ORIGINAL ARTICLE

Check for
updates

Two extensions of consumer surplus

Luis C. Corchón¹ · Ramón J. Torregrosa²

Abstract

We study consumer surplus in a single market when (a) there is a lower bound in the consumption of the outside good and (b) the weights in the social welfare function given to consumers and firms are different. We assume quasilinear utility. When the lower bound constraint on the consumption of the outside good is binding, income effects arise in demand. In some cases, Cournot equilibrium output is below equilibrium output without this constraint because the constraint makes demand less elastic. When the weights given to consumers and firms are not identical, social welfare is not necessarily concave and profits might be negative at the unrestricted optimum. We characterize social welfare optimum with a bound on maximum losses in a class of utility functions. We offer a formula to find the percentage of welfare losses due to oligopoly in this case.

1 Introduction

Since Dupuit (1844) consumer surplus has become a popular method for measuring the social optimality of allocations. It has been used to measure the deadweight loss of monopoly (Harberger 1954), in cost–benefit (Prest and Turvey 1965), macroeconomics (Mankiw 1985), and international trade (Dixit 1984) among many other fields.

The use of consumer surplus has been critiqued from many angles, and the most prevalent criticism is that it requires that either utility is quasilinear in the outside

✉ Ramón J. Torregrosa
rtorregr@usal.es

¹ Department of Economics, Universidad Carlos III de Madrid, Madrid, Spain

² Department of Economic and Economic History, Universidad de Salamanca and IME, Salamanca, Spain

good or that the expense in the good is a tiny fraction of the consumer´s budget (Willig 1976; Hausman 1981).

In this paper we study two extensions of consumer surplus in a market of a homogeneous good, retaining the quasilinearity assumption. On the one hand, we study the effect of a lower bound on the consumption of the outside good on the consumer's decision problem, and, on the other, the effect of different weights of consumers surplus and producers surplus on social welfare. Let us take these two points in turn.

In the standard consumer surplus analysis of a market, for instance electricity, we assume that the utility function is linear in the "outside" good, say housing. Let the utility function of a consumer be $2x^{1/2} + y$, where y is housing, x is the consumption of electricity. The budget constraint is $px + y = 1$ where income and housing price are 1 and p is the price of electricity. A common procedure is to plug the budget constraint in the utility function and getting rid of income to obtain $2x^{1/2} - px$. Maximizing this concave function, first-order condition (FOC) yields the inverse demand function $x^{-1/2} = p$. This implies that electricity expenses are $px = 1/p$ which might exceed income. For those p, $x^{-1/2} = p$ cannot be the inverse demand function because it would imply a negative consumption of housing. And note that, willy-nilly, income effect creeps into the inverse demand function. This situation is likely to arise in poor households in times of crisis. In Sect. 3 we show that a nonnegativity constraint (NNC) represents any lower bound on the consumption of the outside good. And for those consumers for whom NNC bites, demand is unit-elastic in price and income, losing income effect neutrality. Moreover, when demand without NNC is elastic, NNC makes demand less elastic, increasing markups and monopoly power. And the more consumers are affected by NNC, the less elastic demand becomes. Also, NNC makes marginal revenue discontinuous. This implies that Cournot equilibrium may occur in the lower bound of the interval of quantities in which NNC holds. Or even inside this interval. In both cases, equilibrium output is smaller than the one arising when NNC is ignored.

Our second topic arises when the regulator cares about the distribution of surplus between consumers and firms.[1] Usually, it is assumed that regulators are only concerned with efficiency, leaving redistribution to other departments. But there are reasons why regulators may like to take their cut in the redistribution of surplus, ranging from bribes to sheer sympathy and including political support (Hillman 1982).When a planner maximizes a weighted sum of consumer and producer surpluses two problems arise. Firstly, even under constant marginal costs and concave social welfare, profits may be negative at FOC. This occurs when consumer surplus has a large weight compared to profits, because, in this case, an output increase from price equals marginal costs favors the consumer. Secondly, the solution found solving FOC may not satisfy the second-order condition (SOC). This is because when the weights of consumers and the firm are different, the slope of the inverse demand enters in the FOC. Thus, SOC contains the second derivative of the inverse demand function, similar to what happens under profit maximization under oligopoly. Under the assumption of an exogenous constraint on how negative profits might be, we characterize the zones in which FOC satisfy SOC and profits satisfy this constraint for a particular inverse demand

---

[1] See Dworczak et al. (2019) for a study of this question from the point of view of mechanism design.

that generalizes both linear and isoelastic demand functions. We also offer formulae for the percentage of welfare loss due to oligopoly. In our case, loss depends on the weight given to consumers and the firm, as well as the lower bound on profits. Finally, we identify the interval for which an increase in the number of firms decreases social welfare.

As far as we know there is only one paper mentioning the first point above by Amir and Erickson (2017). They note that "when the usual "enough money" condition fails... the resulting demand function is not linear, but hyperbolic (with unitary elasticity)". In their model, there is a single consumer and often utility is quadratic. They show that an equilibrium exists but they do not characterize equilibrium.

This paper is organized in the following manner: Sect. 2 spells the model, Sect. 3 tackles nonnegativity constraint; Sect. 4 studies social welfare when consumers and firms have different weights; and lastly Sect. 5 spells possible applications of our results and suggests new paths of research.

2 The model

We focus on a single market supplied by a given number of identical firms. Let x_i be the output of firm i, x aggregate output and p the market price. The cost function of firm i is a twice differentiable function $c(x_i)$. Profits are $px_i - c(x_i)$. Consumers have identical preferences with a twice differentiable (whenever $x > 0$) utility function $U(x) + y$ where $U(\cdot)$ is strictly concave (with non-vanishing curvature) and strictly increasing in the relevant range.[2] y is the consumption of the outside good and budget constraint is $px + y = M$, where M is the exogenous income. It is assumed that consumers can freely buy as much good as they desire in the market as long as the budget constraint holds. In other words, we assume that once the output is set, the price is the market price that will clear the market.

The following utility function, called beta-linear, will be used extensively in this paper

$$U(x) = Ax - \frac{b}{\beta + 1} x^{\beta+1}, \tag{1}$$

with $b\beta > 0$, $(A - c'(x))b > 0$, $\beta > -1$ and $-A\beta < c'(x)$.[3] These assumptions guarantee that $U(\cdot)$ is concave. The inverse demand function when NNC is not considered, $p = A - bx^\beta$, is decreasing and meets the marginal cost. This inverse demand generalizes both linear ($\beta = 1$, $A > 0$) and constant elasticity ($\beta \in [-1, 0)$, $A = 0$) inverse demands.

---

[2] If the demand curve is well defined when $p = 0$, as it happens with linear demand curves, preferences cannot be strictly monotonic. However, all that is required is that this lack of monotonicity occurs at outputs that are never reached in equilibrium.

[3] This functional form was introduced by Bulow and Pfleiderer (1983) and has been used by González-Maestre (2000), Anderson and Renault (2003), Corchón (2008), Ritz (2018) and Corchón and Torregrosa (2020) in different setups.

3 Nonnegativity of the consumption of the outside good

In this section we assume that the consumption of the outside good is larger or equal than zero. Consumer's maximization problem is

$$\max U(x) + y, \, s.t. \, px + y \le M, \, and \, y \ge 0. \tag{2}$$

We remark that (2) can handle the more general case of a lower bound in the consumption of the outside good, say k, so $y \ge k$. Subtracting k to the utility function does not change underlaying preferences so the consumer maximizes $U(x) + y - k, \, s.t. \, px + y \le M, \, and \, y \ge k$ which setting $z \equiv y - k$ can be rewritten as maximizing $U(x) + z, \, s.t. \, px + z \le M + k, \, and \, z \ge 0$ which is qualitatively identical to (2). Thus, from now on, without loss of generality, we will refer to this problem as the nonnegativity constraint or NNC.

When NNC does not bite, inverse demand is the standard $p = U'(x)$, and demand is $x = [U'(p)]^{-1}$ where $[U'(p)]^{-1}$ denotes the inverse of $U'(\cdot)$ (which under our assumptions exists). On the other hand, when the NNC bites, consumer demand comes from the corner solution $x = M/p$, $y = 0$. Let $\Omega^p = \left\{ p \in \mathbb{R}^+ : M/p \le [U'(p)]^{-1} \right\}$ be the set of prices for which NNC bites.

Henceforth we divide consumers into two categories: rich and poor. Those in the rich group have incomes where NNC never bites. Those in the poor group have identical incomes, denoted by M, where, for some prices, NNC does bite. Let us assume that a proportion $\gamma \in (0, 1]$ of individuals are poor. Hence, aggregate demand can be written as

$$x(p, \gamma) = \begin{cases} [U'(p)]^{-1}, & p \notin \Omega^p \\ \gamma \frac{M}{p} + (1 - \gamma)[U'(p)]^{-1}, & p \in \Omega^p. \end{cases} \tag{3}$$

Note that $x(p, \gamma)$ is continuous but not differentiable.

An initial glimpse of the consequences of the NNC is obtained studying the effect of NNC on the price elasticity of demand. The latter plays an important role in the determination of an equilibrium. Let $\varepsilon(p)$ be the price elasticity of demand when NNC does not bite. Following a standard convention we take this elasticity as positive. Let $E(p, \gamma)$ be the price elasticity of demand when NNC bites to the poor, namely,

$$E(p, \gamma) = \begin{cases} \varepsilon(p), & p \notin \Omega^p \\ \frac{\gamma M + (1-\gamma)R(p)\varepsilon(p)}{\gamma M + (1-\gamma)R(p)}, & p \in \Omega^p. \end{cases} \tag{4}$$

where $R(p) = p[U'(p)]^{-1}$. Hence, price elasticity is the unrestricted-demand price elasticity for the set of prices for which no individual is affected by the NNC, and it is proportional to the unrestricted-demand price elasticity for the set of prices for which a proportion γ of individuals are affected by the NNC. Clearly, $E(p, 1) = 1$ and $E(p, 0) = \varepsilon(p)$. Our first result examines the relationship between $\varepsilon(p)$ and $E(p, \gamma)$ for $p \in \Omega^p$.

Proposition 1 *Suppose $p \in \Omega^p$.*

(a) *When $\varepsilon(p) > 1$, $\varepsilon(p) > E(p, \gamma) > 1$.*
(b) *When $\varepsilon(p) < 1$, $\varepsilon(p) < E(p, \gamma) < 1$.*
(c) *When $\varepsilon(p) = 1$, $\varepsilon(p) = E(p, \gamma)$*

Proof When $\varepsilon(p) > 1$, $\varepsilon(p)\gamma M > \gamma M$, and $\varepsilon(p)\gamma M + (1 - \gamma)R(p)\varepsilon(p) > \gamma M + (1 - \gamma)R(p)\varepsilon(p)$. Hence $\varepsilon(p) > E(p, \gamma)$. Moreover,

$$\frac{\gamma M + (1 - \gamma)R(p)\varepsilon(p)}{\gamma M + (1 - \gamma)R(p)} > 1$$

because, $\gamma M + (1 - \gamma)R(p)\varepsilon(p) > \gamma M + (1 - \gamma)R(p)$, since $\varepsilon(p) > 1$. For $\varepsilon(p) \leq 1$ the proof is analogous. $\qquad \square$

This result indicates that when NNC bites, it changes the elasticity of aggregate demand with respect to aggregate demand without NNC. NNC makes less elastic demands that without NNC are elastic, see Fig. 1 below. The following example illustrates this feature.$\square$

Example 1 Let $U(x) = Ax - \frac{b}{2}x^2$ where the poor have an income such that $M < A^2/4b$. It is readily calculated that the set in which NNC bites is.

$$\Omega^p = \left[\frac{1}{2}\left(A - \sqrt{A^2 - 4bM} \right), \frac{1}{2}\left(A + \sqrt{A^2 - 4bM} \right) \right].$$

Hence, aggregate demand is

$$x(p, \gamma) = \begin{cases} \frac{A-p}{b}, & p \notin \Omega^p \\ \gamma\frac{M}{p} + (1 - \gamma)\frac{A-p}{b}, & p \in \Omega^p. \end{cases}$$
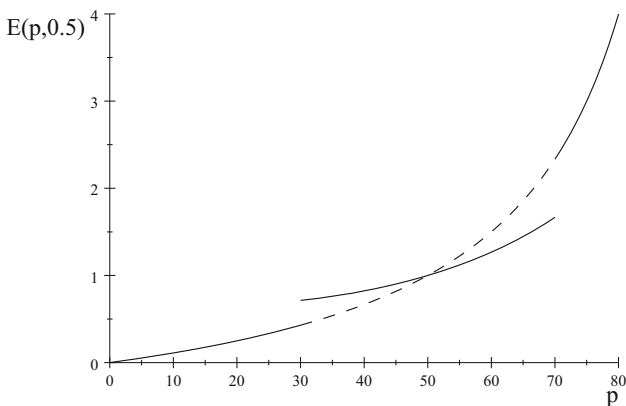
Fig. 1 Price elasticity

And elasticity of demand is

$$E(p, \gamma) = \begin{cases} \frac{p}{A-p}, \, p \notin \Omega^p \\ \frac{\gamma M + (1-\gamma)p^2}{\gamma M + (1-\gamma)(A-p)p}, \, p \in \Omega^p. \end{cases}$$

We see that demand is elastic for $p > A/2$. Figure 1 plots the elasticity as a function of the price when $A = 100$, $b = 1$, $M = 2100$ and $\gamma = 1/2$. For the set of prices for which NNC bites, demand becomes less elastic for $50 < p < 70$, and less inelastic for $3\,0 < p < 50$.

We now study how γ affects $E(p, \gamma)$. The proof is immediate from (4).

Remark 1 For $p \in Interior\,\Omega^p$,

$$\frac{\partial E(p, \gamma)}{\partial \gamma} = \frac{MR(p)(1 - \varepsilon(p))}{(\gamma M + (1 - \gamma) R(p))^2}.$$

Thus, an increase in poor people changes the elasticity of demand, making it less inelastic when it is already inelastic and less elastic when it is already elastic.

This jump in the elasticity creates a kink in the demand function faced by the firm. This kink rests on the presence of consumers whose constraint on the consumption of the outside good binds. The traditional kink in the demand curve, steaming from the work of Sweezy (1939), rests on oligopolistic interactions and an unexplained "prevailing price". Our kink also arises under monopoly or monopolistic competition the latter a usual ingredient of many macro models. Thus, the reasons given by the empirical literature to reject the traditional argument for kinked demand curves (e.g., Dossche et al. 2010) do not apply here.

3.1 Monopoly equilibrium

We start analyzing the impact of NNC on equilibrium allocation by examining the simplest possible case, namely, a market supplied by just one firm. Marginal cost is constant and denoted by c.

We first consider how the classical formula of markups is affected by NNC. Let $\mu(p) = (p - c)/p$ be the markup, also called the Lerner index. It is considered a measure of market power.

Proposition 2 *In the monopoly equilibrium,*

$$\mu(p, \gamma) = \begin{cases} \frac{1}{\varepsilon(p)}, & p \notin \Omega^p, \\ \left(\frac{\gamma}{1-\gamma} \frac{M}{R(p)} \frac{c}{p} + 1\right)\frac{1}{\varepsilon(p)}, & p \in \Omega^p. \end{cases}$$

Proof The monopolist $max_p (p - c)x(p, \gamma)$, whose FOC yields.

$$x(p, \gamma)\left(1 - \frac{p - c}{p} E(p)\right) = 0.$$

Using (4), when $p \notin \Omega^p$ we have the standard markup formula

$$\frac{p-c}{p} = \frac{1}{\varepsilon(p)},$$

but when $p \in \Omega^p$, we have

$$\frac{p-c}{p} = \frac{\gamma M + (1-\gamma) R(p)}{\gamma M + (1-\gamma) R(p)\varepsilon(p)}.$$

Rearranging the terms, $(1-\gamma)R(p)((p-c)\varepsilon(p) - p) = c\gamma M$ or

$$(1-\gamma) R(p) p\varepsilon(p)\left(\frac{p-c}{p} - \frac{1}{\varepsilon(p)}\right) = c\gamma M,$$

and the result is proved.   □

Proposition 2 states that NNC increases markup with respect to its unrestricted counterpart. This implies that NNC increases market power.

From now on, we will work with inverse demand that simplify the study of monopoly equilibrium and is essential to work out the case of Cournot oligopoly in the next subsection. We define $\Omega^x = \{x \in \mathbb{R}^+ : M/x \leq U'(x)\}$ as the set of quantities for which NNC bites.

The following Lemmas study the shape of Ω^x in two different cases: when $xU'(x)$ is hump-shaped and when it is increasing.[4] The first case arises when in the utility functions in (beta-linear utility function), $\beta > 0$, for instance, a linear demand. The second case arises when in the utility functions in (beta-linear utility function), $\beta \in (-1, 0)$ and $A = 0$, i.e. when demand is isoelastic.

Lemma 1 *Suppose $xU'(x)$ is hump-shaped.*

(a) *$xU'(x)$ has, at most, two intersections with $px = M$, denoted by $\underline{x}$ and $\overline{x}$.*
(b) *Let $\hat{x}$ be such that $\hat{x} = \max xU'(x)$. Then if $\Omega^x = [\underline{x}, \overline{x}]$, $\hat{x} \in \Omega^x$.*

For a proof, see "Appendix". Figure 2 plots $xU'(x)$ which is depicted solid-dash-solid for $A = 100$, $b = \beta = 1$ and $M = 2100$. Revenue when the NNC is taken into account and $\gamma = 1/2$ is depicted as a solid line. Lemma 1 part (a) says that when $xU'(x)$ is hump-shaped if Ω^x is nonempty, Ω^x is the interval $[\underline{x}, \overline{x}]$. Lemma 1 part (b) says that when $xU'(x)$ is hump-shaped, the value $\hat{x}$ so that $\varepsilon(\hat{x}) = 1$, belongs to the set in which NNC bites, Ω^x.

Lemma 2 *Suppose $xU'(x)$ is increasing. Then $xU'(x)$ has, at most, one intersection $\underline{x}$ with $px = M$.*

Proof The intersections have to satisfy $xU'(x) = M$. Thus, if there is a solution to this equation, say $\underline{x}$, it exists and is unique.   □

---

[4] A function is hump-shaped when it has a unique global maximum that is interior.
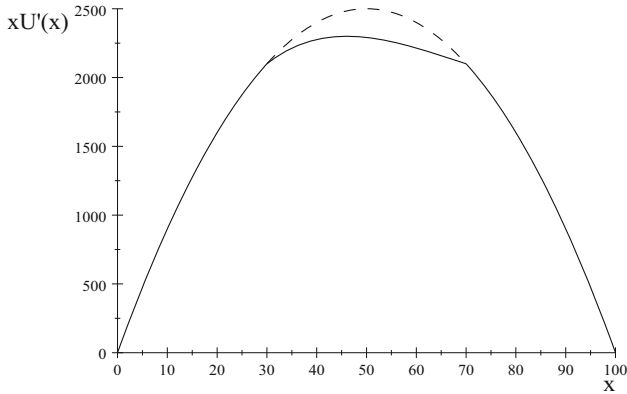
Fig. 2 Total revenue under NNC

The next example shows how marginal revenue looks like when NNC bites.

Example 2 Considering the utility function of Example 1, inverse demand function is given by.

$$p(x, \gamma) = \begin{cases} A - bx, & x \notin \Omega^x \\ \frac{(1-\gamma) A - bx + \sqrt{(bx - (1-\gamma) A)^2 + 4(1-\gamma) \gamma bM}}{2(1-\gamma)}, & x \in \Omega^x. \end{cases},$$

where

$$\Omega^x = \left[\frac{1}{2b} \left(A - \sqrt{A^2 - 4bM} \right), \frac{1}{2b} \left(A + \sqrt{A^2 - 4bM} \right) \right]$$

and marginal revenue is

$$MR(x, \gamma) = \begin{cases} A - 2bx, & x \notin \Omega^x \\ \frac{(1-\gamma) A - 2bx}{2(1-\gamma)} + \frac{(bx - (1-\gamma) A) (2bx - (1-\gamma) A) + 4(1-\gamma) \gamma bM}{2(1-\gamma) \sqrt{(bx - (1-\gamma) A)^2 + 4(1-\gamma) \gamma bM}}, & x \in \Omega^x. \end{cases}$$

Note that while indirect demand function is continuous in the boundary of Ω^x, marginal revenue is not. Figure 3 plots inverse demand and marginal revenue when $A = 100$, $b = 1$, $M = 2100$ and $\gamma = 1/2$. In this case $\Omega^x = [30, 70]$.

For these parameters, if $40 < c < 100$, monopoly equilibrium occurs in the unrestricted part of aggregate demand, equilibrium price is $p^M = 50 + 0.5c$ and markup is

$$\mu \left(50 + \frac{1}{2}c, \frac{1}{2} \right) = \frac{100 - c}{100 + c}, \quad 40 < c < 100.$$

When $0 < c < 28$ monopoly equilibrium occurs where NNC bites aggregate demand. Thus, when, for instance, $c = 10$, $x^M = 39.014$, $p^M = 58.11$, and $\mu(58.11, 0.5) = 0.83$ higher than its counterpart when NNC does not bite.
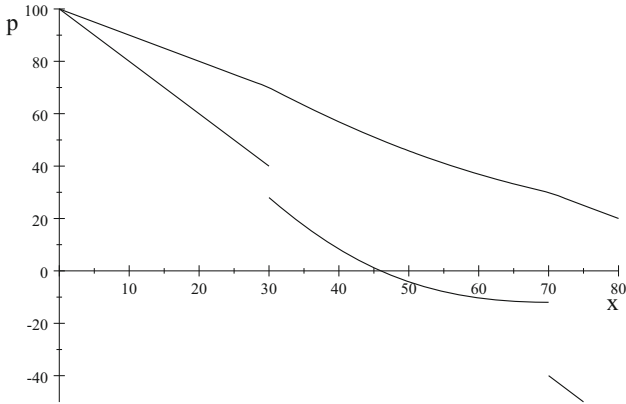
Fig. 3 Inverse demand and marginal revenue under NNC

When $28 < c < 40$, marginal revenue and marginal cost do not intersect. To find the profit-maximizing output, note that for outputs less than 30, marginal profits are positive because marginal revenue is larger than marginal cost. And for outputs larger than 30, marginal profits are negative. Thus the profit-maximizing output, which exist, since profits are continuous in x and bounded, must be 30.

The next proposition generalizes our findings in Example 2.

Proposition 3 *Suppose that marginal revenue is decreasing in x but there is no x such that marginal revenue is equal to c. Then monopoly equilibrium occurs in the lower bound of Ω^x, $\underline{x}$.*

Proof The monopolist profit function is increasing for $x \in (0, \underline{x})$ but decreasing for $x > \underline{x}$. In such a case, $\underline{x}$ maximizes profits, see Fig. 3.

The previous result can be strengthened when $\gamma = 1$. In this case, the marginal revenue of the monopolist when facing a demand curve M/p is zero. □

Remark 2 Suppose $xU'(x)$ is increasing or hump-shaped and $\gamma = 1$, then monopoly equilibrium is given by $x^M = \min\{x^U, \underline{x}\}$, where x^U denote the profit-maximizing choice of the monopolist when facing an inverse demand $p = U'(x)$.

Proof When $xU'(x)$ is increasing or is hump-shaped but $\gamma = 1$, total revenue is increasing until it becomes constant in Ω^x. Thus, when the monopoly choice falls into the interval of output for which NNC bites, monopoly output switches to the minimum of this interval, $\underline{x}$ because in Ω^x, marginal revenue is zero. □

3.2 Several firms

In this subsection, we outline how the conclusions obtained under the assumption of a single firm apply to an oligopolistic market. Our main result here is that under oligopoly, there may be an equilibrium at $(\underline{x}, \overline{x})$ which is impossible under monopoly.

Let us assume that there are n quantity setter firms. To simplify the presentation we assume a hump-shaped $xU'(x)$ with $\gamma = 1$ (only poor consumers) and identical firms with constant marginal costs denoted by c. Note that NNC also increases markups in a symmetric Cournot equilibrium with respect to its unrestricted counterpart when $\varepsilon > 1$ (the proof is a simple generalization of Proposition 2 and it is skipped).

We have three possible kind of equilibria:

(1) Equilibrium in $\left[0, \underline{x}\right) \cup \left[\overline{x}, \infty\right)$.
(2) Equilibrium in $\underline{x}$.
(3) Equilibrium in $\Omega^x = \left(\underline{x}, \overline{x}\right)$

Equilibria in (1) correspond to the standard case in which NNC does not bite. Equilibria in (2) correspond to the equilibrium under monopoly studied in Remark 2. Equilibrium in (3) is a new kind, impossible under monopoly, but possible here.

Let us start with case (1) above. Let x^U be the Cournot equilibrium output when NNC is ignored so inverse demand is $U'(x)$. In this case, we have that:

Proposition 4 *Suppose $x^U \notin \left[\underline{x}, \overline{x}\right]$. Then x^U is a Cournot equilibrium.*

Proof x_i^* is the best reply to $(n - 1)x_i^*$ when the residual demand is $U'(x_i + (n - 1)x_i^*)$. Thus, there are no profitable deviations. Hence, x_i^* must be the best reply when inverse demand is $min\{U'(x), M/x\} \leq U'(x)$, because if a deviation was not profitable before, it cannot be profitable now when demand has shifted downward. Therefore, x_i^* is for each firm the best reply when all other firms produce $(n - 1)x_i^*$ and thus, it is a Cournot equilibrium. $\qquad\square$

Let us now address case 2 above for which we have the following result.

Proposition 5 *Suppose the utility function is beta-linear with $n + \beta > 0$. When $M(n - 1)/cn < \underline{x}$ and $x^U \in \left[\underline{x}, \overline{x}\right]$, there is an equilibrium at $\underline{x}$.*

Proof Suppose all firms produce $\underline{x}/n$ and consider a deviation of firm i. Suppose the firm decreases its output. Thus, it will be on the unrestricted inverse demand $U'(x)$. And, because $x^U \in \left[\underline{x}, \overline{x}\right]$ and is stable according to the gradient dynamics (see Corchón and Torregrosa 2020), it must be that $d\pi_i/dx_i > 0$ that is to say, a decrease in x_i decreases profits. Suppose now that the firm increases its output, and as a result the firm is on the inverse demand M/x. Since the equilibrium exists, because $n + \beta > 0$, and is stable on this inverse demand according to the gradient dynamics (see Corchón and Torregrosa 2020) it must be that $d\pi_i/dx_i < 0$, i.e. the firm is better off by decreasing its output. Thus no deviation is profitable from $\underline{x}/n$ therefore $\underline{x}/n$ is an equilibrium. $\qquad\square$

Next example shows that the conditions of Proposition 5 are possible under oligopoly.

Example 3 Let be the model here as in Example 2. The conditions in Propositio 5 for an equilibrium are

$$\frac{n - 1}{cn}M < \frac{1}{2b}\left(A - \sqrt{A^2 - 4bM}\right),$$

$$\frac{1}{2}\left(A - \sqrt{A^2 - 4bM}\right) < \frac{(A-c)n}{n+1} < \frac{1}{2}\left(A + \sqrt{A^2 - 4bM}\right).$$

The fulfillment of all inequalities can be written as

$$\frac{(n+1)}{2n}A - \frac{(n+1)}{2n}\sqrt{A^2 - 4bM} \le \frac{(n-1)}{n}\frac{2bM}{\left(A - \sqrt{A^2 - 4bM}\right)} < c$$

$$c \le \frac{(n+1)}{2n}A + \frac{(n+1)}{2n}\sqrt{A^2 - 4bM}.$$

The left-hand side of such inequality

$$\frac{(n+1)}{2n}A - \frac{(n+1)}{2n}\sqrt{A^2 - 4bM} \le \frac{(n-1)}{n}\frac{2bM}{\left(A - \sqrt{A^2 - 4bM}\right)},$$

holds whenever $n \ge A/\sqrt{A^2 - 4bM}$ (available under request). Regarding the right-hand side, first it is necessary that

$$\frac{(n-1)}{n}\frac{2bM}{\left(A - \sqrt{A^2 - 4bM}\right)} < \frac{(n+1)}{2n}A + \frac{(n+1)}{2n}\sqrt{A^2 - 4bM}$$

This is a condition that holds for

$$c \in \left[\frac{(n-1)}{n}\frac{2bM}{\left(A - \sqrt{A^2 - 4bM}\right)}, \frac{(n+1)}{2n}A + \frac{(n+1)}{2n}\sqrt{A^2 - 4bM}\right]$$

Thus, there is a set of parameters, in particular, for which Proposition 5 holds. Finally, we proceed to study case 3, showing that it is possible here.

Proposition 6 *Suppose that $M(n-1)/cn \in [\underline{x}, \overline{x}]$. Then, $M(n-1)/cn^2$ is a Cournot equilibrium output for each firm.*

Proof $M(n-1)/cn$ is the equilibrium output when the inverse demand function is M/x. The reasoning is the same as before. Since the true inverse demand is $\min\{U'(x), M/x\}$, if $M(n-1)/cn^2$ is the best reply to $M(n-1)/cn^2$, when the inverse demand is larger or equal to the true demand, it should be the best reply when inverse demand is smaller or equal to $U'(x)$. Thus, $M(n-1)/cn^2$ is the best reply for each firm when all other firms produce this quantity; and thus it is a Cournot equilibrium. $\square$

The condition is key in Proposition 6 and explains why under monopoly ($n = 1$) this kind of equilibrium is impossible.

Next is an example showing that there are parameters fulfilling the conditions of the previous Proposition.

Example 4 Let the model here be as in Examples 2 and 3. Now the condition in Proposition 6 for an equilibrium is

$$\frac{n-1}{cn} M \in \Omega^x,$$

which can be written as

$$\sqrt{A^2 - 4bM} > \max\left\{\frac{n-1}{cn} 2bM - A, \ A - \frac{n-1}{cn} 2bM\right\}.$$

Equilibrium price is $cn/(n-1)$. Ignoring the NNC, equilibrium output and price are, respectively $(A-c)n/b(n+1)$ and $(A+cn)/(n+1)$. The price is now lower than when under no constraint if and only if

$$0 < c < \frac{A(n-1)}{2n}.$$

And NNC decreases output if and only ifwhich can be written as

$$\frac{n-1}{cn} M < \frac{n}{n+1} \frac{A-c}{b}$$

$$n^2 c - An^2 c + b\left(n^2 - 1\right) M < 0. \tag{5}$$

The condition $M < A^2/4b$ guaranties (5) for

$$c \in \left(\frac{1}{2b}\left(An - \sqrt{A^2 n^2 - 4b\left(n^2 - 1\right)M}\right), \frac{1}{2b}\left(An + \sqrt{A^2 n^2 - 4b\left(n^2 - 1\right)M}\right)\right).$$

4 Social welfare when consumers and firms have different weights

In this section, we consider that social welfare is a weighted sum of consumer surplus (CS) and profits (Π),

$$\alpha(U(x) - p(x)x) + (1-\alpha)(p(x)x - c(x)), \alpha \in [0, 1]. \tag{6}$$

where $c(x)$ is the cost of producing x. Social welfare can be thought of as the objective function of a regulator/planner or as the preferences of a society. In the sequel, we will speak of a regulator, but the other interpretation should be kept in mind. The parameter α represents the sympathy (whatever it comes: moral concern, rent seeking, advertising, etc.) of the regulator for the consumer. If $\alpha > 1/2$ we say that the regulator is pro-consumer and if $\alpha < 1/2$ the regulator is pro-firm, see Blackmon and Zeckhauser (1992). When $\alpha = 1/2$ social welfare is the sum of consumer and producer surpluses

$U(x) - c(x)$ (multiplied by α, but this does not affect the maximum of this function). We assume that the regulator is perfectly informed of the costs and utility functions.[5]

An interpretation is that we adjust for surplus disparities generated by the benchmark. Specifically, if $CS > \Pi$ then consumer surplus is the welfare benchmark adjusted by surplus disparity due to the presence of firms in the welfare function, so that $W = CS - (1-\alpha)(CS - \Pi)$. If $\Pi > CS$ then profits are the welfare benchmark adjusted by surplus disparity due to the presence of consumers in the welfare function, so that $W = \Pi - (\Pi - CS)$. These two cases are versions of the utility function with inequity aversion in Fehr and Schmidt (1999), in the former case from the consumers' standpoint, and in the latter from the firms' standpoint.

We present our analysis focusing on quantities because, to compare the choice of the regulator with the market, we will introduce quantity competition later. But identical problems arise in the case in which the regulator sets market price. In this case, social welfare is $\alpha\left(U'(x(p)) - px(p)\right) + (1-\alpha)(px(p) - c(x(p)))$ where $x(p)$ is market demand (see footnote 6 below).

To highlight the points made in this section we disregard NNC. Assuming interiority, FOC of social welfare maximization is

$$\alpha\left(U'(x) - p'(x)x - p(x)\right) + (1-\alpha)(p'(x)x + p(x) - c'(x)) = 0.$$

Taking into account that the consumer maximizes utility $U'(x) = p(x)$ and rearranging we obtain that

$$(1 - 2\alpha)p'(x)x + (1-\alpha)(p(x) - c'(x)) = 0. \tag{7}$$

Second-order condition (SOC) of social welfare maximization is

$$(1 - 2\alpha)(p''(x)x + p'(x)) + (1-\alpha)(p'(x) - c''(x)) \le 0. \tag{8}$$

If $\alpha = 1/2$, equation (SOC) is $U''(x) - c''(x) \le 0$ which holds if $U(\cdot)$ is concave and $c(\cdot)$ is convex. But when $\alpha \ne 1/2$ it is not clear which assumptions on utility and costs guarantee that SOC hold.[6] This is the first problem we face here.

Again, when $\alpha = 1/2$, Eq. (7) says that in the social optimum, price equals marginal cost. Under non-decreasing marginal costs, this implies nonnegative profits. But, when $\alpha \ne 1/2$, even under constant marginal costs, profits might be negative. Thus, when $\alpha \ne 1/2$, a zero-profit constraint might be binding under constant returns. This is the second problem we encounter when $\alpha \ne 1/2$.

---

[5] Our social welfare function is sometimes called α-utilitarian in which a positive scale factor, α, is used in the interpersonal.comparison of weighted utility.

[6] When the regulator chooses prices FOC reads $\alpha\left(\frac{\partial U(x)}{\partial x}\frac{\partial x(p)}{\partial p} - x(p) - p\frac{\partial x(p)}{\partial p}\right) + (1-\alpha)\left(x(p) + p\frac{\partial x(p)}{\partial p} - \frac{\partial c(x)}{\partial x}\frac{\partial x(p)}{\partial p}\right)$ which taking into account that $\frac{\partial U(p)}{\partial p} = p$ simplifies to $-\alpha x(p) + (1-\alpha)\left(x(p) + p\frac{\partial x(p)}{\partial p} - \frac{\partial c(x)}{\partial x}\frac{\partial x(p)}{\partial p}\right)$. Again even if monopoly profits are concave on prices -so the term multiplying $(1-\alpha)$ is concave- social welfare may be not because the term $-x(p)$ in FOC of social welfare maximization.

The remaining part of this section will be devoted to analyzing these two problems under identical firms with constant returns to scale and the beta-linear utility function spelled out at Sect. 2. Under these assumptions, social welfare in can be written as

$$W(x) = \alpha \left(Ax - b\frac{x^{\beta+1}}{\beta+1} - Ax + bx^{\beta+1} \right) + (1-\alpha)\left((A-c)x - bx^{\beta+1} \right), \quad (9)$$

where c as before is marginal cost. Setting $a \equiv A - c$, social welfare is written as

$$W(x) = x\left[(1-\alpha)a - bx^{\beta}\left(\frac{1+\beta-(1+2\beta)\alpha}{\beta+1} \right) \right].$$

The problem of the regulator is to maximize $W(x)$ subject to losses not to exceed a number denoted as $-tx$ with $t \geq 0$. Let us call this constraint the maximum loss condition, written as $px - cx \geq -tx$ which always holds for $x = 0$. If x is positive, dividing by x and inverting $p(\cdot)$ the condition is $x \leq p^{-1}(c-t)$. When $t = 0$, the regulator does not subsidize and the upper bound to output is $p^{-1}(c)$, where profits are zero. When the whole marginal cost is subsidized, $t = c$, and the upper bound to output is $p^{-1}(0)$, where it exhausts demand. We assume that $t \in [0, c]$, i.e. the maximum subsidy is the one that covers the whole production cost. To simplify our analysis, we assume that the regulator does not take into consideration the cost of these funds. You can think of a state regulator receiving federal or EU funds.[7] Our framework is Ramsey pricing with one product in which instead of maximizing the sum of consumer and producer surpluses we maximize a weighted sum.[8]

It is immediate that because $\beta > -1$, W is continuous in x. Also, $x \in [0, p^{-1}(c-t)]$ which is a compact set. Hence Weierestrass theorem guarantees the existence of a socially optimal output; however, we do not know where it is located. There are two candidates only. Either it is located in the FOC or it is determined by the maximum loss condition. Our research strategy is to analyze these two cases separately. First, we check whether the FOC of the social welfare maximization without the maximum loss condition picks a maximum. And then, we study whether this maximum fulfils the maximum loss condition. If this were the case, the procedure for obtaining the optimal output from the FOC is fully justified. In any other case, the maximum is determined by the maximum loss condition.

Let us start by analyzing the fulfilment of the SOC disregarding the maximum loss condition. Under our assumptions, interior FOC of social welfare maximization is

$$-(1-2\alpha)(b\beta x^{\beta}) + (1-\alpha)(a - bx^{\beta}) = 0,$$

$$x^* = \left(\frac{(1-\alpha)a}{b(1+\beta-(1+2\beta)\alpha)} \right)^{\frac{1}{\beta}}, \quad (10)$$

---

[7] When $\alpha = 1/2$ it does not matter who pays taxes as long as they do not distort the market under consideration. But when $\alpha \neq 1/2$, who finances firms' losses matters which introduces a further complication in our analysis. Thus we decide to leave this matter for future research.

[8] The problems of Ramsey pricing under increasing returns are well known, see e.g. Dierker (1991). The problems studied here are caused by the non concavity of profit function. In this sense are more akin to those studied in Roberts and Sonnenschein (1977).

where x^* is the output solving FOC when the maximum loss condition is disregarded. Since $a/b > 0$, to obtain a positive number from (10) we need

$$1 + \beta - (1 + 2\beta)\alpha > 0. \tag{11}$$

Condition (11) makes sure the fulfillment of the SOC of social welfare maximization since

$$\text{sign}\frac{d^2 W(x)}{dx^2} = \text{sign}\{-b\beta x^{\beta-1}(1 + \beta - (1 + 2\beta)\alpha)\}.$$

and this expression is negative if and only if (11) holds. Thus, condition (11) guarantees both a positive solution to the necessary condition of social welfare maximization and that the FOC is sufficient. Also implies that social welfare at x^* is positive ruling out a corner solution in which social welfare is zero. Formally:

Proposition 7 *The output maximizing social welfare disregarding the maximum loss condition is obtained from the FOC of social welfare maximization if and only if $1 + \beta - (1 + 2\beta)\alpha > 0$.*

We study the pairs of (α, β) for which (11) holds. We start by noting that this condition always holds when $1 + 2\beta < 0$. Thus, we can write condition (11) as $\alpha < Z(\beta)$, where $Z(\beta) = (1+\beta)/(1+2\beta)$ with $\beta \neq -1/2$. It is fairly easy to check the following properties of $Z(\beta)$ which are depicted in Fig. 4.

Remark 3

i. $Z(-1) = 0$,
ii. $\lim_{\beta \to -1/2^{\mp}} Z(\beta) = \mp\infty$,
iii. $Z(0) = 1$,
iv. $\lim_{\beta \to \infty} Z(\beta) = 1/2$,
v. $\frac{dZ(\beta)}{d\beta} < 0$ and $\frac{d^2 Z(\beta)}{d\beta^2} < 0$ when $\beta < -\frac{1}{2}$.
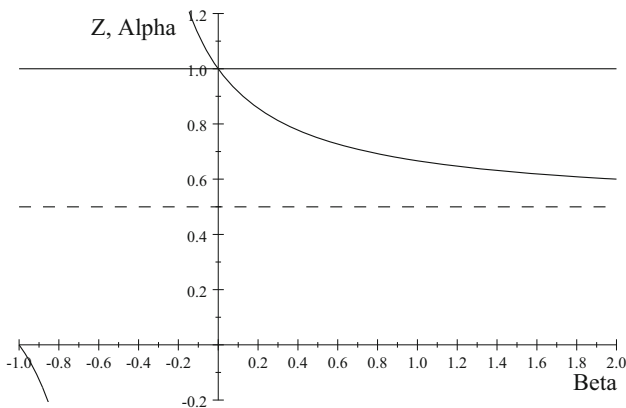
Fig. 4 Condition (11) as a function of beta

Figure 4 plots $Z(\beta)$ as a solid line for $\beta \geq -1$. The dashed line $\alpha = .5$ is the asymptote of $Z(\beta)$ when $\beta \rightarrow \infty$. This figure shows that obtaining the socially optimal output from the FOC is always a good procedure either when $\beta < 0$ or when $\alpha < 1/2$, which remains true when α and β are not very large. However, the combination of relatively high values of α and β makes the FOC to reach a minimum, not a maximum. As an illustration, in the linear case ($\beta = 1$) the weight of consumers in social welfare cannot be larger than 2/3. Also, when demand is quadratic ($\beta = 2$), α cannot be larger than 6.

We now turn to the study of the fulfillment of the maximum loss condition which now can be written as:

$$a + t \geq bx^\beta, \tag{12}$$

where $t \in [0, c]$. We readily see that the upper bound of output is:

$$\overline{x}(t) = \left(\frac{a + t}{b}\right)^{\frac{1}{\beta}}.$$

To have a positive upper bound we assume that $(a + t)$ and b have the same sign. This is a generalization of the assumption made in Sect. 2 that a and b have the same sign.

It is immediate that $\overline{x}(t)$ is increasing in $[0, c]$. Now we investigate whether optimal output calculated in the FOC (10) fulfils the maximum loss condition.

Proposition 8 $x^* \leq \overline{x}(t)$ whenever $\alpha \leq \frac{t+\beta(a+t)}{t+2\beta(a+t)} \equiv Y(\beta, t)$.

Proof When $\beta \in [-1, 0)$, it is convenient to write both optimum and upper bound output raised to a positive power, hence,$x^* \leq \overline{x}(t)$, whenever.

$$\left(\frac{b(1 + \beta - (1 + 2\beta)\alpha)}{(1 - \alpha)a}\right)^{-1/\beta} \geq \left(\frac{b}{a + t}\right)^{-1/\beta},$$

$$\frac{1 + \beta - (1 + 2\beta)\alpha}{1 - \alpha} \geq \frac{a}{a + t},$$

$$t + \beta(a + t) \geq (t + 2\beta(a + t))\alpha.$$

Additionally, when $\beta > 0$, $x^* \leq x(t)$ whenever

$$\left(\frac{(1 - \alpha)a}{b(1 + \beta - (1 + 2\beta)\alpha)}\right)^{1/\beta} \leq \left(\frac{a + t}{b}\right)^{1/\beta},$$

$$\frac{1 - \alpha}{1 + \beta - (1 + 2\beta)\alpha} \leq \frac{a + t}{a},$$

$$(t + 2\beta(a + t))\alpha \leq t + (a + t)\beta.$$

And the claim follows. $\qquad\square$

Unfortunately, the necessary and sufficient condition $\alpha \leq Y(\beta, t)$ cannot be represented graphically as is the case of $\alpha < Z(\beta)$. In "Appendix" we prove two useful properties of $Y(\beta, t)$:

Remark 4

i. $Y(\beta, t) \in (.5, 1)$ for $\beta \in [-1, 0)$,

ii. $Y(\beta, t) \in (.5, Z(\beta))$ for $\beta > 0$.

The first states that when $\beta < 0$, a sufficient condition for the optimal output not to be in the boundary is that the weight given to the firm is less than the weight given to the consumer. The intuition is that when $\alpha < 1/2$, the optimal output without the maximum loss constraint is indeed a maximum since in this case the planner favors firms, bringing the optimal output closer to the monopoly output. The second states that when $\beta > 0$, $Y(\beta, t)$ is located in the area between 0.5 and $Z(\beta)$. This area becomes thinner and thinner with large values of β.

Propositions 7 and 8 allow us to write the maximum value of social welfare, denoted by W^o, as follows:

$$W^o \equiv W(\alpha, \beta, t) = \begin{cases} a\beta \frac{1-\alpha}{1+\beta} x^*, & 0 \leq \alpha \leq Y(\beta, t), \\ \frac{(((1+\beta) t + \beta(a+t)) \alpha - (1+\beta) t)}{1+\beta} \overline{x}(t), & Y(\beta, t) < \alpha \leq 1. \end{cases} \quad (13)$$

Note that $W(\alpha, \beta, t)$ is continuously differentiable in α. In addition, when Proposition 8 applies, the maximum is located at x^*. On the other hand, when $\alpha > Y(\beta, t)$, the maximum is located in $\overline{x}(t)$.

Next, we compare the previous expression with the welfare considering $n \geq 1$ identical firms under Cournot equilibrium. In such a case equilibrium output is

$$\hat{x}(n) = \left(\frac{an}{b(n + \beta)} \right)^{\frac{1}{\beta}}. \quad (14)$$

Note that in our framework, it could be that socially optimal output (10) and equilibrium output (14) coincide. This occurs for $\alpha(n) = (n - 1)/(2n - 1)$. For instance, under monopoly $\alpha(1) = 0$, i.e. for the monopoly output to be socially optimal, we need that only profits count in social welfare. Also, when $n \to \infty, \alpha(n) \to .5$.

Now taking into Eq. (9), social welfare in Cournot equilibrium is

$$\hat{W}(n) = \frac{(1 + \beta + \alpha(n - 1 - \beta))\beta a}{(1 + \beta)(n + \beta)} \hat{x}(n). \quad (15)$$

As in Anderson and Renault (2003), we define the percentage of welfare losses as follows:

$$L = \frac{W^o - \hat{W}(n)}{W^o}. \quad (16)$$

Now, simple calculations show the following result:

Proposition 9 *The percentage of welfare losses is*

$$L(\alpha, \beta, t, n) = \begin{cases} 1 - \dfrac{(\alpha n + (1+\beta)\,(1-\alpha))\,(n(1+\beta-(1+2\beta)\alpha))^{\frac{1}{\beta}}}{(1-\alpha)^{(1+\beta)/\beta}(n+\beta)^{(1+\beta)/\beta}}, & 0 \le \alpha \le Y(\beta, t), \\[2ex] 1 - \dfrac{\beta a(\alpha n + (1+\beta)\,(1-\alpha))\,(n+\beta)^{-(1+\beta)/\beta}}{((1+\beta)\,t + \beta(a+t))\,\alpha - (1+\beta)\,t}\left(\dfrac{na}{a+t}\right)^{\frac{1}{\beta}} & Y(\beta, t) < \alpha \le 1. \end{cases}$$

(17)

Moreover,

(i) $L(\alpha, \beta, t, n)$ *is continuously differentiable* $\forall \alpha \in [0, 1]$.
(ii) $\partial L(\alpha, \beta, t, n)/\partial \alpha < 0 \; \forall \alpha \in [0, \alpha(n)), \; \partial L(\alpha(n), \beta, t, n)/\partial \alpha = 0,$ *and* $\partial L(\alpha, \beta, t, n)/\partial \alpha > 0 \; \forall \alpha \in (\alpha(n), 1]$.
(iii) *When* $\alpha \ge 0.5 \; \partial L(\alpha, \beta, t, n)/\partial n < 0 \; \forall n \ge 1$; *when* $\alpha \in [0, 0.5)$ *there is a value* $n(\alpha) \equiv (1-\alpha)/(1-2\alpha)$ *such that,* $\partial L(\alpha, \beta, t, n)/\partial n < 0$ *for* $n \in [1, n(\alpha)), \partial L(\alpha, \beta, t, n(\alpha))/\partial n = 0,$ *and* $\partial L(\alpha, \beta, t, n)/\partial n > 0,$ *for* $n > n(\alpha)$.
(iv) $L(0, \beta, t, n) \ge 0, L(\alpha(n), \beta, t, n) = 0$ *and* $L(1, \beta, t, n) < 1$.

Proof The calculation of $L(\alpha, \beta, t, n)$ is straightforward by substituting (13) and (15) in (16).

(i)   Is due to $W(\alpha, \beta, t)$ is continuously differentiable $\forall \alpha \in [0, 1]$.

(ii) $\dfrac{\partial L(\alpha, \beta, t, n)}{\partial \alpha} = \begin{cases} \left(\dfrac{(1+2\beta)n}{n+\beta}\right)^{\frac{1}{\beta}}\dfrac{Z(\beta)(2n-1)}{n+\beta}(1-\alpha)^{-\frac{1}{\beta}-2}(Z(\beta)-\alpha)^{\frac{1}{\beta}-1}(\alpha-\alpha(n)), & 0 \le \alpha \le Y(\beta, t), \\[2ex] \left(\dfrac{a}{a+t}\right)^{\frac{1}{\beta}}\left(\dfrac{n}{n+\beta}\right)^{\frac{1}{\beta}}\dfrac{\beta a}{n+\beta}\dfrac{(1+\beta)(\beta(a+t)+nt)}{((1+\beta)t+\beta(a+t))\alpha-(1+\beta)t)^2}, & Y(\beta, t) < \alpha \le 1. \end{cases}$

Hence, the claim holds $\forall \alpha \in [0, 1]$.

(iii) To prove this, let us assume that n is a real number so that $L(\alpha, \beta, t, n)$ is differentiable with respect to n. In such a case, Eq. (17) can be written as

$$L(\alpha, \beta, t, n) = 1 - K(\alpha n + (1+\beta)\,(1-\alpha))\,n^{\frac{1}{\beta}}(n+\beta)^{-\frac{1}{\beta}-1}, \; n \ge 1.$$
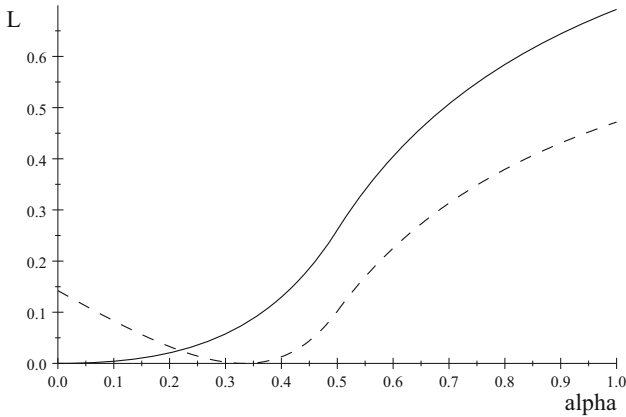
where

$$K = \begin{cases} \dfrac{((1+\beta-(1+2\beta)\alpha))^{\frac{1}{\beta}}}{(1-\alpha)^{(1+\beta)/\beta}}, & 0 \le \alpha \le Y(\beta, t), \\[2ex] \dfrac{\beta a}{((1+\beta)\,t + \beta(a+t))\,\alpha - (1+\beta)\,t}\left(\dfrac{a}{a+t}\right)^{\frac{1}{\beta}} & Y(\beta, t) < \alpha \le 1. \end{cases}$$

Note that K is independent of n. Now compute

$$\frac{\partial L(\alpha, \beta, t, n)}{\partial n} = K(1+\beta)\frac{n^{\frac{1}{\beta}-1}(2n-1)}{(n+\beta)^{\frac{1}{\beta}+2}}(\alpha(n) - \alpha).$$

Which has a unique critical point at $\alpha = \alpha(n)$. It is easily seen that when $\alpha > \alpha(n)$ (resp. $\alpha < \alpha(n)$), $L(\alpha, \beta, t, n)$ is decreasing (resp. increasing) with

Fig. 5 Percentage of welfare losses as a function of α

respect to n. Finally, taking into account that

$$\alpha(n) = \frac{n-1}{2n-1} < 0.5 \leq Y(\beta, t), \forall n \geq 1,$$

The shape of $L(\alpha, \beta, t, n)$ as a function of n is given by the conditions stated in claim (iii).

(iv) $L(\alpha(n), \beta, t, n) = 0$ and $L(1, \beta, t, n) < 1$ are immediate by substitution, and $L(0, \beta, t, n) \geq 0$ follows from tedious calculations.   □

When $\alpha = .5$, (17) collapses in the one by Anderson and Renault (2003). In our case the percentage of welfare losses is a continuous function of α. Also, oligopoly welfare losses are zero when $n = 1$ and consumers do not count on social welfare but are never 100% because the industry always provides some surplus for the consumer. Figure 5 plots the percentage of welfare losses as a function of α for the case of $\beta = 0.4$, $t = 0$ and $a = 1$ when $n = 1$(solid) and $n = 2$(dash). In the second case there are welfare losses for small α due to the profits lost by competition. Note $\alpha(2) = 1/3$.

Figure 6 plots $L(1/3, 1, 1, n)$. $\alpha(= 1/3)$ belongs to the interval considered in the second sentence of claim (iii) of Proposition 9. In this case n(1/3) = 2 and according with the proposition, a change from monopoly to duopoly decreases welfare losses to zero, since duopoly is the preferred number of firms by the regulator. Further increases of competition increase welfare losses given the relatively small appreciation given by the regulator to consumer's welfare.
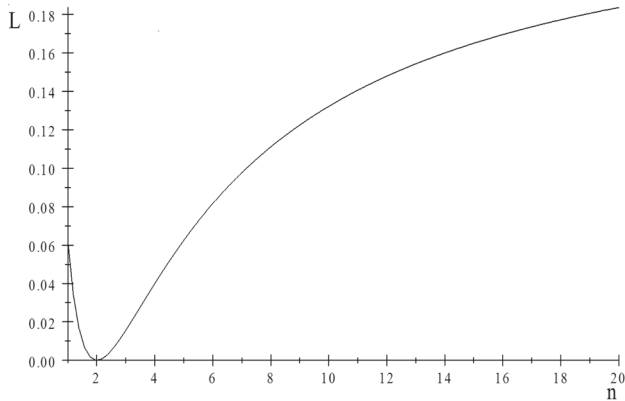
Fig. 6 Percentage of welfare losses as a function of *n*

5 Conclusions

In this paper we study two extensions of standard consumer surplus; namely, the consideration of a nonnegativity constraint (NNC) in the consumption of the outside good and different weights for the consumer and the firm. We have shown that these two extensions are workable so our contribution increases the scope of application of consumer surplus. Our paper can be regarded as a step to the study of welfare losses when consumers have general utility functions, there are several classes of consumers and social welfare is nonlinear on utility and profits.

The existence of a NNC has several important consequences. Firstly, NNC defines an interval of quantities (prices) in which, for those affected consumers, demand becomes unit-elastic and loses income effect neutrality. In other words, quasilinearity is not equivalent to no income effects. In order for income effects to vanish, income must permit the consumer to avoid the NNC. In addition, NNC yields a kink in the demand curve that does not depend on oligopolistic competition or the existence of a "prevailing price". Secondly, when demand is elastic, NNC makes demand less elastic, increasing markups and monopoly power. The more consumers are affected by NNC, the less elastic demand becomes. Thirdly, NNC makes marginal revenue discontinuous, and this affects equilibrium; monopoly equilibrium output may occur at a kink point of marginal revenue and is smaller than the one when NNC does not bite.[9] This suggests that energetic poverty, i.e. when the energy bill meets the minimum consumption of food, shelter, etc., can be used as an explanatory variable of energy prices. A report of the European Commission by Pye and Dobbins (2015) claims that low household income and high energy prices are the basis of energy poverty. Our results show that high energy prices is a cause as well as a consequence. Indeed, it would be difficult to measure if NNC bites or not, but some information can be obtained by looking at demand elasticities before and after a crises and to analyze whether they become significant different other things being equal.

---

[9] This differs from Corchón (2001) in which equilibrium never occurs in not differentiable points.

The introduction of different weights for consumers and firms has significant consequences, as they (1) first-order condition of social welfare maximization is not sufficient, even under standard assumptions on utility and costs; and (2) even if the second-order condition is met, profits can still be negative. Assuming there is an external constraint on how negative profits can be, constant returns and a demand function that generalizes both linear and isoelastic cases, we characterize the zones in which the constraint on profits is satisfied and where first-order condition is sufficient; (3) we offer formulae for the percentage of welfare loss due to oligopoly. Welfare loss depends on the weight given to consumers and the firm and the upper bound to output that satisfies the profit constraint. Using this formula we identify the weight for which an increase in the number of firms decreases social welfare.[10] This may help to explain situations in which the regulator makes entry very difficult.

Our paper opens the door to future research on the role of these two extensions in other fields. We highlight the following possibilities:

(1) To consider these two extensions in other fields like Mechanism Design, International Trade and Natural Resources[11]

(2) To consider the role of consumers affected by NNC in the explanation of countercyclical markups. Specifically, one can conjecture a sequence of events according to which downturns raise the proportion of consumers affected by NNC, which in turn increases demand elasticities and markups.

(3) To reconsider Keynesian multipliers with and without NNC. Our findings suggest that the value of this multiplier depends on the number of consumers affected by NNC and thus it may be significantly different in a recession than in a boom.

(4) To use the model when consumers and producers have different weights to evaluate merger policy when there are distributional concerns. A striking aspect of antitrust policy toward horizontal mergers is that merger reviews usually apply a standard close to consumer surplus assuming that firms only would go for a merger if it is profitable for them. But this procedure disregards entirely the issue of the distribution of the surplus which might matter. In our model the recipe for the optimal merger policy is simple. Allow any merger for $\alpha < \alpha(n)$ and forbid any merger for $\alpha > \alpha(n)$. This follows from proposition 9, part ii). However, the model of our paper is very special. We assume that homogeneous product, identical firms, and constant returns to scale. These conditions have been criticized by the literature on mergers because they yield weird results, see Faulí-Oller and Sandonis (2018), p. 10 and ff. The study of mergers beyond our model is outside the scope of the paper .

---

[10] The traditional reasons for a decrease in competition to be socially optimal, i.e. economies of scale or firms with different technologies do not apply here.

[11] For instance, if the regulator knows the utility function but not the cost function, a mechanism similar to Loeb and Magat (1979) achieves the full optimum when the weights given to consumers and the firm are different. In this case the regulator announces a transfer of $\alpha(U(x) - px)/(1 - \alpha)$ so profits become $(\alpha(U(x) - px) + px(1 - 2\alpha))/(1 - \alpha) - cx$ which is a monotonic transformation of social welfare and thus the maximizer is the same in both cases.

Declarations

Conflicts of interest The authors declare that they have no conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

Appendix

Proof of Lemma 1

a. The intersections depend on the fact that $xU'(x) \geq M$ for any x, and $xU'(x)$ is hump-shaped. Thus, a graphical argument, see Fig. 2, shows that there are, at most, two intersections, $\underline{x}$ *and* $\overline{x}(\underline{x} < \overline{x})$, such that $xU'(x) = M$. This characterizes the set of quantities for which NNC bites as $\Omega^x = \left[\underline{x}, \overline{x}\right]$.

b. If $xU'(x)$ reaches a maximum in $\hat{x}/\hat{x}U'(\hat{x}) \geq xU'(x) \forall x \neq \hat{x}$, On the other hand if $\Omega^x = \left[\underline{x}, \overline{x}\right] \neq \emptyset$, $xU'(x) \geq M$, $\forall x \in \Omega^x$, and the result follows.   □

Proof of Remark 4

$Y(\beta, t) \in (.5, 1)$ for $\beta \in [-1, 0)$, and $Y(\beta, t) \in (.5, Z(\beta))$ for $\beta > 0$.
   Note that

$$Y(\beta, t) = \frac{t + \beta(a + t)}{t + 2\beta(a + t)}.$$

$$\frac{\partial Y(\beta, t)}{\partial \beta} = \frac{-t(a + t)}{(t + 2\beta(a + t))^2}.$$

Thus, $\partial Y(\beta, t)/\partial \beta > 0$ for $\beta \in [-1, 0)$, and $\partial Y(\beta, t)/\partial \beta < 0$ for $\beta > 0$. Moreover, $Y(-1, t) = a/(2a + t) > .5$ because when $\beta = -1$, $a < 0$, and $a + t < 0$; $\lim_{\beta \to 0} Y(\beta, t) = 1$; and, for $\beta > 0$, $.5 < Y(\beta, t) < Z(\beta)$ is immediate.

References

Amir RP, Erickson JJ (2017) On the microeconomic foundations of linear demand for differentiated products. J Econ Theory 169:641–665

Anderson SP, Renault R (2003) Efficiency and surplus bounds in Cournot competition. J Econ Theory 113:253–264

Blackmon G, Zeckhauser R (1992) Fragile commitments and the regulatory process. Yale J Regul 9:73–105

Bulow JI, Pfleiderer P (1983) A note on the effect of cost changes on prices. J Polit Econ 91:182–185

Corchón LC (2001) Differentiable comparative statics with payoff functions not differentiable everywhere. Econ Lett 72:397–401

Corchón LC (2008) Welfare losses under Cournot competition. Int J Ind Organ 26:1120–1131

Corchón LC, Torregrosa RJ (2020) Cournot equilibrium revisited. Math Soc Sci 106:1–10

Dierker E (1991) The optimality of Boiteux-Ramsey pricing. Econometrica 59:99–121

Dixit A (1984) International trade policy for oligopolistic industries. Econ J 94:1–16

Dossche M, Heylen F, Van der Poel D (2010) The kinked demand curve and price rigidity: evidence from scanner data. Scand J Econ 112:723–752

Dupuit AJÉJ (1844) De la mesure de l'utilité des travaux publics. Annales des ponts et chaussées, Second series, 8. Translated by R.H. Barback as "On the measurement of the utility of public works". International Econonomic Papers 1952, 2, 83–110.

Dworczak, P., Kominers, S.K., & Akbarpour, M. (2019). Redistribution through Markets. Becker Friedman Institute for Research in Economics Working Paper No. 2018-16. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3143887.

Faulí-Oller R, Sandonis J (2018) Horizontal mergers in oligopoly. In: Corchón L, Marini M (eds) Chapter 2 in the handbook of game theory and industrial organization. Edward Elgar Publishing, London

Fehr E, Schmidt KM (1999) A theory of fairness, competition, and cooperation. Q J Econ 114(3):817–868

González-Maestre M (2000) Divisionalization and delegation in oligopoly. J Econ Manag Strategy 9:321–338

Harberger AC (1954) Monopoly and resource allocation. Am Econ Rev 5:77–87

Hausman JA (1981) Exact consumer's surplus and deadweight loss. Am Econ Rev 71:662–676

Hillman AL (1982) Declining industries and political-support protectionist motives. Am Econ Rev 72(5):1180–1187

Loeb M, Magat WA (1979) A decentralized method for utility regulation. J Law Econ 22:399–404

Mankiw G (1985) Small menu costs and large business cycles: a macroeconomic model of monopoly. Quart J Econ 100:529–537

Prest AR, Turvey R (1965) Cost-benefit analysis: a survey. Econ J 75:683–735

Pye S, Dobbins A (2015) Energy poverty and vulnerable consumers in the energy sector across the EU: Analysis of policies and measures. European Commission, Energy. https://ec.europa.eu/energy/sites/ener/files/documents/INSIGHT_E_Energy%20Poverty-Main%20Report.pdf

Ritz RA (2018) Oligopolistic competition and welfare. In: Corchón LC, Marini M (eds) Handbook of game theory & industrial organization, vol I. Edward Elgar, Northampton, pp 181–200

Roberts J, Sonnenschein HF (1977) On the foundations of the theory of monopolistic competition. Econometrica 45:101–112

Sweezy P (1939) Demand under conditions of oligopoly. J Polit Econ 47:568–573

Willig RD (1976) Consumer's surplus without apology. Am Econ Rev 66:589–597

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.