A Contribution Margin Approach to Imperfect Competition

Philippe Bontems*

February 28, 2025

Abstract

This paper introduces a novel analysis of symmetric oligopoly equilibrium, focusing on the margin over variable cost. It offers a comprehensive method to determine equilibrium properties across various oligopoly models, emphasizing the profitability ratio's role, distinct from the Lerner index. The study explores short and long-term equilibrium characteristics, highlighting cost shocks transmission, the effects of market expansion and market concentration. Additionally, it extends to regulated oligopolies, delineating the short-run and long-run welfare impact of policy instruments based on price and average variable cost elasticities to regulation.

JEL : D21, H22, H32, L13, L51

Key-words: Pass-Through, Cost Structure, Marginal Value of Public Funds, Tax Incidence, Oligopoly.

*Toulouse School of Economics, INRAE, University of Toulouse Capitole, France. I acknowledges funding from ANR under grant ANR-17-EURE-0010 (Investissements d'Avenir program).

1 Introduction

This article proposes an analysis of the symmetric oligopoly equilibrium using a new approach in terms of the margin on variable cost, also denoted the *contribution margin* approach, as the margin on variable cost usually contributes to cover fixed costs if any. While it is well known that the properties of oligopoly equilibrium depend on the specification of demand, which is usually carefully discussed and chosen, it is striking that the literature usually places less emphasis on the specification of costs, which are often assumed to have constant marginal costs. For instance, Bertoletti and Etro (2016) in their thorough study on how preferences influence entry and market structure in oligopolies and monopolistic competition assume constant marginal costs. In a similar way, Mrázová and Neary (2017) study in depth how the demand structure influences monopolistic competition outcomes while assuming constant marginal costs. To overcome this issue, Bontems (2023a) proposes to describe the equilibrium of the firm in monopoly or monopolistic competition using a small number of demand *and* cost characteristics, namely one elasticity and one curvature measure for both the demand and the cost structures. He shows that the profitability index, also known as the *contribution margin index*, which measures the margin on variable cost *per unit price*, is distinct from the Lerner index measuring market power, except in the case of constant marginal costs. The ratio of the profitability index to the Lerner index is denoted the *profitability scale factor* and is a function of the elasticity of demand and the elasticity of "supply", the value of which depends on the equilibrium economies of scale. Importantly, the elasticity of "supply" is here unconventionally defined from interpreting the average variable cost (hereafter AVC) as if it were a wholesale price that a merchant would obtain from his upstream suppliers before selling in the downstream market. This approach is fruitful to study how market expansion can give rise to increasing concentration in an industry composed of heterogeneous firms in monopolistic competition, and to separate what is due to the demand and the cost structures in the resulting evolution of the industry (see Bontems, 2023a).[1]

Building on these results, I propose extending the above contribution margin approach to consider a wide range of static oligopoly models using a general mode of competition *à la* Weyl and Fabinger (2013), including both price and quantity competition among other possibilities. The framework also considers variety-loving preferences and accounts for endogenous market structures in the presence of firm entry and exit. In this context, I first show that the contribution margin index is inversely

[1] In their in-depth study of size distribution of US corporations for 100 years, Kwon et al. (2023) particularly emphasize the role of economies of scale in explaining rising corporate concentration.

proportional to the effective elasticity of industry's demand factorized by the profitability scale factor. The effective elasticity of industry's demand is the elasticity of industry's demand deflated by an appropriate measure of the intensity of competition within the oligopoly. The profitability scale factor is a function of the effective elasticity of demand and the elasticity of "supply" as defined above. Last, I show that the profitability scale factor is constantly equal to unity if and only if the marginal cost is constant and therefore the contribution margin index and the Lerner index merge. On the contrary, when the AVC is (locally) increasing (decreasing) at the equilibrium then the profitability scale factor is lower (greater) than unity. Overall, the contribution margin approach helps to identify the circumstances in which profitability and market power go hand in hand or not.

If this study of the interplay between profitability and market power in (symmetric) oligopoly models has its own interest, it also appears that it is useful to obtain concise cost pass-through formulas, new but consistent with those that exist in the literature both in the short term and in the long term, when the number of firms is endogenous (Weyl and Fabinger, 2013; Kroft et al., 2021). In particular, these new pass-through formulas highlight the role of economies of scale in the reaction of prices to cost shocks. For instance, the cost pass-through of an infinitesimal additive shock like e.g. a unit tax is shown to be a function of the profitability scale factor, of the market conduct index (and its sensitivity to output) as well as an average measure of curvature of demand and "supply". Moreover, whatever the mode of competition and whether the cost shock is additive (e.g. unit tax) or multiplicative (e.g. ad-valorem tax), cost convexity (concavity) implies reduced (increased) cost pass-through on price or quantity. This result is obtained by comparing the oligopoly equilibrium with an equilibrium obtained by a homologous industry, or *ghost industry*, for which marginal cost is constant but equal to the *equilibrium value of AVC* of the original industry.

I then examine the comparative statics of the model with respect to the number of competitors and the size of the market represented by the number of consumers. I also study the cost pass-through of additive and multiplicative cost shock in the long tun when the number of firms is endogenous. In all these results, it appears that a key factor in determining the evolution of the industry is the size of the pass-through on price of a zero unit tax (also denoted the Absolute Pass-Through or *APT* in short).

To illustrate, consider the impact of an exogenous increase in the number of firms. The resulting impact on market price is decomposed into a lower marginal willingness to pay for each variety, holding consumption per variety constant, and a business stealing/expansion effect majored by *APT*.

Despite the fundamental insight of economics that competition usually lowers prices, the situation of price-increasing competition has been identified as plausible both in the empirical and theoretical literature (see e.g. Pauly and Satterthwaite, 1981; Bresnahan and Reiss, 1990 and 1991; Grabowski and Vernon, 1992; Perloff et al., 2006; Ward et al., 2002; Chen and Savage, 2011; Chen and Riordan, 2008; Bertoletti and Etro, 2016; Mangin, 2022). In the present framework which covers a wide range of oligopoly models with general demand and cost structure, the situation of price-increasing competition is identified through a simple condition that relies on a strong business stealing effect that drives down quantities per firm and that is sufficient to overcome the decreasing marginal willingness to pay for each variety.

Concerning the pass-through of additive and multiplicative cost shocks in the long run, the analysis identifies three possible scenarios with respect to the evolution of the industry, according to the relative value of APT. Scenario 1 characterizes a situation where the shock is pro-competitive but decreases the size of each firm in terms of output at equilibrium. More varieties are sold in equilibrium but the output of each is reduced. In the other two possible scenarios, the shock reduces the number of firms (anti-competitive) but may either reduce (scenario 2: declining industry) or increase the output per firm (scenario 3: fewer varieties but larger firms). Two other results are of interest. First, if marginal costs are constant then the conditions for the different scenarios do not depend on whether the cost shock is additive or multiplicative. Otherwise, when AVC is increasing in output, which implies decreasing return to scale, a multiplicative cost shock is more likely to be pro-competitive and to reduce output per firm than an additive cost shock. Second, it is theoretically possible that an additive cost shock ends up with a negative pass-through in the long run when scenario 3 occurs. This outcome corresponds to a version of Edgeworth's taxation paradox according to which taxation can reduce market price although this paradox is usually developed in the context of a multiproduct monopoly (Edgeworth, 1925; Armstrong and Vickers, 2022).

Furthermore, the contribution margin approach allows to obtain a set of sufficient statistics to describe the profit and output consequences of a raising market size, both in the short and the long run. For this, I compare the oligopoly equilibrium to the equilibrium obtained under another ghost industry where marginal costs are constant and equal to the *equilibrium value of marginal cost* for the original industry. Intuitively, both industries produce the same output per firm in the short run. The incidence of market expansion on market price is non trivial. Actually, market expansion is price-decreasing if and only the elasticity of output per firm with respect to market size is larger than unity.

I show that the latter condition is equivalent to having *APT* for the original industry larger than for the ghost industry. Hence, when the industry shares the same cost structure as the ghost industry then output per firm is directly proportional to market size and market expansion is price-neutral. Outside of this limit case, both the cost convexity/concavity and the sensitivity of market conduct to output influence whether market expansion decreases or raises market price. When the number of firms is endogenous, market expansion can be procompetitive when the output-per-firm reaction is not too large and anticompetitive otherwise.

Finally, the basic framework is extended to consider regulated oligopolies. The motivation is to possibly obtain new insights in situations where the regulatory policy contains many instruments that can be compared from a welfare perspective or in situations where one wishes to assess the welfare impact of a particular instrument given the pre-existing regulations in place. This includes not only the comparison of ad-valorem and unit taxes as in Adachi and Fabinger (2022) and Kroft et al. (2021) but also the welfare ranking between unit and ad-valorem cost subsidies. The notion of regulation has to be understood in a broad sense: the case of exogenous competition (Weyl and Fabinger, 2013; Miklos-Thal and Shaffer, 2021), through the entrance of exogenous quantity on the market, can also be analyzed using the framework.

The contribution margin approach can be extended as long as the *effective* average variable cost is considered, taking into account the regulatory burden. Two other important functions are the first-order sensitivity of the contribution margin to price and the first-order sensitivity of the markup to price. The former is particularly important because it captures how regulation changes the intensity of competition, regardless of the mode of competition. An important step in the analysis is to evaluate the *APT* on price of a zero unit tax added to the existing regulation, which will prove to be a key factor in assessing the impact of the policy under scrutiny. Indeed, the elasticity of market price with respect to any regulatory instrument is expressed as the *APT* factored by a weighted mean of the (partial) elasticity of the effective marginal cost (included the regulation burden) and of the (partial) elasticity of the first-order sensitivity of the contribution margin to price, both with respect to the regulatory instrument. To compare any two policy instruments in terms of market performance, it is convenient to focus on the ratio of elasticities of price with respect to each instrument. When market power is high, this ratio depends mostly on how the two instruments respectively impact the intensity of competition. On the contrary, when market power is low, the relative impact of instruments on market price mainly depends on how these instruments respectively impact the effective marginal cost.

Similarly, when the degree of profitability is high, the comparative effect on profit depends mainly on how the instruments respectively impact the market price. On the contrary, when the degree of profitability is low, the comparative effect on profit depends mainly on how the instruments impact the effective average variable cost respectively.

Finally, the marginal excess burden associated, as well as the marginal value of public funds for the policy instrument considered, are formulated, either in the short run when the number of firms is fixed or in the long run when the number of firms is endogenous. A simple condition for preference reversal on the ranking of two policy instruments between the short run and the long run is established.

The rest of the paper is organized as follows. The last part of this section is devoted to the related literature. Section 2 characterizes the demand and the cost sides of the market modelling, as well as the mode of competition between firms. Following Weyl and Fabinger (2013), the intensity of competition is measured through a market conduct index centered on the focal point of overall industry profit maximization. In Section 3, I study the equilibrium properties of the contribution margin index and I establish the relationship with the Lerner index. I also establish the stability conditions for both the short run and the long-run. Section 4 is devoted to reformulating cost pass-through. Section 5 presents a number of comparative results with respect to the number of firms and the market size. I also study the pass-through of additive and multiplicative cost shocks on price and output when the number of firms is endogenous. In Section 6, I use the above results to study the incidence and the welfare impact of regulation in oligopoly contexts. In particular, I establish simple formulations for the marginal excess burden and the marginal cost of public funds associated with any policy instrument within the regulation policy. Section 7 concludes.

Related literature. The present work contributes to several strands in the literature. First, Weyl and Fabinger (2013) have shown that the cost pass-through is a powerful tool to analyze market performance, building a bridge between the analysis of imperfect competition in industrial organization and the incidence of taxes as examined in public finance (see also Miklos-Thal and Shaffer, 2021; and Ritz et al., 2018, for a survey). Like their study, I consider general demand and cost structures as well as a general mode of competition to study market outcomes under (symmetric) imperfect competition. I extend their work by also studying the incidence of market concentration, of market size and also the long-run cost pass-through.

Second, it contributes to the literature on tax incidence in oligopoly (Delipalla and Keen, 1992;

Anderson et al., 2001a; Anderson et al., 2001b; Häckner and Herzing, 2016; Adachi and Fabinger, 2022; Kroft et al., 2021). The present paper offers a unified framework through which the incidence of a policy instrument can be ascertained given the existing policy, in the short run and in the long run for a wide range of oligopoly models. In particular, the role of the elasticities of price and AVC with respect to the policy instrument is highlighted in forming the social incidence and the marginal value of public funds.

Last, it adds to the literature on endogenous structure of markets like Bertoletti and Etro (2016) or Parenti et al. (2017). While these papers consider environments with non quasi-linear preferences, this work sticks to the quasi linear assumption but considers cost structures other than ones with constant marginal costs.

2 The model

Consider that there are $n \geq 1$ single-product and symmetric firms in the industry, indexed by $i = 1...n$. The market structure is assumed oligopolistic with a general mode of competition along the lines of Weyl and Fabinger (2013) and 2022 that generalize the approach of Genesove and Mullin (1998) and Bresnahan (1989). I also follow the partial equilibrium tradition in assuming that all goods outside the industry are perfectly competitively supplied. In addition, there are no externalities in the economy and information is complete.

Demand side. Even if goods may be distinct in the consumers' eye, the demand system is assumed to be fully symmetric. There are L identical consumers that share the same preferences and thus I abstract from distributional considerations. I denote the individual symmetric demand for variety i as $x_i = D_i(p_1, ..., p_n)$ for all i, where $D_i(.)$ is twice continuously differentiable with $\partial D_i/\partial p_i < 0$ and where p_j is the price of variety j for any j. The demand addressed to firm/variety i is then $q_i = Lx_i = LD_i(p_1, ..., p_n)$ for all i. I also denote the individual symmetric inverse demand for variety i as $p_i = P_i(x_1, ..., x_n)$ for all i, where $P_i(.)$ is twice continuously differentiable with $\partial P_i/\partial x_i < 0$. Furthermore, I consider demand systems that satisfy the following assumption.

Assumption 1. *For all i, the own-price or own-quantity effect strictly dominates the cross-price or cross-quantity effects:*

$$\sum_j \frac{\partial D_i}{\partial p_j} < 0 \text{ and } \sum_j \frac{\partial P_i}{\partial x_j} < 0.$$

Denote the industry production/total sales as $Q = \sum_i q_i$. Moreover, denote ε_d as the *elasticity of industry demand*.[2] To derive this elasticity, consider a symmetric equilibrium with price $p_i = p$, production per variety $q_i = q = Q/n$ and consumption per head and per variety as $x_i = x = Q/nL = q/L$ for all i. Define the industry demand as $Q = nLD(p, n)$ where $D(p, n) \equiv D_i(p, ..., p)$ for any i. Define also the inverse of $D(p, n)$ with respect to price as the industry inverse demand, i.e. $p = P(x, n)$ where $P(x, n) \equiv P_i(x, ..., x)$ for any i.

For the analysis to follow, it is convenient to define the elasticity of industry demand as a function of x by:

$$\varepsilon_d(x, n) = -\frac{P(x, n)}{xP_x(x, n)} > 0$$

where $P_x(x, n) = \sum_j \frac{\partial P_i}{\partial x_j} = \left(\sum_j \frac{\partial D_i}{\partial p_j}\right)^{-1} < 0$ is the slope of the industry inverse demand.[3] The fact that the industry inverse demand is downward sloping follows directly from the Assumption 1 made above on the demand system.

Furthermore, I consider the elasticity of the slope $P_x(x, n)$ as a measure of curvature for the industry inverse demand:

$$\rho_d(x, n) = -\frac{xP_{xx}(x, n)}{P_x(x, n)}$$

where $P_{xx}(x, n) = \sum_{j,k} \frac{\partial^2 P_i}{\partial x_j \partial x_k}$.

As is common in the literature (Seade, 1980a; Delipalla and Keen, 1992; and Kroft et al., 2021), I ignore the integer constraint on n and treat it as a continuous variable for simplicity. I also impose the following assumption on all demand systems considered.

Assumption 2. *Holding the consumption per variety x constant, adding new varieties (weakly) reduces the inverse demand for each variety, i.e. $P_n(x, n) \leq 0$. Holding price constant, adding new varieties (weakly) reduces the demand for each variety, i.e. $D_n(p, n) \leq 0$.*

This assumption holds for standard demand and inverse demand functions like the linear demand and inverse demand function, the demand system generated by additive preferences à la Dixit and Stiglitz (1977) or the Logit demand and inverse demand.[4]

---

[2] This is also referred to the price elasticity of market-wide demand in the literature.

[3] Writing the price elasticity of industry demand as a function of x stems from $P_x(x, n) = \left.\frac{1}{D_p(p,n)}\right|_{P(x,n)=p}$. By contrast, the elasticity of the industry inverse demand function $P(x, n)$ is $-xP_x(x, n)/P(x, n) = 1/\varepsilon_d(x, n)$. Starting with industry demand $D(p, n)$, the price elasticity of industry demand writes as $-pD_p(p, n)/D(p)$ and the elasticity of the industry inverse demand writes as a function of price according to $-D(p, n)/(pD_p(p, n))$.

[4] See Appendix A.

When dealing with welfare, I will consider that the demand system originates from individual preferences represented by a quasi-linear utility function $U^n(x_1, ...x_n) + z$ where z is the consumption of the outside good.[5] Considering a symmetric equilibrium, let us denote $U(x, n) \equiv U^n(x, ..., x)$ as the utility level for a consumption x per variety and with n varieties. Marginal utility w.r.t. quantity is $U_x(x, n) = \sum_i \frac{\partial U^n}{\partial x_i} = nP(x, n)$.

Also, the marginal utility w.r.t. the number of varieties is given by:

$$\begin{aligned} U_n(x,n) &= \int_0^x \left[P(s,n) + nP_n(s,n)\right] ds \\ &= \frac{U(x,n)}{n} + \int_0^x nP_n(s,n)ds. \end{aligned} \tag{1}$$

using $U_x(x, n) = nP(x, n)$. I assume that preferences are characterized by love-for-variety but at a decreasing rate, i.e. U is increasing concave in n. Observe from (1) that Assumption 2 is a sufficient condition for the latter property.

Quasi-linearity of preferences ensures that the consumers' surplus given by $LU^n - L\sum_i p_i x_i$ constitutes an adequate measure of consumers' welfare changes. Moreover, at a symmetric equilibrium, consumers' surplus writes:

$$CS(x, n) = LU(x, n) - LnxP(x, n). \tag{2}$$

Let us define

$$\mathcal{V}_0 \equiv \frac{LU_n - pq}{p} \tag{3}$$

as a unit-free measure of the "variety effect" which captures the effect of a marginal change in the number of varieties n on consumers' surplus *keeping price and consumption per variety fixed*, per unit of price.[6] It is also convenient to define, for further reference, the marginal consumers' surplus w.r.t. n (again per unit of price), denoted $\mathcal{V}_1 \equiv CS_n/p = (LU_n - pq - QP_n)/p > \mathcal{V}_0$ given Assumption 2. $\mathcal{V}_1$ captures the effect of a marginal change in n on consumers' surplus, per unit of price and holding only consumption per variety fixed.

Cost side. I assume that all firms face the same cost function that depends only on each own output q. More precisely, let us denote $c(q)$ the variable cost (VC) function for any output q with $c(0) = 0$. For the purpose of clarity and because the analysis to follow focus on variable profits at the equilibrium, let us assume that there are no fixed costs for the moment.[7]

[5]The subutility function $U^n(.)$ is assumed thrice continuously differentiable, symmetric and strictly concave.

[6]Note that Kroft et al. (2021) define the variety effect as $LU_n - pq$. Here it is defined per unit of price to obtain a unit-free measure useful in the analysis of regulated oligopolies in Section 6.

[7]Fixed costs will be reintroduced later when considering entry on the market.

Importantly, I also denote $w(q) \equiv c(q)/q$ as the average variable cost (AVC) function. In the following, I focus on all VC smooth functions that are (weakly) increasing in quantity, $c'(q) = w'(q)q + w(q) \geq 0$. I also allow for convex and concave cost functions. Indeed, concave costs represent important cases because, as suggested by Bykadorov et al. (2015), endogenous technology choice can yield situations where higher output fosters investment in marginal cost reduction which implies concave cost.[8] The empirical literature has also documented the possibility of concave costs for some manufacturing industries (Diewert and Wales, 1987; Friedlander *et al.*, 1983; and Ramey, 1991).

Note that the AVC function $w(.)$ could also be interpreted as a price per unit to be paid to a supplier in order to be able to deliver the good to final consumers. This is particularly meaningful when the firm is a pure merchant that acts as an intermediary between some buyers and some sellers, as analyzed by e.g. Stahl (1988) and Hamilton et al. (2015). This merchant may benefit from exploiting market power not only downstream towards buyers but also upstream where $w'(q) > 0$ would correspond to an upwards sloping inverse supply from sellers.

Keeping this analogy with the merchant's problem, let us interpret the AVC function $w(.)$ *as if* it were an "inverse supply" so that the elasticity of "supply" is indexed with s and defined by:[9]

$$\varepsilon_s(q) = \frac{w(q)}{qw'(q)}$$

Importantly, our assumption of increasing cost does not constrain the sign of w'. More precisely, the elasticity ε_s is related to the elasticity ε_c of VC according to $\varepsilon_c(q) = \frac{qc'(q)}{c(q)} = \frac{\varepsilon_s(q)+1}{\varepsilon_s(q)}$. An increasing cost ($\varepsilon_c \geq 0$) translates into either $\varepsilon_s(q) \geq 0$ or $\varepsilon_s(q) \leq -1$. In the limit case where the marginal cost is constant ($c'(q) = c > 0$ and thus $\varepsilon_c = 1$) then $w'(q) = 0$ and consequently $\varepsilon_s = \infty$. Let us denote this situation as *Constant AVC* (hereafter CAVC).

Consider next the case where the firm's equilibrium lies in a point where $\varepsilon_s(q) \geq 0$. This means that locally the AVC function is increasing ($w'(q) \geq 0$) or equivalently that VC is convex. Clearly, the convexity of VC is a necessary but non sufficient condition to the presence of decreasing returns to scale as long as there are fixed costs of production.

Now consider the opposite situation where at the equilibrium $\varepsilon_s(q) \leq -1$. This means that locally the AVC function is decreasing ($w'(q) \leq 0$) or equivalently that VC is concave. Note that the concavity

[8] Consider the example provided by Bykadorov et al. (2015), where $c(q) \equiv \min_e e + \tilde{c}(e)q$ and $e \geq 0$ is a (continuous) technology choice that decreases $\tilde{c}(e)$ but a decreasing rate ($\tilde{c}''(e) > 0$). The FOC w.r.t e is $\tilde{c}'(e) = -1/q$ which is sufficient thanks to convexity of $\tilde{c}$ and it defines the optimal tech choice $e(q)$. Totally differentiating the FOC indicates that more production fosters investment, i.e. $e'(q) > 0$. It follows that the resulting production cost $c(q)$ following the endogenous technology choice is such that $c'(q) = \tilde{c}(e(q))$ and $c''(q) = \tilde{c}'(e(q))e'(q) < 0$. Hence, $c(q)$ is concave.

[9] In the merchant case, the "supply" is a function $q(w)$ and it describes the quantity that the merchant's suppliers are ready to supply given the wholesale price w. The function $w(q)$ is the corresponding inverse supply.

of VC implies that the average cost, that would include any fixed cost, is locally decreasing and thus it is a sufficient but not necessary condition to the presence of increasing returns to scale (IRS). While imperfect competition is often justified by the presence of IRS, this feature of the cost structure is compatible with either convex or concave VC.

Finally, I denote $\rho_s(q)$ the following measure of the curvature of the supply function (i.e. the elasticity of $w'(q)$):

$$\rho_s(q) = -\frac{qw''(q)}{w'(q)}$$

Both ε_s and ρ_s vary with output unless the AVC function is iso-elastic (or CES) and writes as follows:[10]

$$w(q) = \beta q^{1/\varepsilon_s} \Rightarrow \rho_s = \frac{\varepsilon_s - 1}{\varepsilon_s} \leq 2.$$

The special case of Cobb-Douglas is obtained when $\varepsilon_s = -1$ and $\rho_s = 2$. Finally, note that ε_s and ρ_s can be related to their counterparts for the VC function $c(.)$. Indeed, it is immediate to check that $\varepsilon_c = qc'/c = (\varepsilon_s + 1)/\varepsilon_s > 0$ and $\rho_c = -qc''/c' = -(2-\rho_s)/(1+\varepsilon_s)$. In particular, the curvature measure ρ_c of the cost function c is negative if and only if c is convex.

Mode of competition. In the analysis below, I follow the general approach developed by Weyl and Fabinger (2013) (see also Adachi and Fabinger (2022)) by not specifying a particular mode of interaction between firms.

To proceed, it is useful to construct an index of competitive intensity centered on the focal point of overall industry profit maximization. Let us write the profit for any industry's member i as $\pi_i = (p_i - w_i(q_i))\, q_i$, keeping for the moment the possibility of asymmetry in costs and in demand system, and consider the impact of a change $d\sigma_i$ in the strategic variable σ_i of firm i on industry's aggregate profit, which can be decomposed as the sum of a pecuniary effect and a real effect as follows:

$$\frac{d}{d\sigma_i}\left(\sum_j \pi_j\right) = \underbrace{\sum_j q_j \frac{dp_j}{d\sigma_i}}_{\text{pecuniary effect}} + \underbrace{\sum_j (p_j - w_j - q_j w_j')\frac{dq_j}{d\sigma_i}}_{\text{real effect}} \tag{4}$$

denoting $w_j = w(q_j)$. Let us define θ_i as the ratio between the real effect and the pecuniary effect:

$$\theta_i = \frac{\sum_j (p_j - w_j - q_j w_j')\frac{dq_j}{d\sigma_i}}{-\sum_j q_j \frac{dp_j}{d\sigma_i}} \tag{5}$$

[10]The proof that the elasticity ε_s and curvature measure ρ_s are constant if and only if $w(.)$ is iso-elastic is as follows.The fact that iso-elasticity of w is sufficient for constant elasticity is obvious and necessity comes from setting $\frac{w(q)}{qw'(q)}$ equal to a constant ε_s and integrating.

which allows to rewrite (4) as follows:

$$\frac{d}{d\sigma_i}\left(\sum_j \pi_j\right) = (1-\theta_i)\sum_j q_j \frac{dp_j}{d\sigma_i}$$

Observe that if the strategic choice σ_i maximizes industry's profit then this change $d\sigma_i$ in strategy would be perfectly internalized so that $\theta_i = 1$ in that case. Besides this focal point, θ_i diverges in general from unity and it can be lower or greater than 1 as will be clearer below.

When the strategic choice maximizes firm i's profit for any i, then using the corresponding first-order condition, we can obtain an expression of the mark-up $m_i = p_i - w_i - q_i w_i'$. Indeed

$$\frac{d\pi_i}{d\sigma_i} = 0 \Leftrightarrow m_i = -q_i \frac{dp_i}{d\sigma_i} / \frac{dq_i}{d\sigma_i}.$$

Substituting in (5) and rearranging, one can rewrite θ_i as follows:

$$\theta_i = \sum_j \left(\frac{q_j \frac{dp_j}{d\sigma_j}}{\sum_j q_j \frac{dp_j}{d\sigma_i}} \right) \frac{\frac{dq_j}{d\sigma_i}}{\frac{dq_j}{d\sigma_j}}. \tag{6}$$

When considering a symmetric equilibrium with $p_i = p$, $q_i = q = Q/n$ and $x_i = x = q/L$ as well as $w_i(q_i) = w(q)$ for all i, let us denote $\theta \equiv \theta_i$ for any i as the market conduct index that captures the intensity of competition in the game for profit-maximising firms. A more concise expression for θ can be obtained by denoting the own price derivative $\frac{dp_j}{d\sigma_j} = p_\sigma^o$ for all j and the cross price derivative $\frac{dp_j}{d\sigma_i} = p_\sigma^c$ for all $i \neq j$ and using similar notations for the own and cross quantity derivative, i.e. $\frac{dq_j}{d\sigma_j} = q_\sigma^o$ and $\frac{dq_j}{d\sigma_i} = q_\sigma^c$. Then (6) becomes under symmetry:

$$\theta = \frac{1 + (n-1)q_\sigma^c / q_\sigma^o}{1 + (n-1)p_\sigma^c / p_\sigma^o}. \tag{7}$$

In general, θ depends on q (or equivalently on x) except in some particular cases as shown below. It also depends on n directly. However, *it does not depend directly on market size L or on the cost side*, but only indirectly through the quantities/prices levels. This would prove important when dealing with market expansion and regulation incidence (see Sections 5.3 and 6 respectively).

Modelling the nature of competition using the market conduct index θ allows to consider simultaneously a wide range of competition models (Weyl and Fabinger, 2013; Adachi and Fabinger, 2022). This approach allows to describe oligopoly market conduct in a very concise way, without specifying precisely the mode and the nature of competition (e.g. whether strategic variables are substitutes or complements). Indeed, from (5), a symmetric equilibrium is characterized by an elasticity-adjusted Lerner index $\frac{m}{p}\varepsilon_d$ that is equal to θ for all i, which implies that all the first-order conditions at a symmetric equilibrium reduce to the following unique condition:

$$MP \equiv P(x,n) + \theta x P_x(x,n) - w(q) - qw'(q) = 0 \tag{8}$$

such that the *perceived* marginal profit with respect to output, and denoted MP, equals zero.

With substitute varieties which we assume in the following, the market conduct index θ belongs to $(0,1)$.[11] A value for θ close to zero means that competition is very intense and thus the equilibrium price is close to marginal cost. Conversely, when θ is close to 1, then the industry almost behaves like a colluding one or a monopoly. Hence, θ measures the degree of market monopolization. Similarly, $\theta = 1$ when products are independent.

Consider the case of a homogenous product oligopoly. In that case, when a firm i chooses its quantity, it assumes that a change dq_i will induce a change $dQ = \nu dq_i$ in aggregate output (or equivalently that each other firm j will change its quantity by $dq_j = \frac{\nu-1}{n-1} dq_i$ in response). Cournot competition corresponds to $\nu = 1$, Bertrand competition to $\nu = 0$ and perfect collusion to $\nu = n$. More generally, when ν is a constant that belongs to $[0,n]$ (as in Delipalla and Keen (1992) conjectural variation model also denoted "Generalized Cournot model"), then (7) gives $\theta = \nu/n$.[12]

Now, consider a symmetrically differentiated product oligopoly where firms compete in quantities. The inverse demand for a variety i is $p_i = P_i(x_i, x_{-i})$ and from (7), one obtains:[13]

$$\theta = \frac{\partial P_i}{\partial x_i} \Big/ \sum_j \frac{\partial P_j}{\partial x_i} > 0. \tag{9}$$

If the inverse demand system is linear then θ is constant.

Finally, for a symmetrically differentiated product oligopoly where firms compete in prices, the demand for variety i is $q_i = LD_i(p_1, ..., p_n)$ and from (6), one gets:

$$\theta = \sum_j \frac{\partial D_j}{\partial p_i} \Big/ \frac{\partial D_i}{\partial p_i} = 1 - A > 0 \tag{10}$$

where $A = \sum_{j \neq i} d_{ij}$ where $d_{ij} = -\frac{\partial D_j}{\partial p_i} / \frac{\partial D_i}{\partial p_i} \geq 0$ is the price diversion ratio from i to j. A is the aggregate diversion ratio from any individual firm to the rest of the industry, and it represents the fraction of sales lost by a firm when it increases its price and that is captured by the competitors (Shapiro, 1995). If the demand system is linear then A and thus θ are constant.[14]

[11] With complement varieties, $\theta > 1$ (Weyl and Fabinger, 2013).

[12] Taking $\sigma_i = q_i$ for all i, we have $q^o_\sigma = 1$ and $q^c_\sigma = (\nu-1)/(n-1)$. Also, homogeneity implies $p^o_\sigma = p^c_\sigma$. Hence, replacing in (7) leads to $\theta = \nu/n$.

[13] Taking $\sigma_i = q_i$ for all i, we have $q^c_\sigma = 0$ (Cournot behavior) and $q^o_\sigma = 1$. Also $p^o_\sigma = (1/L)\partial P_i/\partial x_i$ and $p^c_\sigma = (1/L)\partial P_j/\partial x_i$. Replacing in (7) leads to (9).

[14] The market conduct index approach can also accommodate the case of monopolistic competition (see Weyl and Fabinger, 2013) and the case of competition in supply functions à la Klemperer and Meyer (1989) (see Mahoney and Weyl, 2017).

As noted earlier, θ is a function of q and n in general. For further reference in the analysis, I denote $\varepsilon_{\theta,q} \equiv \partial \log \theta / \partial \log q$. I define similarly $\varepsilon_{\theta,n}$. Both measures reflect the sensitivity of the market conduct index w.r.t respectively q and n. For instance, in the Generalized Cournot model, $\theta = \nu/n$ and thus $\varepsilon_{\theta,q} = 0$ and $\varepsilon_{\theta,n} = -1$.

3 A contribution margin approach to the oligopoly equilibrium

Analysis. Bontems (2023a) introduces a contribution margin approach to the monopoly's problem by showing how the firm's optimum can be described using only the elasticities of demand and "supply" (ε_d and ε_s), as well as their curvature measures (ρ_d and ρ_s). In this section, I extend this approach to consider the case of a symmetric oligopoly. To proceed, I first focus on a *profitability ratio* built on the comparison between price and AVC for each firm.

Definition 1. *The Contribution Margin Index, or in short the $\mathcal{C}$-index, is defined at the firm's level by:*

$$\mathcal{C}_i = \frac{p_i - w_i}{p_i}.$$

Obviously, at the symmetric equilibrium, $\mathcal{C}_i = \mathcal{C} \equiv (P(x,n) - w(q))/P(x,n)$ for all i. The $\mathcal{C}$-index represents the portion of total sales revenues not used to cover the aggregate variable cost of the industry and thus that contributes to covering the industry's aggregate fixed cost. Like the Lerner index ($\mathcal{L} = (p - c')/p$), the $\mathcal{C}$-index lies in the range (0,1) at the equilibrium. At the firm's level, we have $\pi = \mathcal{C}r$ where $r = pq$ is the revenue and π the variable profit. And at the industry's level, we have $\Pi = \mathcal{C}R$ where $R = pQ$ is the industry's revenue and Π the industry's variable profit. Hence, while the Lerner index is a usual measure of market power, the $\mathcal{C}$-index is a measure of profitability in terms of sales revenues, both at the firm's and industry's levels.

Second, let us introduce the useful notion of *effective* (industry) demand elasticity.

Definition 2. *The effective elasticity of industry demand is defined by:*

$$\tilde{\varepsilon}_d = \varepsilon_d / \theta$$

that is the industry demand elasticity deflated by the intensity of competition as measured by θ.

The following Proposition indicates how the profit-maximizing oligopoly symmetric equilibrium can be described using the $\mathcal{C}$-index, instead of using the usual Lerner index.

Proposition 1. *At a symmetric equilibrium under imperfect competition, the $\mathcal{C}$-index obeys to an inverse pseudo-elasticity rule:*

$$\mathcal{C} = \frac{p - w}{p} = \frac{1}{\varepsilon}$$

where the pseudo-elasticity ε is proportional to the effective demand elasticity:

$$\varepsilon \equiv \lambda \tilde{\varepsilon}_d \qquad \text{with} \qquad \lambda = \frac{\varepsilon_s + 1}{\varepsilon_s + \tilde{\varepsilon}_d},$$

and $\varepsilon \geq 1$ as well as $\tilde{\varepsilon}_d \geq 1$ are required at the equilibrium.

Proof. To obtain the result, rewrite (8) as $p(1 - 1/\tilde{\varepsilon}_d) = w(1 + 1/\varepsilon_s)$ by using the definition of the effective elasticity of industry demand and of the elasticity of "supply". Then, rearranging to form $\mathcal{C}$ yields the desired formulation. An alternative formulation of the $\mathcal{C}$-index can be derived by relying on revenue and cost, i.e. $\mathcal{C} = (r(q) - c(q))/r(q) = 1/\varepsilon$ where ε (and thus λ) can be written as function of the elasticities ε_c, ε_r and θ. See Appendix B for details. ■

This extends the result obtained by Bontems (2023a) in the case of monopoly, to the oligopoly framework. Indeed, in the monopoly case ($n = 1$ and thus $\theta = 1$), then the effective demand elasticity is simply the monopolist's demand elasticity and the $\mathcal{C}$-index is inversely proportional to a pseudo-elasticity ε which represents the demand elasticity factorized by the profitability scale factor λ. This profitability scale factor λ is itself a function of both the demand and "supply" elasticities. More generally, in the oligopoly case ($n \geq 2$), the $\mathcal{C}$-index can be expressed in a similar way, except that the demand elasticity should be replaced with the effective demand elasticity $\tilde{\varepsilon}_d$, thus reflecting the various possibilities in terms of competition intensity in the range of oligopolies models permitted by the framework.

The equilibrium restriction on the pseudo-elasticity ε also requires that λ should be strictly positive. Furthermore, excluding the Bertrand/perfect competition case ($\theta = 0$) for the moment, the following Proposition indicates that the properties of the profitability scale factor λ are intimately related to whether the firms operate under CAVC, convex or concave VC at the equilibrium.

Proposition 2. *Assume $\theta > 0$. If at the equilibrium the variable cost is convex ($\varepsilon_s > 0$) then $\lambda \in (0, 1)$. If on the contrary the variable cost is concave ($\varepsilon_s < -1$) then $\lambda > 1$. Last, under CAVC, then $\lambda = 1$ for any equilibrium output.*

Proof. Note first that when the marginal cost is constant (CAVC), i.e. $\varepsilon_s = \infty$, then λ is equal to 1 for any output level. Now assume that the VC is convex at the equilibrium, then the restriction on ε

implies that:

$$\varepsilon \geq 1 \Leftrightarrow \frac{\varepsilon_s + 1}{\varepsilon_s + \tilde{\varepsilon}_d}\tilde{\varepsilon}_d \geq 1$$

which is equivalent to $\tilde{\varepsilon}_d \geq 1$ or equivalently $\varepsilon_d \geq \theta$ as $\varepsilon_s > 0$. Then we have

$$\lambda - 1 = \frac{\varepsilon_s + 1}{\varepsilon_s + \tilde{\varepsilon}_d} - 1 = \frac{1 - \tilde{\varepsilon}_d}{\varepsilon_s + \tilde{\varepsilon}_d} < 0.$$

As λ is also positive, we thus have $\lambda \in (0, 1)$ for the convex VC regime. Similarly, under concavity of VC at the equilibrium ($\varepsilon_s \leq -1$), $\varepsilon \geq 1 \Leftrightarrow \tilde{\varepsilon}_d \geq 1$ and thus $\lambda > 1$. ■

As shown in Proposition 2, the profitability scale factor λ is constant precisely under CAVC. The non constancy of marginal cost on the contrary opens up the possibility for λ to depend on equilibrium output.

To complete the Proposition, note that under Bertrand or perfect competition, $\theta \to 0$ and then $\tilde{\varepsilon}_d \to \infty$ (if ε_d is finite) and $\lambda \to 0$ (if ε_s is finite). One gets $\lim_{\theta \to 0} \frac{\varepsilon_s + 1}{\varepsilon_s + \tilde{\varepsilon}_d}\tilde{\varepsilon}_d = \varepsilon_s + 1$ and thus $\lim_{\theta \to 0} \mathcal{C} = \frac{p-w}{p} = \left.\frac{1}{\varepsilon_s + 1}\right|_{\theta=0}$. Hence, under Bertrand or perfect competition, clearly only the regimes CAVC or convex VC are possible as $\varepsilon_s \geq 0$ is needed. If in addition marginal cost is constant, then $\varepsilon_s = \infty$ and consequently $P = w$.

Interpreting the profitability scale factor λ. Proposition 2 implies a straightforward relationship between the Lerner index $\mathcal{L} = (p - c')/p$ and the $\mathcal{C}$-index at the symmetric equilibrium:

$$\mathcal{L} = \lambda \mathcal{C}.$$

In other words, under CAVC ($\lambda = 1$) the two indexes confound and accordingly, measuring profitability is equivalent to measure the degree of market power. This no longer holds when marginal cost is non constant. Indeed, the Lerner index still estimates the degree of market power but underestimates the degree of profitability under convex VC ($\lambda < 1$) while it overestimates it under concave VC ($\lambda > 1$). A higher λ means a larger discrepancy between market power and profitability.

It is also interesting to compare the industry with n firms under scrutiny with non constant marginal cost with its counterpart for which marginal cost would be constant but equal to the *equilibrium value of* w of the original industry. This *ghost* industry faces the same demand, has the same number of members and the same mode of competition, but differs in that its profitability scale factor λ is 1 as marginal cost is constant, and hence its Lerner index and its $\mathcal{C}$-index are identical. More precisely, the original industry with non constant marginal cost produces a quantity x^* given by

AVC.pdf

Figure 1: Equilibrium output for the industry and its ghost. MR_i is the perceived marginal revenue curve for $i = 1, 2$, MC is the marginal cost curve, AVC is the average variable cost curve. Black dots indicate output equilibria for an industry with flexible marginal cost while white dots indicate output equilibria for the corresponding ghost industry that has constant marginal cost equal to the equilibrium value of AVC of the original industry.

(8) and that corresponds to a black dot in Figure 1 at the intersect between the perceived marginal revenue curve, $P(x,n)+\theta x P_x(x,n)$, and the marginal cost curve $w(q)+qw'(q)$.[15] The two black dots illustrates two possible different equilibria, one under convex VC and the other one under concave VC. By contrast, the ghost industry has a constant marginal cost equal precisely to $w(q^*)$ and produces a quantity that corresponds to a white dot in Figure 1, at the intersect between the perceived marginal revenue curve and the value $w(q^*)$. Clearly, at an equilibrium under convex VC, the industry with non constant marginal cost produces less and has a larger profitability ratio than its ghost industry, because it benefits from a larger price while having the same AVC. Conversely, at an equilibrium under concave VC, the industry with non constant marginal cost produces more than the ghost industry and has a lower profitability ratio.

Intuitively, under convex VC, $\lambda \in (0,1)$ and it represents a profitability scale factor that partially absorbs the importance of the effective elasticity of demand in the calculation of profitability. In other words, holding the effective demand elasticity constant, the industry benefits from the non constancy of marginal cost (compared to its ghost) as the presence of convex VC constitutes an additional motive for a representative firm to reduce production because this allows to reach a lower average variable cost. Under concave VC, $\lambda > 1$ and the profitability of the industry with non constant marginal cost is reduced compared to the ghost industry because of the presence of countervailing incentives: on the one hand, each firm is willing to increase output to benefit from a lower AVC and thus from the presence of increasing return to scale, but on the other hand, each firm would like to reduce output to better extract consumer surplus.

Note that another type of ghost industry can be defined and will prove useful in the following analysis. Indeed, consider the ghost industry defined by a constant marginal cost equal to the *equilibrium marginal cost* of the original industry. Both industries produce the same output, but the ghost industry benefits (suffers) from a lower (higher) AVC under concave VC (convex VC) and hence earns more (less) in terms of profit.

Existence, unicity and stability. Like Weyl and Fabinger (2013), Adachi and Fabinger (2022) and Kroft et al. (2021), I assume that the conditions for existence and uniqueness of the interior symmetric equilibrium are satisfied. In particular, following Seade (1980b), I assume that the stability condition holds in the sense that the perceived marginal profit at the symmetric equilibrium is decreasing in q.

[15]To illustrate the different equilibrium regimes, the Figure 1 considers a classic specification with a U-shaped AVC function (based on a bipower function which in turn determines the marginal cost curve).

More precisely, given (8), the stability condition writes as follows:

$$MP_q \equiv \frac{1}{L}(1+\theta+q\theta_q)P_x + \frac{1}{L}\theta x P_{xx} - \frac{d}{dq}\left[w(q)+qw'(q)\right] < 0$$

where $MP_q = \partial MP/\partial q$ and $\theta_q = \partial\theta/\partial q$ are the partial derivatives of respectively MP and θ w.r.t q.

For further reference in the analysis to follow, it is convenient here to consider an *average measure of curvature* for both demand and "supply" as follows.

Definition 3. *The "average curvature" measure $\rho \equiv \lambda\rho_d + (1-\lambda)\rho_s$ is the average of curvature measures for the demand and the "supply" with weight λ and $1-\lambda$ respectively.*

When marginal cost is constant ($\lambda = 1$), then observe that ρ is simply the curvature of demand ρ_d. It is well known that under CAVC the stability condition puts an upper bound on ρ_d whose value depends on the competition model studied, as will be clear below. When marginal cost is not constant, the new insight brought by our analysis is that the stability condition amounts to put an upper bound on the "average curvature" ρ defined above. Indeed, straightforward manipulations allow to rewrite MP_q as follows:[16]

$$MP_q = \frac{\theta P_x}{\lambda L}\left[\lambda\left(\frac{1}{\theta} - 1 + \varepsilon_{\theta,q}\right) + 2 - \rho\right] \tag{11}$$

Throughout the paper, I thus assume that demand and cost functions satisfy the following condition globally.

Assumption 3 (Stability condition). *$MP_q < 0$ or equivalently,*

$$\rho < 2 + \lambda\left(\frac{1}{\theta} - 1 + \varepsilon_{\theta,q}\right). \tag{12}$$

This general stability condition includes many specific conditions corresponding to particular cases. Observe that the term between brackets on the right hand side of the inequality depends only on θ and on $\varepsilon_{\theta,q}$, i.e. it depends on the characteristics of the mode of competition.

First, consider the case of a monopoly ($n = 1$ and thus $\theta = 1$). In that case, (12) reduces to $\rho < 2$, which represents the second-order condition for the monopolist's profit maximization problem exhibited by Bontems (2023a).[17] If we further assume CAVC ($\lambda = 1$), then (12) amounts to the well known second-order condition $\rho_d < 2$.

[16] See Appendix C.

[17] The stability condition is similar in the case of a perfectly colluding industry with n members.

Now consider a n symmetric Cournot oligopoly ($\theta = 1/n$) with CAVC ($\lambda = 1$) and observe that (12) reduces to the condition exhibited by Seade (1980b): $\rho_d < n + 1$.[18] When marginal cost is non constant, (12) provides a direct generalization of Seade (1980b): $\rho < 2 + \lambda(n-1)$. Finally, consider a differentiated product oligopoly where firms compete in price and where demands are linear, then $\theta = 1 - A$ and is constant w.r.t q, and (12) amounts to $\rho < 2 + \lambda \frac{A}{1-A}$.

Stability in the long run. In the long run, the number of firms is endogenous and is driven by free exit/entry on the market. To enter the market, a fixed cost f has to be spent by each entrant and the zero profit condition writes:

$$(P(x,n) - w(q))q - f = 0 \tag{13}$$

Condition (13) as well as the behavioral response of symmetric firms given by (8) allow to determine the long run equilibrium in terms of output per firm and number of varieties. At a long run equilibrium, the $\mathcal{C}$-index satisfies $\mathcal{C}r = f$ or equivalently $r = \varepsilon f$. Also, it is clear that condition (13) implies that only the regimes CAVC or convex VC are possible in the long run. Indeed the long run equilibrium lies at the intersect of AC and MC if any and AVC crosses MC before AC.

As usual, the stability condition in the long run amounts to have the following matrix M as being definite negative:

$$M = \begin{pmatrix} MP_q & MP_n \\ (1-\theta)xP_x & qP_n \end{pmatrix} \tag{14}$$

where MP_n denotes the marginal impact of n on the perceived marginal profit holding x constant:

$$MP_n = \frac{\partial}{\partial n}\left(P + \theta x P_x - w - qw'\right) = P_n + \theta x P_{xn} + \theta_n x P_x. \tag{15}$$

Note that, although $P_n \leq 0$ under Assumption 2 and $\theta_n < 0$, the sign of MP_n remains ambiguous at this level of generality.[19] Nevertheless, the stability of the long run equilibrium imposes that MP_n cannot be too negative, as seen below.

Assumption 4 (Long run stability conditions)**.**

$$MP_q + qP_n < 0 \textit{ and } MP_q qP_n - (1-\theta)xP_x MP_n > 0. \tag{16}$$

[18]Note that the formulation of the stability condition in Seade (1980b) is $(1+1/n)\hat{P}' + x\hat{P}'' < 0$ where the inverse demand is $\hat{P}(nx)$. With our notation, $P(x,n) = \hat{P}(nx)$ and thus $P_x = n\hat{P}'$ and $P_{xx} = n^2\hat{P}''$. Then the condition of Seade is equivalent to $(1+1/n)P_x + (1/n)xP_{xx} < 0$, which is equivalent to the condition in the text, using $P_x < 0$ and the definition of ρ_d.

[19]The sign of MP_n and its interpretation will be discussed later in Section and will be shown to be related to whether business stealing or expansion occurs at the equilibrium.

In particular, for the perfectly colluding oligopoly ($\theta = 1$), Assumption 2, i.e. $P_n < 0$, is necessary for the long run stability conditions (16) to hold.

4 Reformulating cost pass-through

We examine in this section the incidence in terms on price, quantity and profit of some cost shock in the short run.[20] This allows to show that the expression of absolute and relative cost pass-through can be solely determined through the values of the elasticity λ and the "average" convexity measure ρ as well as the intensity of competition θ and its sensitivity to output, $\varepsilon_{\theta,q}$. For the relative cost pass-through, the pseudo elasticity ε is also involved. The new expressions obtained are compared with those of the literature (in particular to Weyl and Fabinger, 2013). As will be shown in Sections 5.1 and 5.3, these cost pass-through measures are key elements to explain how market concentration and market expansion affect market outcomes.

Cost pass-through in the short run. Let us start with defining the absolute and relative cost pass-through in our setting. The absolute pass-through (hereafter *APT*) corresponds to the absolute impact on price of an infinitesimal additive cost shock t (e.g. a specific tax/subsidy). For this, let us write the profit for any firm i as $\pi_i = (p_i - w(q_i) - t)q_i$ and at the symmetric equilibrium, $APT = \left.\frac{dp}{dt}\right|_{t=0}$. Similarly, the relative pass-through (hereafter *RPT*) corresponds to the relative impact on price of a multiplicative cost shock τ (e.g. an ad-valorem tax or subsidy on cost) valued in $\tau = 1$. In this case, profit writes $\pi_i = (p_i - \tau w(q_i))q_i$ and $RPT = \left.\frac{d\log p}{d\log \tau}\right|_{\tau=1}$.[21]

Proposition 3. *The absolute and relative pass-through are positive and given respectively by:*

$$APT = \frac{\lambda}{\theta(2-\rho) + \lambda\left(1 - \theta + \theta\varepsilon_{\theta,q}\right)} \tag{17}$$

and

$$RPT = \frac{\varepsilon - \lambda}{\varepsilon} APT. \tag{18}$$

Proof. See Appendix D. ■

The stability condition (12) ensures that *APT* and *RPT* are positive. Note also that the relationship (18) also writes $RPT = (1-\mathcal{L})APT$ or equivalently $\mathcal{L} = 1 - RPT/APT$. This formulation is equivalent to

[20] Long-run consequences will be examined later in Section 5.2.

[21] Because it is taken in $\tau = 1$, RPT is also equivalent to $d\log P/d\tau$, that is the semi-elasticity of P w.r.t τ.

the one proposed by Adachi and Fabinger (2022, Proposition 3).[22] Because the Lerner index belongs to $(0, 1)$ under imperfect competition, then clearly $RPT < APT$. It follows that the multiplicative cost shock is less likely to be overshifted, i.e. $RPT > 1$, than the additive one, i.e. $APT > 1$.[23] Also under perfect competition, we get the well-known result that the pass-through of both shocks are the same.

The novelty here is the formulation of APT which is based solely on λ and ρ as well as θ and $\varepsilon_{\theta,q}$. To better understand the role of the different components of APT, it is instructive to consider some specific cases of the general formula (17).

- Consider first the case of a monopoly ($n = 1$), the case of a perfectly colluding n-industry and the case of independent products. In all of these cases, $\theta = 1$ and thus $\varepsilon_{\theta,q} = 0$, so that (17) reduces to

  $$APT|_{\theta=1} = \frac{\lambda}{2-\rho}. \tag{19}$$

  This is the APT formulation obtained by Bontems (2023a) in his study of monopolistic competition and it is a straightforward extension of the well-known formulation one gets when marginal cost is constant. Indeed, under CAVC ($\lambda = 1$), $APT = 1/(2 - \rho_d)$ which is the classic APT formulation obtained by Bulow and Pfleiderer (1983), and there is overshifting of the additive cost shock if and only if demand is strictly log-convex at the equilibrium, i.e. $\rho_d > 1$ (Seade, 1985; Stern, 1987).[24]

  When marginal cost is not constant, (19) reveals that there is undershifting of the cost shock if and only if $\rho < 2 - \lambda$. Conversely, there is overshifting if and only if $2 - \lambda < \rho < 2$.

  Finally, we obtain $RPT = \frac{\varepsilon-\lambda}{\varepsilon}\frac{\lambda}{2-\rho}$ which is a generalization of the formulation $\frac{\varepsilon_d-1}{\varepsilon_d(2-\rho_d)}$ proposed by Mrázová and Neary (2017) in the particular case of CAVC.

- Consider now the oligopoly case but where θ is constant w.r.t. output (Generalized Cournot model or Nash in price/quantity with differentiated products and linear direct or inverse demands). Then $\varepsilon_{\theta,q} = 0$ and the cost pass-through simplifies to:

  $$APT|_{\theta=\text{cst}} = \frac{\lambda}{\theta(2-\rho) + \lambda(1-\theta)}$$

[22] Actually, Proposition 3 in Adachi and Fabinger (2022) is more general in that it shows that this relationship between the pass-through of a specific tax and the pass-through of an ad-valorem tax, measured by the tax pass-through semi-elasticity $d \log P/d\tau$, holds when both taxes coexist and are non zero and non unitary respectively. See Section 6 where I consider the contribution margin approach to oligopolies in regulation contexts.

[23] This result extends to differentiated products the result obtained by Delipalla and Keen (1992) in the context of the Generalized Cournot model.

[24] As argued by Delipalla and Keen (1992), this condition can even be traced back to Cournot (1960).

Observe that there is overshifting of the cost shock if and only if $APT|_{\theta=\text{cst}} > 1$ which yields the same condition on ρ as for the monopoly above: $\rho > 2 - \lambda$. This condition generalizes the condition $\rho_d > 1$ obtained by e.g. Seade (1985) and Stern (1987) in oligopoly contexts under CAVC.

In the general case, if θ increases in q then lower quantities/higher prices create more competitive market conduct and pass-through tends to be smaller (compared to the case with constant θ). And conversely.

- Finally, consider now the case of price competition with homogenous product (Bertrand), then $\theta = 0$ and thus $\lambda \to 0$. Also,

$$APT|_{\theta=0} = RPT|_{\theta=0} = \frac{1}{1 + \frac{\varepsilon_d}{\varepsilon_s+1}(2-\rho_s)} = \frac{1}{1+\frac{\varepsilon_d}{\varepsilon_{c'}}} \in (0,1)$$

where $\varepsilon_{c'} = c'/qc'' = (1+\varepsilon_s)/(2-\rho_s)$ is the inverse of the elasticity of marginal cost or equivalently the elasticity of the supply function defined by the marginal cost curve. Hence the pass-through decreases in the ratio of the elasticity of demand to that of supply.[25]

Importantly, note that, in any case, (17) provides a consistent formulation but alternative to the one proposed by Weyl and Fabinger (2013):

$$APT = \frac{1}{1 + \frac{\theta}{\epsilon_\theta} + \frac{\varepsilon_d - \theta}{\varepsilon_{c'}} + \frac{\theta}{\varepsilon_{ms}}}$$

which relies on the elasticity of inverse marginal surplus $ms = -xP_x$, i.e. $\varepsilon_{ms} = \frac{xP_x}{x(P_x + xP_{xx})}$, the elasticity of supply, i.e. $\varepsilon_{c'}$ and the sensitivity of θ to output as measured by Weyl and Fabinger through $\epsilon_\theta = 1/\varepsilon_{\theta,q}$.[26]

Another interesting result relates to the impact of cost curvature on the cost pass-through. For this, consider the ghost industry where each firm has a constant marginal cost equal to the equilibrium marginal cost of the original industry and where the mode of competition is the same. As argued above, the two industries produce the same output. Nevertheless, the impacts of additive and multiplicative cost shocks differ in the two industries depending on the convexity or concavity of variable cost. The

[25]The proof can be established directly like Weyl and Fabinger by considering perfect competition and the equality between demand and supply under the tax intervention (see Weyl and Fabinger (2013), Principle of Incidence under perfect competition) or indirectly from our general formulation (17) when $\theta \to 0$ and using L'Hospital rule (see Appendix D for details).

[26]To see the equivalence, note that $\varepsilon_{ms} = \frac{xP_x}{x(P_x+xP_{xx})} = \frac{1}{1-\rho_d}$ with our notations. Moreover, note that $\varepsilon_{c'} = \frac{1+\varepsilon_s}{2-\rho_s}$ and note that $\frac{\tilde{\varepsilon}_d - 1}{\varepsilon_s+1} = \frac{1-\lambda}{\lambda}$. We thus have that $1 + \frac{\theta}{\epsilon_\theta} + \frac{\varepsilon_d-\theta}{\varepsilon_{c'}} + \frac{\theta}{\varepsilon_{ms}} = 1 + \theta\varepsilon_{\theta,q} + \theta\frac{\varepsilon-1}{1+\varepsilon_s}(2-\rho_s) + \theta(1-\rho_d)$ and this further simplifies into $1 - \theta + \theta\varepsilon_{\theta,q} + \theta\frac{(2-\rho)}{\lambda}$.

following Proposition extends to oligopoly models a result obtained for the monopoly by Bontems (2023a).

Proposition 4. *Whatever the mode of competition and whether the cost shock is additive or multiplicative, the non constancy of marginal cost entails reduced (increased) cost pass-through on price or quantity (in absolute value) under variable cost convexity (concavity).*

Proof. By comparing APT for $\lambda = 1$ and APT for $\lambda \neq 1$ at the equilibrium, we have the following inequality

$$APT = \frac{\lambda}{\theta(2-\rho) + \lambda\,(1 - \theta + \theta\varepsilon_{\theta,q})} < APT|_{\lambda=1} = \frac{1}{1 + \theta - \theta\rho_d + \theta\varepsilon_{\theta,q}}$$

that is equivalent to $\lambda(2-\rho_d) < 2-\rho$ which in turn is equivalent to $(1-\lambda)(2-\rho_s) > 0$. From the definition of $c = qw$, we have $c'' = 2w' + qw'' = (2-\rho_s)w'$. As $w' > 0 \Leftrightarrow \lambda < 1$, the sign of c'' is also the sign of $(1-\lambda)(2-\rho_s)$. Hence the result follows. Also, because $RPT = (1-\mathcal{L})APT$ and the Lerner index is the same for both industries, the convexity/concavity of c also determines the comparison of RPT in the same way. Finally, because $APT = \left.\frac{dp}{dt}\right|_{t=0} = \frac{1}{L}P_x \left.\frac{dq}{dt}\right|_{t=0}$ and $RPT = \left.\frac{d\log p}{d\log\tau}\right|_{\tau=1} = -\frac{1}{\theta\tilde{\varepsilon}_d}\frac{d\log q}{d\log\tau}$, the result also extends to the absolute value of pass-through on quantity, whether we consider an additive or multiplicative cost shock. ∎

Before studying the long term consequences of the cost shock, it is also instructive to look at the profit impact in the short run, because, as will be clear below, it drives the consequences in the long run in terms of entry/exit on the market.

Proposition 5. *The additive cost shock is profit-enhancing if and only if $(1-\theta)APT > 1$ while the multiplicative cost shock is profit-enhancing if and only if $(1-\theta)RPT > 1-\mathcal{C}$.*

Proof. See Appendix E. ∎

Intuitively, in both cases, the cost shock brings two effects, a negative direct effect on profit and a positive effect in terms of reducing output which brings aggregate output closer to the collusive one and thereby enhancing profits holding the number of firms fixed. The condition for the additive cost shock is due to Weyl and Fabinger (2013). As $\theta < 1$ (substitute products), it follows that overshifting, i.e. $APT > 1$, is needed for the shock to be profit-increasing. The present condition for the multiplicative cost shock is new and allows to draw new insights on the comparison between both types of shocks. In particular, Proposition 5 suggests that overshifting, i.e. $RPT > 1$, is *not* needed for the multiplicative shock to be profit-increasing.

Moreover, using Proposition 3 and $\mathcal{C} = 1/\varepsilon$, the second condition in Proposition 5 can be rewritten as $(1-\theta)APT > (\varepsilon - 1)/(\varepsilon - \lambda)$. Observe that under CAVC ($\lambda = 1$) then the two conditions in Proposition 5 coincide. On the contrary, when $\lambda > 1$ (concave VC) then an additive cost shock is more likely to enhance profits than a multiplicative cost shock. Conversely, under convex VC ($\lambda < 1$), a multiplicative cost shock is more likely to enhance profits than the additive shock.

5 Comparative statics

In this section, I successively: (i) examine the comparative statics of the oligopoly model with respect to the number of firms n (section 5.1), (ii) analyze the cost pass-through in the long run when the number of firms is endogenous (section 5.2) and (iii) consider the comparative statics of the oligopoly model when the number of consumers L varies (section 5.3).

5.1 Market concentration incidence

Let us first analyze the incidence of market concentration. As we will see, a key factor in determining the evolution of the industry when n changes is the relative size of the *APT*, i.e. the strength of the price response to an infinitesimal additive cost shock (e.g., a unit tax). This also provides a number of useful results for the sequel.

First, the sign of MP_n, i.e. which measures the marginal impact of n on the perceived marginal profit, determines whether there is a business stealing or a business expansion externality in the industry. Indeed, total differentiation of (8) gives $dq/dn = -MP_n/MP_q$ and thus *business stealing, i.e.* $dq/dn < 0$, *occurs if and only if* $MP_n < 0$. From (15), recall that MP_n is given by:

$$MP_n = P_n + \theta x P_{xn} + \theta_n x P_x,$$

the sign of which is ambiguous in general and hence whether there is business stealing or expansion depends on the demand and the market conduct index properties.[27]

Second, observe that *the equilibrium profit decreases in* n if and only if the long run stability conditions hold. Indeed, totally differentiating $\pi = (P(x,n) - w(q))q$ yields

$$d\pi = \frac{\partial \pi}{\partial n} dn + \frac{\partial \pi}{\partial q} dq = qP_n dn + (1-\theta) x P_x dq$$

Using $dq/dn = -MP_n/MP_q$ and replacing, one obtains:

$$\frac{d\pi}{dn} = qP_n - (1-\theta)xP_x \frac{MP_n}{MP_q} = \frac{1}{MP_q}\left(MP_q q P_n - MP_n (1-\theta) x P_x\right) < 0$$

[27]See Cao et al. (2021) for a theoretical and empirical study of competition between dockless bikesharing firms in China that suggests that entry of a new firm can generate business extension for an incumbent.

under (16). Observe that by using Proposition which shows that $MP_q = (P_x/L)/APT$, one can rewrite the long-run stability condition (16) as follows:

$$P_n - MP_n(1-\theta)APT < 0 \tag{20}$$

Finally, the impact of market concentration on price can be measured through:

$$\frac{dp}{dn} = P_x\frac{dx}{dn} + P_n = P_n - MP_nAPT$$

using the total derivation of (8) which gives $dq/dn = -MP_n/MP_q$ and $MP_q = (P_x/L)/APT$.

For further reference, I gather all these results in the following Proposition.

Proposition 6. *When the number n of competitors increases, then*

(i) *there is business stealing ($dq/dn < 0$) if and only if $MP_n < 0$,*

(ii) *the equilibrium profit decreases if and only if $P_n - MP_n(1-\theta)APT < 0$ (long-run stability condition),*

(iii) *the equilibrium price decreases if and only if $P_n - MP_nAPT < 0$.*

The combination of the conditions expressed in Proposition 6 results in three different scenarios highlighted in Figure 2. Each scenario occurs based on only two conditions: (i) whether there is business stealing or expansion, and (ii) whether the *APT* is strong, moderate or weak. As an example, scenario 3 is characterized by the usual situation of downward price competition, which unambiguously benefits consumers. This actually occurs under two different contexts. First, observe that the presence of a *business expansion* effect, i.e. $MP_n > 0$, is a sufficient condition for the equilibrium price to be decreasing in n. However, despite increasing output per firm, profit is unambiguously decreasing. Profit per firm is also decreasing when there is business stealing and the *APT* is weak, i.e. $APT < P_n/MP_n$. The industry in scenario 3 is thus characterized by less profitable firms even though their size in terms of output may increase.

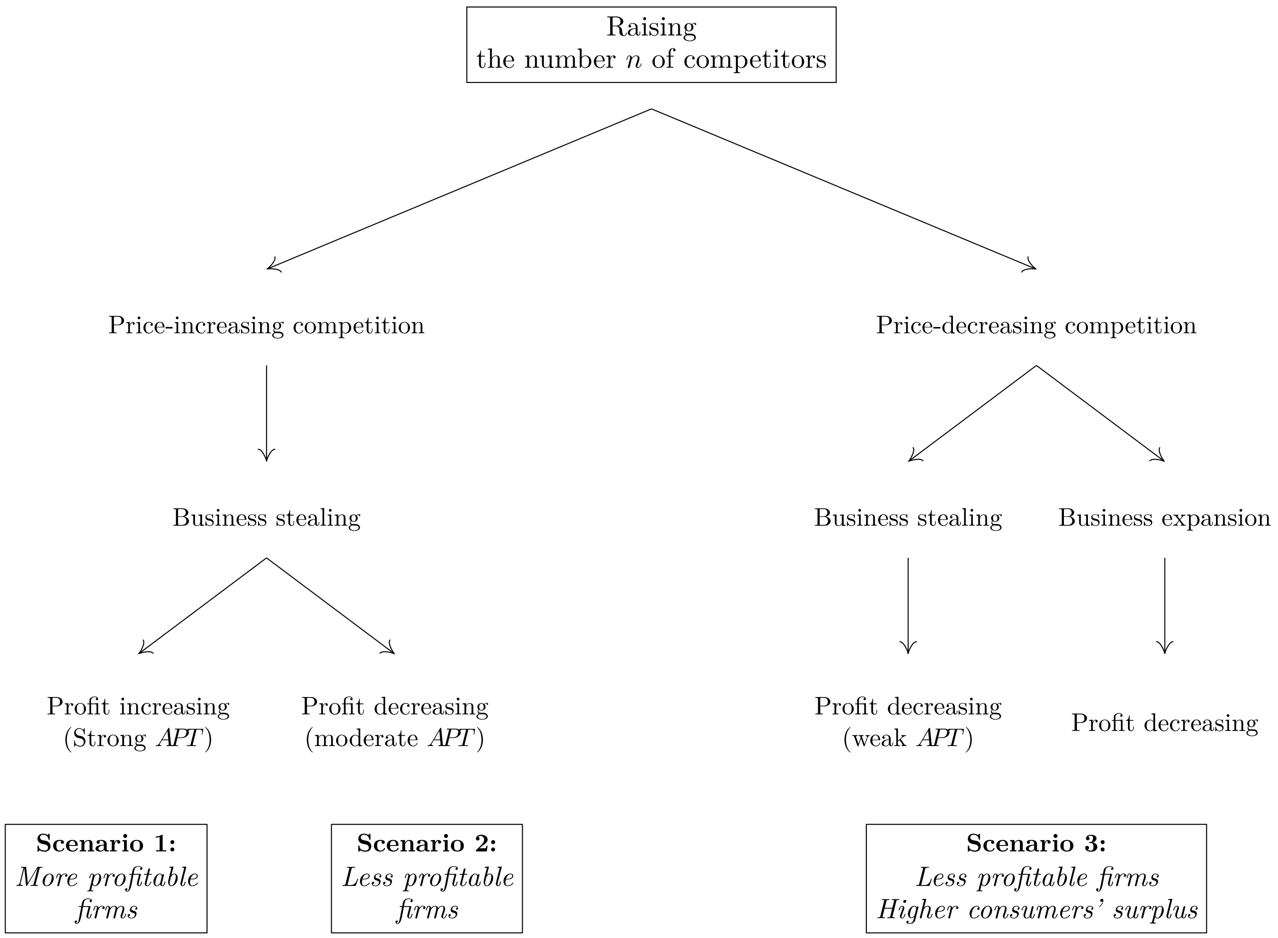
By contrast, under scenarios 1 and 2, competition drives up prices. More precisely, scenario 1 occurs because there is business stealing and the *APT* is strong. In this scenario, competition results in higher prices and profits, resulting in a larger industry with more profitable but smaller firms. In scenario 2, competition also increases prices, but because the *APT* is moderate, the industry evolves into a larger industry with less profitable and smaller firms. In both scenarios 1 and 2, consumers suffer from higher prices but benefit from greater variety, so that the impact of competition on consumers remains ambiguous at this stage.

Despite the fundamental insight of economics that competition usually lowers prices, the situation of price-increasing competition has already been identified as plausible both in the empirical and theoretical literature.[28] In particular, Chen and Riordan (2008) shows that, in a discrete choice model of product differentiation, the symmetric duopoly price can be higher than the single-product monopoly price. Specifically, because of competition for market share, duopolists have an incentive to price below the monopoly price, but in some situations, greater consumer choice may steepen the demand curve for each firm, and the latter incentive to raise the price may dominate. Similarly, Bertoletti and Etro (2016) show that, in a representative consumer model with preferences characterized by generalized linear direct utility, Nash in price or in quantity equilibria could be characterized by increasing prices in n, especially when the number of firms is large enough.

The present analysis covers a wide range of oligopoly models and, like in Bertoletti and Etro (2016), it treats n as a continuous variable. The impact of a greater number of firms on the market price results from the confrontation of two effects. On the one hand, at constant consumption per capita and per variety, the increase in n exerts a downward pressure on the price ($P_n < 0$). On the other hand, holding n constant, the possible business stealing effect ($MP_n < 0$) lowers the equilibrium quantity per variety and therefore it increases the price. When the latter dominates the former, especially when APT is sufficiently large, then competition drives up prices. In addition, this analysis also identifies circumstances in which competition also increases profits in the short run. Of course, in the long run when the market structure is endogenous, the stability conditions preclude the possibility that equilibrium profits may increase with competition.

Private versus social incentives to enter the market. It is well known that whether oligopolistic competition does result in too few or too many products from a welfare point of view depends on the relative forces of two effects (Mankiw and Whinston, 1986). On the one hand, firms do not take into account the business stealing/expansion externality and on the other hand, firms usually do not fully internalize consumer surplus. Hence, each potential entrant only considers the net profit $\pi - f$ to be

[28] As early as Satterthwaite (1979) and Rosenthal (1980), it has been shown that the equilibrium price can raise following an increase in the number of firms. Empirical evidence has been found in medical services (Pauly and Satterthwaite, 1981), in the automobile retail and automobile tire markets (Bresnahan and Reiss, 1990 and 1991), in the drug market (Grabowski and Vernon, 1992; Perloff et al., 2006), and private labels in the food industry (Ward et al., 2002) and Internet access (Chen and Savage, 2011). More recently, Mangin (2022) theoretically examines the conditions for price-increasing competition in a random utility model where the number of firms is ex ante uncertain.



.

Figure 2: Scenarios following an increase in the number n of competitors. Business stealing (expansion) means $MP_n < (>) 0$. Strong *APT* means $APT > P_n/((1-\theta)MP_n)$. Moderate *APT* means $P_n/MP_n < APT < P_n/((1-\theta)MP_n)$. Weak *APT* means $APT < P_n/MP_n$.

made while a social planner unable to control the behavior of active firms would consider the marginal impact of entry on welfare. Let us take welfare as the sum of consumers and producers surplus as follows:

$$W = LU(x, n) - nqw(q) - nf.$$

Computing the derivative of welfare with respect to n yields:

$$\begin{aligned} W'(n) &= LU_n + LU_x \frac{dx}{dn} - qw - n(w + qw')\frac{dq}{dn} - f \\ &= \underbrace{p\mathcal{V}_0}_{\text{taste for diversity}} + \underbrace{\pi - f}_{\text{net profit}} + \underbrace{n(p - w - qw')\frac{dq}{dn}}_{\text{business stealing/expansion}} \end{aligned}$$

recalling that $U_x = np$ and that $\mathcal{V}_0 = (LU_n - pq)/p$ denotes the variety effect. The marginal impact of an additional variety on welfare can be decomposed into three terms: (i) a taste for diversity term which is positive when $\mathcal{V}_0 > 0$, (ii) a net profit term because a new variety means additional profit in the industry and (iii) a business stealing/expansion term that appears as long as price is above marginal cost and that is negative if and only if there is business stealing, i.e. $dq/dn < 0$. In other words, the presence of a taste for diversity and of a business stealing/expansion effect drives a wedge between the marginal entrant's evaluation of the interest of entry and the social planner's.

If the sum of the taste for diversity and business stealing/expansion terms is negative then $\pi - f > W'(n)$ and private incentives to enter exceeds the social planner's. Using $\mathcal{L} = (p - w - qw')/p = \theta/\varepsilon_d$, $dq/dn = -MP_n/MP_q$ and $MP_q = (P_x/L)/APT$, the latter condition rewrites

$$p\mathcal{V}_0 + Q\theta MP_n APT < 0. \tag{21}$$

In particular, when the latter condition holds at the free entry equilibrium, then there is excess entry on the market from a welfare point of view.[29] Under product homogeneity, the variety effect vanishes ($\mathcal{V}_0 = 0$) and excess entry arises if and only if there is business stealing, i.e. $MP_n < 0$ (see Amir et al. (2014) for establishing rigorously this result in the case of Cournot oligopolies and for any cost function). When products are differentiated, excess entry remains as long as the variety effect $\mathcal{V}_0$ is not too strongly positive. Mankiw and Whinston (1986) have shown this result for the class of direct additive preferences *à la* Dixit and Stiglitz (1977).[30]

[29]Indeed, denoting n^* the social optimum and n^e the equilibrium number of firms under free entry then $\pi - f > W'(n) \Rightarrow \pi(n^*) - f > W'(n^*) = 0 = \pi(n^e) - f$. Because the long run stability condition implies that $\pi'(n) < 0$ (see Proposition 6), then $n^e > n^*$.

[30]Bertoletti and Etro (2016) also study optimal market structures in a fairly general model with respect to preferences, but where (i) marginal cost is constant and more importantly (ii) labor is the only input and labor supply is inelastic, while in the present setting, labor supply is implicitly assumed to be elastic which is a wide-spread assumption in standard oligopoly models. See also Parenti et al. (2017) for a study of imperfect competition with labor as the sole input and a fixed labor supply.

5.2 Cost pass-through in the long run

Consider the pass-through of an additive cost shock t in the long run. The equilibrium is characterized by the behavioral response of firms and the zero profit condition:

$$P(x,n) + \theta x P_x(x,n) - w(q) - qw'(q) - t = 0 \quad (22)$$

$$(P(x,n) - w(q) - t)q - f = 0. \quad (23)$$

Similarly, for a multiplicative cost shock, the long run equilibrium is characterized by:

$$P(x,n) + \theta x P_x(x,n) - \tau(w(q) + qw'(q)) = 0 \quad (24)$$

$$(P(x,n) - \tau w(q))q - f = 0. \quad (25)$$

Applying the Implicit Function Theorem to both cases, I identify the circumstances under which a cost shock is procompetitive or anticompetitive. The Proposition below also indicates the circumstances under which output per firm raises or not following the cost shock.

Proposition 7. *In the long run,*

(i) *an additive cost shock is procompetitive if and only if* $(1-\theta)APT > 1$ *while a multiplicative cost shock is procompetitive if and only if* $(1-\theta)APT > (\varepsilon - 1)/(\varepsilon - \lambda)$*,*

(ii) *an additive cost shock decreases the output per firm if and only if* $P_n - MP_n < 0$ *while a multiplicative cost shock decreases output per firm if and only if* $P_n < MP_n(\varepsilon - 1)/(\varepsilon - \lambda)$*,*

(iii) *Under convex VC (CAVC), a multiplicative cost shock is more (equally) likely to be pro-competitive and to reduce output per firm than an additive cost shock.*

Proof. See Appendix F. ∎

Consider first the additive cost shock. Recall that $(1-\theta)APT > 1$ is precisely the condition given in Proposition 5, under which an additive cost shock raises profit keeping n constant. Intuitively, the same condition evaluated at the long run equilibrium reveals that the additive cost shock is procompetitive (part (i)). For this to hold, a sufficiently strong overshifting of the cost shock is needed as $\theta < 1$, and an equivalent formulation of the condition is $\rho > 2 + \lambda\varepsilon_{\theta,q}$ using the definition of APT contained in Proposition 3.

As indicated in part (ii), the change in output q depends on the sign of $P_n - MP_n$, which describes how the difference between the price and the perceived marginal profit changes with the number of competitors, holding x constant. Note that $P_n - MP_n$ equivalently represents how the *perceived inverse marginal surplus per consumer*, i.e. $-\theta x P_x$, changes in n holding x constant. Hence, we have:

$$P_n - MP_n = \frac{\partial}{\partial n}\left(-\theta x P_x\right) = -\frac{\theta}{n} x P_x \left(\varepsilon_{P_x,n} + \varepsilon_{\theta,n}\right),$$

where $\varepsilon_{P_x,n} = nP_{xn}/P_x$ is the partial elasticity of the slope of the inverse demand w.r.t n. It follows that $P_n - MP_n < 0$ if and only if $\varepsilon_{P_x,n} + \varepsilon_{\theta,n} < 0$. Clearly, more competitors naturally means a raising competition intensity and thus a lower market conduct index ($\varepsilon_{\theta,n} < 0$), but the elasticity value depends on the mode of competition. The demand term $\varepsilon_{P_x,n}$ results of the consumers' preferences independently of the specification for the mode of competition. Moreover, the additive cost shock is output-neutral in the long-run if and only if $\varepsilon_{P_x,n} + \varepsilon_{\theta,n} = 0$. To illustrate, assuming linear demands, this is indeed the case under Nash in quantities competition and the competition effect and the demand effect are perfectly counterbalanced. On the contrary, under Nash in prices competition, we have $P_n - MP_n < 0$, thereby indicating that the additive cost shock decreases output per firm in the long-run.

By combining the two results above, let us now evaluate the price change by computing the long run absolute pass-through denoted as APT^{LR}. As shown below, it can be decomposed into the sum of an intensive and extensive margin terms:[31]

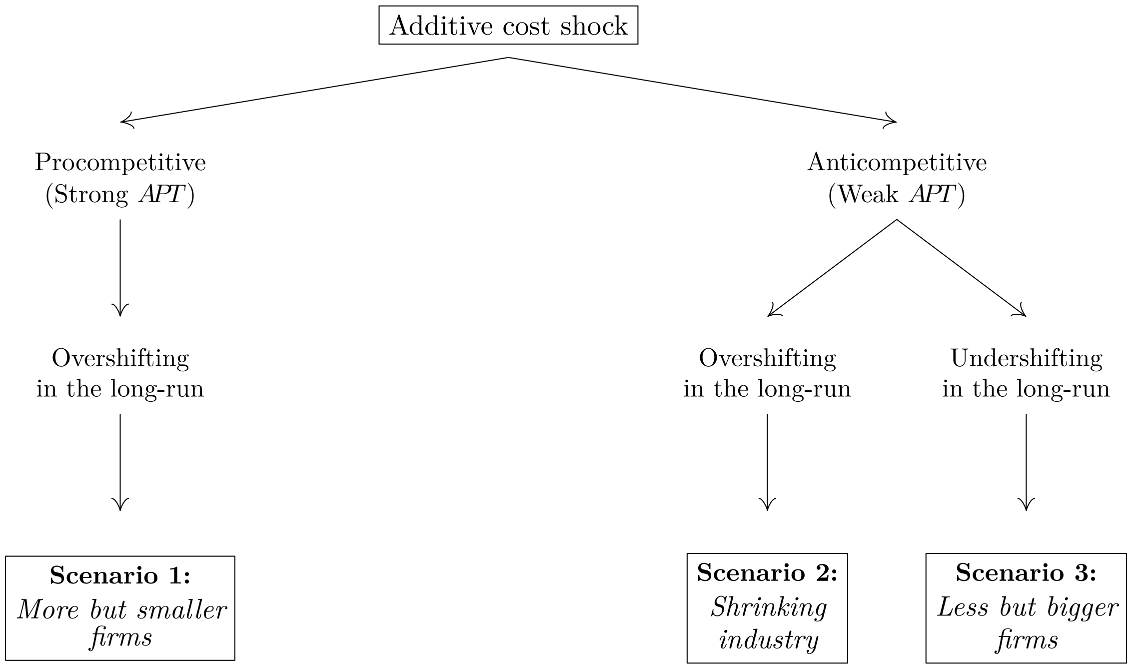
$$\begin{aligned} APT^{LR} &= \left.\frac{dp}{dt}\right|_{t=0} = \underbrace{\frac{P_x}{L}\left.\frac{dq}{dt}\right|_{t=0}}_{\text{Intensive margin}} + \underbrace{P_n \left.\frac{dn}{dt}\right|_{t=0}}_{\text{Extensive margin}} \\ &= 1 + \frac{\theta(P_n - MP_n)APT}{P_n - MP_n(1-\theta)APT}. \end{aligned} \tag{26}$$

It follows that overshifting of the additive cost shock occurs in the long run if and only if output per firm decreases, i.e. iff $P_n - MP_n < 0$. Note that perfect competition ($\theta = 0$) entails that $APT^{LR} = 1$ meaning that under free entry the cost shock is fully transmitted to consumers. Under imperfect competition, this situation also occurs whenever the cost shock is output-neutral in the long-run.

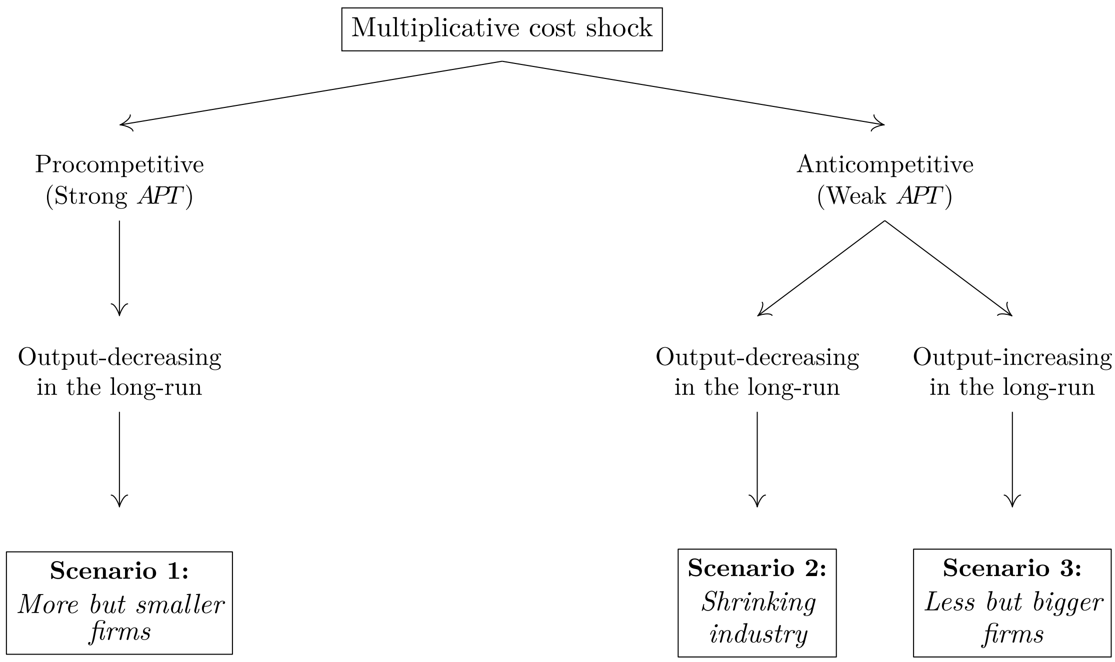
The possible scenarios for the industry in the long run resulting from the above observations are depicted in panel (a) of Figure 3. First, when the additive shock is procompetitive, it is also characterized by overshifting and this results in an industry with more varieties/firms but where each firm is smaller in terms of output (scenario 1).[32]

[31] See Appendix F for details of the derivation.

[32] Indeed, there are two possible cases. Either $MP_n > 0$ (business stealing), and then overshifting occurs as $P_n - MP_n < 0$. Or $MP_n < 0$ (business creation) and the stability condition $P_n < MP_n(1-\theta)APT$ implies that $P_n - MP_n < 0$ as $1 < (1-\theta)APT$ in scenario 1. Overall, cost overshifting always occurs in scenario 1.



(a): Strong (weak) *APT* means $APT > (<) 1/(1-\theta)$.
Overshifting (undershifting) in the long-run means $MP_n > (<) P_n$.



(b): Strong (weak) *APT* means $APT > (\varepsilon - 1)/[(\varepsilon - \lambda)(1-\theta)]$.
Output decreasing (increasing) in the long-run means $MP_n(\varepsilon - 1)/(\varepsilon - \lambda) > (<) P_n$.

Figure 3: Long-run scenarios following an additive or a multiplicative cost shock.

Second, when the additive cost shock is anticompetitive, the final outcome depends on whether there is over or undershifting (scenarios 2 or 3 respectively). Note that in scenario 3 (fewer varieties but each in larger quantities), it is theoretically possible that the additive cost shock ends up with a *negative pass-through on price*, i.e. $APT^{LR} < 0$. This outcome would correspond to a version of Edgeworth's taxation paradox according to which taxation can reduce market price of the good.[33] What are the conditions for a negative pass-through on price? First, observe that business stealing ($MP_n < 0$) is necessary in scenario 3 for output per firm to increase in the long run, i.e. for $P_n - MP_n > 0$ to hold. Moreover, following a straightforward manipulation of (26) and using the stability condition, we obtain that $APT^{LR} < 0$ if and only if $0 > MP_n(1-\theta)APT > P_n > \max\left(MP_n, MP_n\frac{APT}{1+\theta}\right)$.[34] Hence, when APT valued at the long run equilibrium is greater than $1+\theta$, the conditions for scenario 3 highlighted in Figure 3 already ensure that the long run pass-through is negative. When $APT < 1+\theta$, the condition $P_n > MP_n\frac{APT}{1+\theta}$ has to be added to obtain that $APT^{LR} < 0$.[35]

Let us now turn to the case of a multiplicative cost shock. Proposition 7 suggests that when marginal cost is constant (CAVC, $\lambda = 1$), there is no difference between an additive and a multiplicative cost shock in terms of the conditions describing the change in output per firm and number of firms under free entry/exit. However, under convex VC ($\lambda < 1$), both conditions (i) and (ii) now involve the factor $(\varepsilon - 1)/(\varepsilon - \lambda) < 1$ and a multiplicative cost shock is more likely to be pro-competitive and to reduce output per firm than an additive cost shock. The panel (b) of Figure 3 describes the different scenarios in the same way as for the additive cost shock. Importantly, the equivalence between overshifting and output reduction found for the additive cost shock does not hold for the multiplicative cost shock.

Finally, the long-run relative cost pass-through denoted RPT^{LR} can be evaluated as follows:[36]

$$\begin{aligned} RPT^{LR} &= \frac{\tau}{P}\frac{dp}{d\tau}\bigg|_{\tau=1} = \frac{P_x/L}{P}\frac{dq}{d\tau}\bigg|_{\tau=1} + \frac{P_n}{P}\frac{dn}{d\tau}\bigg|_{\tau=1} \\ &= \frac{\varepsilon-\lambda}{\varepsilon}APT^{LR} + \frac{\lambda-1}{\varepsilon}(P_n - MP_nAPT) \end{aligned}$$

Under CAVC, the relationship between APT^{LR} and RPT^{LR} is $RPT^{LR} = \frac{\varepsilon-\lambda}{\varepsilon}APT^{LR}$, which is the same

[33]More precisely, Edgeworth's taxation paradox (1925) states that a unit tax can decrease the price of the taxed good in the context of a multiproduct monopoly under conditions of complementarity between goods.

[34]As $\theta \in (0,1)$, we can check that $1-\theta < 1/(1+\theta)$.

[35]Note that the characteristics of scenario 3 where cost pass-through can become negative are opposite to those obtained by Ritz (2015). In that paper, Ritz considers firms that are heterogeneous in terms of constant marginal cost and shows that a unit tax can induce the entry of a new firm in the face of an incumbent monopoly and ultimately reduce the price within the resulting duopoly a la Cournot. The key conditions for this result are (i) the unitary tax can increase profit and (ii) entry reduces the price. Here, within symmetric oligopolies with endogenous entry, a cost pass-through can only become negative if the cost shock reduces profit and firm exit reduces price.

[36]See Appendix F for details of the derivation.

as the one that prevails in the short term and which is given by (18). Under convex VC ($\lambda < 1$), an additional term appears, the sign of which depends on the sign of $P_n - MP_n APT$. Using the results from Proposition 6, it follows that $RPT^{LR} > \frac{\varepsilon-\lambda}{\varepsilon} APT^{LR}$ if and only if competition is price-decreasing.

5.3 Market expansion incidence

Let us now turn to the comparative statics of the short run and the long run equilibria w.r.t. market size L. As will be clear below, the contribution margin approach allows to obtain a set of sufficient statistics to describe the profit and output consequences in the short run as well as the consequences on the intensive and extensive margins in the industry in the long run.

Short-run analysis. To proceed, it is convenient to define the *ghost industry* here as follows.

Definition 4. *The ghost industry is the industry with constant marginal cost equal to the equilibrium value of marginal cost and with θ being constant in output and equal to the equilibrium value of θ. It follows that the APT for the ghost industry writes as:*

$$APT_{ghost} = APT|_{\lambda=1,\ \theta=cst} = \frac{1}{1+\theta-\theta\rho_d}.$$

As argued above, both industries produce the same amount. Let us denote the relative change in (variable) profit due to market expansion as $\varepsilon_{\pi,L} \equiv \frac{L}{\pi}\frac{d\pi}{dL}$ and accordingly the relative change in output as $\varepsilon_{q,L} = \frac{L}{q}\frac{dq}{dL}$. Straightforward computations lead to the results contained in the following Proposition.

Proposition 8. *Holding the number of firms constant,*

(i) *market expansion always raises output per firm according to*

$$\varepsilon_{q,L} = APT/APT_{ghost} > 0.$$

(ii) *market expansion is price-decreasing if and only if $\varepsilon_{q,L} > 1$ or equivalently $APT > APT_{ghost}$.*

(iii) *the relative change in equilibrium variable profit resulting from changes in L is given by*

$$\varepsilon_{\pi,L} = \frac{\lambda}{\theta}\left[1 - (1-\theta)\varepsilon_{q,L}\right].$$

(iv) *market expansion is profit-enhancing if and only if $(1-\theta)APT < APT_{ghost}$.*

Proof. See Appendix G. ■

The elasticity of output per firm w.r.t market size is governed by the comparison between the pass-through *APT* for the oligopoly and the corresponding pass-through APT_{ghost} for the ghost industry. The relative size of $\varepsilon_{q,L}$ with respect to 1 also determines whether market expansion is price-decreasing or price-increasing in the short run, and this actually depends on two potentially conflicting terms. To show this, let us compute the following quantity:

$$\frac{1}{APT} - \frac{1}{APT_{ghost}} = \frac{\theta}{\lambda}\left[\underbrace{\lambda\varepsilon_{\theta,q}}_{\text{competition intensity}} + \underbrace{(1-\lambda)(2-\rho_s)}_{\text{cost convexity/concavity}}\right]$$

which can be decomposed into a competition intensity term and a cost term whose sign depends on the convexity/concavity of $c(q)$. Consider first that θ is constant in output ($\varepsilon_{\theta,q} = 0$). The above comparison is such that $\varepsilon_{q,L} < 1$ if and only if the variable cost function $c(q)$ is convex.[37] Intuitively, with convex variable cost, market expansion raises output but at a moderate pace. Conversely, concave variable costs ensure a strong output reaction to market expansion ($\varepsilon_{q,L} > 1$). Finally, under CAVC, then $\lambda = 1$ and thus $\varepsilon_{q,L} = 1$. In this case, output per firm is directly proportional to market size. Now, consider that higher quantities/lower prices create more competitive market conduct, i.e. $\varepsilon_{\theta,q} < 0$, then we have seen that this tends to make *APT* larger in Section 4 and in particular compared to APT_{ghost}, thereby favoring a strong output reaction market expansion ($\varepsilon_{q,L} > 1$).

Furthermore, Proposition 8 suggests that the profitability scale factor λ has an interesting interpretation when the industry is perfectly colluding ($\theta = 1$) which echoes the result obtained by Bontems (2023a) for the monopoly and monopolistic competition cases. Indeed, when $\theta = 1$ then λ *represents the elasticity of (variable) profit with respect to market size at the equilibrium.* When $\theta \neq 1$ then λ/θ represents the partial elasticity of profit w.r.t. market size, holding q constant.

Also, parts (iii) and (iv) state that $\varepsilon_{\pi,L} < 0$ if and only if $\varepsilon_{q,L} > 1/(1-\theta) > 1$ as $\theta \in (0, 1)$ given the assumption of substitute products. It follows that market expansion is profit decreasing (holding constant the number of firms) whenever the induced expansion of production is large enough to cause a deleterious drop in prices. This situation is all the more plausible as the competition is strong, i.e. when θ is low.

Market expansion in the long run. As earlier, with free entry, the long run equilibrium is described by the zero profit condition (13) as well as the first-order condition (8). Applying the

[37] Recall that $c(q)$ is convex if and only if $(1-\lambda)(2-\rho_s) > 0$.

Implicit Function Theorem to this system of equations yields the following result.

Proposition 9. *At the long run equilibrium, market expansion*

(i) *raises output per firm if and only if* $P_n - MP_n APT_{ghost} < 0$,

(ii) *is price-decreasing if* $P_n - MP_n APT_{ghost} < 0$,

(iii) *encourages entry if and only if* $(1-\theta)APT < APT_{ghost}$.

Proof. See Appendix H. ■

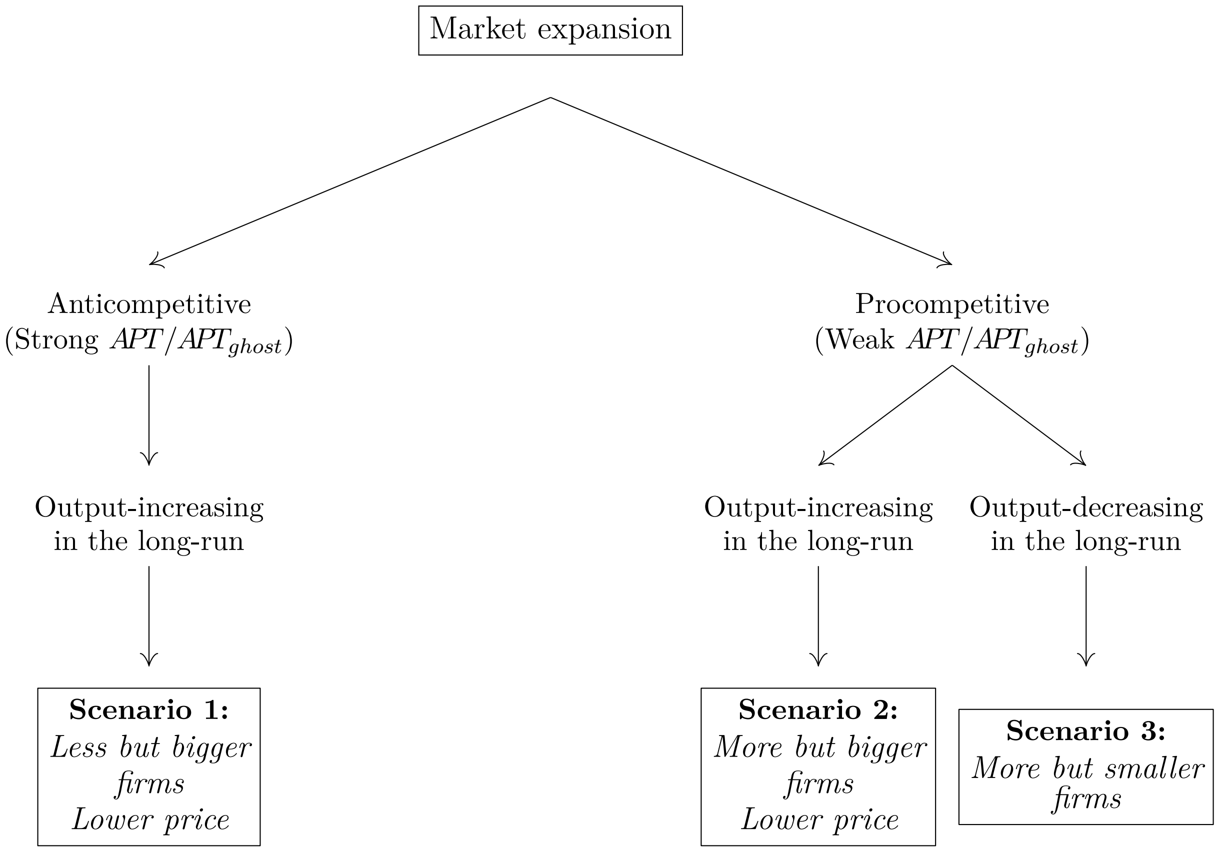
Part (iii) indicates that the condition under which market expansion raises profits holding n constant intuitively implies that, in the long run, the number of firms increases with the size of the market. Figure 4 describes the two possibilities in which increasing the size of the market is pro- or anti-competitive. The different scenarios that follow depend on the effect of market size on output per firm which is given by the condition in part (i). Note that the condition according to which the increase in market size makes firms grow in terms of output and that reduces price is also the condition according to which competition is price decreasing in the ghost industry (see part (iii) in Proposition 6).

Three scenarios emerge from the conditions proposed in Proposition 9. Scenario 1 is characterized by an anticompetitive effect of an increase in L appearing in the case of a sufficiently high value of the APT/APT_{ghost} ratio. In this case, only an increase in output per firm is possible and the industry is composed in the long run of fewer but larger firms. The effect on consumer surplus remains ambiguous between a lower price but a smaller variety of products. Scenarios 2 and 3 arise from a pro-competitive effect of market size and are distinguished by the effect on the size of each firm. When firms grow in size (Scenario 2), consumer surplus increases due to the dual effect of lower prices and access to a greater number of varieties. When firms decrease in size (Scenario 3), the effect of increasing L on consumer surplus remains ambiguous at this level of generality.[38]

6 Regulation incidence made simple

The cost pass-through formulas we have derived above are obtained by assuming zero initial cost shock like e.g. Weyl and Fabinger (2013). Adachi and Fabinger (2022) and Kroft et al. (2021) show how to extend the formulas obtained by Weyl and Fabinger in contexts where taxation exists initially

[38]The impact of market expansion on welfare can also be studied in detail (see Appendix I).



.

Figure 4: Scenarios following market expansion. Strong (weak) APT/APT_{ghost} means $APT > (<)1/(1-\theta)$. Output-decreasing (expanding) means $MP_n < (>) P_n/APT_{ghost}$.

before the marginal (or discrete) change. This is particularly useful when one wants to compare the relative merit of different taxes like e.g. ad-valorem versus specific taxes. Following these analysis, I extend my approach to consider regulation pass-through with non-zero initial regulation (that may not take the form of taxation). I will first show how to extend the contribution margin approach to this regulatory context and then examine the impact of regulation on prices and profits in the short run within the industry. I also propose a formulation of the incidence, marginal excess burden, and marginal value of public funds associated with each instrument in the regulatory system, both in the short and long run.

For this, I follow Adachi and Fabinger (2022) by assuming that any firm i is subject to a common regulation system $\boldsymbol{t}$ whose cost, *per unit of output*, borne by each firm is denoted $\psi(q_i, p_i, \boldsymbol{t})$ and represents the *Average Regulation Burden* function (hereafter ARB):[39]

$$\pi_i = (p_i - w(q_i) - \psi(q_i, p_i, \boldsymbol{t}))q_i.$$

Hence, if the regulation only entails some taxation scheme, $q_i\psi(q_i, p_i, \boldsymbol{t})$ represents not only the tax burden for firm i, but also the tax revenue made on firm i for the tax authority.[40] I assume that all profits as well as regulation revenues are redistributed to consumers as a lump-sum transfer. Let us also denote:

$$\tilde{w}(q_i, p_i, \boldsymbol{t}) = w(q_i) + \psi(q_i, p_i, \boldsymbol{t}),$$

as the *effective* Average Variable Cost function, i.e. the AVC adjusted for the regulation burden as represented by the ARB. Similarly, the *effective* Variable Cost function is denoted $\tilde{c}(q_i, p_i, \boldsymbol{t}) \equiv q_i\tilde{w}(q_i, p_i, \boldsymbol{t})$.

To illustrate, consider the following examples of such policies.

Examples.

(i) Ad-valorem cum specific taxation. Suppose that each firm is subject to an ad valorem tax $1 - t_v$ on price and a unit tax t_s so that the regulation is parameterized with $\boldsymbol{t} = (t_v, t_s)$. In this case, $\pi_i = ((1 - t_v)p_i - w(q_i) - t_s)q_i$ so that the effective AVC is $\tilde{w}(q_i, p_i, \boldsymbol{t}) = w(q_i) + t_v p_i + t_s$. Here, $\tilde{w}$ is an affine function of price p only as long as $t_v \neq 0$.

[39] One could add easily n as an argument of the ARB to represent regulation policies that depend on e.g. aggregate output (for instance the case of a tax refunding rule proportional to aggregate output as studied by Bontems (2019) in the context of pollution). For the sake of clarity/brevity this is not pursued here.

[40] This cost would only contain the variable cost of regulation and not its fixed part (independent of price or quantities) that would be relegated into the fixed cost f. For clarity, in the present context of an exogenous number of firms, I will assume that there is no such lump sum taxes or subsidies on firms.

(ii) Cost subsidies. Suppose that firms now receive cost subsidies. Denoting the rate of cost subsidy by τ and the specific subsidy s, one can write profit $\pi_i = (p_i - (1-\tau)w(q_i) + s)q_i$ and the effective AVC is $\tilde{w}(q_i, p_i, \boldsymbol{t}) = (1-\tau)w(q_i) - s$ where $\boldsymbol{t} = (\tau, s)$.

(ii) Exogenous Competition. Weyl and Fabinger (2013) and Miklos-Thal and Shaffer (2021) analyze the impact of competition on firms and consumers by looking at the effect of a hypothetical entrance of exogenous quantity into a market on the total output exchanged. Let us denote $\tilde{q}$ the exogenous quantity added to each variety/firm on the market so that aggregate output is now $Q = n\tilde{q} + \sum_i q_i$, and each firm i now produces $q_i - \tilde{q}$. The intervention is parameterized with $\boldsymbol{t} = \tilde{q}$. Profit writes now $\pi_i = (p_i - w(q_i - \tilde{q}))(q_i - \tilde{q})$ and the effective AVC is $\tilde{w}(q_i, p_i, \boldsymbol{t}) = (\tilde{q}/q_i)p_i + (1 - \tilde{q}/q_i)\, w(q_i - \tilde{q})$.

The contribution margin approach under regulation. Let us start by considering the first-order condition at the symmetric equilibrium which now writes:

$$P + \theta x P_x - \tilde{w} - q\tilde{w}_q - \theta x P_x \tilde{w}_p = 0. \tag{27}$$

where the last term is new and involves $\tilde{w}_p = \psi_p$ which represents the *first-order price sensitivity of the average regulation burden.* It is convenient to define both the first-order sensitivity of the margin $P - \tilde{w}$ (or equivalently the variable profit) with respect to price and similarly for the mark-up $P - \tilde{w} - q\tilde{w}_q$ w.r.t. price as follows.

Definition 5. *The first-order sensitivity of the margin $P - \tilde{w}$ and of the mark-up $P - \tilde{w} - q\tilde{w}_q$, with respect to price at the symmetric equilibrium are respectively:*

$$\gamma(q, n, \boldsymbol{t}) \equiv 1 - \tilde{w}_p(q, P(q/L, n), \boldsymbol{t})$$

and

$$\delta(q, n, \boldsymbol{t}) \equiv 1 - \tilde{w}_p(q, P(q/L, n), \boldsymbol{t}) - q\tilde{w}_{pq}(q, P(q/L, n), \boldsymbol{t}).$$

Both functions depend on the equilibrium output q, on the number of firms n and on the regulation parameters vector $\boldsymbol{t}$. In the following, I consider all regulations that leave γ and δ strictly positive. Observe that under laissez-faire, $\tilde{w}_p = 0$ and consequently $\gamma = \delta = 1$. This is also the case for the additive and multiplicative cost shocks studied in Section 4. There, the ARB does not depend on price either and hence $\gamma = \delta = 1$. By contrast, under some other regulation systems, the ARB could depend potentially on price and thus $\gamma \neq 1$ and $\delta \neq 1$ in general. For instance, in the ad-valorem cum specific tax system example, recall that $\psi = t_v p + t_s$, $\psi_p = t_v$ and hence $\gamma = \delta = 1 - t_v \ \in (0,1)$. In the exogenous competition example, $\psi_p = \tilde{q}/q$ and thus $\gamma = 1 - \tilde{q}/q \in (0,1)$ while $\delta = 1 - 2\tilde{q}/q$.

Clearly, (27) can be rewritten as follows:

$$P + \gamma\theta x P_x - \tilde{w} - q\tilde{w}_q = 0.$$

The regulation system brings two consequences on the different terms that explain the firm's decision compared to the laissez-faire situation: first, each firm now takes into account the *effective cost* that includes the regulation burden ($\tilde{w}$ instead of w) and second, the market conduct index θ has to be replaced with the *effective market conduct index*, i.e. θ has to be factorized by the scale factor γ. However, it can be checked that, because of symmetry, the market conduct index θ has the same formulation as before (see Appendix J). This implies that both θ and γ depend on q and n, but unlike γ, θ does not depend directly on the regulation system $\boldsymbol{t}$. In other words, the effective market conduct index $\gamma\theta$ captures the competition intensity. This measure can be broken down into two parts: (i) θ measures the competition intensity due to the underlying nature of competition (e.g. price or quantity competition) and (ii) γ captures how regulation structurally changes the competition intensity.

It is then immediate to extend our results on the contribution margin approach to the present context of regulation. For this, let us denote $\tilde{\varepsilon}_s$ as the partial elasticity of the effective "supply" w.r.t. q, i.e. $\tilde{\varepsilon}_s = \tilde{w}/q\tilde{w}_q$ and its associated curvature measure $\tilde{\rho}_s = -q\tilde{w}_{qq}/\tilde{w}_q$. Also define $\tilde{\varepsilon}_d = \varepsilon_d/\gamma\theta$ as the *effective elasticity of industry's demand,* where ε_d is now deflated by $\gamma\theta$.[41]

Proposition 10. *At a symmetric equilibrium under imperfect competition and under a regulation system* $\boldsymbol{t}$ *and the corresponding regulation burden leading to the effective AVC* $\tilde{w}$*, the Lerner index is* $\mathcal{L} = 1/\tilde{\varepsilon}_d$ *and the* $\mathcal{C}$*-index obeys to an inverse pseudo-elasticity rule:*

$$\mathcal{C} = \frac{P - \tilde{w}}{P} = \frac{1}{\lambda\tilde{\varepsilon}_d} = \frac{1}{\varepsilon}$$

where

$$\lambda = \frac{\tilde{\varepsilon}_s + 1}{\tilde{\varepsilon}_d + \tilde{\varepsilon}_s}.$$

[41] Note that the elasticity of the effective "supply" $\tilde{\varepsilon}_s$ is the harmonic mean of the elasticity of "supply" ε_s and of the output sensitivity of the ARB function, weighted by their respective share in the effective AVC. Indeed, $\tilde{\varepsilon}_s = \tilde{w}/q\tilde{w}_q = \frac{w+\psi}{q(w'+\psi_q)} = \left(\frac{w}{\tilde{w}}\frac{1}{\varepsilon_s} + \frac{\psi}{\tilde{w}}\frac{1}{\varepsilon_{\psi,q}}\right)^{-1}$, with $\varepsilon_{\psi,q} = \psi/q\psi_q$ is the inverse of the output partial elasticity of $\psi(q,p,t)$. In case where ψ is independent of q as in the ad-valorem cum specific taxes example, then $\varepsilon_\psi = \infty$ and hence $\tilde{\varepsilon}_s = \frac{\tilde{w}}{w}\varepsilon_s$.

Also, the (short run) stability condition under the regulation system can be derived similarly to condition (12) by assuming that the marginal profit with respect to output is decreasing in output.[42] For this, consider now the "average curvature" measure $\rho \equiv \lambda\rho_d + (1-\lambda)\tilde{\rho}_s$ as the average of curvature measures for the demand and the effective AVC with weight λ and $1-\lambda$ respectively.

Assumption 5 (Stability condition). *Let us assume that the following stability condition under the regulation policy* $\boldsymbol{t}$ *holds:*

$$\rho < 2 + \lambda\left(\frac{\delta}{\gamma\theta} - 1 + \varepsilon_{\gamma\theta,q}\right). \tag{28}$$

where $\varepsilon_{\gamma\theta,q} = \varepsilon_{\theta,q} + \varepsilon_{\gamma,q}$ *is the output elasticity of* $\gamma\theta$.

Observe that, compared to condition (12), apart from the changes on effective AVC elasticity and curvature, two new terms also appears in (28) with γ and δ as the first-order sensitivity of respectively the margin and the mark-up to price.

Regulation pass-through on price and profit. Consider an element t of the policy vector $\boldsymbol{t}$. For convenience and without ambiguity, I denote

$$APT = \frac{\lambda}{\gamma\theta\left[2 - \rho + \lambda\left(\frac{\delta}{\gamma\theta} - 1 + \varepsilon_{\gamma\theta,q}\right)\right]} > 0 \tag{29}$$

as the expression of the *absolute pass-through on price of a zero unit tax in this context of oligopoly competition under the existing regulation* t, characterized by the effective AVC $\tilde{w}$ and the effective market conduct index $\gamma\theta$. Clearly, under laissez-faire, $\gamma = \delta = 1$, $\tilde{w} = w$ and the expression above boils down to the *APT* expression formulated in Proposition 3.

The following Proposition exhibits the formula for the relative pass-through of t on price given the regulation characterized by the policy vector $\boldsymbol{t}$.

Proposition 11. *Holding the number of firms constant,*

(i) *the relative pass-through of* t *on price is given by:*

$$\varepsilon_{p,t} = APT\left[(1-\mathcal{L})\,\varepsilon_{\tilde{c}_q,t} + \mathcal{L}\varepsilon_{\gamma,t}\right] \tag{30}$$

where $\varepsilon_{\tilde{c}_q,t} = \partial\log\tilde{c}_q/\partial\log t$ *is the partial elasticity of the effective marginal cost w.r.t* t *and* $\varepsilon_{\gamma,t} = \partial\log\gamma/\partial\log t$ *is the partial elasticity of* γ *w.r.t* t.

[42] See Appendix K.

(ii) *The change in t is profit-enhancing, i.e. $\varepsilon_{\pi,t} > 0$, if and only if:*

$$\gamma(1-\theta)\varepsilon_{p,t} > (1-\mathcal{C})\varepsilon_{\tilde{w},t} \tag{31}$$

where $\varepsilon_{\tilde{w},t} = \partial \log \tilde{w} / \partial \log t$ is the partial elasticity of the effective AVC w.r.t t.

Proof. See Appendix L. ■

Hence, part (i) indicates that the relative pass-through on price of any policy element t is driven by a weighted mean of $\varepsilon_{\tilde{c}_q,t}$ and $\varepsilon_{\gamma,t}$ (weighted by $1-\mathcal{L}$ and $\mathcal{L}$ respectively), which is magnified by a positive scale factor equal to *APT*. On the one hand, the partial elasticity $\varepsilon_{\tilde{c}_q,t}$ assesses how much the effective marginal cost is sensitive to t and on the other hand, the partial elasticity $\varepsilon_{\gamma,t}$ measures how much γ the sensitivity of margin (and thus profit) to price is impacted by a change in t.

Also, a direct implication of part (i) is that when comparing two policy instruments, say t_1 and t_2 then the ratio of price elasticities w.r.t regulation reduces to:

$$\frac{\varepsilon_{p,t_1}}{\varepsilon_{p,t_2}} = \frac{(1-\mathcal{L})\,\varepsilon_{\tilde{c}_q,t_1} + \mathcal{L}\varepsilon_{\gamma,t_1}}{(1-\mathcal{L})\,\varepsilon_{\tilde{c}_q,t_2} + \mathcal{L}\varepsilon_{\gamma,t_2}}. \tag{32}$$

When the degree of market power is high ceteris paribus, i.e. $\mathcal{L}$ closed to 1, the comparative effect on price of changes in t_1 or t_2 depends mainly on how these policy instruments respectively impact γ, i.e. on the ratio $\varepsilon_{\gamma,t_1}/\varepsilon_{\gamma,t_2}$. In this situation, the main concern when explaining the relative impact on price is how policy instruments affect competition intensity measured through the effective market conduct index $\gamma\theta$. Conversely, when the degree of market power is low, i.e. $\mathcal{L}$ closed to 0, the comparative effect on price of changes in t_1 or t_2 depends mainly on how these policy instruments respectively impact the effective marginal cost, i.e. on the ratio $\varepsilon_{\tilde{c}_q,t_1}/\varepsilon_{\tilde{c}_q,t_2}$. Intuitively, how marginal cost is affected by regulation is more important in explaining price changes when market power is low.

Condition (31) in part (ii) suggests that a change in t is profit-enhancing if and only if its marginal benefit at the firm's level outweighs its marginal cost. To illustrate, consider that t is a tax that raises price and the effective AVC. Not surprisingly, when competition intensity is low, i.e. θ is close to 1, the marginal benefit tends to be low ceteris paribus as the benefit of raising price is expected to be small. Nevertheless, when the regulation makes profit less sensitive to price, i.e. γ is small, then it is less likely that the marginal benefit can be substantial. Finally, the tax is more likely to be profit enhancing when the industry is highly profitable (large $\mathcal{C}$-index) because then the impact of raising AVC is of less importance from a profit viewpoint. Similarly to the price impact of regulation, a direct implication of part (ii) is that when comparing two policy instruments, say t_1 and t_2 then the ratio of profit elasticities w.r.t. regulation is simply:

$$\frac{\varepsilon_{\pi,t_1}}{\varepsilon_{\pi,t_2}} = \frac{\gamma(1-\theta)\varepsilon_{p,t_1} - (1-\mathcal{C})\varepsilon_{\tilde{w},t_1}}{\gamma(1-\theta)\varepsilon_{p,t_2} - (1-\mathcal{C})\varepsilon_{\tilde{w},t_2}} \tag{33}$$

Ceteris paribus, when the degree of profitability is high, i.e. $\mathcal{C}$ close to 1, the comparative effect on profit of changes in t_1 or t_2 depends mainly on how these policy instruments respectively impact price, i.e. on the ratio $\varepsilon_{p,t_1}/\varepsilon_{p,t_2}$. Conversely, when the degree of profitability is low, i.e. $\mathcal{C}$ closed to 0, the comparative effect on profit of changes in t_1 or t_2 depends mainly on how these policy instruments respectively impact the effective AVC, i.e. on the ratio $\varepsilon_{\tilde{w},t_1}/\varepsilon_{\tilde{w},t_2}$.

Under CAVC, clearly $\varepsilon_{\tilde{c}_q,t} = \varepsilon_{\tilde{w},t}$ and $\mathcal{C} = \mathcal{L}$. [43] In this context, when the degree of market power/profitability is high ceteris paribus, (32) and (33) suggest that both the ratios of profit and price elasticities w.r.t. regulation mainly depend on how policy instruments respectively impact γ, i.e. on the ratio $\varepsilon_{\gamma,t_1}/\varepsilon_{\gamma,t_2}$. Conversely, ceteris paribus, when the degree of market power/profitability is low, (32) and (33) show that both the ratios of profit and price elasticities w.r.t. regulation mainly depend on how policy instruments respectively impact the effective marginal cost , i.e. on the ratio $\varepsilon_{\tilde{c}_q,t_1}/\varepsilon_{\tilde{c}_q,t_2}$.

Marginal excess burden, incidence and MVPF. To appreciate the welfare incidence of regulation, let us assume that the regulation $\boldsymbol{t}$ is composed only of taxes/subsidies to ease the exposition. Following Weyl and Fabinger (2013), let us consider the incidence of the policy instrument t as the ratio of marginal change in consumer surplus relative to marginal change in producer surplus:

$$I_t = \frac{dCS}{dt} / \frac{dPS}{dt}$$

where $CS = LU(x,n) - nPq$ is the consumers' surplus and $PS = n(\pi - f)$ denotes the producers' surplus. Also the marginal excess burden of t is the marginal change in welfare denoted $\frac{dW}{dt}$ where welfare W is taken as the sum of consumers' surplus, producers surplus and taxpayer surplus $TS = n(\tilde{w} - w)q$ and is given by:

$$\begin{aligned} W &= CS + PS + TS \\ &= LU(x,n) - nqw(q) - nf. \end{aligned}$$

[43]When marginal cost is non constant, then the two elasticities differ. Nevertheless, it is straightforward to show that $\varepsilon_{\tilde{c}_q,t}$ is a weighted average of $\varepsilon_{\tilde{w},t}$ and of the partial elasticity of the marginal AVC w.r.t t, i.e. $\varepsilon_{\tilde{w}_q,t}$:

$$\varepsilon_{\tilde{c}_q,t} = \frac{\varepsilon - 1}{\varepsilon - \lambda}\varepsilon_{\tilde{w},t} + \frac{1-\lambda}{\varepsilon - \lambda}\varepsilon_{\tilde{w}_q,t}.$$

Furthermore, let us consider the marginal value of public funds associated to t which is denoted $MVPF_t = -\frac{dW}{dt}/\frac{dTS}{dt}$.

As it will be clear in the following analysis, it is convenient to define the following notions of average and marginal regulation burden ratios. These ratios are actually useful to express the marginal excess burden of regulation and the associated marginal cost of public funds.

Definition 6. *Let $\chi = \frac{\tilde{c}_q - c_q}{p}$ be the "marginal regulation burden ratio", i.e. the regulation burden in terms of marginal cost increase per unit of price. Similarly, let $\phi = \frac{\tilde{w}-w}{p}$ be the "average regulation burden ratio", i.e. the ARB (average regulation burden) per unit of price.*

Importantly, the average regulation burden ratio ϕ and the marginal regulation burden ratio χ differ in general, except in some specific cases.[44] In the example of ad-valorem cum specific taxation, we have $\phi = \chi = t_v + \frac{t_s}{p}$, while in the cost subsidies example, we get $\phi = -\frac{\tau w + s}{p}$ and $\chi = -\frac{\tau c_q + s}{p}$ and $\phi \neq \chi$ unless marginal cost is constant (CAVC). In the exogenous competition example, we have $\chi = (1/p)\left[(w(q-\tilde{q}) - w(q))(1+q) - \tilde{q}w'(q-\tilde{q})\right]$ while $\phi = (1/p)\left[w(q-\tilde{q}) - w(q) - (\tilde{q}/q)(p - w(q-\tilde{q}))\right]$.

Proposition 12. *Holding the number of firms constant, the marginal excess burden, the incidence and the marginal value of public funds for the policy instrument t in the regulation system $\boldsymbol{t}$ are given respectively by:*

$$\frac{dW}{dt} = -Q\varepsilon_d\left(\mathcal{L} + \chi\right)\frac{dp}{dt} = p\left(\mathcal{L}+\chi\right)\frac{dQ}{dt}$$

and

$$I_t = \frac{1}{(1-\mathcal{C})\frac{\varepsilon_{\tilde{w},t}}{\varepsilon_{p,t}} - \gamma(1-\theta)} \qquad \text{and} \qquad MVPF_t = \frac{\varepsilon_d\left(\mathcal{L}+\chi\right)}{(1-\mathcal{C})\frac{\varepsilon_{\tilde{w},t}}{\varepsilon_{p,t}} + 1 - \gamma - \varepsilon_d\chi}.$$

Proof. See Appendix M. ■

Any price increase following the raise in t is welfare decreasing if and only if $\mathcal{L} + \chi > 0$. If $p < c_q$ then any price increase is welfare enhancing and conversely.

Under perfect competition, $\mathcal{L} = 0$ and thus $\frac{dW}{dt} = p\chi\frac{dQ}{dt}$. This formula nests in particular the classic formulation of Harberger (1964) for the marginal excess burden of a unit tax t_s. Indeed, when the regulation is only represented by a unit tax t_s, then $\chi = \frac{\tilde{c}_q - c_q}{p} = \frac{t_s}{p}$ and thus $\frac{dW}{dt_s} = t_s\frac{dQ}{dt_s}$. Moreover, under perfect competition, welfare is maximized under laissez-faire ($\chi = 0$).

Under imperfect competition, $\mathcal{L} > 0$ and thus welfare is maximized when the regulation is such that $\mathcal{L} + \chi = 0$. That is when the regulation is made of explicit or implicit subsidies which ensures

[44] Note that χ and ϕ are related through $\chi = \phi(1+\varepsilon_{\psi,q})$ where $\varepsilon_{\psi,q} = q\psi_q/\psi$ is the partial elasticity of the ARB w.r.t output q.

that $p = c_q$ at the equilibrium.

Proposition 12 also suggests that, similarly to the profit incidence, the way the price and the effective AVC change in t is crucial to evaluate the welfare and surplus incidence of the policy instrument under scrutiny. Indeed, when comparing two policy instruments, the relative incidence as well as the relative marginal value of public funds only depend on the evaluation of the ratio $\frac{\varepsilon_{\tilde{w},t}}{\varepsilon_{p,t}}$ across instruments, because both I_t and $MVPF_t$ are monotonic in this ratio. More precisely, I_t is always strictly decreasing in $\frac{\varepsilon_{\tilde{w},t}}{\varepsilon_{p,t}}$ and $MVPF_t$ is strictly decreasing in $\frac{\varepsilon_{\tilde{w},t}}{\varepsilon_{p,t}}$ if and only if $\mathcal{L} + \chi > 0$. In particular, in this context where the equilibrium price is too high from the point of view of well-being, a low relative pass-through $\varepsilon_{p,t}$ leads ceteris paribus to a low value of the MVPF associated with the policy instrument t.

Finally, note that when the degree of profitability is very high ceteris paribus, then the incidence tends to vary little from one policy instrument to another. The same remark applies to the marginal value of public funds.

Regulation incidence in the long run. To complete the analysis, I now turn to the long run impacts of regulation. For the sake of simplicity in the exposition, I assume that regulation does not change the fixed cost f.[45] Let us denote $\varepsilon_{p,t}^{SR} = APT\left(\mathcal{L}\varepsilon_{\gamma,t} + (1-\mathcal{L})\varepsilon_{\tilde{c}_q,t}\right)$ as the short run elasticity of price given in (30) but valued at the long run equilibrium.

Proposition 13. *At the long run equilibrium, regulation incidence is such that an increase in t*

(i) *raises output per firm if and only if* $P_n\gamma\varepsilon_{p,t}^{SR} > MP_n APT(1-\mathcal{C})\varepsilon_{\tilde{w},t}$,

(ii) *is procompetitive if and only if* $(1-\mathcal{C})\varepsilon_{\tilde{w},t} < (1-\theta)\gamma\varepsilon_{p,t}^{SR}$,

(iii) *raises price if and only if* $P_n\gamma\theta\varepsilon_{p,t}^{SR} < -\left[P_n - MP_n APT\right](1-\mathcal{C})\varepsilon_{\tilde{w},t}$

where $MP_n = \delta P_n + \gamma\theta x P_{xn} + (\gamma\theta)_n x P_x$ *is the partial derivative of marginal profit under regulation w.r.t. n.*

Proof. See Appendix N. ∎

In all three statements contained in the Proposition 13, the (short run) elasticity of price and the partial elasticity of effective AVC with respect to the policy instrument are the key elements for judging changes in, respectively, output, long-run price, and number of active firms. These statements

[45] It would be interesting to extend the analysis in this direction but this is devoted to future research.

constitute the generalization of parts (i) and (ii) of Proposition 7 to the regulated oligopoly when the regulation may not take the simple form of an additive cost shock like e.g. a unit tax.

In the long run, the marginal excess burden for element t of the regulation in the long run is given by:

$$\begin{aligned}\frac{dW}{dt} &= LU_x\frac{dx}{dt}+LU_n\frac{dn}{dt}-qw\frac{dn}{dt}-n(w+qw')\frac{dq}{dt}-f\frac{dn}{dt}\\ &= n(p-w-qw')\frac{dq}{dt}+(LU_n-qw-f)\frac{dn}{dt}.\end{aligned}$$

Note that dividing the gross markup $p-w-qw'$ by p allows to get $(p-w-qw')/p=\mathcal{L}+\chi$ by using the definitions of $\chi=(\tilde{c}_q-c_q)/p$ and of the Lerner index $\mathcal{L}=(p-\tilde{c}_q)/p$. Moreover, using $\phi=\frac{\tilde{w}-w}{p}$, it follows that the welfare gain from an additional variety net of its total cost, i.e. LU_n-qw-f, can be rewritten as $p(\mathcal{V}_0+\phi)$ by also using the zero profit condition, i.e. $(p-\tilde{w})q-f=0$ and the definition of the variety effect $\mathcal{V}_0$= $(LU_n-pq)/p$. Overall, the marginal excess burden writes:

$$\frac{dW}{dt}=\underbrace{np\left(\mathcal{L}+\chi\right)\frac{dq}{dt}}_{\text{quantity term}}+\underbrace{p(\mathcal{V}_0+\phi)\frac{dn}{dt}}_{\text{variety term}}. \tag{34}$$

The expression of the marginal excess burden associated to the policy instrument t in (34) can be viewed as a generalization to any regulation policy of the expression obtained by Kroft *et al.* (2021) in the case of ad-valorem cum specific taxation. As in Kroft *et al.*, the marginal excess burden is the combination of a quantity term and a variety term reflecting the distortionary effect of regulation respectively on output and on variety. First, on the intensive margin, the gross mark-up ratio $\mathcal{L}+\chi$ can be positive or negative and it is the distortionary wedge on output. When $\mathcal{L}+\chi$ is positive then any additional reduction in quantity created by the increase in t would be welfare decreasing. Second, on the extensive margin, the net welfare gain of an additional variety measured through $\mathcal{V}_0+\phi$ can also be positive or negative. When $\mathcal{V}_0+\phi$ is positive, there is too few entry and if t is anticompetitive any increase in t reduces welfare. Conversely, either on the intensive or extensive margins, regulation can increase welfare. Overall, in the long run, any optimal regulation $\boldsymbol{t}$ must be such that $\mathcal{L}+\chi=0$ and $\mathcal{V}_0+\phi=0$. In other words, an optimal regulation aims at neutralizing the negative consequences of market power expressed by $\mathcal{L}>0$ and at the same time, it aims to regulate the entry of varieties on the market so that the average regulation burden balances the variety effect.[46]

The next Proposition gathers not only the above expression on the marginal excess burden but also the expression for the marginal value of public funds for the policy instrument t.

[46]Equation (34) nests several well-known formulas. To be completed.

Proposition 14. *In the long run, the marginal excess burden and the marginal value of public funds for the policy instrument t in the regulation system $\boldsymbol{t}$ are given by respectively:*

$$\frac{dW}{dt} = np\left(\mathcal{L}+\chi\right)\frac{dq}{dt} + p(\mathcal{V}_0+\phi)\frac{dn}{dt}$$

and

$$MVPF_t = -\frac{Q\left(\mathcal{L}+\chi\right)\frac{\varepsilon_{q,t}}{\varepsilon_{n,t}} + n(\mathcal{V}_0+\phi)}{Q\left(\mathcal{L}+\chi-\frac{1}{\varepsilon_d}\right)\frac{\varepsilon_{q,t}}{\varepsilon_{n,t}} + n\left(\mathcal{V}_0+\phi-\mathcal{V}_1\right)}. \tag{35}$$

where $\phi = \frac{\tilde{w}-w}{p}$ is the ARB (average regulation burden) per unit of price.

Proof. See Appendix O. ■

Not surprisingly, similarly to the marginal excess burden, the MVPF in the long run now depends not only on χ but also on ϕ. Furthermore, I show in Appendix O that the ratio of elasticities $\frac{\varepsilon_{q,t}}{\varepsilon_{n,t}}$ is a function of the ratio $\frac{\varepsilon_{\tilde{w},t}}{\varepsilon_{p,t}}$:

$$\frac{\varepsilon_{q,t}}{\varepsilon_{n,t}} = \varepsilon_d \frac{nP_n}{p} \frac{(1-\mathcal{C})\frac{\varepsilon_{\tilde{w},t}}{\varepsilon_{p,t}} - \gamma}{(1-\mathcal{C})\frac{\varepsilon_{\tilde{w},t}}{\varepsilon_{p,t}} - \gamma\left(1-\theta\right)}.$$

This relationship is useful because, under Assumption 2, it shows that $\frac{\varepsilon_{q,t}}{\varepsilon_{n,t}}$ is decreasing in $\frac{\varepsilon_{\tilde{w},t}}{\varepsilon_{p,t}}$. Hence, similarly to what happens in the short run as described in Proposition 12, the marginal value of public funds associated to a policy instrument is a function on how this instrument influences both the market price and the effective AVC through the ratio $\frac{\varepsilon_{\tilde{w},t}}{\varepsilon_{p,t}}$.

Given this result, expression (35) is also useful to identify environments where the ranking between two policy instruments in the short run would be different in the long run when the extensive margin also matters. This would be a case of preferences reversal between the short run and the long run. To illustrate, consider two taxes, say t_1 and t_2, such that that when increased marginally would both reduce the equilibrium quantity per firm q in the short and in the long run. Assume also that when increased marginally, both policy instruments also reduce the equilibrium number of firms n in the long run. Furthermore, assume that the inequality $\frac{\varepsilon_{\tilde{w},t_1}}{\varepsilon_{p,t_1}} > \frac{\varepsilon_{\tilde{w},t_2}}{\varepsilon_{p,t_2}}$ holds both in the short run and in the long run. Finally, assume that, with the existing regulation, the short run equilibrium as well as the long run equilibrium are such that the market price is too high in the sense that $\mathcal{L}+\chi > 0$. And, there are too few varieties at the long run equilibrium in the sense that $\mathcal{V}_0+\phi > 0$.

From Proposition 12, the policy instrument t_1 is preferred in the short run to the policy instrument t_2 in the sense that $MVPF_{t_1} < MVPF_{t_2}$, because $\frac{\varepsilon_{\tilde{w},t_1}}{\varepsilon_{p,t_1}} > \frac{\varepsilon_{\tilde{w},t_2}}{\varepsilon_{p,t_2}}$. The latter inequality would induce some preference reversal in the long-run if and only if $MVPF_t$ is monotonically increasing in $\varepsilon_{\tilde{w},t}/\varepsilon_{p,t}$.[47]

[47] Or equivalently monotonically decreasing in $\varepsilon_{q,t}/\varepsilon_{n,t}$.

From (35), this holds if and only if:

$$nQ(\mathcal{V}_0 + \phi)\left(\mathcal{L} + \chi - \frac{1}{\varepsilon_d}\right) - nQ\left(\mathcal{L} + \chi\right)\left(\mathcal{V}_0 + \phi - \mathcal{V}_1\right) < 0.$$

Rearranging, we get the following condition:

$$(\mathcal{V}_0 + \phi)\frac{1}{\varepsilon_d} > \left(\mathcal{L} + \chi\right)\mathcal{V}_1. \tag{36}$$

Condition (36) reveals that for a given positive distortionary wedge on output as measured by $\mathcal{L} + \chi$, the tax t_2 is associated with a lower *MVPF* than for the tax t_1 when $\mathcal{V}_0 + \phi$ is sufficiently large, that is when there is a sufficiently large problem of too few entry at the long run equilibrium. Intuitively, the tax t_2 is favored as it is more favorable to entry than t_1, because $\frac{\varepsilon_{\tilde{w},t_1}}{\varepsilon_{p,t_1}} > \frac{\varepsilon_{\tilde{w},t_2}}{\varepsilon_{p,t_2}}$.

7 Conclusion

This paper has proposed an analysis of the symmetric oligopoly equilibrium using a new approach in terms of the margin over variable cost. It allows to take into account the main characteristics of demand and costs in a symmetric way to establish the equilibrium properties within a large class of oligopoly models. It also sheds light on several issues: (i) obtaining new formulations of short-run and long-run cost transmission, (ii) comparing the short-run and long-run effects of additive and multiplicative cost shocks, (iii) characterizing the short-run and long-run consequences of market expansion. The impact of market concentration is also characterized. Last, the contribution margin approach is extended to consider a regulated oligopoly and to show that the incidence and the welfare impact of a policy instrument can be easily characterized as a function of the ratio of price and average variable cost elasticities to regulation. Among other results, the analysis provides simple formulations of the marginal cost of public funds associated with different regulatory instruments and identifies a condition under which the order of preference between e.g. two taxes can be changed when moving from a short-run analysis (fixed number of firms) to a long-run analysis where the number of firms is endogenous.

Natural extensions of this modelling are as follows. First, considering multimarkets competition and introducing trade frictions between markets would help understand the comparative statics of trade equilibrium when affected by shocks on trade costs. Another potential and related extension is to derive conditions under which group pricing is welfare improving. Indeed, the existing literature (e.g. Aguirre et al., 2010, Cowan, 2016; Adachi, 2023) assumes the constancy of marginal cost which simplifies the analysis as the different decision problems of which price for which market are independent. This is no longer the case when marginal cost is non constant (Bontems, 2023b).

Second, it is worth investigating the case of oligopolies that are asymmetric. For instance, work in progress suggests that in an heterogenous industry according to cost structures and competing à la Cournot, the contribution margin helps to derive an extended version of the Herfindhal concentration index. Third, the case of multiproduct oligopolies (Armstrong and Vickers, 2022, 2018; Hamilton, 2009) can also be investigated to analyze cost pass-through in this relevant empirically context.

References

Adachi, T. (2023). A sufficient statistics approach for welfare analysis of oligopolistic third-degree price discrimination. *International Journal of Industrial Organization*, 86:102893.

Adachi, T. and Fabinger, M. (2022). Pass-through, welfare, and incidence under imperfect competition. *Journal of Public Economics*, 211:104589.

Aguirre, I., Cowan, S., and Vickers, J. (2010). Monopoly price discrimination and demand curvature. *American Economic Review*, 100(4):1601–1615.

Amir, R., De Castro, L., and Koutsougeras, L. (2014). Free entry versus socially optimal entry. *Journal of Economic Theory*, 154:112–125.

Anderson, S. P., De Palma, A., and Kreider, B. (2001a). The efficiency of indirect taxes under imperfect competition. *Journal of Public Economics*, 81(2):231–251.

Anderson, S. P., De Palma, A., and Kreider, B. (2001b). Tax incidence in differentiated product oligopoly. *Journal of Public Economics*, 81(2):173–192.

Armstrong, M. and Vickers, J. (2018). Multiproduct pricing made simple. *Journal of Political Economy*, 126(4):1444–1471.

Armstrong, M. and Vickers, J. (2022). Multiproduct cost passthrough: Edgeworth's paradox revisited.

Bertoletti, P. and Etro, F. (2016). Preferences, entry, and market structure. *The RAND Journal of Economics*, 47(4):792–821.

Bontems, P. (2019). Refunding emissions taxes: the case for a three-part policy. *The BE Journal of Economic Analysis & Policy*, 19(2):20180123.

Bontems, P. (2023a). Not so costly: Cost structure and firm behavior.

Bontems, P. (2023b). Trade frictions, group pricing and the contribution margin approach to the firm.

Bresnahan, T. F. (1989). Empirical studies of industries with market power. *Handbook of industrial organization*, 2:1011–1057.

Bresnahan, T. F. and Reiss, P. C. (1990). Entry in monopoly market. *The Review of Economic Studies*, 57(4):531–553.

Bresnahan, T. F. and Reiss, P. C. (1991). Entry and competition in concentrated markets. *Journal of political economy*, 99(5):977–1009.

Bulow, J. I. and Pfleiderer, P. (1983). A note on the effect of cost changes on prices. *Journal of political Economy*, 91(1):182–185.

Bykadorov, I., Gorn, A., Kokovin, S., and Zhelobodko, E. (2015). Why are losses from trade unlikely? *Economics Letters*, 129:35–38.

Cao, G., Jin, G. Z., Weng, X., and Zhou, L.-A. (2021). Market-expanding or market-stealing? competition with network effects in bike-sharing. *The RAND Journal of Economics*, 52(4):778–814.

Chen, Y. and Riordan, M. H. (2008). Price-increasing competition. *The RAND Journal of Economics*, 39(4):1042–1058.

Chen, Y. and Savage, S. J. (2011). The effects of competition on the price for cable modem internet access. *The Review of Economics and Statistics*, 93(1):201–217.

Cournot, A. (1960). Recherches sur les principes mathematiques de la theorie des richesses, translated into english by n. *Bacon as Researches in the Mathematical Principles of the Theory of Wealth, London: Haffner*.

Cowan, S. (2016). Welfare-increasing third-degree price discrimination. *The RAND Journal of Economics*, 47(2):326–340.

Delipalla, S. and Keen, M. (1992). The comparison between ad valorem and specific taxation under imperfect competition. *Journal of Public Economics*, 49(3):351–367.

Diewert, W. E. and Wales, T. J. (1987). Flexible functional forms and global curvature conditions. *Econometrica: Journal of the Econometric Society*, 55:43–68.

Dixit, A. K. and Stiglitz, J. E. (1977). Monopolistic competition and optimum product diversity. *The American economic review*, 67(3):297–308.

Genesove, D. and Mullin, W. P. (1998). Testing static oligopoly models: conduct and cost in the sugar industry, 1890-1914. *The RAND Journal of Economics*, pages 355–377.

Grabowski, H. G. and Vernon, J. M. (1992). Brand loyalty, entry, and price competition in pharmaceuticals after the 1984 drug act. *The journal of law and economics*, 35(2):331–350.

Häckner, J. and Herzing, M. (2016). Welfare effects of taxation in oligopolistic markets. *Journal of Economic Theory*, 163:141–166.

Hamilton, S. F. (2009). Excise taxes with multiproduct transactions. *American Economic Review*, 99(1):458–471.

Hamilton, S. F., Bontems, P., and Lepore, J. (2015). Oligopoly intermediation, relative rivalry and market conduct. *International journal of industrial organization*, 40:49–59.

Harberger, A. C. (1964). The measurement of waste. *The American Economic Review*, pages 58–76.

Klemperer, P. D. and Meyer, M. A. (1989). Supply function equilibria in oligopoly under uncertainty. *Econometrica: Journal of the Econometric Society*, pages 1243–1277.

Kroft, K., Laliberté, J.-W. P., Vizcaíno, R. L., and Notowidigdo, M. J. (2021). Efficiency and incidence of taxation with free entry and love-of-variety preferences. Technical report, National Bureau of Economic Research.

Kwon, S. Y., Ma, Y., and Zimmermann, K. (2023). 100 years of rising corporate concentration. *University of Chicago, Becker Friedman Institute for Economics Working Paper*, (2023-20).

Mahoney, N. and Weyl, E. G. (2017). Imperfect competition in selection markets. *Review of Economics and Statistics*, 99(4):637–651.

Mangin, S. (2022). When is competition price-increasing? the impact of expected competition on prices. Technical report, Working paper.

Mankiw, N. G. and Whinston, M. D. (1986). Free entry and social inefficiency. *The RAND Journal of Economics*, pages 48–58.

Miklos-Thal, J. and Shaffer, G. (2021). Pass-through as an economic tool: on exogenous competition, social incidence, and price discrimination. *Journal of Political Economy*, 129(1):323–335.

Mrázová, M. and Neary, J. P. (2017). Not so demanding: Demand structure and firm behavior. *American Economic Review*, 107(12):3835–3874.

Parenti, M., Sidorov, A. V., Thisse, J.-F., and Zhelobodko, E. V. (2017). Cournot, bertrand or chamberlin: toward a reconciliation. *International Journal of Economic Theory*, 13(1):29–45.

Pauly, M. V. and Satterthwaite, M. A. (1981). The pricing of primary care physicians services: a test of the role of consumer information. *The Bell Journal of Economics*, pages 488–506.

Perloff, J., Suslow, V., and Seguin, P. (2006). Higher prices from entry: Pricing of brand-name drugs. u of california, berkeley. Technical report, Competition Policy Working Paper No. CPC99-03.

Ramey, V. A. (1991). Nonconvex costs and the behavior of inventories. *Journal of Political Economy*, 99(2):306–334.

Ritz, R. A. (2015). *The simple economics of asymmetric cost pass-through*. JSTOR.

Ritz, R. A. et al. (2018). Oligopolistic competition and welfare. *Handbook of game theory & industrial organization*, 1:181–200.

Seade, J. (1980a). On the effects of entry. *Econometrica: Journal of the Econometric Society*, pages 479–489.

Seade, J. (1980b). The stability of cournot revisited. *Journal of Economic Theory*, 23(1):15–27.

Seade, J. (1985). Profitable cost increases and the shifting of taxation: equilibrium responses of markets in oligopoly. Technical report.

Shapiro, C. (1995). Mergers with differentiated products. *Antitrust*, 10:23.

Stahl, D. O. (1988). Bertrand competition for inputs and walrasian outcomes. *The American Economic Review*, pages 189–201.

Stern, N. (1987). The effects of taxation, price control and government contracts in oligopoly and monopolistic competition. *Journal of Public Economics*, 32(2):133–158.

Ward, M. B., Shimshack, J. P., Perloff, J. M., and Harris, J. M. (2002). Effects of the private-label invasion in food industries. *American Journal of Agricultural Economics*, 84(4):961–973.

Weyl, E. G. and Fabinger, M. (2013). Pass-through as an economic tool: Principles of incidence under imperfect competition. *Journal of Political Economy*, 121(3):528–583.

Appendix

A Examples of inverse demand functions satisfying Assumption 2

From $x = D(p, n)$ and $p = P(x, n)$, we have $D_n = -P_n/P_x$ and, as thanks to Assumption 1 $P_x < 0$, it is sufficient to check that $P_n \leq 0$.

Linear inverse demands. Consider the symmetric demand system based on linear inverse demands, $P_i(x_1, ..., x_n) = \alpha - \beta x_i - \gamma \sum_{j \neq i} x_j$ with $\beta \geq \gamma > 0$. It follows that under symmetry, $P(x, n) = \alpha - \beta x - \gamma(n-1)x$ and $P_n(x, n) = -\gamma x < 0$.

Inverse demands based on additive preferences. Consider the additive and quasi linear preferences given by $z + V\left(\sum_i u(x_i)\right)$ where z is the consumption of the outside good taken as the numeraire, $V(.)$ and $u(.)$ are increasing concave, so that inverse demand given by utility maximization under the budget constraint is $P_i(x_1, ..., x_n) = u'(x_i)V'\left(\sum_i u(x_i)\right)$. Under symmetry, define $P(x, n) = u'(x)V'(nu(x))$ and it follows that $P_n(x, n) = u'(x)u(x)V''(nu(x)) < 0$.

Inverse Logit demands. Let us posit the inverse demand function for variety i as a function of market shares given by:

$$\hat{P}_i(s_1, ...s_n) = \frac{1}{\alpha}\left[\beta - \log\frac{s_i}{s_0}\right]$$

where $s_i = x_i/(z + X)$ is the market share of variety i, $X = \sum_i x_i$ is the aggregate consumption of differentiated products per head and $s_0 + \sum_i s_i = 1$ with s_0 being the market share of the outside good consumed in quantity z, i.e. $s_0 = z/(z + X)$. Under symmetry, $s_i = s = x/(z + nx)$ for all i and we have

$$\hat{P}(s, n) = \frac{1}{\alpha}\left[\beta - \log\frac{s}{1 - ns}\right]$$

and it follows that $\hat{P}_n(s, n) < 0$.

B Relationships between revenue, cost, demand and AVC elasticities

In this section, I connect the elasticities measures for industry demand and AVC to the ones one can define for the revenue and cost. Let us start by the demand side. From the equilibrium revenue function (per firm) $r(q) \equiv qP(q/L, n)$, straightforward manipulations allow to derive the elasticity ε_r as a function of ε_d:

$$\varepsilon_r = \frac{qr'}{r} = \frac{\varepsilon_d - 1}{\varepsilon_d} \in (0, 1) \tag{37}$$

Now, let us define the *effective elasticity of revenue* by replacing ε_d with the effective elasticity of demand $\tilde{\varepsilon}_d$ in (37):

$$\tilde{\varepsilon}_r = \frac{\tilde{\varepsilon}_d - 1}{\tilde{\varepsilon}_d} = 1 - \mathcal{L}. \tag{38}$$

Given that at the equilibrium necessarily the Lerner index belongs to $(0, 1)$ then the equilibrium value of the effective elasticity of revenue $\tilde{\varepsilon}_r$ is lower than unity.

For the cost side, I similarly obtain the following relationships between the elasticity of cost ε_c and the elasticity of "supply" ε_s:

$$\varepsilon_c = \frac{qc'}{c} = \frac{\varepsilon_s + 1}{\varepsilon_s} > 0 \tag{39}$$

The consequence of this connection is that the contribution margin approach can be alternatively expressed in function of revenue and cost instead of demand and AVC. From (38) and (39) I respectively deduce that

$$\tilde{\varepsilon}_d = \frac{1}{1 - \tilde{\varepsilon}_r} > 1 \text{ and } \varepsilon_s = \frac{1}{\varepsilon_c - 1}.$$

Replacing in the expression of $\lambda = (\varepsilon_s + 1)/(\varepsilon_s + \tilde{\varepsilon}_d)$, I obtain

$$\lambda = \frac{\varepsilon_c (1 - \tilde{\varepsilon}_r)}{\varepsilon_c - \tilde{\varepsilon}_r}$$

As λ is strictly positive and $\tilde{\varepsilon}_r \in (0,1)$ at the equilibrium, this implies that $\varepsilon_c > \tilde{\varepsilon}_r$. Moreover, under concave VC, $\lambda > 1$ entails that $\varepsilon_c < 1$. Similarly under convex VC, $\lambda < 1$ entails that $\varepsilon_c > 1$. And finally, a constant marginal cost is equivalent to have ε_c constant and equals to unity as well as λ. The contribution margin ratio writes

$$\mathcal{C} = \frac{p - w}{p} = \frac{r - c}{r} = \frac{1}{\varepsilon}$$

with

$$\varepsilon = \lambda \tilde{\varepsilon}_d = \frac{\varepsilon_c}{\varepsilon_c - \tilde{\varepsilon}_r} > 1.$$

C The stability condition

We have

$$MP_q = \frac{1}{L}(1 + \theta + q\theta_q)P_x(x,n) + \frac{1}{L}\theta x P_{xx}(x,n) - 2w'(q) - qw''(q) \tag{40}$$

Dropping arguments for the sake of simplicity, and using the definition of ρ_d and ε_s, we get

$$MP_q = \frac{1}{L}(1 + \theta + q\theta_q)P_x - \frac{1}{L}\theta\rho_d P_x - (2 - \rho_s)\, w'$$

Recall that the first-order condition for firms writes $P - w + \theta x P_x - qw' = 0$. As $P - w = P/(\lambda\tilde{\varepsilon}_d) = P/(-\lambda P/\theta x P_x) = -\theta x P_x/\lambda$ and replacing in the first-order condition, we get

$$\frac{1}{L}\theta\left(1 - \frac{1}{\lambda}\right)P_x = w' \tag{41}$$

Replacing w' in (40) leads to

$$MP_q = \frac{1}{L}(1 + \theta + q\theta_q)P_x - \frac{1}{L}\theta\rho_d P_x - \frac{1}{L}\theta\left(1 - \frac{1}{\lambda}\right)P_x\,(2 - \rho_s)$$

and rearranging we finally get

$$MP_q = \frac{\theta P_x}{\lambda L}\left[\lambda\left(\frac{1}{\theta} - 1 + \varepsilon_{\theta,q}\right) + 2 - \rho\right]$$

D Cost pass-through formulas

APT derivation: The first-order conditions at the symmetric equilibrium with a specific tax t all write $P + \theta(q/L)P_x(q/L) = w(q) + qw'(q) + t$ and their total differentiation subsequently taken in $t = 0$ yields $MP_q dq = dt$ and MP_q takes the form given by (11) when $t = 0$. As $APT = \left.\frac{dp}{dt}\right|_{t=0} = \frac{1}{L}P_x \left.\frac{dq}{dt}\right|_{t=0} = \frac{P_x/L}{MP_q} > 0$ following the stability condition (12), we obtain the desired result.

RPT derivation: The first-order conditions at the symmetric equilibrium with an ad-valorem tax τ all write $P + \theta(q/L)P_x(q/L) = \tau w(q) + \tau qw'(q)$ and their total differentiation subsequently taken in $\tau = 1$ yields $MP_q dq = (w(q) + qw'(q))\, d\tau$. We thus have:

$$\frac{d\log q}{d\tau} = \frac{w(1 + 1/\varepsilon_s)}{qMP_q} = -\tilde{\varepsilon}_d \frac{w}{P} \frac{1 + 1/\varepsilon_s}{\frac{1}{\lambda}\left[\lambda\left(\frac{1}{\theta} - 1 + \varepsilon_{\theta,q}\right) + 2 - \rho\right]} \tag{42}$$

using $\varepsilon_s = w/qw'$, $\tilde{\varepsilon}_d = -P/\theta(q/L)P_x$ and (11). Recall that $\lambda\tilde{\varepsilon}_d = \varepsilon$ and that, from the definition of the $\mathcal{C}$-index in Proposition 1, we get $w/P = (\varepsilon - 1)/\varepsilon$. Also, from the definition of λ in Proposition 1, we have $\varepsilon_s = (\varepsilon - 1)/(1 - \lambda)$. Replacing all these terms in (42) and rearranging, we obtain that

$$\frac{d\log q}{d\tau} = -\frac{\varepsilon - \lambda}{\left[\lambda\left(\frac{1}{\theta} - 1 + \varepsilon_{\theta,q}\right) + 2 - \rho\right]}$$

and from

$$\left.\frac{d\log P}{d\log\tau}\right|_{\tau=1}=\frac{1}{P}\frac{dp}{d\tau}=\frac{P_x}{LP}\frac{dq}{d\tau}=-\frac{1}{\theta\tilde{\varepsilon}_d}\frac{d\log q}{d\log\tau} \tag{43}$$

we obtain the desired result:

$$\left.\frac{d\log P}{d\log\tau}\right|_{\tau=1}=\frac{\varepsilon-\lambda}{\varepsilon}\frac{\lambda}{\theta\,(2-\rho)+\lambda\,(1-\theta+\theta\varepsilon_{\theta,q})}>0.$$

Derivation for perfect competition/Bertrand: note that for constant θ, we have

$$APT=\frac{\lambda}{\theta(2-\rho)+\lambda\,(1-\theta)}$$

Under perfect competition/Bertrand with homogenous products, $\theta=0$ and when $\theta\to0$, then $\lambda\to0$. Also the denominator $\theta(2-\rho)+\lambda\,(1-\theta)$ tends towards 0. Applying L'Hospital rule, we get

$$\frac{\partial\lambda}{\partial\theta}=\frac{\partial\lambda}{\partial\tilde{\varepsilon}_d}\frac{\partial\tilde{\varepsilon}_d}{\partial\theta}=-\frac{(\varepsilon_s+1)}{(\varepsilon_s+\tilde{\varepsilon}_d)^2}\left(-\frac{\varepsilon_d}{\theta^2}\right)=\frac{(\varepsilon_s+1)\varepsilon_d}{(\theta\varepsilon_s+\varepsilon_d)^2}\underset{\theta\to0}{\to}\frac{\varepsilon_s+1}{\varepsilon_d}$$

and

$$\begin{aligned}\frac{\partial}{\partial\theta}\left(\theta(2-\rho)+\lambda\,(1-\theta)\right) &= 2-\rho+\theta\left(2-\frac{\partial\lambda}{\partial\theta}\rho_d+\frac{\partial\lambda}{\partial\theta}\rho_s\right)-\lambda+\frac{\partial\lambda}{\partial\theta}\,(1-\theta)\\ \underset{\theta\to0}{\to}2-\rho_s+\left.\frac{\partial\lambda}{\partial\theta}\right|_{\theta=0} &= 2-\rho_s+\frac{\varepsilon_s+1}{\varepsilon_d}\end{aligned}$$

We thus obtain that $APT\underset{\theta\to0}{\to}\frac{\left.\frac{\partial\lambda}{\partial\theta}\right|_{\theta=0}}{\left.\frac{\partial}{\partial\theta}(\theta(2-\rho)+\lambda(1-\theta))\right|_{\theta=0}}=\frac{\frac{\varepsilon_s+1}{\varepsilon_d}}{2-\rho_s+\frac{\varepsilon_s+1}{\varepsilon_d}}$ which simplifies into the formula in the text.

E Proof of Proposition 5

Consider the additive cost shock t. In this case, equilibrium profit writes $\pi=(P(x,n)-w(q)-t)q$. Differentiating totally, we have $d\pi=(\partial\pi/\partial q)dq+(\partial\pi/\partial t)dt$ and thus

$$\left.\frac{d\pi}{dt}\right|_{t=0}=(1-\theta)xP_x\left.\frac{dq}{dt}\right|_{t=0}-q=(1-\theta)xP_x\frac{1}{MP_q}-q=q\left[(1-\theta)APT-1\right]$$

and the result follows.

Now consider the multiplicative cost shock τ, then $\pi=(P(x,n)-\tau w(q))q$. Differentiating totally, we have $d\pi=(\partial\pi/\partial q)dq+(\partial\pi/\partial\tau)d\tau$ and the relative change of profit taken in $\tau=1$ is such that:

$$\begin{aligned}\left.\frac{\tau}{\pi}\frac{d\pi}{d\tau}\right|_{\tau=1} &= \frac{1}{\pi}(1-\theta)xP_x\left.\frac{dq}{d\tau}\right|_{\tau=1}-\frac{qw}{\pi}\\ &= \frac{1}{P-w}(1-\theta)xP_x\left.\left(\frac{\tau}{q}\frac{dq}{d\tau}\right)\right|_{\tau=1}-\frac{w}{P-w}\end{aligned}$$

Recall that

$$P-w=-\frac{\theta}{\lambda}xP_x$$

and

$$\frac{w}{P-w}=\varepsilon-1.$$

Replacing and rearranging, we get

$$\begin{aligned}\left.\frac{\tau}{\pi}\frac{d\pi}{d\tau}\right|_{\tau=1} &= -\frac{\lambda}{\theta}(1-\theta)\left.\left(\frac{\tau}{q}\frac{dq}{d\tau}\right)\right|_{\tau=1}-(\varepsilon-1)\\ &= (1-\theta)\varepsilon RPT-(\varepsilon-1)\end{aligned}$$

where the second line uses (43) and hence the result follows, recalling that $\mathcal{C}=1/\varepsilon$.

F Proof of Proposition 7

Consider first the system (22)-(23) corresponding to the additive cost shock. Differentiating it totally and taking $t = 0$ yields:

$$M \begin{pmatrix} dq \\ dn \end{pmatrix} = \begin{pmatrix} 1 \\ q \end{pmatrix} dt \tag{44}$$

Note that from (17) and (11) we have that $MP_q = \frac{P_x/L}{APT}$, and thus the matrix M can be rewritten as:

$$M = \begin{pmatrix} \frac{P_x/L}{APT} & MP_n \\ (1-\theta)xP_x & qP_n \end{pmatrix}.$$

Solving the linear system (44), we find that the entry effect is given by:

$$\frac{dn}{dt} = \frac{\frac{1}{APT}xP_x - (1-\theta)xP_x}{\frac{1}{APT}xP_xP_n - (1-\theta)xP_xMP_n} = \frac{1-(1-\theta)APT}{P_n - MP_n(1-\theta)APT}$$

where the denominator is negative following the stability condition (16). Indeed, using $MP_q = \frac{P_x/L}{APT}$ and rearranging, the second condition in (16) rewrites

$$P_n - MP_n(1-\theta)APT < 0.$$

It follows that the introduction of the tax is procompetitive if and only if $1 - (1-\theta)APT < 0$. Also the output effect is given by:

$$\frac{dq}{dt} = \frac{qP_n - qMP_n}{\frac{1}{APT}xP_xP_n - (1-\theta)xP_xMP_n} = \frac{APT}{P_x/L}\frac{P_n - MP_n}{P_n - MP_n(1-\theta)APT}$$

and t increases the output per firm if and only if $P_n - MP_n > 0$.

The long run price effect of the additive cost shock is evaluated through:

$$\begin{aligned} APT^{LR} &= \left.\frac{dp}{dt}\right|_{t=0} = \underbrace{\frac{P_x}{L}\left.\frac{dq}{dt}\right|_{t=0}}_{\text{Intensive margin}} + \underbrace{P_n\left.\frac{dn}{dt}\right|_{t=0}}_{\text{Extensive margin}} \\ &= APT\frac{P_n - MP_n}{P_n - MP_n(1-\theta)APT} + P_n\frac{1-(1-\theta)APT}{P_n - MP_n(1-\theta)APT} \\ &= 1 + \frac{\theta APT(P_n - MP_n)}{P_n - MP_n(1-\theta)APT} \end{aligned} \tag{45}$$

Now consider the multiplicative cost shock and the associated system of equations (24)-(25). Totally differentiating it and taking $\tau = 1$ yields:

$$M \begin{pmatrix} dq \\ dn \end{pmatrix} = \begin{pmatrix} w + qw' \\ qw \end{pmatrix} d\tau$$

and hence the entry effect is given by:

$$\frac{dn}{d\tau} = \frac{\frac{1}{APT}\frac{qw}{L}P_x - (1-\theta)\frac{q(w+qw')}{L}P_x}{\frac{1}{APT}\frac{q}{L}P_xP_n - (1-\theta)\frac{q}{L}P_xMP_n}$$

Recalling that $\varepsilon_s = qw'(q)/w(q) = (\varepsilon-1)/(1-\lambda)$ and rearranging yields

$$\frac{dn}{d\tau} = w\frac{1-(1-\theta)\frac{\varepsilon-\lambda}{\varepsilon-1}APT}{P_n - MP_n(1-\theta)APT}$$

Hence, the multiplicative cost shock increases the number of firms if and only if $\frac{\varepsilon-\lambda}{\varepsilon-1}APT > 1/(1-\theta)$ which is exactly the condition for having a profit-enhancing cost shock in the short run.

Also the output effect is given by:

$$\frac{dq}{d\tau}=\frac{(w+qw')qP_n-qwMP_n}{\frac{1}{APT}\frac{q}{L}P_xP_n-(1-\theta)\frac{q}{L}P_xMP_n}=w\frac{APT}{P_x/L}\frac{\frac{\varepsilon-\lambda}{\varepsilon-1}P_n-MP_n}{P_n-MP_n(1-\theta)APT}$$

The multiplicative cost shock τ increases the output per firm if and only if $\frac{\varepsilon-\lambda}{\varepsilon-1}P_n-MP_n>0$.

Finally, the long-run price effect of the multiplicative cost shock is evaluated through:

$$\begin{aligned}
RPT^{LR} &= \left.\frac{\tau}{P}\frac{dp}{d\tau}\right|_{\tau=1}=\left.\frac{P_x/L}{P}\frac{dq}{d\tau}\right|_{\tau=1}+\left.\frac{P_n}{P}\frac{dn}{d\tau}\right|_{\tau=1}\\
&= \frac{wAPT}{P}\frac{\frac{\varepsilon-\lambda}{\varepsilon-1}P_n-MP_n}{P_n-MP_n(1-\theta)APT}+\frac{wP_n}{P}\frac{1-(1-\theta)\frac{\varepsilon-\lambda}{\varepsilon-1}APT}{P_n-MP_n(1-\theta)APT}\\
&= \frac{\varepsilon-\lambda}{\varepsilon}\left(APT\frac{P_n-\frac{\varepsilon-1}{\varepsilon-\lambda}MP_n}{P_n-MP_n(1-\theta)APT}+P_n\frac{\frac{\varepsilon-1}{\varepsilon-\lambda}-(1-\theta)APT}{P_n-MP_n(1-\theta)APT}\right)
\end{aligned}$$

Rearranging and using (45), we obtain that:

$$RPT^{LR}=\frac{\varepsilon-\lambda}{\varepsilon}APT^{LR}+\frac{\lambda-1}{\varepsilon}(P_n-MP_nAPT)$$

Part (i) and (ii) of Proposition 7 follow from the above results. Part (iii) is obtained by noting that $\frac{\varepsilon-\lambda}{\varepsilon-1}>1$ under convex VC which implies that the multiplicative cost shock is more often pro-competitive. Also, consider first the case where $APT<\frac{1}{1-\theta}\frac{\varepsilon-1}{\varepsilon-\lambda}$. In this situation, both shocks are anticompetitive, but output per firm q decreases in the long run for $MP_n>P_n$ in the additive cost case while q decreases for $MP_n>\frac{\varepsilon-\lambda}{\varepsilon-1}P_n$ for the multiplicative one. As $\frac{\varepsilon-\lambda}{\varepsilon-1}>1$, it follows that q decreases more often under the multiplicative cost shock. Now consider that $\frac{1}{1-\theta}>APT>\frac{1}{1-\theta}\frac{\varepsilon-1}{\varepsilon-\lambda}$. In this situation, the multiplicative cost shock always reduces q whatever MP_n greater than $P_n/((1-\theta)APT)$ while the additive one reduces q only if $MP_n>P_n$. Hence, as $P_n>P_n/((1-\theta)APT)$, the same conclusion applies w.r.t. reduction in q. Finally, consider the last case where $APT>\frac{1}{1-\theta}$. In this situation, both cost shocks always reduce q whatever MP_n greater than $P_n/((1-\theta)APT)$. Overall, the multiplicative cost shock reduces q more often.

G Proof of Proposition 8

Part (i): The first order condition for a symmetric equilibrium writes

$$P(q/L,n)+\theta(q/L)P_x(q/L,n)-w(q)-qw'(q)=0 \tag{46}$$

Let us differentiate totally (46) to obtain:

$$P_xd(q/L)+(q/L)P_x\theta_qdq+\theta\left(P_x+(q/L)P_{xx}\right)d(q/L)=\left(2w'+qw''\right)dq$$

which becomes, using $d(q/L)=(1/L)dq-(q/L^2)dL$ and the definition of MP_q:

$$MP_qdq=(q/L^2)\left[P_x+\theta\left(P_x+(q/L)P_{xx}\right)\right]dL$$

Recalling that $MP_q=\frac{P_x/L}{APT}$ and rearranging, we get:

$$\varepsilon_{q,L}=\frac{L}{q}\frac{dq}{dL}=\left(1+\theta-\theta\rho_d\right)APT.$$

Part (ii): it follows from computing the market size elasticity of price as:

$$\begin{aligned}
\varepsilon_{p,L} &= \frac{L}{p}\left(\frac{dP(q/L,n)}{dL}\right)=\frac{LP_x}{p}\left(\frac{1}{L}\frac{dq}{dL}-\frac{q}{L^2}\right)\\
&= \frac{xP_x}{p}\left(\varepsilon_{q,L}-1\right)
\end{aligned}$$

and the desired result follows.

Part (iii): Differentiating profit totally, we obtain $d\pi = (\partial\pi/\partial q)dq + (\partial\pi/\partial L)dL$, where

$$\frac{\partial\pi}{\partial q} = xP_x + P - qw' - w \tag{47}$$

$$\frac{\partial\pi}{\partial L} = -(q^2/L^2)P_x \tag{48}$$

Substituting (46) into (47) yields

$$\frac{\partial\pi}{\partial q} = (1-\theta)xP_x$$

Hence, we have

$$\begin{aligned}\varepsilon_{\pi,L} &= (1-\theta)\frac{L}{\pi}xP_x\frac{dq}{dL} - \frac{L}{\pi}\frac{q^2}{L^2}P_x = \frac{L}{\pi}\frac{q^2P_x}{L^2}\left[-1+(1-\theta)\frac{L}{q}\frac{dq}{dL}\right] \\ &= \frac{xP_x}{P-w}\left[-1+(1-\theta)\varepsilon_{q,L}\right]\end{aligned}$$

From the first order condition, we also have

$$P - w = qw' - \theta(q/L)P_x = \frac{q}{L}\left[\frac{qw'}{(q/L)P_x} - \theta\right]P_x = \frac{q}{L}\left[\frac{qw'}{w}\frac{w}{P}\frac{P}{(q/L)P_x} - \theta\right]P_x$$

Recall that $\varepsilon_s = w/qw' = (\varepsilon-1)/(1-\lambda)$, $w/P = (\varepsilon-1)/\varepsilon$ and that $\varepsilon_d = -P/(q/L)P_x = \theta\varepsilon/\lambda$. Hence, replacing and rearranging, we have

$$P - w = -\theta\left[\frac{1-\lambda}{\lambda} + 1\right]xP_x$$

so that finally, the relative change in profit writes:

$$\varepsilon_{\pi,L} = -\frac{1}{\theta\left[\frac{1-\lambda}{\lambda}+1\right]}\left[-1+(1-\theta)\varepsilon_{q,L}\right] = \frac{\lambda}{\theta}\left[1-(1-\theta)\varepsilon_{q,L}\right].$$

Part (iv) follows straightforwardly from parts (i) and (iii).

H Proof of Proposition 9

Let us differentiate totally the system above to see how q and n change in L at the long run equilibrium. We get

$$M\begin{pmatrix} dq \\ dn \end{pmatrix} = \begin{pmatrix} \frac{q}{L^2}\left(P_x + \theta(P_x + \frac{q}{L}P_{xx})\right) \\ \frac{q^2}{L^2}P_x \end{pmatrix} dL \tag{49}$$

Observe that

$$\frac{q}{L^2}\left(P_x + \theta(P_x + \frac{q}{L}P_{xx})\right) = \frac{q}{L^2}P_x\left(1+\theta-\theta\rho_d\right) = \frac{q}{L^2}\frac{P_x}{APT_{ghost}}$$

Solving the system (49), we get

$$\begin{aligned}\varepsilon^{LR}_{q,L} &= \frac{APT}{APT_{ghost}}\frac{P_n - MP_nAPT_{ghost}}{P_n - MP_n(1-\theta)APT} \\ &= \varepsilon_{q,L}\frac{P_n - MP_nAPT_{ghost}}{P_n - MP_n(1-\theta)APT}\end{aligned} \tag{50}$$

Hence, the long run elasticity of output w.r.t market size is the short run elasticity factorized by $\frac{P_n - MP_nAPT_{ghost}}{P_n - MP_n(1-\theta)APT}$. And we have

$$\begin{aligned}\varepsilon^{LR}_{q,L} &> \varepsilon_{q,L} \Leftrightarrow \frac{P_n - MP_nAPT_{ghost}}{P_n - MP_n(1-\theta)APT} > 1 \Leftrightarrow P_n - MP_nAPT_{ghost} < P_n - MP_n(1-\theta)APT \\ &\Leftrightarrow -MP_nAPT_{ghost} < -MP_n(1-\theta)APT\end{aligned}$$

Moreover, we get

$$\varepsilon_{n,L} = \frac{xP_x}{nP_n}\left(1-(1-\theta)\varepsilon_{q,L}\right) \tag{51}$$

Finally, the impact on price is such that:

$$\frac{dp}{dL} = P_x\frac{dx}{dL} + P_n\frac{dn}{dL} = P_x\left(\frac{1}{L}\frac{dq}{dL} - \frac{q}{L^2}\right) + P_n\frac{dn}{dL}$$

and thus in terms of elasticities,

$$\varepsilon_{p,L} = \frac{1}{\varepsilon_d}\left(1-\varepsilon_{q,L}\right) + \frac{nP_n}{p}\varepsilon_{n,L} \tag{52}$$

Using (50) and (51) and replacing in (52), we get

$$\begin{aligned}
\varepsilon_{p,L} &= \frac{1}{\varepsilon_d}\left(1 - \frac{APT}{APT_{ghost}}\frac{P_n - MP_n APT_{ghost}}{P_n - MP_n(1-\theta)APT}\right) + \frac{nP_n}{p}\frac{xP_x}{nP_n}\left(1-(1-\theta)\frac{APT}{APT_{ghost}}\right) \\
&= \frac{1}{\varepsilon_d}\frac{APT}{APT_{ghost}}\left(1-\theta - \frac{P_n - MP_n APT_{ghost}}{P_n - MP_n(1-\theta)APT}\right)
\end{aligned}$$

Hence, if $P_n - MP_n APT_{ghost} > 0$ then $\varepsilon_{p,L} > 0$.

I Welfare incidence of market expansion

The welfare incidence of market expansion in the long run is given by:

$$\begin{aligned}
\frac{dW}{dL} &= \frac{d}{dL}\left(LU(x,n) - nqw(q) - nf\right) \\
&= U + LU_x\frac{dx}{dL} + LU_n\frac{dn}{dL} - qw(q)\frac{dn}{dL} - nc'(q)\frac{dq}{dL} - f\frac{dn}{dL}.
\end{aligned}$$

Using $U_x = np$ and $\frac{dx}{dL} = \frac{1}{L}\frac{dq}{dL} - \frac{q}{L^2}$ and replacing, we get

$$\frac{dW}{dL} = U - p\frac{Q}{L} + n(p - c'(q))\frac{dq}{dL} + (LU_n - qw - f)\frac{dn}{dL}$$

Using the zero-profit condition, $(p-w)q = f$, and the definition of consumer surplus, $CS = LU - pQ$, we get

$$\begin{aligned}
\frac{dW}{dL} &= \frac{CS}{L} + n(p-c'(q))\frac{dq}{dL} + (LU_n - pq)\frac{dn}{dL} \\
&= \frac{CS}{L} + np\mathcal{L}\frac{dq}{dL} + p\mathcal{V}_0\frac{dn}{dL} \\
&= \frac{CS}{L} + \frac{np}{L}\left(\mathcal{L}q\varepsilon_{q,L} + \mathcal{V}_0\varepsilon_{n,L}\right)
\end{aligned}$$

The welfare incidence of market expansion exceeds the additional consumer surplus, CS/L, if and only if $\mathcal{L}q\varepsilon_{q,L} + \mathcal{V}_0\varepsilon_{n,L} > 0$.

J Mode of competition under regulation

Consider the profit expression under regulation keeping the possibility of asymmetry in the demand system and/or costs:

$$\pi_i = (p_i - \tilde{w}_i(q_i, p_i, \boldsymbol{t}))q_i$$

Then the derivative of joint profits w.r.t. strategy σ_i for firm i is:

$$\frac{d}{d\sigma_i}\left(\sum_j \pi_j\right) = \sum_j q_j(1-\tilde{w}_{jp})\frac{dp_j}{d\sigma_i} + \sum_j (p_j - \tilde{w}_j - q_j\tilde{w}_{jq})\frac{dq_j}{d\sigma_i}$$

Denoting $\gamma_j = 1 - \tilde{w}_{jp}$ we can define

$$\theta_i = -\frac{\sum_j (p_j - \tilde{w}_j - q_j \tilde{w}_{jq}) \frac{dq_j}{d\sigma_i}}{\sum_j q_j \gamma_j \frac{dp_j}{d\sigma_i}}$$

Also profit maximization implies that

$$\frac{d\pi_i}{d\sigma_i} = 0 \Rightarrow q_i(1 - \tilde{w}_{ip})\frac{dp_i}{d\sigma_i} + m_i \frac{dq_i}{d\sigma_i} = 0$$

And hence,

$$\theta_i = \sum_j \frac{q_j \gamma_j \frac{dp_j}{d\sigma_j}}{\sum_j q_j \gamma_j \frac{dp_j}{d\sigma_i}} \frac{\frac{dq_j}{d\sigma_i}}{\frac{dq_j}{d\sigma_j}}$$

Under symmetry of demands and costs, $\gamma_j = \gamma$ and thus it disappears from the formula that gives θ. But, under asymmetry, this also occurs when a common ad valorem tax rate applies to all firms, i.e. $\gamma_j = 1 - t$ for any j. In these cases, θ or θ_i do not depend on the regulation directly, only indirectly through prices or quantities.

K The stability condition under regulation

At the equilibrium under regulation, the marginal profit with respect to output is such that:

$$MP = P + \gamma\theta x P_x - \tilde{w} - q\tilde{w}_q = 0 \tag{53}$$

Differentiating MP w.r.t. q yields:

$$MP_q = \frac{P_x}{L} + (\gamma\theta)_q x P_x + \gamma\theta \frac{P_x}{L} + \gamma\theta x \frac{P_{xx}}{L} - 2\tilde{w}_q - \tilde{w}_p \frac{P_x}{L} - q\tilde{w}_{qq} - q\tilde{w}_{qp} \frac{P_x}{L}$$

where $(\gamma\theta)_q$ designates the partial derivative of $\gamma\theta$ w.r.t. q. Collecting terms, using the definition of $\rho_d = -xP_{xx}/P_x$, $\tilde{\rho}_s = -q\tilde{w}_{qq}/\tilde{w}_q$ and recalling that $\delta = 1 - \tilde{w}_p - q\tilde{w}_{qp}$, MP_q becomes:

$$MP_q = \delta \frac{P_x}{L} + \gamma\theta \frac{P_x}{L}(1 - \rho_d) - \tilde{w}_q(2 - \tilde{\rho}_s) + (\gamma\theta)_q x P_x. \tag{54}$$

From the definition of $\mathcal{C}$, $P - \tilde{w} = P/(\lambda\tilde{\varepsilon}_d) = P/(-\lambda P/\gamma\theta x P_x) = -\gamma\theta x P_x/\lambda$ and replacing in (53), we get:

$$\left(1 - \frac{1}{\lambda}\right) \gamma\theta \frac{P_x}{L} = \tilde{w}_q \tag{55}$$

Replacing $\tilde{w}_q$ in (54) and collecting terms, we obtain

$$MP_q = \frac{\gamma\theta P_x}{\lambda L} \left[\lambda \left(\frac{\delta}{\gamma\theta} - 1 + \varepsilon_{\gamma\theta,q}\right) + 2 - \rho\right]$$

and the stability condition follows from the assumption $MP_q < 0$.

L Proof of Proposition 11

Let us differentiate totally the following equilibrium first-order condition:

$$P + \gamma\theta x P_x - \tilde{w} - q\tilde{w}_q = 0 \tag{56}$$

by considering a change in component t of the regulation vector $\boldsymbol{t}$ to obtain:

$$MP_q dq = [\tilde{w}_t + q\tilde{w}_{qt} - \gamma_t \theta x P_x]\, dt \tag{57}$$

where $MP_q = \frac{P_x/L}{APT}$ consistently with the definition of APT in (29). Also, from (55) we deduce that

$$\theta x P_x = \frac{1}{\gamma} \frac{\lambda}{\lambda - 1} q\tilde{w}_q$$

Replacing in (57) both in the denominator and the numerator, and rearranging, we get:

$$\frac{dq}{dt} = APT\frac{\tilde{w}_t + q\tilde{w}_{qt} - \gamma_t\theta x P_x}{P_x/L} = APT\frac{\tilde{w}_t + q\tilde{w}_{qt} - \frac{\lambda}{\lambda-1}\frac{\gamma_t}{\gamma}q\tilde{w}_q}{\frac{1}{\gamma\theta}\frac{\lambda}{\lambda-1}\tilde{w}_q}. \tag{58}$$

Forming the elasticity of output w.r.t t, we obtain:

$$\varepsilon_{q,t} = \frac{t}{q}\frac{dq}{dt} = -\gamma\theta APT\left(\frac{1-\lambda}{\lambda}t\left(\frac{\tilde{w}_t + q\tilde{w}_{qt}}{q\tilde{w}_q}\right) + t\frac{\gamma_t}{\gamma}\right) \tag{59}$$

Denoting $\varepsilon_{\tilde{c}_q,t} = \partial\log\tilde{c}_q/\partial\log t$ as the partial elasticity of the effective marginal cost w.r.t t, observe that the first term in the bracket of the RHS in (59) can be rewritten as follows:

$$\frac{1-\lambda}{\lambda}t\left(\frac{\tilde{w}_t + q\tilde{w}_{qt}}{q\tilde{w}_q}\right) = \frac{1-\lambda}{\lambda}\frac{t}{\tilde{c}_q}\frac{\partial\tilde{c}_q}{\partial t}\frac{\tilde{c}_q}{q\tilde{w}_q} = \frac{1-\lambda}{\lambda}\frac{t}{\tilde{c}_q}\frac{\partial\tilde{c}_q}{\partial t}(1+\tilde{\varepsilon}_s) = \frac{\varepsilon-\lambda}{\lambda}\varepsilon_{\tilde{c}_q,t} \tag{60}$$

recalling that $\tilde{\varepsilon}_s = (\varepsilon-1)/(1-\lambda)$. Finally, using (60) in (59) and denoting $\varepsilon_{\gamma,t} = \partial\log\gamma/\partial\log t$ is the partial elasticity of γ w.r.t t, (59) becomes

$$\varepsilon_{q,t} = -\gamma\theta APT\left(\frac{\varepsilon-\lambda}{\lambda}\varepsilon_{\tilde{c}_q,t} + \varepsilon_{\gamma,t}\right).$$

It follows that

$$\varepsilon_{p,t} = \frac{t}{P}\frac{dp}{dt} = \frac{q}{P}\frac{P_x}{L}\varepsilon_{q,t} = -\gamma\theta\frac{xP_x}{P}APT\left(\frac{\varepsilon-\lambda}{\lambda}\varepsilon_{\tilde{c}_q,t} + \varepsilon_{\gamma,t}\right)$$

Using $-\gamma\theta\frac{xP_x}{P} = 1/\tilde{\varepsilon}_d = \mathcal{L}$ we obtain that $\varepsilon_{p,t}/\varepsilon_{q,t} = -1/\varepsilon_d$ and the relative pass through on price is a weighted sum of $\varepsilon_{\tilde{c}_q,t}$ and $\varepsilon_{\gamma,t}$ (weighted by $1-\mathcal{L}$ and $\mathcal{L}$ respectively) factorized by APT:

$$\varepsilon_{p,t} = APT\left[(1-\mathcal{L})\,\varepsilon_{\tilde{c}_q,t} + \mathcal{L}\varepsilon_{\gamma,t}\right].$$

Given that profit writes $\pi = (P(x,n) - \tilde{w}(q,P(x,n),t))q$, total differentiation yields:

$$\begin{aligned} d\pi &= \frac{\partial\pi}{\partial q}dq + \frac{\partial\pi}{\partial t}dt \\ &= \left[\left(\frac{P_x}{L} - \tilde{w}_q - \tilde{w}_p\frac{P_x}{L}\right)q + P - \tilde{w}\right]dq - \tilde{w}_t q dt \end{aligned}$$

Using (56), we obtain

$$d\pi = (1-\theta)\gamma x P_x dq - \tilde{w}_t q dt$$

and the elasticity of profit w.r.t. t writes:

$$\varepsilon_{\pi,t} = \frac{t}{\pi}\frac{d\pi}{dt} = \frac{t(1-\theta)\gamma x P_x}{(P-\tilde{w})\,q}\frac{dq}{dt} - \frac{t\tilde{w}_t q}{(P-\tilde{w})q} \tag{61}$$

As above, using the definition of $\mathcal{C}$ gives $P - \tilde{w} = -\gamma\theta x P_x/\lambda$ and replacing in (61) yields:

$$\begin{aligned} \varepsilon_{\pi,t} &= -\frac{t(1-\theta)\gamma x P_x}{\gamma\theta x P_x/\lambda q}\frac{dq}{dt} - \frac{t\tilde{w}_t q}{(P-\tilde{w})q} \\ &= -\frac{\lambda\gamma(1-\theta)}{\gamma\theta}\varepsilon_{q,t} - \frac{\tilde{w}}{(P-\tilde{w})}\frac{t\tilde{w}_t}{\tilde{w}} \end{aligned}$$

As $\frac{\tilde{w}}{(P-\tilde{w})} = \varepsilon - 1$ and $\varepsilon_{q,t} = -\varepsilon_d\varepsilon_{p,t}$, this yields

$$\begin{aligned} \varepsilon_{\pi,t} &= \frac{\lambda\gamma(1-\theta)}{\gamma\theta}\varepsilon_d\varepsilon_{p,t} - (\varepsilon-1)\varepsilon_{\tilde{w},t} \\ &= \gamma(1-\theta)\varepsilon\varepsilon_{p,t} - (\varepsilon-1)\varepsilon_{\tilde{w},t} \end{aligned}$$

where I denote $\varepsilon_{\tilde{w},t} = \frac{t\tilde{w}_t}{\tilde{w}}$ as the partial elasticity of effective AVC w.r.t. t. It follows that $\varepsilon_{\pi,t} > 0$ if and only if

$$\gamma(1-\theta)\varepsilon_{p,t} > (\frac{\varepsilon-1}{\varepsilon})\varepsilon_{\tilde{w},t} = (1-\mathcal{C})\varepsilon_{\tilde{w},t}$$

M Proof of Proposition 12

Let us start with the marginal excess burden. We have

$$\frac{dW}{dt} = \frac{d}{dt}\left(LU(x,n) - nqw(q) - nf\right) = n(p - c_q)\frac{dq}{dt}. \tag{62}$$

Introducing $\tilde{c}_q$ in (62), using $dp/dt = (P_x/L)dq/dt$ and rearranging, we obtain

$$\begin{aligned}\frac{dW}{dt} &= n\left(\frac{p-\tilde{c}_q}{p} + \frac{\tilde{c}_q - c_q}{p}\right)\frac{p}{P_x/L}\frac{dp}{dt} \\ &= n\left(\mathcal{L} + \chi\right)\frac{p}{P_x/L}\frac{dp}{dt} \end{aligned} \tag{63}$$

Introducing q and using $\mathcal{L} = 1/\tilde{\varepsilon}_d = \gamma\theta/\varepsilon_d$ with $\varepsilon_d = -P/xP_x$, (63) writes

$$\frac{dW}{dt} = -\gamma\theta Q\left(1 + \frac{\chi}{\mathcal{L}}\right)\frac{dp}{dt}. \tag{64}$$

Now let us consider the incidence I_t. We have

$$\begin{aligned}\frac{dCS}{dt} &= \frac{d}{dt}\left(LU(x,n) - npq\right) = LU_x\frac{dx}{dt} - Q\frac{dp}{dt} - np\frac{dq}{dt} \\ &= -Q\frac{dp}{dt}\end{aligned} \tag{65}$$

using $U_x = np$. Moreover,

$$\begin{aligned}\frac{dPS}{dt} &= \frac{d}{dt}\left(n\left(\pi - f\right)\right) = n\frac{\pi}{t}\varepsilon_{\pi,t} \\ &= \frac{Q(p-\tilde{w})}{t}\left[\gamma(1-\theta)\varepsilon\varepsilon_{p,t} - (\varepsilon - 1)\varepsilon_{\tilde{w},t}\right]\end{aligned} \tag{66}$$

where the last line follows from Proposition 11. Hence,

$$\begin{aligned}I_t &= -\frac{Q\frac{dp}{dt}}{\frac{Q(p-\tilde{w})}{t}\left[\gamma(1-\theta)\varepsilon\varepsilon_{p,t} - (\varepsilon-1)\varepsilon_{\tilde{w},t}\right]} \\ &= \frac{\varepsilon_{p,t}}{(1-\mathcal{C})\varepsilon_{\tilde{w},t} - \gamma(1-\theta)\varepsilon_{p,t}}.\end{aligned}$$

Finally, using $\frac{dW}{dt} = \frac{dCS}{dt} + \frac{dPS}{dt} + \frac{dTS}{dt}$, we get

$$MVPF_t = -\frac{\frac{dW}{dt}}{\frac{dTS}{dt}} = -\frac{\frac{dW}{dt}}{\frac{dW}{dt} - \frac{dCS}{dt} - \frac{dPS}{dt}} \tag{67}$$

Using (64), (65) and (66) and replacing in (67), the marginal value of public funds attached to t writes:

$$MVPF_t = \frac{\gamma\theta Q\left(1+\frac{\chi}{\mathcal{L}}\right)\frac{dp}{dt}}{-\gamma\theta Q\left(1+\frac{\chi}{\mathcal{L}}\right)\frac{dp}{dt} + Q\frac{dp}{dt} - \frac{Q(p-\tilde{w})}{t}\left[\gamma(1-\theta)\varepsilon\varepsilon_{p,t} - (\varepsilon-1)\varepsilon_{\tilde{w},t}\right]}$$

Multiplying by t/p the numerator and the denominator to introduce $\varepsilon_{p,t}$, simplifying by Q and rearranging, $MVPF_t$ writes

$$\begin{aligned}MVPF_t &= \frac{\gamma\theta\left(1+\frac{\chi}{\mathcal{L}}\right)}{-\gamma\theta\left(1+\frac{\chi}{\mathcal{L}}\right) + 1 - \left[\gamma(1-\theta) - (1-\mathcal{C})\frac{\varepsilon_{\tilde{w},t}}{\varepsilon_{p,t}}\right]} \\ &= \frac{\gamma\theta\left(1+\frac{\chi}{\mathcal{L}}\right)}{(1-\mathcal{C})\frac{\varepsilon_{\tilde{w},t}}{\varepsilon_{p,t}} + 1 - \gamma - \gamma\theta\frac{\chi}{\mathcal{L}}}.\end{aligned}$$

N Proof of Proposition 13

Once again, the system to be totally differentiated is composed of the first-order condition describing the behavior of firms and the zero profit condition:

$$\begin{aligned} P + \gamma\theta x P_x - \tilde{w} - q\tilde{w}_q &= 0 && (68) \\ (P - \tilde{w})q - f &= 0. && (69) \end{aligned}$$

Total differentiation yields:

$$\begin{pmatrix} MP_q & \frac{\partial}{\partial n}\left(P + \gamma\theta x P_x - \tilde{w} - q\tilde{w}_q\right) \\ \frac{\partial}{\partial q}\left[(P - \tilde{w})q\right] & q\frac{\partial}{\partial n}(P - \tilde{w}) \end{pmatrix} \begin{pmatrix} dq \\ dn \end{pmatrix} = -\begin{pmatrix} \frac{\partial}{\partial t}\left(P + \gamma\theta x P_x - \tilde{w} - q\tilde{w}_q\right) \\ \frac{\partial}{\partial t}\left[(P - \tilde{w})q - f\right] \end{pmatrix} dt \tag{70}$$

Let us denote MP_n as the partial derivative w.r.t n of the marginal profit w.r.t output given by (68):

$$\begin{aligned} MP_n &\equiv \frac{\partial}{\partial n}\left(P + \gamma\theta x P_x - \tilde{w} - q\tilde{w}_q\right) = P_n + \gamma\theta x P_{xn} + (\gamma\theta)_n x P_x - \tilde{w}_p P_n - q\tilde{w}_{qp} P_n \\ &= \delta P_n + \gamma\theta x P_{xn} + (\gamma\theta)_n x P_x \end{aligned}$$

This is the analogous quantity to MP_n defined earlier in (15), with now γ and δ being involved in the expression.

Similarly, let us denote MP_t as the partial derivative of MP w.r.t. t :

$$MP_t \equiv \frac{\partial}{\partial t}\left(P + \gamma\theta x P_x - \tilde{w} - q\tilde{w}_q\right) = \gamma_t \theta x P_x - \tilde{w}_t - q\tilde{w}_{qt}$$

Also the term $\frac{\partial}{\partial q}\left[(P - \tilde{w})q\right]$ can be rewritten as:

$$\begin{aligned} \frac{\partial}{\partial q}\left[(P - \tilde{w})q\right] &= \gamma\frac{q}{L}P_x - q\tilde{w}_q + P - \tilde{w} \\ &= \gamma(1-\theta)xP_x \end{aligned}$$

using (68).

Furthermore, the term $q\frac{\partial}{\partial n}(P - \tilde{w})$ can be rewritten as:

$$q\frac{\partial}{\partial n}(P - \tilde{w}) = q\left(P_n - \tilde{w}_p P_n\right) = \gamma q P_n$$

Finally, we have

$$\frac{\partial}{\partial t}\left[(P - \tilde{w})q - f\right] = -q\tilde{w}_t$$

Hence solving the system (70), we obtain:

$$\frac{dq}{dt} = -\frac{\gamma q P_n MP_t + q\tilde{w}_t MP_n}{MP_q \gamma q P_n - \gamma(1-\theta)xP_x MP_n}$$

and

$$\frac{dn}{dt} = \frac{q\tilde{w}_t MP_q + \gamma(1-\theta)xP_x MP_t}{MP_q \gamma q P_n - \gamma(1-\theta)xP_x MP_n}.$$

Note that the long run stability conditions imply that

$$\begin{aligned} MP_q + \gamma q P_n &< 0 \\ MP_q \gamma q P_n - \gamma(1-\theta)xP_x MP_n &> 0 \end{aligned}$$

Using $MP_q = \frac{P_x}{L} APT^{-1}$, the latter condition rewrites

$$P_n - MP_n(1-\theta)APT < 0.$$

Furthermore, the term MP_t can be rewritten as follows:

$$\begin{aligned} MP_t &= \gamma_t \theta x P_x - \tilde{w}_t - q\tilde{w}_{qt} \\ &= \frac{\gamma_t}{\gamma}\frac{\lambda}{\lambda-1} q\tilde{w}_q - (\tilde{w}+q\tilde{w}_q)\frac{\tilde{w}_t + q\tilde{w}_{qt}}{\tilde{w}+q\tilde{w}_q} \end{aligned}$$

using once again (55). It follows that

$$\begin{aligned} MP_t &= \frac{1}{t}\left(\varepsilon_{\gamma,t}\frac{\lambda}{\lambda-1} q\tilde{w}_q - (\tilde{w}+q\tilde{w}_q)\varepsilon_{\tilde{c}_q,t}\right) \\ &= \frac{\tilde{w}}{t}\left(\varepsilon_{\gamma,t}\frac{\lambda}{\lambda-1}\frac{1}{\tilde{\varepsilon}_s} - (1+\frac{1}{\tilde{\varepsilon}_s})\varepsilon_{\tilde{c}_q,t}\right) \end{aligned}$$

and using $\tilde{\varepsilon}_s = (\varepsilon-1)/(1-\lambda)$, and rearranging, we obtain that:

$$MP_t = -\frac{\tilde{w}}{t}\frac{1}{\varepsilon-1}\left(\lambda\varepsilon_{\gamma,t} + (\varepsilon-\lambda)\varepsilon_{\tilde{c}_q,t}\right)$$

Finally, we obtain the output effect as:

$$\begin{aligned} \frac{dq}{dt} &= -\frac{APT}{\gamma P_x/L}\frac{\tilde{w}_t MP_n + \gamma P_n MP_t}{P_n - MP_n(1-\theta)APT} \\ &= \frac{\tilde{w}}{t}\frac{1}{\varepsilon-1}\frac{APT}{\gamma P_x/L}\frac{\gamma P_n\left(\lambda\varepsilon_{\gamma,t}+(\varepsilon-\lambda)\varepsilon_{\tilde{c}_q,t}\right) - (\varepsilon-1)MP_n\varepsilon_{\tilde{w},t}}{P_n - MP_n(1-\theta)APT} \end{aligned}$$

and forming the elasticity, we get, recalling that $\frac{\tilde{w}}{p} = 1-\mathcal{C} = (\varepsilon-1)/\varepsilon$ and $\varepsilon_d = -p/xP_x$:

$$\varepsilon_{q,t} = -\frac{\theta APT}{\mathcal{L}}\frac{\gamma P_n\left(\mathcal{L}\varepsilon_{\gamma,t}+(1-\mathcal{L})\varepsilon_{\tilde{c}_q,t}\right) - MP_n(1-\mathcal{C})\varepsilon_{\tilde{w},t}}{P_n - MP_n(1-\theta)APT}. \tag{71}$$

Hence $\varepsilon_{q,t} > 0$ if and only if

$$\gamma P_n\left(\mathcal{L}\varepsilon_{\gamma,t}+(1-\mathcal{L})\varepsilon_{\tilde{c}_q,t}\right) > MP_n(1-\mathcal{C})\varepsilon_{\tilde{w},t}$$

If I denote $\varepsilon_{p,t}^{SR} = APT\left(\mathcal{L}\varepsilon_{\gamma,t}+(1-\mathcal{L})\varepsilon_{\tilde{c}_q,t}\right)$ as the short run elasticity of price but valued at the long run equilibrium, then this condition can be rewritten as in part (i) of the Proposition.

Similarly, the impact of t on entry is given by:

$$\begin{aligned} \frac{dn}{dt} &= \frac{1}{\gamma}\frac{\tilde{w}_t + \gamma MP_t(1-\theta)APT}{P_n - MP_n(1-\theta)APT} \\ &= \frac{1}{\gamma}\frac{\tilde{w}_t + \gamma(1-\theta)APT\left(-\frac{\tilde{w}}{t}\frac{1}{\varepsilon-1}\left(\lambda\varepsilon_{\gamma,t}+(\varepsilon-\lambda)\varepsilon_{\tilde{c}_q,t}\right)\right)}{P_n - MP_n(1-\theta)APT} \end{aligned}$$

and forming the corresponding elasticity we get

$$\begin{aligned} \varepsilon_{n,t} &= \frac{p}{n\gamma}\frac{(1-\mathcal{C})\varepsilon_{\tilde{w},t} - \gamma(1-\theta)APT\left(\mathcal{L}\varepsilon_{\gamma,t}+(1-\mathcal{L})\varepsilon_{\tilde{c}_q,t}\right)}{P_n - MP_n(1-\theta)APT} \\ &= \frac{p}{n\gamma}\frac{(1-\mathcal{C})\varepsilon_{\tilde{w},t} - (1-\theta)\gamma\varepsilon_{p,t}^{SR}}{P_n - MP_n(1-\theta)APT} \end{aligned} \tag{72}$$

using the above definition of $\varepsilon_{p,t}^{SR}$. Consequently, $\varepsilon_{n,t} > 0$ if and only if $(1-\mathcal{C})\varepsilon_{\tilde{w},t} < (1-\theta)\gamma\varepsilon_{p,t}^{SR}$ as indicated in part (ii) of the Proposition.

The impact of t on price is given by:

$$\varepsilon_{p,t} = \frac{t}{p}\frac{dp}{dt} = \frac{t}{p}\frac{P_x}{L}\frac{dq}{dt} + \frac{t}{p}P_n\frac{dn}{dt} = \frac{xP_x}{p}\varepsilon_{q,t} + \frac{nP_n}{p}\varepsilon_{n,t}$$

which using (71) and (72) rewrites

$$\begin{aligned}
\varepsilon_{p,t} &= \frac{APT}{\gamma}\left(\frac{\gamma P_n\left(\mathcal{L}\varepsilon_{\gamma,t}+(1-\mathcal{L})\varepsilon_{\tilde{c}_q,t}\right)-MP_n(1-\mathcal{C})\varepsilon_{\tilde{w},t}}{P_n-MP_n(1-\theta)APT}\right)+\frac{P_n}{\gamma}\left(\frac{(1-\mathcal{C})\varepsilon_{\tilde{w},t}-\gamma(1-\theta)\varepsilon_{p,t}^{SR}}{P_n-MP_n(1-\theta)APT}\right)\\
&= \frac{1}{\gamma}\frac{P_n\gamma\theta\varepsilon_{p,t}^{SR}+\left[P_n-MP_nAPT\right](1-\mathcal{C})\varepsilon_{\tilde{w},t}}{P_n-MP_n(1-\theta)APT}
\end{aligned}$$

by using the definition of $\varepsilon_{p,t}^{SR}$. It follows that $\varepsilon_{p,t}>0$ if and only if

$$P_n\gamma\theta\varepsilon_{p,t}^{SR}+\left[P_n-MP_nAPT\right](1-\mathcal{C})\varepsilon_{\tilde{w},t}<0$$

O Proof of Proposition 14

Let us start with the derivation of consumer surplus:

$$\begin{aligned}
\frac{dCS}{dt} &= \frac{d}{dt}(LU-npq)\\
&= LU_x\frac{dx}{dt}+LU_n\frac{dn}{dt}-pq\frac{dn}{dt}-nq\frac{dp}{dt}-np\frac{dq}{dt}\\
&= p\mathcal{V}_0\frac{dn}{dt}-nq\frac{dp}{dt}
\end{aligned}$$

using $U_x=np$ and $\mathcal{V}_0=(LU_n-pq)/p$.

Also we obtain for the producer surplus:

$$\begin{aligned}
\frac{dPS}{dt} &= \frac{d}{dt}(n(\pi-f))\\
&= n\frac{d\pi}{dt}+(\pi-f)\frac{dn}{dt}\\
&= 0
\end{aligned}$$

using the zero profit condition $\pi-f=0$ that implies $\frac{d\pi}{dt}=0$. Furthermore, $\frac{d\pi}{dt}=0$ produces a relationship between the price and the quantity effect:

$$\begin{aligned}
\frac{d\pi}{dt} &= \frac{d}{dt}((p-\tilde{w})q)=q\frac{dp}{dt}-q\left(\tilde{w}_q\frac{dq}{dt}+\tilde{w}_p\frac{dp}{dt}+\tilde{w}_t\right)+(p-\tilde{w})\frac{dq}{dt}\\
&= \gamma q\frac{dp}{dt}+(p-\tilde{w}-q\tilde{w}_q)\frac{dq}{dt}-q\tilde{w}_t\\
&= \gamma q\frac{dp}{dt}+p\mathcal{L}\frac{dq}{dt}-q\tilde{w}_t
\end{aligned}$$

using $\gamma=1-\tilde{w}_p$ and $\mathcal{L}=(p-\tilde{w}-q\tilde{w}_q)/p$. Hence,

$$\gamma Q\frac{dp}{dt}+np\mathcal{L}\frac{dq}{dt}-Q\tilde{w}_t=0$$

or equivalently in terms of elasticities obtained by multiplying by t/p:

$$\gamma\varepsilon_{p,t}+\mathcal{L}\varepsilon_{q,t}-(1-\mathcal{C})\varepsilon_{\tilde{w},t}=0$$

using $t\tilde{w}_t/p=(\tilde{w}/p)\varepsilon_{\tilde{w},t}=(1-\mathcal{C})\varepsilon_{\tilde{w},t}$.

Totally differentiating $p=P(x,n)$ also yields:

$$\begin{aligned}
\frac{dp}{dt} &= \frac{P_x}{L}\frac{dq}{dt}+P_n\frac{dn}{dt}\\
\varepsilon_{p,t} &= -\frac{\varepsilon_{q,t}}{\varepsilon_d}+\frac{nP_n}{p}\varepsilon_{n,t}
\end{aligned}$$

These two equations allows to obtain:

$$\begin{aligned}\varepsilon_{q,t} &= \frac{1}{\mathcal{L}}\left((1-\mathcal{C})\varepsilon_{\tilde{w},t}-\gamma\varepsilon_{p,t}\right)\\ \varepsilon_{n,t} &= \frac{p}{\gamma\theta nP_n}\left((1-\mathcal{C})\varepsilon_{\tilde{w},t}-\gamma\left(1-\theta\right)\varepsilon_{p,t}\right)\end{aligned}$$

and thus

$$\frac{\varepsilon_{q,t}}{\varepsilon_{n,t}}=\varepsilon_d\frac{nP_n}{p}\frac{(1-\mathcal{C})\frac{\varepsilon_{\tilde{w},t}}{\varepsilon_{p,t}}-\gamma}{(1-\mathcal{C})\frac{\varepsilon_{\tilde{w},t}}{\varepsilon_{p,t}}-\gamma\left(1-\theta\right)}$$

Under Assumption 2, it is straightforward to check that $\frac{\varepsilon_{q,t}}{\varepsilon_{n,t}}$ is decreasing in $\frac{\varepsilon_{\tilde{w},t}}{\varepsilon_{p,t}}$ as the ratio above is increasing in $\frac{\varepsilon_{\tilde{w},t}}{\varepsilon_{p,t}}$.

It follows that $MVPF_t$ is given by:

$$\begin{aligned}MVPF_t &= -\frac{dW}{dt}/\frac{dTS}{dt}=-\frac{dW}{dt}/(\frac{dW}{dt}-\frac{dCS}{dt}-\frac{dPS}{dt})\\ &= -\frac{np\left(\mathcal{L}+\chi\right)\frac{dq}{dt}+p(\mathcal{V}_0+\phi)\frac{dn}{dt}}{np\left(\mathcal{L}+\chi\right)\frac{dq}{dt}+p(\mathcal{V}_0+\phi)\frac{dn}{dt}-p\mathcal{V}_0\frac{dn}{dt}+nq\frac{dp}{dt}}\\ &= -\frac{np\left(\mathcal{L}+\chi\right)\frac{dq}{dt}+p(\mathcal{V}_0+\phi)\frac{dn}{dt}}{np\left(\mathcal{L}+\chi\right)\frac{dq}{dt}+p\phi\frac{dn}{dt}+nq\frac{dp}{dt}}\\ &= -\frac{Q\left(\mathcal{L}+\chi\right)\varepsilon_{q,t}+n(\mathcal{V}_0+\phi)\varepsilon_{n,t}}{Q\left(\mathcal{L}+\chi\right)\varepsilon_{q,t}+n\phi\varepsilon_{n,t}+Q\varepsilon_{p,t}}\\ &= -\frac{Q\left(\mathcal{L}+\chi\right)\varepsilon_{q,t}+n(\mathcal{V}_0+\phi)\varepsilon_{n,t}}{Q\left(\mathcal{L}+\chi\right)\varepsilon_{q,t}+n\phi\varepsilon_{n,t}+Q\left(-\frac{\varepsilon_{q,t}}{\varepsilon_d}+\frac{nP_n}{p}\varepsilon_{n,t}\right)}\\ &= -\frac{Q\left(\mathcal{L}+\chi\right)\varepsilon_{q,t}+n(\mathcal{V}_0+\phi)\varepsilon_{n,t}}{Q\left(\frac{\gamma\theta-1}{\gamma\theta}\mathcal{L}+\chi\right)\varepsilon_{q,t}+n\left(\frac{QP_n}{p}+\phi\right)\varepsilon_{n,t}}\end{aligned}$$

Recall that we have:

$$CS_n=\frac{\partial}{\partial n}\left(LU(x,n)-nqP(x,n)\right)=LU_n-pq-QP_n$$

Observe also that because $U_{nx}=U_x/n+nP_n=p+nP_n$ then

$$CS_n=LU_n\left(1-\frac{xU_{nx}}{U_n}\right).$$

Therefore, $\frac{QP_n}{p}=\mathcal{V}_0-\mathcal{V}_1$ and it follows that:

$$MVPF_t=-\frac{Q\left(\mathcal{L}+\chi\right)\frac{\varepsilon_{q,t}}{\varepsilon_{n,t}}+n(\mathcal{V}_0+\phi)}{Q\left(\mathcal{L}+\chi-\frac{1}{\varepsilon_d}\right)\frac{\varepsilon_{q,t}}{\varepsilon_{n,t}}+n\left(\mathcal{V}_0+\phi-\mathcal{V}_1\right)}.$$