilon)\, d\varepsilon = \frac{e^R}{\gamma} \tag{53}$$

In the absence of trade, and assuming that the emission cap binds for all firms, input employment follows

$$y(n,\varepsilon) = \frac{e^R}{\gamma}$$

and once again (53) holds. Here, a firm with higher productivity employs a smaller amount. The choice problem is trivial.

Allocations implied by (52) and (53) replicate the allocations of a planner who only faces aggregate constraint $\gamma y^a = e^R$, whereas no-trade outcome implies a constraint on each firm. Thus, the two outcomes are different – producing the same output would utilize more inputs in the absence of trade.

8.2.1 Proportional transactions costs

Assume that all firms in equilibrium hold $e^0 = e^R$ units of permits before realizing their productivity. After their productivity realization, firms employ inputs by solving their profit maximization problem:

$$\max \left\{ \begin{array}{c} \max_{n^s} \left\{py\left(n^s,\varepsilon\right) - \left(r-\kappa^s\right)\left(\gamma y\left(n^s,\varepsilon\right) - e^R\right) - wn^s\right\}, \\ \max_{n^b} \left\{py\left(n^b,\varepsilon\right) - \left(r+\kappa^b\right)\left(\gamma y\left(n,\varepsilon\right) - e^R\right) - wn^b\right\}, \\ py\left(n^a,\varepsilon\right) - wn^a, \end{array} \right\}$$

A firm here faces a discrete choice of whether to enter the permit market or stay out. If it stays out, it just produces its autarkic amount $y(n^a,\varepsilon) = \frac{e^R}{\gamma}$. Otherwise, it either sells or buys given its optimal input employment in either case.

Sellers' and buyers' optimal input employments are respectively characterized by

$$y_n(n^s,\varepsilon) = \frac{w}{p-\gamma(r-\kappa^s)} \; ; \; y_n\left(n^b,\varepsilon\right) = \frac{w}{p-\gamma\left(r+\kappa^b\right)}. \tag{54}$$

For a seller $y(n^s,\varepsilon) \leq \frac{e^R}{\gamma}$, whereas for a buyer $y\left(n^b,\varepsilon\right) \geq \frac{e^R}{\gamma}$. Further, since $y_{nn} < 0$ and $y_{n\varepsilon} \geq 0$, (54) implies that there exist $\varepsilon^b > \varepsilon^s$ such that[29]

$$y_n(n^s,\varepsilon) = \frac{w}{p-\gamma(r-\kappa^s)} \text{ for all } \varepsilon \leq \varepsilon^s; \tag{55a}$$

$$y_n\left(n^b,\varepsilon\right) = \frac{w}{p-\gamma\left(r+\kappa^b\right)} \text{ for all } \varepsilon \geq \varepsilon^b. \tag{55b}$$

[29] A proof is offered by contradiction. Suppose $\varepsilon^b \leq \varepsilon^s$. Since $y_n\left(n^b,\varepsilon^b\right) > y_n(n^s,\varepsilon^s)$ for either κ^b or $\kappa^s > 0$. Since $y_{nn} < 0$ and $y_{n\varepsilon} \geq 0$, $\varepsilon^s \geq \varepsilon^b$ requires $n^s > n^b$, which invalidates the equilibrium $y\left(n^b,\varepsilon^b\right) \geq \frac{e^0}{\gamma} \geq y(n^s,\varepsilon^s)$.

Furthermore,

$$y\left(n^{a},\varepsilon\right)=\frac{e^{R}}{\gamma}\text{ for all }\varepsilon\in\left[\varepsilon^{s},\varepsilon^{b}\right] \tag{56}$$

Finally, the marginal firm at either ε^s and ε^b must be indifferent between being in or out of the permit market:

$$y\left(n^{b},\varepsilon^{b}\right)=y\left(n^{s},\varepsilon^{s}\right)=\frac{e^{R}}{\gamma}. \tag{57}$$

When trade occurs in equilibrium, we also require

$$\int_{\varepsilon^{\min}}^{\varepsilon^{s}}\left(e^{R}-\gamma\ y\left(n^{s},\varepsilon\right)\right)g\left(\varepsilon\right)d\varepsilon=\int_{\varepsilon^{b}}^{\varepsilon^{\max}}\left(\gamma\ y\left(n^{b},\varepsilon\right)-e^{R}\right)g\left(\varepsilon\right)d\varepsilon \tag{58}$$

Thus given $\left\{\kappa^s,\kappa^b\right\}$ equations (55a) - (57) determine $r, n^s\left(\varepsilon\right), n^b\left(\varepsilon\right), \varepsilon^b$, and ε^s in equilibrium.

To illustrate how trade costs $\left\{\kappa^s,\kappa^b\right\}$ impede/preclude permit trade, we resort to the parametric example studied in the main text.

A parametric Example Let

$$y\left(n,\varepsilon\right)=\sqrt{n}+\varepsilon,\quad \varepsilon\sim U\left[-c,c\right]$$

We restrict $e^R\leq\gamma\left(\frac{p}{2w}-c\right)$ to ensure that the cap is utilized by all firms even in the absence of trade. We further need $\frac{e^R}{\gamma}>c$ to make sure that

Then choices n^s and n^b are independent of ε. Specifically,

$$\sqrt{n^{s}}=\frac{p-\gamma\left(r-\kappa^{s}\right)}{2w};\sqrt{n^{b}}=\frac{p-\gamma\left(r+\kappa^{b}\right)}{2w}, \tag{59}$$

which using (56) gets

$$\sqrt{n^{s}}-\sqrt{n^{b}}=\varepsilon^{b}-\varepsilon^{s}=\gamma\frac{\kappa^{b}+\kappa^{s}}{2w}=\gamma\frac{\kappa^{T}}{2w}. \tag{60}$$

On the other hand, (58) along with the preceding result gets

$$\frac{\sqrt{n^{s}}+\sqrt{n^{b}}}{2}=\frac{e^{R}}{\gamma},$$

The preceding two results obtain sellers' and buyers' input employment as

$$\sqrt{n^{s}}=\frac{e^{R}}{\gamma}+\gamma\frac{\kappa^{T}}{4w};\quad\sqrt{n^{b}}=\frac{e^{R}}{\gamma}-\gamma\frac{\kappa^{T}}{4w}.$$

while permit price is obtained using (59) as

$$r=\frac{p}{\gamma}-2w\frac{e^{R}}{\gamma^{2}}+\frac{\kappa^{s}-\kappa^{b}}{2}$$

Allocations thus depend solely on the total cost κ^T and not on κ^b and κ^s separately. Trading costs create a wedge between sellers' and buyers' opportunity costs of permits and distort input employment accordingly. Sellers end up employing a higher amount of inputs while buyers' input employment is depressed. When $\kappa^T=0$, $n^s=n^b$, i.e., all firms employ an identical amount. On the other hand, when

$\kappa^T = \frac{4wc}{\gamma}$, all firms produce equal output and emissions and their input employment trivially follows from

$$\sqrt{n^0} = \frac{e^R}{\gamma} - \varepsilon$$

Figure 5 below illustrates how input allocations distort with increase in total cost. The horizontal line represents $\kappa^T = 0$, z-shaped represents $\kappa^T = \frac{2wc}{\gamma}$. The allocations return to autarky with $\kappa^T = \frac{4wc}{\gamma}$ shown with dashed line.

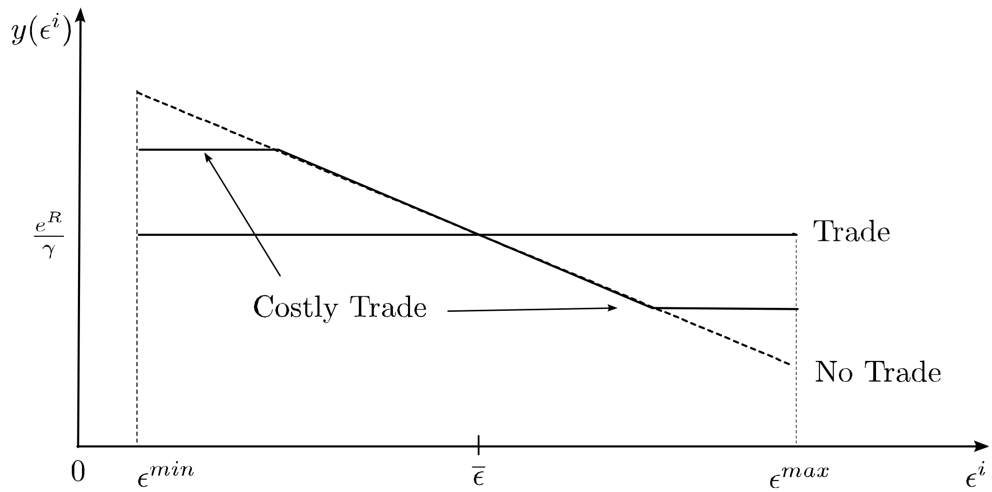


Figure 5

In addition, trade requires

$$\begin{aligned} r + \kappa^b &= \frac{p}{\gamma} - 2w\frac{e^R}{\gamma^2} + \frac{\kappa^T}{2} < \frac{p}{\gamma} \Rightarrow \frac{\gamma\kappa^T}{4w} < 2\frac{e^R}{\gamma}; \\ r - \kappa^s &= \frac{p}{\gamma} - 2w\frac{e^R}{\gamma^2} - \frac{\kappa^T}{2} > 0 \Rightarrow \frac{\kappa^T}{4w} < \frac{p}{4w} - 2\frac{e^R}{\gamma}. \end{aligned}$$

Under an equilibrium with trade, both conditions must hold. Similar to irreversible inputs, there is an upper bound on total costs which precludes trade:

$$\bar{\kappa}^T\left(e^R\right) \equiv \begin{cases} \frac{4e^R w}{\gamma^2} & \text{for} \quad e^R < \frac{\gamma\, p}{4w} \\ \frac{2p}{\gamma} - \frac{4e^R w}{\gamma^2} & \text{for} \quad e^R \in \left[\frac{\gamma\, p}{4w}, \frac{\gamma\, p}{2w}\right] \end{cases} \tag{61}$$

The above bounds are identical to those for irreversible inputs studied in the main text (see equation (14)). There, these bounds impede trade partially. Here, in contrast, trade ceases completely once these bounds are crossed. The first bound for $e^R < \frac{\gamma\, p}{4w}$ ensures a buyer's participation, while the one for $e^R > \frac{\gamma\, p}{4w}$ ensures non-negative surplus for the sellers.

However, (61) has to be consistent with (60), which requires $\kappa^T < \frac{4wc}{\gamma}$. Suppose $p < 4wc$. Then, $\bar{\kappa}^T\left(e^R\right) = \frac{4e^R w}{\gamma^2}$ for all $e^R < \frac{\gamma\, p}{4w}$. If instead $p > 4wc$, $\bar{\kappa}^T\left(e^R\right) = \frac{4e^R w}{\gamma^2}$ for all $e^R < \gamma c$ and $\bar{\kappa}^T\left(e^R\right) = \frac{4wc}{\gamma}$ for all $e^R \in \left[\gamma c, \frac{\gamma\, p}{4w}\right]$. Similarly, if $p < 4wc$, $\bar{\kappa}^T\left(e^R\right) = \frac{2p}{\gamma} - \frac{4e^R w}{\gamma^2}$ for all $e^R > \frac{\gamma\, p}{4w}$. Instead if $p > 4wc$, $\bar{\kappa}^T\left(e^R\right) = \frac{4wc}{\gamma}$ for all $e^R < \gamma\left(\frac{p}{2w} - c\right)$ and $\bar{\kappa}^T\left(e^R\right) = \frac{2p}{\gamma} - \frac{4e^R w}{\gamma^2}$ for all $e^R \in \left[\gamma\left(\frac{p}{2w} - c\right), \frac{\gamma\, p}{2w}\right]$.

Efficiency Total output produced is $\frac{e^R}{\gamma}$. Total inputs employed equal $\frac{1}{2c}$ times

$$\begin{aligned} & 2\left(\left(\frac{e^R}{\gamma}\right)^2+\left(\gamma\frac{\kappa^T}{4w}\right)^2\right)\left(c-\gamma\frac{\kappa^T}{4w}\right)+\left(\frac{e^R}{\gamma}\right)^2 2\gamma\frac{\kappa^T}{4w}+\frac{1}{3}2\left(\gamma\frac{\kappa^T}{4w}\right)^3 \\ = \; & 2c\left(\left(\frac{e^R}{\gamma}\right)^2+\left(\gamma\frac{\kappa^T}{4w}\right)^2\right)-\frac{4}{3}\left(\gamma\frac{\kappa^T}{4w}\right)^3 \end{aligned}$$

Excess input employment is measured by

$$\frac{\left(\gamma\frac{\kappa^T}{4w}\right)^2-\frac{2}{3c}\left(\gamma\frac{\kappa^T}{4w}\right)^3}{\left(\frac{e^R}{\gamma}\right)^2}$$

which increases in $\gamma\frac{\kappa^T}{4w}\in[0,c]$, with its maximum at $\frac{c^2}{3}/\left(\frac{e^R}{\gamma}\right)^2$

8.2.2 Fixed costs

With fixed costs, all trading firms that are in the market face the same marginal opportunity cost and their input employment follows

$$y_n\left(n^T,\varepsilon\right)=\frac{w}{p-\gamma r}$$

Here, superscript T denotes trading firms. Firms with sufficiently low productivities have excess permits and those with high productivities fall short of them, and there again exist ε^s and ε^b such that

$$\begin{aligned} r &= \frac{\kappa^s}{e^R-\gamma y\left(n,\varepsilon^s\right)} \\ p-\gamma r &= \frac{\kappa^b}{y\left(n,\varepsilon^b\right)-\frac{e^R}{\gamma}} \end{aligned}$$

Again, permit market equilibrium requires

$$\int_{\varepsilon^{\min}}^{\varepsilon^s}\left(e^R-\gamma y\left(n,\varepsilon\right)\right)\,g(\varepsilon)\,d\varepsilon=\int_{\varepsilon^b}^{\varepsilon^{\max}}\left(\gamma y\left(n,\varepsilon\right)-e^R\right)\,g(\varepsilon)\,d\varepsilon.$$

Firms that are not in the market produce their autarkic amount:

$$y\left(n^a,\varepsilon\right)=\frac{e^R}{\gamma}\text{ for all }\varepsilon\in\left[\varepsilon^s,\varepsilon^b\right]$$

Thus given $\left\{\kappa^s,\kappa^b\right\}$ and input/output prices the preceding system of equations determines $r, n\left(\varepsilon\right), \varepsilon^b$, and ε^s in equilibrium.

To illustrate how trade costs $\left\{\kappa^s,\kappa^b\right\}$ impede/preclude permit trade, we resort to the parametric example studied for the proportional case.

8.2.3 A parametric example

Let

$$y\left(n,\varepsilon\right)=\sqrt{n}+\varepsilon,\quad\varepsilon\sim U\left[-c,c\right]$$

Then, for trading firms, denoted by superscript T, we have

$$\sqrt{n^T} = \frac{p - \gamma r}{2w};$$

The cut off productivities are given by

$$\varepsilon^s = \frac{e^R}{\gamma} - \frac{p - \gamma r}{2w} - \frac{\kappa^s}{\gamma r};\ \varepsilon^b = \frac{e^R}{\gamma} - \frac{p - \gamma r}{2w} + \frac{\kappa^b}{p - \gamma r} \tag{62}$$

Using above results with the permit market clearing obtains (identical to that obtained in the main text for irreversible investments):

$$\sqrt{n^T} = \frac{p - \gamma r}{2w} = \frac{e^R}{\gamma} + \frac{1}{4c}\left(\left(\frac{\kappa^b}{p - \gamma r}\right)^2 - \left(\frac{\kappa^s}{\gamma r}\right)^2\right) \tag{63}$$

Those under autarky, denoted by superscript a, choose

$$\sqrt{n^a} = \frac{e^R}{\gamma} - \varepsilon$$

Using (62) and (63) It can be shown that $\varepsilon^b \to c \Leftrightarrow \varepsilon^s \to -c$. For the symmetric case (as in the main text) $\kappa^b = \kappa^s = \kappa$. The value of κ that obtains $\varepsilon^b \to c$ and $\varepsilon^s \to -c$ is

$$\bar{\kappa} = \frac{pc}{2}\left[1 - \left(\frac{e^R - \gamma\frac{p}{4w}}{\gamma\frac{p}{4w} + c}\right)^2\right],$$

thus obtaining exactly the same cost threshold as under irreversible investments in the main text.

8.3 Emission/abatement proportional to inputs

Suppose $z = \theta m$, where $m \leq n$ is input used in production, where n is the input irreversibly employed before ε is realized. First, consider the equilibrium under costless trade. When $z = \gamma y$, all firms have a positive net benefit from input use at the production stage, i.e., $(p - \gamma r)\, y_m > 0$, and therefore utilize all their inputs. Firms with high productivity produce and emit more than the average by buying permits, while those with low productivity produce and emit less. With $z = \theta m$, the net marginal benefit $p\, y_m - \theta\, r$ may be negative for some firms with sufficiently low productivity if they attempt to utilize their full capacity. In this case, they may choose to utilize $m < n$, and some of them may even benefit by selling their permits off by employing $m < \frac{e^R}{\theta}$. On the other hand, firms with sufficiently large productivity may employ $m \in (\frac{e^R}{\theta}, n]$ by purchasing additional permits in the market. This ex post scenario may lead to firms ex ante employing inputs $n > \frac{e^R}{\theta}$. Of course, if ex ante $n \leq \frac{e^R}{\theta}$, there is no ex post trade in equilibrium. In general, as we find below, $n > \frac{e^R}{\theta}$ is likely to hold only when $\frac{e^R}{\theta} << \tilde{n}$.

When $z = \theta m$, an output technology of the form $y\,(m, \varepsilon) = y\,(m) + \varepsilon$ rules out any ex-post trade. To see this, note that the marginal input cost (opportunity cost of emissions) is θr whereas its marginal value product is $p y_m\,(m, \varepsilon)$. A firm will employ all its inputs as long as $p y_m\,(n, \varepsilon) \geq \theta r$. If not, it will use $m < n : y_m\,(m, \varepsilon) = \theta r$. Since all firms begin with n, all firms will use same amount of input regardless of their ε and generate the same amount of emissions. Therefore, $y\,(m, \varepsilon) \equiv \varepsilon y\,(m)$ is assumed in what follows. Below we summarize the key features of equilibria under (a) no permit trade (b) costless trade, and (c) with proportional trading costs.

8.3.1 No trade in permits

Here, the problem is trivial: choose $n = m$ from

$$py_n^e(n) = w$$

if

$$py_n^e\left(\frac{e^R}{\theta}\right) < w$$

Otherwise

$$n = \frac{e^R}{\theta}$$

8.3.2 Costless trade in permits

Assuming that a firm enters the production stage with $e^0 = e^R$ units of emission permits, the profit maximization problem at the production stage is

$$py(m,\varepsilon) - r\left(\theta m - e^R\right) \text{ subject to } m \leq n$$

The optimality demands that for all firms with $\varepsilon < \hat{\varepsilon}$, m is given by

$$py_m(m,\varepsilon) = \theta r \tag{64}$$

whereas for all $\varepsilon > \hat{\varepsilon} : py_n(n,\hat{\varepsilon}) = \theta r$

$$py_n(n,\varepsilon) > \theta r$$

and $m = n$. These are the firms that would have liked to employ more inputs if they could.

Finally, aggregate permit market clearing requires that

$$\int_{\varepsilon_{\min}}^{\hat{\varepsilon}} m(\varepsilon)\, g(\varepsilon)\, d\varepsilon + n\left(1 - G(\hat{\varepsilon})\right) = \frac{e^R}{\theta} \tag{65}$$

In the planning stage, a firm knows its state-contingent choices after ε is realized, can rationally forecast the equilibrium r, and therefore ex ante chooses e^0 and n to maximize its expected profit

$$\int_{\varepsilon_{\min}}^{\hat{\varepsilon}} \left[py(m(\varepsilon),\varepsilon) - \theta r' m(\varepsilon)\right] g(\varepsilon)\, d\varepsilon + \int_{\hat{\varepsilon}}^{\varepsilon_{\max}} \left[py(n,\varepsilon) - \theta r' n\right] g(\varepsilon)\, d\varepsilon - wn + re^0 - \rho e^0$$

which gets by arbitrage

$$r = \rho.$$

The optimal choice of n follows

$$p\int_{\hat{\varepsilon}}^{\varepsilon_{\max}} y_n(n,\varepsilon)\, g(\varepsilon)\, d\varepsilon = w + \theta r\left(1 - G(\hat{\varepsilon})\right) \tag{66}$$

Recall that $\hat{\varepsilon}$ solves $y_n(n,\hat{\varepsilon}) = \frac{\theta r}{p}$. There are two unknowns: r and n. They are obtained from (65) and (66).

A parametric example with trade An additive productivity shock would imply that firms ex-post employ an identical amount of inputs and generate identical emissions which would obviate any ex post permit trade. Therefore, for this part we resort to

$$y(n,\varepsilon)=\varepsilon\sqrt{n};\quad \ln\varepsilon\tilde{\ }N(0,\sigma).$$

The production stage choice (64) gets

$$\frac{\varepsilon}{2\sqrt{m}}=\frac{\theta r}{p};\qquad m=\frac{1}{4}\frac{\varepsilon^2p^2}{\theta^2r^2}.$$

The cutoff is given by

$$\hat{\varepsilon}=2\sqrt{n}\frac{\theta r}{p}$$

And the input choice follows from (66):

$$\sqrt{n}=\frac{p}{2}\frac{\exp\left[\frac{1}{2}\sigma^2\right]\left(1-\Phi\left[\frac{\ln 2\sqrt{n}\frac{\theta r}{p}-\sigma^2}{\sigma}\right]\right)}{w+\theta r\left(1-\Phi\left[\frac{\ln 2\sqrt{n}\frac{\theta r}{p}}{\sigma}\right]\right)} \tag{67}$$

and the permit market equilibrium requires

$$\frac{1}{4}\frac{p^2}{\theta^2r^2}\exp\left[2\sigma^2\right]\Phi\left[\frac{\ln 2\sqrt{n}\frac{\theta r}{p}-2\sigma^2}{\sigma}\right]+n\left(1-\Phi\left[\frac{\ln 2\sqrt{n}\frac{\theta r}{p}}{\sigma}\right]\right)=\frac{e^R}{\theta} \tag{68}$$

Thus, n and r are jointly determined by (67) and (68).

Finally, the aggregate output produced under trade equals

$$y^*=\frac{1}{2}\frac{p}{\theta r}\exp\left[2\sigma^2\right]\Phi\left[\frac{\ln 2\sqrt{n}\frac{\theta r}{p}-2\sigma^2}{\sigma}\right]+\sqrt{n}\exp\left[\frac{1}{2}\sigma^2\right]\left(1-\Phi\left[\frac{\ln 2\sqrt{n}\frac{\theta r}{p}-\sigma^2}{\sigma}\right]\right)$$

8.3.3 Proportional trading costs

Trading costs, once again, place a wedge between buyers and sellers and generate respective *ex post* output/abatement choices. As a result, there are three productivity thresholds: $\{\varepsilon^s,\varepsilon^b,\hat{\varepsilon}\}$ such that firms below ε^s sell, those above ε^b buy, those with $\varepsilon\in(\varepsilon^s,\varepsilon^b)$ stay out of the market, and those above $\hat{\varepsilon}$ not only buy but also use all inputs to produce and do no abatement. These cutoffs follow from

$$\frac{\varepsilon^s}{2\sqrt{m^s}}=\frac{\theta(r-\kappa^s)}{p};\ \frac{\varepsilon^b}{2\sqrt{m^b}}=\frac{\theta\left(r+\kappa^b\right)}{p};\ \hat{\varepsilon}=\varepsilon^b\sqrt{\frac{n}{m^b}}.$$

Since at the margin $m^s=m^b=\frac{e^R}{\theta}$,

$$\frac{\varepsilon^s}{\varepsilon^b}=\frac{r-\kappa^s}{r+\kappa^b}.$$

Again, the two unknowns r and n are determined jointly from *ex ante* input choice:

$$\frac{p}{2\sqrt{n}}\int_{\hat{\varepsilon}}^{\varepsilon_{\max}}\varepsilon g(\varepsilon)\,d\varepsilon=w+\theta\left(r+\kappa^b\right)(1-G(\hat{\varepsilon}))$$

and the permit market equilibrium

$$\int_{\varepsilon_{\min}}^{\varepsilon^s} m(\varepsilon)\, g(\varepsilon)\, d\varepsilon + \int_{\varepsilon^b}^{\hat{\varepsilon}} m(\varepsilon)\, g(\varepsilon)\, d\varepsilon + n\left(1 - G(\hat{\varepsilon})\right) = \frac{e^R}{\theta}\left(1 - \int_{\varepsilon^s}^{\varepsilon^b} g(\varepsilon)\, d\varepsilon\right).$$

With our parametric assumptions, these are, respectively:

$$\sqrt{n} = \frac{p}{2}\frac{\exp\left[\frac{1}{2}\sigma^2\right]\left(1 - \Phi\left[\frac{\ln 2\sqrt{n}\frac{\theta\left(r+\kappa^b\right)}{p} - \sigma^2}{\sigma}\right]\right)}{w + \theta\left(r + \kappa^b\right)\left(1 - \Phi\left[\frac{\ln 2\sqrt{n}\frac{\theta\left(r+\kappa^b\right)}{p}}{\sigma}\right]\right)} \tag{69}$$

and

$$\frac{e^R}{\theta} = \frac{\begin{array}{c} \frac{1}{4}\frac{p^2}{\theta^2(r-\kappa^s)^2}\exp\left[2\sigma^2\right]\Phi\left[\frac{\ln\left(2\sqrt{\frac{\theta}{\theta}}\frac{\theta\left(r-\kappa^s\right)}{p}\right) - 2\sigma^2}{\sigma}\right] \\ +\frac{1}{4}\frac{p^2}{\theta^2(r+\kappa^b)^2}\exp\left[2\sigma^2\right]\left(\begin{array}{c} \Phi\left[\frac{\ln\left(2\sqrt{n}\frac{\theta\left(r+\kappa^b\right)}{p}\right) - 2\sigma^2}{\sigma}\right] \\ -\Phi\left[\frac{\ln\left(2\sqrt{\frac{\theta}{\theta}}\frac{\theta\left(r+\kappa^b\right)}{p}\right) - 2\sigma^2}{\sigma}\right] \end{array}\right) \\ +n\left(1 - \Phi\left[\frac{\ln\left(2\sqrt{n}\frac{\theta\left(r+\kappa^b\right)}{p}\right)}{\sigma}\right]\right) \end{array}}{\left[1 - \Phi\left[\ln\left(2\sqrt{\frac{\theta}{\theta}}\frac{\theta(r+\kappa^b)}{p}\right)\right] + \Phi\left[\ln\left(2\sqrt{\frac{\theta}{\theta}}\frac{\theta(r-\kappa^s)}{p}\right)\right]\right]}$$

where n follows from (69).

Figure 6 below shows how the mass of various trading and non-trading groups vary with symmetric trading costs $\kappa^b = \kappa^s = \kappa$. For this example, we assume $w = 1$, $p = 100$, and $\sigma = 0.2$. With $\theta = 1$, unconstrained amount of emissions are $\tilde{z}^a = \left(\frac{p}{2w}\right)^2 \exp\left[\sigma^2\right]$. Let $e^R = 0.01\tilde{z}^a$. The figure below exhibits how trading costs separate firms population in four groups. "Full-emitters" are firms that utilize 100% inputs and emit to their full capacity. "Sellers" are those who utilize less than their full capacity and sell a part of their permits to full-emitters as well as to some other "buyers", i.e., firms that do not utilize full capacity yet need extra permits to produce at their optimum. Finally, there are "Non-traders", i.e., firms that neither buy nor sell but simply use $m = \frac{e^R}{\theta}$ to fully utilize their permit holdings.

Intuitively, as trading costs increase, the mass of traders – buyers that produce less than capacity and sellers - shrink to zero. The entire population of firms is either out of market or producing and emitting at their full capacity. A natural question to ask is that if the mass of sellers is shrinking, how are full-emitters, i.e., firms that were utilizing full capacity by buying permits from sellers, could still be about the same? This is explained by the ex ante input employment that takes into account the ex post market equilibrium. As costs increase, in the absence of trade, all firms indeed employ the same amount of input. As a result, n approaches $\frac{e^R}{\theta}$.

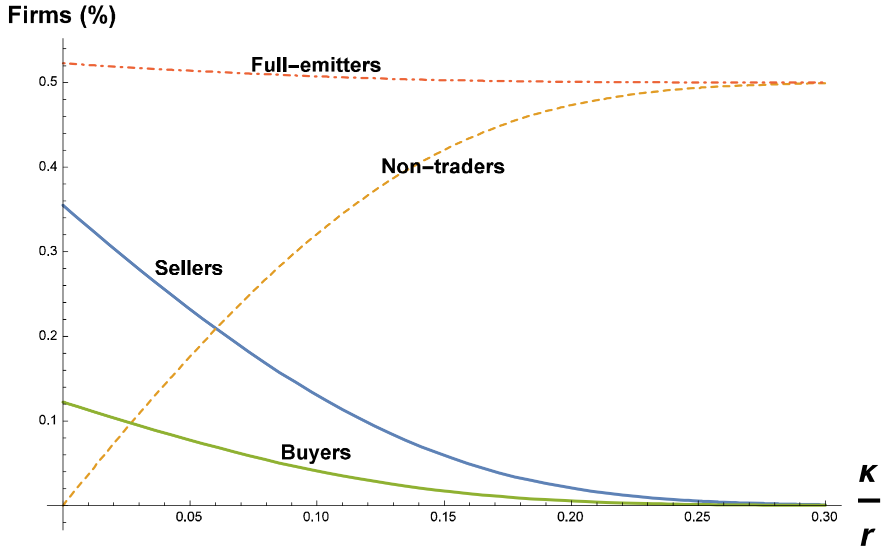
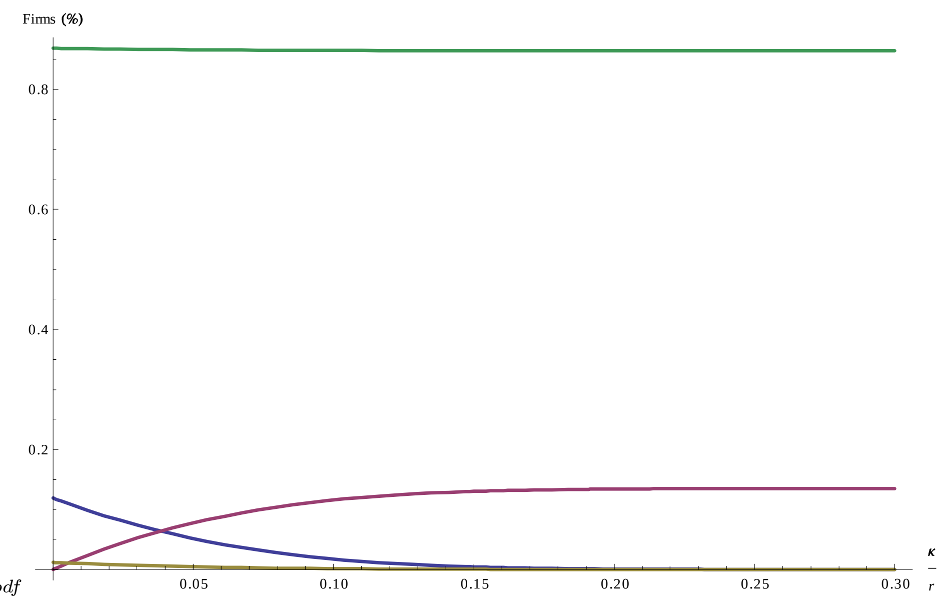


Figure 6

The distribution changes to for $e^R = 0.1\tilde{z}^a$ for varying trading costs are exhibited below (labels correspond to those in Figure 6 above)

7



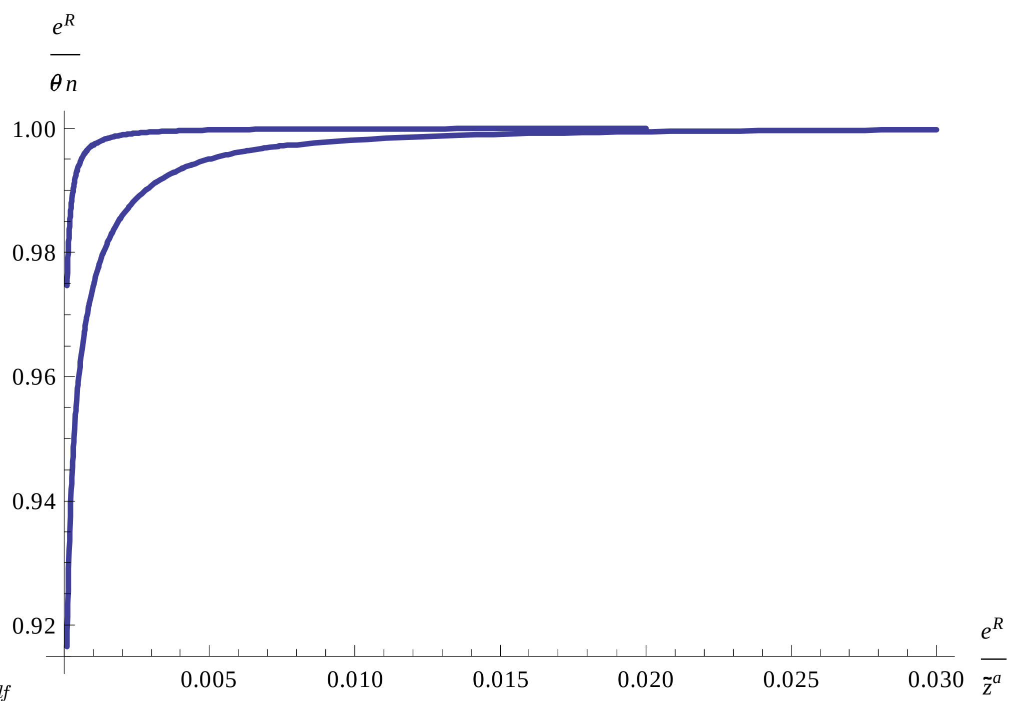
$_1$3.pdf

Figure 7

Clearly, about 90% of the population utilizes all of its inputs while about 10% sell some of their permits to these full-emitters when the trade is costless. As costs increase sellers shrink and non-traders increase. In fact, the input used by non-traders and full emitters coincides as can be seen in the Figure 8 below that shows how $\frac{e^R}{\theta n}$ changes with $\frac{e^R}{\tilde{z}^a}$. The curve that lies below represents trading costs of 20%.

It illustrates how quickly the input choice approach its common level of $\frac{e^R}{\theta}$ when the cost goes up from 20% to 30% (represented by the curve that lies above)

8



$_1$4.pdf

Figure 8