Quality distortions in monopolistic competition*

Sergei Kichko†, Alina Ozhegova‡, and Alexander Tarasov§

November 6, 2024

Abstract

In this paper, we explore how heterogeneous firms decide on vertical and horizontal qualities of their products. We show that if increasing the product qualities appears to be relatively costly, more productive firms choose higher vertical quality but lower horizontal quality. We also document distortions that arise in our framework. Specifically, we find that in the market equilibrium, firms tend to underinvest in horizontal quality but overinvest in vertical quality compared to the first best allocation. Using data from pizzerias in Oslo, Norway, we provide a calibration exercise to estimate welfare losses due to the quality distortions.

Keywords: Monopolistic competition, vertical quality, horizontal quality, welfare.
JEL codes: D43, L13.

*We thank Pierre M. Picard for helpful comments and suggestions. The authors have no relevant financial or non-financial interests to disclose. All remaining errors are our own.

†Department of Economics and Management, University of Trento, Italy; e-mail: sergei.kichko@unitn.it, sergey.kichko@gmail.com

‡Department of Economics, Aalto University, Finland; e-mail: alina.ozhegova@aalto.fi, alinozhegova@gmail.com

§Faculty of Economic Sciences, HSE University, Russia; e-mail: atarasov@hse.ru

1 Introduction

One of the robust observations based on micro-level data is a strong within-industry firm heterogeneity (Schott, 2004; Khandelwal, 2010). Firms differ in many respects, including productivities, sales, employment, or, the central focus of this paper, quality of products. Among different sources of firm heterogeneity, quality differentiation across firms plays an important role as it directly affects consumer decision making. This paper considers two product quality dimensions: vertical and horizontal. For instance, in the car industry, vertical quality can be represented by car's reliability, while horizontal quality may be captured by variations in configurations of the same model, such as hatchback or sedan. While the literature on product quality choice is quite extensive, not much attention has been paid to how firms strategically choose both horizontal and vertical quality of a product. In this paper, we aim to fill this gap by studying how firms simultaneously choose both types of product quality, how these decisions depend on firm productivity heterogeneity, and the resulting implications for social welfare.

We develop a general equilibrium single-sector model of monopolistic competition. Firms - heterogeneous in productivity - choose not only the price-quantity pair, but also two types of product quality. Importantly, both qualities are costly for firms, which allows modeling the interplay between exogenous firm productivity and the choice of quality levels. We then determine welfare distortions arising in the model and quantify them.

Our framework builds on Eaton and Fieler (2024) and relies on consumer preferences similar to the constant elasticity of substitution (CES) preferences. In our framework, horizontal quality acts as a product demand shifter in the utility function, functioning as a perfect substitute for the quantity consumed. In contrast, vertical quality complements the quantity, enhancing consumer utility. We assume that production costs rises with both types of quality. However, given the distinction between the two quality types, we differentiate between the overall cost of quality and the cost of horizontal quality, relative to the cost of vertical quality. As is typical in monopolistically competitive with free entry, the zero profit condition and zero expected profit condition determine the production cost cutoff and the number of firms active in the market. We then investigate firms' choice of the qualities, the impact of market size expansion on that choice, and provide normative analysis to uncover distortions associated with the endogenous choice of the qualities.

To isolate the effects of two qualities, we work with CES alike preferences. First, CES demand leads to the optimum product diversity in a monopolistically competitive market (Dixit and Stiglitz, 1977). This result stems from the fact that two distortions of imperfect markets, the price distortion due to pricing above marginal costs and the entry distortion caused by the business stealing effect, are balanced. Demand systems with the variable elasticity of substitution do not feature this property neither for symmetric firms nor for heterogeneous firms (Dhingra and Morrow, 2019). Thus, to encapsulate the pure effect of the quality distortions, we assume away the other distortions by working with CES alike preferences, which still feature the balance between price and entry distortions. Second, the demand specification leads to constant markups in the equilibrium. Thus, price variation across firms is not affected by pro-competitive effects and is solely explained by differences in qualities choice based on firms' productivity.

Our main results can be summarized as follows. First, we show that in the equilibrium, the vertical quality of products increases with firm productivity. At the same time, we find that the relationship between firm productivity and the level of horizontal quality is ambiguous and depends on the parameters in the model. In particular, if increasing the product qualities appears to be relatively costly, more productive firms tend to produce goods with lower horizontal quality. Conversely, when the cost of increasing quality levels is not too high, more productive firms find it profitable to keep high quality standards for both vertical and horizontal qualities. We also show that in our framework, more productive firms can set higher prices for their products. This happens when the quality component in the marginal production cost is relatively substantial, which in turn takes place under relatively low product differentiation in the market. Finally, we find that market size matters for the choice of the qualities. Specifically, market size expansion fosters each firm to produce goods with higher horizontal but lower vertical qualities.

Second, we investigate a problem of utilitarian social planner who chooses the number of products, production amounts, and the qualities of each product. We show that, similar to the standard CES based models, the market provides the optimum levels of firm outputs, the number of entrants, and the production cost cutoff. However, each firm chooses non-optimal levels of both horizontal and vertical qualities. To be precise, in the equilibrium, each firm chooses lower horizontal quality and higher vertical quality compared to the socially optimal levels. These deviations from the optimal quality levels lead to welfare losses.

Third, to quantify the welfare losses caused by the quality distortions, we calibrate some parameters in the model using data from the pizza restaurant industry in Oslo, Norway. The characteristics of the pizza restaurant sector closely align with our theoretical framework. Specifically, we consider a pizza restaurant as a firm producing a differentiated product, where the horizontal quality is interpreted as the number of different pizza varieties offered by the restaurant. Meanwhile, the vertical quality is interpreted in the usual way and proxied by the restaurant's rating. Using the collected data, we uncover all key for the welfare losses parameters besides the elasticity of substitution between the two qualities. We then estimate welfare losses for different values of this elasticity and find that for the moderate values of the elasticity, the losses are somewhat between 30% and 50%: the welfare level in the market equilibrium can be up to the half of the social optimum. These quantitative findings support the idea that the considered quality distortions can be substantial and, therefore, play an important role in analyzing market outcomes in monopolistically competitive markets.

This paper contributes to the growing literature on exogenous and endogenous product quality. In this literature, product quality is typically modeled as a demand shifter in the consumer utility function. For instance, Schott (2004), Hummels and Klenow (2005), Hallak (2006) and many others consider frameworks where product quality is an exogenous parameter. Another strand of this literature (Kugler and Verhoogen, 2012; Feenstra and Romalis, 2014; Fan *et al.*, 2015; Flach and Unger, 2022) considers monopolistically competitive markets with CES consumer preferences and firm heterogeneity where firms choose quality of their products depending on their productivity. Antoniades (2015) follows the same approach in determining endogenous quality by heterogeneous firms, but relies on a non-CES linear demand system. In Picard (2015), the choice of product quality is examined within a framework with a regional dimension: firm's quality choice depends on the region of the firm origin. The quality choice problem is also explored within other market structures (see, for instance, Ma and Burgess Jr., 1993, and Wauthy, 1996, for the analysis of duopoly). The most of this literature agree on a positive correlation between firm's productivity and its quality, which is in turn supported by the empirical evidence (see, for instance, Manova and Zhang, 2012; Kugler and Verhoogen, 2012). Our paper contributes to this literature by qualitatively and quantitatively analyzing the framework, in which firms endogenously choose not only horizontal quality (a demand shifter), but also vertical quality, which complements quantity and horizontal quality.

By introducing different quality types, the present paper also contributes to the literature on monopolistic competition with firm heterogeneity (see Melitz, 2003; Bernard, Redding, and Schott, 2007; Melitz and Ottaviano, 2008, among many others). The paper potentially contributes to the trade literature, since, as it is common in this literature (see, for instance, Krugman, 1979; Zhelobodko *et al.*, 2012), we examine the impact of market size expansion that can be treated as the first approximation for opening up to free trade between countries. However, in contrast to those papers, we assume away the pro-competitive effects such that in our framework markups are invariant to market size (as, for instance, in Krugman, 1980). To some extent, this allows us to disentangle the impact of endogenous choice of qualities on product prices and to uncover distortions associated with the quality choice problem. Our paper is also close to normative studies, which deal with various distortions arising in imperfectly competitive markets (Spence, 1976; Dixit and Stiglitz, 1977; Mankiw and Whinston, 1986; Dhingra and Morrow, 2019).

As noted earlier, we employ the structure of consumer preferences with two qualities from Eaton and Fieler (2024). They use this framework to explore quantitatively the role of quality in the welfare gains from international trade and economic growth. Our paper focuses instead on the closed economy general equilibrium and provides one more step towards understanding the interplay between firm heterogeneity and product quality choice as well as the welfare distortions that can arise in such frameworks.

The remainder of the paper is organized as follows. Section 2 develops the model. Section 3 discusses the market equilibrium with free entry, while Section 4 examines the optimal allocation and compares it with the market outcome. Section 5 calibrates the model. Section 6 concludes.

2 The model

We consider a single-sector economy, which produces a set Ω of varieties of the differentiated good. Following Eaton and Fieler (2024), we distinguish between two types of quality of a certain variety. Horizontal quality reflects the direct characteristics of the product: for instance, the size of the product unit (the weight of a dish served in a certain restaurant or the capacity of a wine bottle). Vertical quality can be considered as a proxy for indirect product features, such as the restaurant's non-food characteristics (the level of service, location, in-door amenities) or the position of the wine in global wine rankings.

2.1 Consumption

There are L identical consumers with the following utility function:

$$U = \int_{\Omega} \left[(Q(\omega)x(\omega))^{\rho} + q(\omega)^{\rho} \right]^{\frac{\beta}{\rho}} \mathrm{d}\omega, \tag{1}$$

where $x(\omega)$ is the quantity of variety $\omega \in \Omega$, $Q(\omega)$ is the horizontal quality, and $q(\omega)$ is the vertical quality. Horizontal quality is the perfect substitute for quantity, whereas vertical quality complements $x(\omega)$. Parameter ρ represents the elasticity of substitution between the two types of quality, while β measures for the elasticity of substitution between different varieties, which is given by $1/(1-\beta)$. We assume that $\beta \in (0, 1)$ and $\rho < 0$, which implies complementarity between the two qualities in consumption.

Consumers choose optimal quantities subject to the budget constraint: $\int p(\omega)x(\omega)d\omega = w$, where w is the wage level per unit of labor, which is chosen as a numéraire. The first order condition of the consumer problem implies the inverse demand function for variety ω:

$$p(\omega) = \frac{\beta}{\lambda} \left((Q(\omega)x(\omega))^{\rho} + q(\omega)^{\rho} \right)^{\frac{\beta}{\rho}-1} Q^{\rho}(\omega)x^{\rho-1}(\omega), \tag{2}$$

where λ is the Lagrange multiplier given by

$$\lambda = \beta \int \left[(Q(\omega)x(\omega))^{\rho} + q(\omega)^{\rho} \right]^{\frac{\beta}{\rho}-1} (Q(\omega)x(\omega))^{\rho} \, \mathrm{d}\omega. \tag{3}$$

2.2 Production

Each firm draws its cost of production c from the cumulative distribution $G(c)$. Hereafter, we suppress the variety index ω for notation convenience. The profit function of a firm with cost c

(as a function of its output, price, and two qualities) is given by

$$\pi(p, y, Q, q, c) = py - T(Q,\, q,\, y,\, c), \tag{4}$$

where $T(Q,\, q,\, y,\, c)$ is the total cost of production that depends on two qualities of the variety, marginal cost c and firm output y. To make the model analytically tractable, we assume that the total cost function has the following form:

$$T(Q,\, q,\, y,\, c) \equiv F + (f + yc)(Q^{\alpha} q)^{\gamma}, \tag{5}$$

with α,$\gamma > 0$. Thus, we distinguish between fixed costs which are related to the qualities (represented by f) and those which do not (represented by F). The above specification means that f and the marginal costs c depend on the qualities of the product in the same positive way represented by $(Q^{\alpha} q)^{\gamma}$. The latter stands for the costs of qualities, where parameter α reflects the cost of production of the horizontal quality Q relative to the vertical quality q. In other words, higher α means higher cost of improvements in horizontal quality relative to vertical quality. Note that the cost function (5) implies the supermodularity property of the two types of quality: $T_{Qq} \equiv \partial^2 T / \partial Q \partial q > 0$; which captures the idea that the cost of quality improvement is greater when the level of the other quality is higher. Note that the assumed specification is rich enough to generate a number of interesting implications, while assuming a more general total cost function leads to unnecessary complexity of the analysis.

Plugging the demand function (2) into the profit (4), we obtain

$$\pi(y, Q, q, c) = \frac{\beta}{\lambda} L \left[\left(\frac{Qy}{L} \right)^{\rho} + q^{\rho} \right]^{\frac{\beta}{\rho} - 1} \left(\frac{Qy}{L} \right)^{\rho} - (f + yc)(Q^{\alpha} q)^{\gamma} - F, \tag{6}$$

where λ is taken by firms as given. Each firm simultaneously chooses output y and two qualities Q and q. Let us assume for the moment that the triple $\{y, Q, q\}$ maximizing the profit function is strictly positive. In this case, it solves the following first order conditions:

$$\frac{\partial \pi}{\partial y} = 0 \Longleftrightarrow \left[\beta \left(\frac{Qy}{L} \right)^{\rho} + \rho q^{\rho} \right] G(y, Q, q) = c,$$

$$\frac{\partial \pi}{\partial Q} = 0 \Longleftrightarrow y \left[\beta \left(\frac{Qy}{L} \right)^{\rho} + \rho q^{\rho} \right] G(y, Q, q) = \alpha\gamma(f + yc),$$

$$\frac{\partial \pi}{\partial q} = 0 \Longleftrightarrow (\beta - \rho) y q^{\rho} G(y, Q, q) = \gamma(f + yc),$$

where

$$G(y, Q, q) \equiv \frac{\beta Q}{\lambda (Q^{\alpha} q)^{\gamma}} \left(\frac{Qy}{L} \right)^{\rho - 1} \left[\left(\frac{Qy}{L} \right)^{\rho} + q^{\rho} \right]^{\frac{\beta}{\rho} - 2}.$$

After some simplifications, the first order conditions can be rewritten in the following way:

$$yc = \frac{\alpha\gamma}{1-\alpha\gamma} f, \tag{7}$$

$$\beta \left(\frac{Qy}{qL}\right)^{\rho} = \alpha(\beta - \rho) - \rho, \tag{8}$$

$$(\beta - \rho) y q^{\rho} G(y, Q, q) = \gamma(f + yc). \tag{9}$$

Combining (7)-(9), it is straightforward to show that the above first-order conditions imply that the price of the variety is given by

$$p(c) = \frac{\beta}{\lambda} \left[\left(\frac{Qy}{L}\right)^{\rho} + q^{\rho}\right]^{\frac{\beta}{\rho}-1} Q^{\rho} \left(\frac{y}{L}\right)^{\rho-1} = \frac{1+\alpha}{\alpha\beta} c(Q^{\alpha}q)^{\gamma}, \tag{10}$$

while the profit function can be written as follows:

$$\pi = py - (f + yc)(Q^{\alpha}q)^{\gamma} - F = \frac{\gamma(1+\alpha) - \beta}{(1-\alpha\gamma)\beta} f(Q^{\alpha}q)^{\gamma} - F. \tag{11}$$

Note that if $\beta - \gamma < \alpha\gamma < 1$, the first term in the right-hand side of (11) is strictly positive and increasing in both qualities q and Q. Next, we formulate a lemma regarding the existence of optimal $\{y, Q, q\}$.

Lemma 1. *If $\beta - \gamma < \alpha\gamma < 1$, then, for any $c > 0$, there exists a unique strictly positive and finite triple $\{y, Q, q\}$ that maximizes the firm profit $\pi(y, Q, q, c)$.*

Proof. In the Appendix A. □

Notice that the markup is equal to

$$m(c) = \frac{p(c) - c(Q^{\alpha}q)^{\gamma}}{p(c)} = 1 - \frac{\alpha\beta}{1+\alpha}. \tag{12}$$

As can be seen, the markup does not depend on the marginal costs c and costs of qualities $(Q^{\alpha}q)^{\gamma}$. That is, our framework resembles the monopolistic competition model with CES preferences. As in the standard monopolistic competition models, the markup is lower when varieties are less differentiated (higher β). This is the standard effect of monopolisitcally competitive markets where tougher competition leads to reduction in firms' markup. However, there is an additional term in the markup, $\alpha/(1+\alpha)$, that stands for the presence of qualities in the utility function. In particular, higher α leads to lower firms' markups. The intuition behind this additional term is as follows. Consider the price elasticity of demand that can be derived from (2):

$$\frac{\partial x}{\partial p}\frac{p}{x} = -\frac{1 + \left(\frac{Qx}{q}\right)^{\rho}}{1 - \rho + (1-\beta)\left(\frac{Qx}{q}\right)^{\rho}}.$$

As can be seen, the absolute value of the elasticity is increasing in $(Qx/q)^{\rho}$. For instance, if $(Qx/q)^{\rho}$ is close to infinity, the price elasticity is close to that derived in the standard monopolistic competition framework with CES preference, $-1/(1-\beta)$. Since parameter α represents, to some extent, the cost of production of the horizontal quality Q relative to the vertical quality q, an increase in this parameter decreases the ratio Qx/q in the market equilibrium (this can be seen from (8)) and, therefore, raises $(Qx/q)^{\rho}$ (as ρ is negative). Hence, a rise in α makes demand for a certain product more elastic, which in turn reduces firms' markups.

3 The free entry equilibrium

As in a standard Melitz type model of monopolistic competition with firm heterogeneity, firms pay the sunk entry cost F_e to enter the market and learn their cost of production c. The least productive firms, with $c > c^*$, leave the market, as their potential profits are negative. In this case, we define the free entry equilibrium in the following way. Given the Lagrange multiplier λ, firms choose the triple $\{y(c), Q(c,\lambda), q(c,\lambda)\}$ that solves (7), (8), and (9). The cutoff cost of production $c^*(\lambda)$ is determined from the zero profit condition (the profit of a marginal firm is equal to zero):

$$(Q^{\alpha}(c^*,\lambda)q(c^*,\lambda))^{\gamma} = \frac{F}{f}\frac{(1-\alpha\gamma)\beta}{\gamma(1+\alpha)-\beta}. \tag{13}$$

The Lagrange multiplier λ that can be found from the free entry condition:

$$\int_0^{c^*(\lambda)} \left[\frac{\gamma(1+\alpha)-\beta}{(1-\alpha\gamma)\beta} f(Q^{\alpha}(c,\lambda)q(c,\lambda))^{\gamma} - F\right] \mathrm{d}G(c) = F_e. \tag{14}$$

Finally, the number of entrants N is pined down by the labor market clearing condition:

$$L = N\left[\int_0^{c^*(\lambda)} \left(F + (f + yc)(Q^{\alpha}(c,\lambda)q(c,\lambda))^{\gamma}\right) \mathrm{d}G(c) + F_e\right].$$

In particular, using (7), one can obtain the following expression for N:

$$N = \frac{L}{FG(c^*(\lambda)) + \frac{f}{1-\alpha\gamma}\int_0^{c^*(\lambda)}(Q^{\alpha}(c,\lambda)q(c,\lambda))^{\gamma}\mathrm{d}G(c) + F_e}. \tag{15}$$

3.1 Choice of qualities and production costs

We start with exploring how the cost of production c affects the equilibrium choice of the qualities. In particular, we consider the dependence of Q and q on c. The first thing to notice is that due to the envelope theorem

$$\frac{\partial \pi(y, Q, q, c)}{\partial c} = -y(Q^\alpha q)^\gamma < 0.$$

Taking into account (11), this immediately implies that costs of qualities $(Q^\alpha q)^\gamma$ is decreasing in c. Note also that by plugging (7) and (8) into (9), one can derive that

$$\frac{L}{\lambda}(\beta - \rho)\left(\alpha(\beta - \rho) - \rho\right)\left(\frac{(1+\alpha)(\beta - \rho)}{\beta}\right)^{\frac{\beta}{\rho}-2} q^\beta = \frac{f\gamma}{1-\alpha\gamma}(Q^\alpha q)^\gamma. \tag{16}$$

Since the right-hand side in the above equation is decreasing in c, while the left-hand side increases with q due to positive β, the vertical quality q is unambiguously decreasing in c. Moreover, the latter equation can be rewritten as follows:

$$\frac{L}{\lambda}(\beta - \rho)\left(\alpha(\beta - \rho) - \rho\right)\left(\frac{(1+\alpha)(\beta - \rho)}{\beta}\right)^{\frac{\beta}{\rho}-2} q^{\beta-\gamma} = \frac{f\gamma}{1-\alpha\gamma}Q^{\alpha\gamma}.$$

As $\alpha\gamma$ is positive, the horizontal quality Q is decreasing in c if and only if $\beta - \gamma$ is positive. Hence, the following proposition holds.

Proposition 1. *The vertical quality q decreases with the firm marginal cost c. The horizontal quality Q increases with c if and only if $\gamma > \beta$.*

The above results imply that more productive firms choose a higher or lower horizontal quality depending on the interplay between the elasticity of substitution among varieties, captured by β, and the cost of quality represented by γ. If the cost of quality is high, $\gamma > \beta$, more productive firms find it profitable to keep higher standards only for the vertical quality q.[1] In this case, less productive firms compete with more productive ones by choosing a higher horizontal quality and, thereby, differentiating their products from those produced by more productive firms. Indeed, the quality difference is a source for vertical product differentiation, which complements horizontal product differentiation captured by β. When the cost of quality is relatively low, $\gamma < \beta$, more productive firms choose both higher horizontal and vertical qualities for their products. To differentiate their goods, less productive firms find it optimal to produce goods with both qualities being lower. Note that the level of horizontal product differentiation captured by β plays a role here. Lower product differentiation (higher β) leads to tougher competition in the market and makes the inequality $\gamma > \beta$ less likely to hold. As a result, tougher

[1] In particular, this case takes place, when $\gamma > 1$, as β is strictly lower than unity.

competition can foster more productive firms to produce goods with higher vertical and horizontal qualities.

Finally, combining (7) and (8), and substituting for q in (10), we derive that

$$p(c) = C_1 c^{1-\beta} Q^{\beta}(c),$$

where C_1 is a positive constant. In words, the product price is determined by the exogenous production cost c and the endogenous choice of the horizontal quality Q. The relative importance of the ingredients is represented by the elasticity of substitution β. In particular, if β is sufficiently close to unity (varieties are close substitutes), the price is mostly determined by the horizontal quality. In this case, the price can decrease or increase as a function of c depending on the sign of $\beta - \gamma$. If, however, β is close to zero, then the price is mainly determined by the production cost c. In the next section, we derive an explicit expression for the product price in the equilibrium and show exactly how the price depends on the production cost c.

3.2 Market size effects

We now turn our attention to the effects of market size. To this end, we formulate next lemma which helps to describe the properties of the equilibrium.

Lemma 2. *In the equilibrium,*

$$\frac{Q^{\alpha}(c,\lambda)q(c,\lambda)}{Q^{\alpha}(c^*,\lambda)q(c^*,\lambda)} = \left(\frac{c}{c^*}\right)^{\frac{\alpha\beta}{\beta-\gamma(1+\alpha)}},$$

with the cutoff c^* solving

$$(c^*)^{\frac{\alpha\beta\gamma}{\gamma(1+\alpha)-\beta}} \int_0^{c^*} c^{-\frac{\alpha\beta\gamma}{\gamma(1+\alpha)-\beta}} dG(c) = \frac{F_e}{F} + G(c^*). \tag{17}$$

Proof. In the Appendix B. □

As can be seen from Lemma 2, the cutoff c^* does not depend on the Lagrange multiplier λ and the market size L. This outcome is similar to that in the standard Melitz type model of monopolistic competition with CES preferences.

Using Lemma 2, we show in the same Appendix that the horizontal and vertical qualities in the equilibrium are given by

$$q(c) = \left(\frac{L}{f}\right)^{-\frac{\alpha}{1+\alpha}} (c^*B)^{-\frac{\alpha}{1+\alpha}} \left(\frac{c}{c^*}\right)^{-\frac{\alpha\gamma}{(1+\alpha)\gamma-\beta}} \left(\frac{F}{f}D\right)^{\frac{1}{\gamma(1+\alpha)}}, \tag{18}$$

$$Q(c) = \left(\frac{L}{f}\right)^{\frac{1}{1+\alpha}} (c^* B)^{\frac{1}{1+\alpha}} \left(\frac{c}{c^*}\right)^{-\frac{\beta-\gamma}{(1+\alpha)\gamma-\beta}} \left(\frac{F}{f} D\right)^{\frac{1}{\gamma(1+\alpha)}}, \tag{19}$$

where B and D are positive constants. It is straightforward to see that, for each product, a rise in the market size L decreases the vertical quality and increases the horizontal one. To provide the intuition for these findings, we consider the profit function in (11) that can be rewritten as follows:

$$\pi(y, Q, q, c) = \frac{\beta}{\lambda} L^{1-\beta} \left((Qy)^\rho + (Lq)^\rho\right)^{\frac{\beta}{\rho}-1} (Qy)^\rho - (f + yc)(Q^\alpha q)^\gamma - F.$$

As can be seen, the impact of L on the profit can be decomposed into two effects. First, a rise in $L^{1-\beta}$ (recall that β is strictly less than unity) increases the first term in the profit function, making investment in both qualities more profitable. Second, a rise in L decreases L^ρ, which also in turn increases the first term in the profit function, but at the same time reducing the returns from investment in the vertical quality. The latter is explained by the complementarity between individual consumption x, which falls under a rise in L (given the same total output y), and vertical quality q in the utility function. As a result, the overall effect of a rise in L is positive for the horizontal quality and negative for the vertical one. Note that one needs to take into account the effect of L on qualities through entry, which is captured by λ in the profit function. However, changes in λ "uniformly" shift the first term in the profit function, which does not affect our intuition. We summarize our findings in the following proposition.

Proposition 2. *Each firm increases horizontal quality Q and decreases vertical quality q of its product with larger market size L.*

It is worth noting that the trade literature acknowledges the positive effect of market size on horizontal quality (Antoniades, 2015; Picard, 2015). The above proposition in turn states that an improvement in horizontal quality comes at the cost of lower vertical quality. In other words, it reflects the idea that firms in larger markets gain more from mass-market production rather than from hi-end products.

The findings in Lemma 2 also allow deriving an explicit expression for the product price. In particular,

$$p(c) = \frac{1}{\beta}\left(1 + \frac{1}{\alpha}\right) c(Q^\alpha q)^\gamma = \frac{1+\alpha}{\alpha\beta}\frac{F}{f} D c^{1-\frac{\alpha\beta\gamma}{\gamma(1+\alpha)-\beta}} (c^*)^{\frac{\alpha\beta\gamma}{\gamma(1+\alpha)-\beta}}. \tag{20}$$

In standard Melitz type models, prices increase with production cost c. However, the presence of the qualities leads to more complex price patterns. As can be inferred, the price is increasing in c if and only if $1 + \alpha\beta\gamma/(\beta - \gamma - \alpha\gamma) > 0$. The idea behind is straightforward. Recall that $p(c) = C_1 c^{1-\beta} Q^\beta(c)$. When $\gamma > \beta$, $Q(c)$ is increasing in c and, as a result, $p(c)$ is increasing.[2] If, however, the opposite holds, then $Q(c)$ is decreasing and the behavior of $p(c)$ depends on the sigh of $1 + \alpha\beta\gamma/(\beta - \gamma - \alpha\gamma)$. As $\gamma < \beta$ is a necessary condition for $p(c)$ to decrease with c, the

[2] It is easy to check that if $\gamma > \beta$, $1 + \alpha\beta\gamma/(\beta - \gamma - \alpha\gamma) > 0$.

latter takes place only when less productive firms produce varieties with both lower horizontal and vertical quality. It is worth noting that $1 + \alpha\beta\gamma/(\beta - \gamma - \alpha\gamma) > 0$ is equivalent to

$$\beta < \frac{\gamma(1+\alpha)}{1+\alpha\gamma}. \tag{21}$$

Condition (21) implies that the level of product differentiation is high enough, implying the sufficiently low level of competition between firms. Furthermore, (21) shows that prices increase with c for any level of product differentiation if and only if $\gamma \geq 1$. This is because the right hand side of (21) in this case is larger than or equal to unity. We summarize our findings in the following proposition.

Proposition 3. *Equilibrium prices decrease with production cost c if and only if the level of production differentiation is sufficiently low: $\beta > (\gamma(1+\alpha)) / (1+\alpha\gamma)$.*

Finally, note that, since c^* is not affected by the market size, firms' prices (20) are not affected by it as well. This is again similar to the corresponding outcome in a standard monopolistic competition framework with CES preferences.

4 Social optimum

In this section, we find the first best outcome in the model assuming that the social planner can choose the number of products in the market and production amounts and qualities of each product to maximize the consumer utility. Specifically, the social planner solves the following optimization problem:

$$\max_{N_o, c_o^*, x_o(c), q_o(c), Q_o(c)} N_o \int_0^{c_o^*} \left((Q_o(c)x_o(c))^\rho + q_o(c)^\rho\right)^{\frac{\beta}{\rho}} \mathrm{d}G(c)$$

subject to the resource constraint

$$L = N_o \left[\int_0^{c_o^*} \left(F + (f + Lcx_o(c))(Q_o^\alpha(c)q_o(c))^\gamma\right) \mathrm{d}G(c) + F_e\right]. \tag{22}$$

In the Appendix C, we show that the equilibrium results in socially optimal total firm outputs, the number of entrants, and the production cost cutoffs. However, the qualities firms choose are not optimal. In particular, the following proposition holds.

Proposition 4. *In the market equilibrium, total firm outputs, the number of entrants, and the production cost cutoffs are socially optimal, whereas each firm chooses lower horizontal quality and higher vertical quality compared to the socially optimal levels.*

Proof. In the Appendix C. □

As can be inferred, the comparison of the market equilibrium with the social planner outcome resembles that in the traditional model of monopolistic competition with CES preferences, where the market equilibrium is socially optimal (see Dixit and Stiglitz, 1979, and Dhingra and Morrow, 2019). However, in our framework, the presence of product qualities leads to a market distortion associated with inefficient investment in these qualities.

In particular, in the Appendix C we show that each firm employs the socially optimal amount of labor, as $Q^\alpha q = Q_0^\alpha q_0$ holds for any marginal cost c. However, the "combination" of the qualities is determined by

$$\left(\frac{Q_o(c)x_o(c)}{q_o(c)}\right)^\rho = \alpha \tag{23}$$

in the social optimum (see the Appendix C) and by

$$\left(\frac{Q(c)x(c)}{q(c)}\right)^\rho = \frac{\alpha(\beta-\rho)-\rho}{\beta}$$

in the market equilibrium (see (8)). Since $\rho < 0$, $(\alpha(\beta-\rho)-\rho)/\beta > \alpha$, and $Q^\alpha q = Q_o^\alpha q_o$, we have that $Q(c) < Q_o(c)$ and $q(c) > q_o(c)$ as stated in Proposition 3. The inefficient investment in the qualities in the market outcome are explained by different objective functions the social planner and firms in the market equilibrium face. Indeed, in the social planner problem, $Q_o(c)$ and $q_o(c)$ are relatively symmetric: the relative size of $Q_o(c)$ and $q_o(c)$ is mainly determined by $x_o(c)$ and α (see (23)). For instance, if α and $x_o(c)$ are equal to unity, the qualities are equal to each other in the social optimum. In contrast, in the market equilibrium, firms maximize their profits, which follows from the consumer demand function and represented by (6). As can be seen, the qualities in this function are not symmetric. Specifically, there is term $(Q(c)x(c))^\rho$ in the revenue part of the profit function, which is decreasing in $Q(c)$ and makes the role of qualities for firms asymmetric compared to the social optimum. As a result, in the market equilibrium, there are fewer incentives to invest in the horizontal quality.

4.1 Welfare losses

In this section, we quantitatively compare consumer welfare in the market equilibrium with that in the social optimum. As the measure of the individual welfare, we consider

$$W = \left(\int \left((Q(\omega)x(\omega))^\rho + q(\omega)^\rho\right)^{\frac{\beta}{\rho}} d\omega\right)^{\frac{1}{\beta}}$$

evaluated at the market and social optimum outcomes (which is in fact the indirect utility derived from (1) to the power of $1/\beta$).

Using (8), in the market equilibrium, the latter can be rewritten as

$$W = \left(\frac{(1+\alpha)(\beta-\rho)}{\beta}\right)^{\frac{1}{\rho}} \left(N \int_0^{c^*} q(c)^\beta \mathrm{d}G(c)\right)^{\frac{1}{\beta}}.$$

Similarly, in the social optimum, taking into account (23), we derive that

$$W_o = (1+\alpha)^{\frac{1}{\rho}} \left(N_o \int_0^{c_o^*} q_o(c)^\beta \mathrm{d}G(c)\right)^{\frac{1}{\beta}}.$$

Thus, the relative welfare is given by (remember that the number of firms in the market equilibrium is the same as that in the social optimum)

$$\frac{W}{W_o} = \left(\frac{\beta-\rho}{\beta}\right)^{\frac{1}{\rho}} \left(\frac{\int_0^{c^*} q(c)^\beta \mathrm{d}G(c)}{\int_0^{c_o^*} q_o(c)^\beta \mathrm{d}G(c)}\right)^{\frac{1}{\beta}}. \tag{24}$$

As can be seen, to evaluate the relative welfare one needs to know $q_o(c)$. The following lemma holds.

Lemma 3. *In the social optimum, the vertical quality is given by*

$$q_o(c) = \left(\frac{L}{f}\right)^{-\frac{\alpha}{1+\alpha}} \left(\frac{1-\alpha\gamma}{\alpha\gamma}\alpha^{\frac{1}{\rho}} c_o^*\right)^{-\frac{\alpha}{1+\alpha}} \left(\frac{c}{c_o^*}\right)^{-\frac{\alpha\gamma}{\gamma(1+\alpha)-\beta}} \left(\frac{F}{f}D\right)^{\frac{1}{\gamma(1+\alpha)}},$$

where D is the same constant as in (18).

Proof. In the Appendix D. □

Using the result in Lemma 3, (18), and that $c^* = c_o^*$, we derive

$$\frac{W}{W_o} = \left(\frac{\beta-\rho}{\beta}\right)^{\frac{1}{\rho}} \left(\frac{\alpha\gamma B}{(1-\alpha\gamma)\alpha^{\frac{1}{\rho}}}\right)^{-\frac{\alpha}{1+\alpha}}.$$

Taking into account that (see the proof of Lemma 2 in the Appendix B)

$$B = \frac{1-\alpha\gamma}{\alpha\gamma}\left[\frac{\alpha(\beta-\rho)-\rho}{\beta}\right]^{1/\rho},$$

we can formulate the following proposition.

Proposition 5. *The individual welfare in the market equilibrium relative to that in the social optimum is given by*

$$\frac{W}{W_o} = \left(\frac{\beta - \rho}{\beta}\right)^{\frac{1}{\rho}} \left(\frac{\alpha\beta}{\alpha(\beta - \rho) - \rho}\right)^{\frac{\alpha}{(1+\alpha)\rho}}. \tag{25}$$

Proof. The proof directly follows from (25). □

It is worth noting that the relative welfare in the above proposition does not depend on γ, which captures the joint impact of the qualities on the production costs, but depends on α that stands for the relative role of the horizontal quality with respect to the vertical one in the production.

5 Calibration

In this section, we calibrate the model to assess the welfare losses caused by the quality distortions, focusing on the pizza restaurant industry in Oslo, Norway. The characteristics of the pizza restaurant sector closely align with our theoretical framework. Specifically, we consider a pizza restaurant as a firm producing a differentiated product, where the horizontal quality is interpreted as the number of different pizza varieties offered by the restaurant. Meanwhile, the vertical quality is interpreted in the usual way and proxied by the restaurant's rating.

5.1 Data description

The data we use for calibrating the model come from different sources. First, we leverage Google Maps to compile the list of restaurants in Oslo available in May 2023. In particular, we collect the name of a restaurant, rating, and the number of reviews, geographical location, and website URL. Further, we collect the information on pizza prices and menu offerings. This step requires collecting data from multiple sources, such as visiting the restaurants' official websites, scrutinizing photos of their menus available on TripAdvisor and Google Maps, and, in some cases, conducting in-person visits. For each pizzeria, we calculate the number of pizza varieties offered and the average price of a pizza. Finally, we augment our dataset with data on each restaurant's revenue using the information on the website proff.no.[3] During the data collection process, we intentionally exclude certain establishments that, although serving pizza, do not meet the criteria for traditional pizzerias. For example, we exclude a restaurant Olivia, which offers a wide variety of dishes but only a few pizza options. Ultimately, we obtain data on 56 pizza restaurants, forming our final dataset. Descriptive statistics for this dataset are presented in Table 1.

[3] `proff.no` is a Norwegian business information website that provides comprehensive data on firms operating in Norway. It offers detailed financial information, including revenue, profit, and other key business metrics. The platform is widely used for business analysis and market research, offering information about the financial health and performance of firms across various industries in Norway.

The Google Maps ratings for pizza restaurants range from 2.7 to 5, with an average rating of 4.27. On average, a restaurant offers around 21 different pizza varieties, though this number varies significantly, ranging from 7 to 47 different types, highlighting a wide dispersion in menu diversity. The average price of a pizza is NOK 184.45, with prices spanning from NOK 110.3 to NOK 234.4. Restaurant revenues also exhibit substantial variation, with values ranging from NOK 0.50 million to NOK 40.03 million. Finally, the quantity metric, defined as revenue divided by the average price, varies between 2.5 to 180 thousand pizzas and, on average, amounts to 47.

Table 1: Descriptive Statistics

	mean	std	min	median	max
Google Maps rating	4.27	0.39	2.70	4.30	5.00
Number of offerings in menu	21.61	11.54	7.00	18.00	47.00
Average price, NOK	184.45	32.56	110.33	184.70	234.39
Revenue, mln NOK	8.77	8.17	0.50	6.16	40.03
Quantity, thou.	47.24	38.13	2.47	38.40	180.04

Note: For the rating of Google Maps, 5 is the highest rating.
Quantity is calculated as revenue divided by average price

5.2 Calibration strategy and results

To calibrate the parameters that are necessary to compute the welfare losses, we start with the firm revenue $r(c) = p(c)y(c)$. Using (7) and (10), we derive that

$$r(c) = \frac{f\gamma(1+\alpha)}{(1-\alpha\gamma)\beta}(Q(c)^{\alpha}q(c))^{\gamma}. \tag{26}$$

Taking logs results in

$$\log(r(c)) = \text{const}_1 + \alpha\log(Q(c)) + \alpha\gamma\log(q(c)), \tag{27}$$

where

$$\text{const}_1 = \log\left(\frac{(1+\alpha)\gamma f}{(1-\alpha\gamma)\beta}\right).$$

Since we observe firms' revenue and both types of their product quality, we can run the across firms OLS regression in (27) and then set α and γ in the model to the corresponding OLS estimates.[4] Specifically, we have that $\alpha = 0.47$ and $\gamma = 0.23$. Notice that in this case, $\alpha\gamma < 1$, fitting the condition on the parameters formulated in Lemma 1.

Next, from 7) and (20) one can derive that in the equilibrium,

[4]Note that we use OLS estimation to identify correlation rather than to establish a causal link.

$$\log(p(c)) = \text{const}_2 - \left(1 - \frac{\alpha\beta\gamma}{(1+\alpha)\gamma - \beta}\right)\log(y(c)). \tag{28}$$

To calibrate β, we estimate the above OLS regression and then substitute for the derived values of α and γ. As a result, $\beta = 0.31$.[5] Again, the condition in Lemma 1, $\beta - \gamma < \alpha\gamma$, is satisfied.

Note that to assess the welfare losses caused by the quality distortions, we also need to calibrate the elasticity of substitution between the two types of quality, ρ. Unfortunately, the employed data set does not allow us to do that. To handle this issue, we build a graph that illustrates the dependence of the welfare losses on the value of ρ. Specifically, we substitute the calibrated values of α and β into (25) and then depict the following function of ρ:

$$\frac{W}{W_o}(\rho) = \left(\frac{0.31-\rho}{0.31}\left(\frac{0.31 * 0.47}{0.47\,(0.31-\rho)-\rho}\right)^{\frac{0.47}{1+0.47}}\right)^{\frac{1}{\rho}}. \tag{29}$$

The above function is plotted in Figure 1.

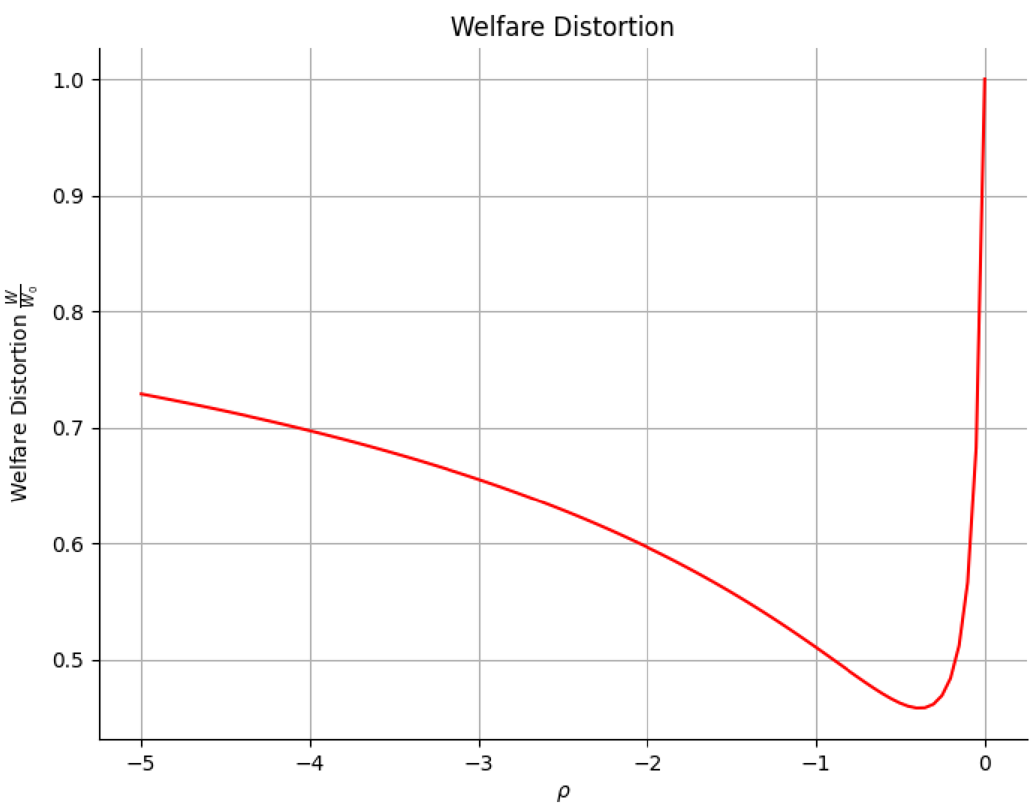

Figure 1: Welfare Distortion

As can be seen, the dependence of the welfare losses on ρ is not monotonic. If ρ is close to

[5] The estimate of β remains 0.31 if we use price of Margarita instead of average price of pizza.

zero or sufficiently large (in the absolute value), then the welfare losses are close to zero: the market equilibrium is close to the socially optimal allocation. The losses are the highest when ρ is around -0.4 and constitute more than 50%: the welfare level in the market equilibrium is more than two times lower than that in the social optimum. For higher (in the absolute terms) values of ρ, the losses are smaller, but still substantial. For instance, when ρ is around -4,the losses constitute around 30%. These quantitative findings support the idea that the considered quality distortions can be substantial and, therefore, play an important role in analyzing market outcomes in monopolistically competitive frameworks.

6 Conclusion

This paper considers a monopolistic competition model where quality of products produced by heterogeneous firms has two dimensions: horizontal and vertical. The horizontal product quality is modeled as a demand shifter in the CES alike consumer utility function, which substitutes the quantity consumed. The vertical quality in turn complements the quantity in the utility function. We find that, in the general equilibrium with free entry, more productive firms choose higher vertical quality of their products, while the choice of horizontal quality depends on the cost of the qualities in the production. In particular, if the cost of increasing the qualities is relatively high, more productive firms find it optimal to choose lower horizontal quality for their products. If, however, increasing the qualities is not too costly, more productive firms tend to keep high standards for both vertical and horizontal qualities. In the paper, we also consider the impact of market size expansion on the quality level, which can be interpreted as the effects of opening up to free international trade. We show that larger market size results in higher horizontal quality of all products, but comes at the cost of lower vertical quality. In other words, firms in larger markets may gain more from mass-market production rather than from hi-end products.

We then compare the market equilibrium with the first best allocation and show that, like in standard CES based models, the market provides the optimum levels of firm outputs, the number of entrants, and the production cost cutoff. However, in the market equilibrium, each firm chooses lower horizontal quality and higher vertical quality compared to the socially optimal levels. Using data from the pizza restaurant industry in Oslo, Norway, we calibrate the model to assess the size of the quality distortions. We find that the welfare losses associated with the quality distortions may be substantial: the welfare level in the market equilibrium can be up two times lower than that in the social optimum.

There is a number of potential extensions of the considered framework. As there are distortions in the market equilibrium, one of the next steps can be the policy analysis aimed at eliminating these distortions. Another direction of further research can be related to the role of income distribution in firms' choice of horizontal and vertical quality, which is assumed away in the present paper. Finally, generalizing consumer preferences seems to be an important step forward. It will allow considering the interplay of the quality distortions with the price and entry distortions.

References

[1] Antoniades, A. 2015. Heterogeneous firms, quality, and trade. *Journal of International Economics*, Vol. 95(2): 263-273.

[2] Bernard, A. B., Redding, S. J., and Schott, P. K. 2007. Comparative advantage and heterogeneous firms. *The Review of Economic Studies*, Vol. 74(1): 31-66.

[3] Dhingra, S., and Morrow, J. 2019. Monopolistic competition and optimum product diversity under firm heterogeneity. *Journal of Political Economy*, Vol. 127(1): 196-232.

[4] Dixit, A. K., and Stiglitz J. E. 1977. Monopolistic competition and optimum product diversity. *The American Economic Review*, Vol. 67(3): 297-308.

[5] Eaton, J., and Fieler, A. C. 2024. The margins of trade. *Econometrica, forthcoming*.

[6] Fan, H., Li, Y. A., and Yeaple, S. R. 2015. Trade liberalization, quality, and export prices. *Review of Economics and Statistics*, Vol. 97(5): 1033-1051.

[7] Feenstra, R. C., and Romalis, J. 2014. International prices and endogenous quality. *The Quarterly Journal of Economics,* Vol. 129(2): 477-527.

[8] Flach, L., and Unger, F. 2022. Quality and gravity in international trade. *Journal of International Economics*, Vol. 137: 103578.

[9] Hallak, J. C. 2006. Product quality and the direction of trade. *Journal of international Economics*, Vol. 68(1): 238-265.

[10] Hummels, D., and Klenow, P. J. 2005. The variety and quality of a nation's exports. *American Economic Review*, Vol. 95(3): 704-723.

[11] Khandelwal, A. 2010. The long and short (of) quality ladders. *The Review of Economic Studies*, Vol. 77(4): 1450-1476.

[12] Kugler, M., and Verhoogen, E. 2012. Prices, plant size, and product quality. *The Review of Economic Studies*, Vol. 79(1): 307-339.

[13] Krugman, P. R. 1979. Increasing returns, monopolistic competition, and international trade. *Journal of international Economics*, Vol. 9(4): 469-479.

[14] Krugman, P. 1980. Scale economies, product differentiation, and the pattern of trade. *American Economic Review*, Vol. 70(5): 950-959.

[15] Ma, Ching-to A., and J. F. Burgess Jr. 1993. Quality competition, welfare, and regulation. *Journal of Economics*, Vol. 58(2): 153-173.

[16] Mankiw, N. G., and Whinston, M. D. 1986. Free entry and social inefficiency. *The RAND Journal of Economics*, 48-58.

[17] Manova, K., and Zhang, Z. 2012. Export prices across firms and destinations. *The Quarterly Journal of Economics*, Vol. 127(1): 379-436.

[18] Melitz, M. J. 2003. The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, Vol. 71(6): 1695-1725.

[19] Melitz, M. J., and Ottaviano, G. I. 2008. Market size, trade, and productivity. *The review of economic studies*, Vol. 75(1): 295-316.

[20] Picard, P. M. 2015. Trade, economic geography and the choice of product quality. *Regional Science and Urban Economics*, Vol. 54: 18-27.

[21] Schott, P. K. 2004. Across-product versus within-product specialization in international trade. *The Quarterly Journal of Economics*, Vol. 119(2): 647-678.

[22] Spence, M, 1976. Product Selection, Fixed Costs, and Monopolistic Competition, *Review of Economic Studies*, Vol. 43: 217–235.

[23] Wauthy, X. 1996. Quality choice in models of vertical differentiation. *The journal of industrial economics*, 345-353.

[24] Zhelobodko, E., Kokovin, S., Parenti, M., and Thisse J.-F. 2012. Monopolistic competition: Beyond the constant elasticity of substitution. *Econometrica*, Vol. 80(6): 2765-2784.

Appendices

In this Appendices, we provide the proofs of some lemmas and propositions stated in the paper.

Appendix A. The proof of Lemma 1

The first thing to notice is that, under the restrictions on the parameters formulated in the lemma, the solution of the first order conditions (7) - (9) is unique. The proof follows from simple and straightforward algebra. Now, let us turn to the existence of the internal optimum.

We have the following maximization problem:

$$\max_{Q,q,y}\left[\frac{\beta}{\lambda}L\left[\left(\frac{Qy}{L}\right)^{\rho}+q^{\rho}\right]^{\frac{\beta}{\rho}-1}\left(\frac{Qy}{L}\right)^{\rho}-(f+yc)(Q^{\alpha}q)^{\gamma}-F\right]$$

where $\alpha>0$, $0<\beta<1$,ρ< 0, and $\gamma>0$. Let us introduce two new variables:

$$\left(\frac{Qy}{L}\right)^{\rho}=t\Longleftrightarrow Q=\frac{Lt^{\frac{1}{\rho}}}{y},$$

and

$$z=q^{\rho}.$$

Then, the maximization problem can be rewritten as (we omit parameter F in the below)

$$\max_{z,x,y}\left[\frac{\beta}{\lambda}L\,(t+z)^{\frac{\beta}{\rho}-1}\,t-L^{\alpha\gamma}\frac{f+yc}{y^{\alpha\gamma}}(zt^{\alpha})^{\frac{\gamma}{\rho}}\right].$$

The above immediately implies that the optimal value of y is

$$y=\frac{f}{c}\frac{\alpha\gamma}{1-\alpha\gamma},$$

which is positive if and only if $1-\alpha\gamma>0$.

Substituting then for the optimal y, we derive the following optimization problem:

$$\max_{z,x}\left\{C_1\,(t+z)^{\frac{\beta}{\rho}-1}\,t-C_2(zt^{\alpha})^{\frac{\gamma}{\rho}}\right\},$$

where C_1 and C_2 are strictly positive. Note that the function,

$$\pi(t,z)=C_1\,(t+z)^{\frac{\beta}{\rho}-1}\,t-C_2(zt^{\alpha})^{\frac{\gamma}{\rho}},$$

has the following properties:

1) If $z\in(0,\infty)$,then $\pi(0,z)=-\infty$ and $\pi(\infty,z)=0$

2) If $t\in(0,\infty)$,then $\pi(x,0)=-\infty$ and $\pi(t,\infty)=0$

3) $\pi(0,\infty)\leq 0$, $\pi(\infty,0)\leq 0$, $\pi(\infty,\infty)=0$.

Next, we explore the value and sign of $\pi(0,0)$. We have

$$\pi(t,z)=C_1\frac{1}{t^{-\frac{\beta}{\rho}}\,(1+z/t)^{1-\frac{\beta}{\rho}}}-C_2\left(\frac{z}{t}\right)^{\frac{\gamma}{\rho}}t^{\frac{(1+\alpha)\gamma}{\rho}}.$$

Assume that $z/t=B\geq 0$ - some constant. When if, as assumed,

$$-\frac{\gamma}{\rho}-\frac{\alpha\gamma}{\rho}>-\frac{\beta}{\rho},$$

we have that $\pi(0,0) = -\infty$. If $z/t = \infty$, $\pi(t,z)$ around $(0,0)$ behaves like

$$\pi(t,z) = C_1 z^{\frac{\beta}{\rho}-1} t - C_2 (z t^{\alpha})^{\frac{\gamma}{\rho}} = z^{\frac{\beta}{\rho}} \left[C_1 \frac{t}{z} - C_2 z^{\frac{\gamma(1+\alpha)-\beta}{\rho}} \left(\frac{t}{z}\right)^{\frac{\alpha\gamma}{\rho}} \right]$$

It is straightforward to see that $C_1 t/z$ converges to zero (as $t/x = \infty$). If $-\gamma/\rho - \alpha\gamma/\rho > -\frac{\beta}{\rho}$, then $C_2 z^{(\gamma(1+\alpha)-\beta)/\rho} (t/z)^{\alpha\gamma/\rho}$ converges to infinity. As a result, $\pi(0,0) \leq 0$.

Recall that we show in the main text that the function, $\pi(t,z)$, can be strictly positive under the lemma conditions. Based on that and the findings above, we can choose a compact subset of $[0,\infty] \times [0,\infty]$, on which there exists a global internal maximum. Then, the uniqueness of the solution of the FOC's immediately implies that this solution is the global maximum.

Appendix B. The proof of Lemma 2

Taking into account (7), we have

$$\beta \left(\frac{\alpha\gamma}{1-\alpha\gamma} \frac{Qf}{cL} \right)^{\rho} = \left(\alpha(\beta-\rho)-\rho\right) q^{\rho} \Longleftrightarrow Q = \frac{cL}{f} \frac{1-\alpha\gamma}{\alpha\gamma} \left[\frac{\alpha\left(\beta-\rho\right)-\rho}{\beta} \right]^{\frac{1}{\rho}} q. \tag{30}$$

We also have (see (16)) that

$$\frac{L}{\lambda} (\beta-\rho) \left(\alpha(\beta-\rho)-\rho\right) q^{\beta-\gamma} \left(\frac{(1+\alpha)(\beta-\rho)}{\beta} \right)^{\frac{\beta}{\rho}-2} = \frac{\gamma f}{1-\alpha\gamma} Q^{\alpha\gamma}. \tag{31}$$

If we define

$$A = \left(\frac{1}{\gamma} - \alpha\right) (\beta-\rho) \left(\alpha(\beta-\rho)-\rho\right) \left(\frac{(1+\alpha)(\beta-\rho)}{\beta} \right)^{\frac{\beta}{\rho}-2},$$

and

$$B = \frac{1-\alpha\gamma}{\alpha\gamma} \left(\frac{\alpha(\beta-\rho)-\rho}{\beta} \right)^{\frac{1}{\rho}},$$

then we derive a simple system of two equations:

$$\frac{L}{\lambda} A q^{\beta-\gamma} = f Q^{\alpha\gamma}, \qquad \text{and} \qquad Q = \frac{cL}{f} B q,$$

which implies that

$$q = \left(\frac{L}{f}\right)^{\frac{1-\alpha\gamma}{(1+\alpha)\gamma-\beta}} (Bc)^{-\frac{\alpha\gamma}{(1+\alpha)\gamma-\beta}} \left(\frac{\lambda}{A}\right)^{-\frac{1}{(1+\alpha)\gamma-\beta}} \tag{32}$$

and

$$Q = \left(\frac{L}{f}\right)^{\frac{1+\gamma-\beta}{(1+\alpha)\gamma-\beta}} (Bc)^{\frac{\gamma-\beta}{(1+\alpha)\gamma-\beta}} \left(\frac{\lambda}{A}\right)^{-\frac{1}{\beta(1+\alpha)\gamma-\beta}}. \tag{33}$$

Thus,

$$Q^{\alpha}q = \left(\frac{L}{f}\right)^{\frac{1+(1-\beta)\alpha}{(1+\alpha)\gamma-\beta}} (Bc)^{-\frac{\alpha\beta}{(1+\alpha)\gamma-\beta}} \left(\frac{\lambda}{A}\right)^{-\frac{1+\alpha}{(1+\alpha)\gamma-\beta}}. \tag{34}$$

As a result,

$$\frac{Q^{\alpha}(c)q(c)}{Q^{\alpha}(c^*)q(c^*)} = \left(\frac{c}{c^*}\right)^{-\frac{\alpha\beta}{(1+\alpha)\gamma-\beta}}.$$

Recall that c^* solves

$$\int_0^{c^*(\lambda)} \left[\frac{1}{\beta}\frac{(1+\alpha)\gamma-\beta}{1-\alpha\gamma} f(Q^{\alpha}(c,\lambda)q(c,\lambda))^{\gamma} - F\right] \mathrm{d}G(c) = F_e.$$

Substituting for $Q^{\alpha}(c,\lambda)q(c,\lambda)$ and rearranging result in

$$\frac{1}{\beta}\frac{(1+\alpha)\gamma-\beta}{1-\alpha\gamma} f(Q^{\alpha}(c^*)q(c^*))^{\gamma} \int_0^{c^*} \left(\frac{c}{c^*}\right)^{-\frac{\alpha\beta\gamma}{(1+\alpha)\gamma-\beta}} \mathrm{d}G(c) = FG(c^*) + F_e.$$

Taking into account the zero profit condition (13), we derive

$$\int_0^{c^*} \left(\frac{c}{c^*}\right)^{-\frac{\alpha\beta\gamma}{(1+\alpha)\gamma-\beta}} \mathrm{d}G(c) = G(c^*) + \frac{F_e}{F}.$$

Using (34), the zero profit condition can be written as

$$\left(\frac{L}{f}\right)^{\gamma\frac{1+(1-\beta)\alpha}{(1+\alpha)\gamma-\beta}} (Bc^*)^{-\frac{\alpha\beta\gamma}{(1+\alpha)\gamma-\beta}} \left(\frac{\lambda}{A}\right)^{-\gamma\frac{1+\alpha}{(1+\alpha)\gamma-\beta}} = \frac{F}{f}D,$$

where

$$A = \left(\frac{1}{\gamma}-\alpha\right)(\beta-\rho)\left(\alpha(\beta-\rho)-\rho\right)\left(\frac{(1+\alpha)(\beta-\rho)}{\beta}\right)^{\frac{\beta}{\rho}-2},$$

$$B = \frac{1-\alpha\gamma}{\alpha\gamma}\left[\frac{\alpha(\beta-\rho)-\rho}{\beta}\right]^{\frac{1}{\rho}},\ D = \frac{(1-\alpha\gamma)\beta}{\gamma(1+\alpha)-\beta}.$$

Thus,

$$\lambda = A\left(\frac{L}{f}\right)^{\frac{1+(1-\beta)\alpha}{1+\alpha}} (Bc^*)^{-\frac{\alpha\beta}{1+\alpha}} \left(\frac{F}{f}D\right)^{-\frac{(1+\alpha)\gamma-\beta}{\gamma(1+\alpha)}}.$$

Finally, substituting the expression for λ into (32) and (33), we obtain

$$q(c) = \left(\frac{L}{f}\right)^{-\frac{\alpha}{1+\alpha}} (c^*B)^{-\frac{\alpha}{1+\alpha}} \left(\frac{c}{c^*}\right)^{-\frac{\alpha\gamma}{(1+\alpha)\gamma-\beta}} \left(\frac{F}{f}D\right)^{\frac{1}{\gamma(1+\alpha)}},$$

$$Q(c) = \left(\frac{L}{f}\right)^{\frac{1}{1+\alpha}} (c^*B)^{\frac{1}{1+\alpha}} \left(\frac{c}{c^*}\right)^{-\frac{\beta-\gamma}{(1+\alpha)\gamma-\beta}} \left(\frac{F}{f}D\right)^{\frac{1}{\gamma(1+\alpha)}}.$$

Appendix C. The Proof of Proposition 3

Substituting the resource constraint into the objective function, we can rewrite the social planer's problem in the following way:

$$\max_{c_o^*, x_o(c), q_o(c), Q_o(c)} \frac{\int_0^{c_o^*} \left[(Q_o(c)x_o(c))^\rho + q_o(c)^\rho\right]^{\frac{\beta}{\rho}} \mathrm{d}G(c)}{\int_0^{c_o^*} \left(F + (f + Lcx_o(c))(Q_o^\alpha(c)q_o(c))^\gamma\right) \mathrm{d}G(c) + F_e}.$$

The corresponding first order conditions are given by

$$\frac{\beta\,(Q_o(c)x_o(c))^\rho \left((Q_o(c)x_o(c))^\rho + q_o(c)^\rho\right)^{\frac{\beta}{\rho}-1}}{Lcx_o(c)(Q_o^\alpha(c)q_o(c))^\gamma} = \frac{\int_0^{c_o^*} \left((Q_o(c)x_o(c))^\rho + q_o(c)^\rho\right)^{\frac{\beta}{\rho}} \mathrm{d}G(c)}{\int_0^{c_o^*} \left(F + (f + Lcx_o(c))(Q_o^\alpha(c)q_o(c))^\gamma\right) \mathrm{d}G(c) + F_e}, \tag{35}$$

$$\frac{\beta\,(Q_o(c)x_o(c))^\rho \left((Q_o(c)x_o(c))^\rho + q_o(c)^\rho\right)^{\frac{\beta}{\rho}-1}}{\alpha\gamma(f + Lcx_o(c))(Q_o^\alpha(c)q_o(c))^\gamma} = \frac{\int_0^{c_o^*} \left((Q_o(c)x_o(c))^\rho + q_o(c)^\rho\right)^{\frac{\beta}{\rho}} \mathrm{d}G(c)}{\int_0^{c_o^*} \left(F + (f + Lcx_o(c))(Q_o^\alpha(c)q_o(c))^\gamma\right) \mathrm{d}G(c) + F_e}, \tag{36}$$

$$\frac{\beta q_o(c)^\rho \left((Q_o(c)x_o(c))^\rho + q_o(c)^\rho\right)^{\frac{\beta}{\rho}-1}}{\gamma(f + Lcx_o(c))(Q_o^\alpha(c)q_o(c))^\gamma} = \frac{\int_0^{c_o^*} \left((Q_o(c)x_o(c))^\rho + q_o(c)^\rho\right)^{\frac{\beta}{\rho}} \mathrm{d}G(c)}{\int_0^{c_o^*} \left(F + (f + Lcx_o(c))(Q_o^\alpha(c)q_o(c))^\gamma\right) \mathrm{d}G(c) + F_e}, \tag{37}$$

$$\frac{\left((Q_o(c_o^*)x_o(c_o^*))^\rho + q_o(c_o^*)^\rho\right)^{\frac{\beta}{\rho}}}{F + (f + Lc_o^* x_o(c_o^*))(Q_o^\alpha(c_o^*)q_o(c_o^*))^\gamma} = \frac{\int_0^{c_o^*} \left((Q_o(c)x_o(c))^\rho + q_o(c)^\rho\right)^{\frac{\beta}{\rho}} \mathrm{d}G(c)}{\int_0^{c_o^*} \left(F + (f + Lcx_o(c))(Q_o^\alpha(c)q_o(c))^\gamma\right) \mathrm{d}G(c) + F_e}. \tag{38}$$

The first thing to notice is that the above equations yield (it is sufficient to divide (35) by (36))

$$y_o(c)c = \frac{\alpha\gamma}{1 - \alpha\gamma} f, \tag{39}$$

where $y_o(c) = Lx_o(c)$. Dividing (36) by (37), we derive that

$$\left(\frac{Q_o(c)y_o(c)}{q_o(c)L}\right)^\rho = \alpha. \tag{40}$$

Thus, comparison between (40) with its equilibrium counterpart (8) yields

$$\frac{Q_o(c)}{q_o(c)} < \frac{Q(c)}{q(c)}, \tag{41}$$

as

$$\frac{\alpha(\beta - \rho) - \rho}{\beta} > \alpha$$

since $\rho < 0$. From (37) by (38), we also have

$$\frac{(Q_o(c_o^*)x_o(c_o^*))^\rho + q_o(c_o^*)^\rho}{F + (f + Lc_o^* x_o(c_o^*))(Q_o^\alpha(c_o^*)q_o(c_o^*))^\gamma} = \frac{\beta q_o(c_o^*)^\rho}{\gamma(f + Lcx_o(c_o^*))(Q_o^\alpha(c_o^*)q_o(c_o^*))^\gamma} \iff$$

$$\left(\frac{Q_o(c_o^*)x_o(c_o^*)}{q_o(c_o^*)}\right)^\rho + 1 = \frac{\beta\left[F + (f + Lc_o^*x_o(c_o^*))(Q_o^\alpha(c_o^*)q_o(c_o^*))^\gamma\right]}{\gamma(f + Lcx_o(c_o^*))(Q_o^\alpha(c_o^*)q_o(c_o^*))^\gamma} \iff$$
$$\left(\frac{Q_o(c_o^*)x_o(c_o^*)}{q_o(c_o^*)}\right)^\rho + 1 = \frac{\beta}{\gamma}\left[\frac{F}{(f + Lcx_o(c_o^*))(Q_o^\alpha(c_o^*)q_o(c_o^*))^\gamma} + 1\right]$$

Taking into account that

$$\left(\frac{Q_o(c_o^*)x_o(c_o^*)}{q_o(c_o^*)}\right)^\rho = \alpha$$

and

$$Lc_o^*x_o(c_o^*) = \frac{\alpha\gamma}{1-\alpha\gamma}f,$$

we derive that

$$(Q_o^\alpha(c_o^*)q_o(c_o^*))^\gamma = \beta\frac{1-\alpha\gamma}{(1+\alpha)\gamma - \beta}\frac{F}{f}. \tag{42}$$

The above equation is the same as that in (13), meaning that $Q_o^\alpha(c_o^*)q_o(c_o^*) = Q^\alpha(c^*)q(c^*)$, where c^* corresponds to the market equilibrium cutoff. Combining this with (41), one can see that $Q_o(c_o^*) > Q(c^*)$ and $q_o(c_o^*) < q(c^*)$.

We also have from (35) that

$$\frac{(Q_o(c)x_o(c))^\rho\left((Q_o(c)x_o(c))^\rho + q_o(c)^\rho\right)^{\frac{\beta}{\rho}-1}}{cx_o(c)(Q_o^\alpha(c)q_o(c))^\gamma} = \frac{(Q_o(c^*)x_o(c^*))^\rho\left((Q_o(c^*)x_o(c^*))^\rho + q_o(c^*)^\rho\right)^{\frac{\beta}{\rho}-1}}{c^*x_o(c^*)(Q_o^\alpha(c^*)q_o(c^*))^\gamma} \iff$$

$$\frac{(Q_o(c)x_o(c))^\rho q_o(c)^{\beta-\rho}\left(\left(\frac{Q_o(c)x_o(c)}{q_o(c)}\right)^\rho + 1\right)^{\frac{\beta}{\rho}-1}}{(Q_o^\alpha(c)q_o(c))^\gamma} = \frac{q_o(c^*)^{\beta-\rho}(Q_o(c^*)x_o(c^*))^\rho\left(\left(\frac{Q_o(c^*)x_o(c^*)}{q_o(c^*)}\right)^\rho + 1\right)^{\frac{\beta}{\rho}-1}}{(Q_o^\alpha(c^*)q_o(c^*))^\gamma}.$$

As a result, from (40) we derive

$$\frac{q_o(c)^\beta}{(Q_o^\alpha(c)q_o(c))^\gamma} = \frac{q_o(c^*)^\beta}{(Q_o^\alpha(c^*)q_o(c^*))^\gamma}. \tag{43}$$

Taking into account (39) and (40), we obtain that

$$\left(\frac{\alpha\gamma}{1-\alpha\gamma}\frac{f}{Lc}\frac{Q_o(c)}{q_o(c)}\right)^\rho = \alpha \iff q_o(c) = \frac{\alpha\gamma}{1-\alpha\gamma}\frac{f}{\alpha^{\frac{1}{\rho}}cL}Q_o(c).$$

Combining with (43), we have

$$\frac{\left(\frac{Q_o(c)}{c}\right)^\beta}{\left(Q_o^\alpha(c)\frac{Q_o(c)}{c}\right)^\gamma} = \frac{\left(\frac{Q_o(c^*)}{c^*}\right)^\beta}{\left(Q_o^\alpha(c^*)\frac{Q_o(c^*)}{c}\right)^\gamma}$$

implying that

$$\frac{Q_o(c)}{Q_o(c^*)} = \left(\frac{c}{c^*}\right)^{\frac{\gamma-\beta}{\gamma(1+\alpha)-\beta}}$$

and

$$\frac{Q_o^\alpha(c)q_o(c)}{Q_o^\alpha(c^*)q_o(c^*)} = \left(\frac{c}{c^*}\right)^{-\frac{\alpha\beta}{\gamma(1+\alpha)-\beta}}. \tag{44}$$

The latter equation is the same as that in Lemma 2.

Finally, taking into account the results above, we can rewrite (35) as follows:

$$\frac{\beta q_o(c)^\beta(1-\alpha\gamma)}{\gamma f} = \frac{(Q_o^\alpha(c)q_o(c))^\gamma(1+\alpha)\int_0^{c_o^*} q_o(c)^\beta \mathrm{d}G(c)}{FG(c_o^*) + \frac{f}{1-\alpha\gamma}\int_0^{c_o^*}(Q_o^\alpha(c)q_o(c))^\gamma \mathrm{d}G(c) + F_e}.$$

As the latter holds for any c, we can integrate on $[0, c_o^*]$ deriving

$$\frac{\beta(1-\alpha\gamma)}{(1+\alpha)\gamma f}\int_0^{c_o^*} q_o(c)^\beta \mathrm{d}G(c) = \frac{\int_0^{c_o^*}(Q_o^\alpha(c)q_o(c))^\gamma \mathrm{d}G(c)\int_0^{c_o^*} q_o(c)^\beta \mathrm{d}G(c)}{FG(c_o^*) + \frac{f}{1-\alpha\gamma}\int_0^{c_o^*}(Q_o^\alpha(c)q_o(c))^\gamma \mathrm{d}G(c) + F_e},$$

which in turn implies that

$$\frac{\beta(1-\alpha\gamma)}{(1+\alpha)\gamma f} = \frac{\int_0^{c_o^*}(Q_o^\alpha(c)q_o(c))^\gamma \mathrm{d}G(c)}{FG(c_o^*) + \frac{f}{1-\alpha\gamma}\int_0^{c_o^*}(Q_o^\alpha(c)q_o(c))^\gamma \mathrm{d}G(c) + F_e}.$$

Using (44) and (42), we obtain

$$\frac{\beta(1-\alpha\gamma)}{(1+\alpha)\gamma f} = \frac{(1-\alpha\gamma)\beta}{(1+\alpha)\gamma-\beta}\frac{F}{f}\frac{\int_0^{c_o^*}\left(\frac{c}{c^*}\right)^{-\frac{\alpha\beta\gamma}{\gamma(1+\alpha)-\beta}}\mathrm{d}G(c)}{FG(c_o^*) + \frac{\beta F}{(1+\alpha)\gamma-\beta}\int_0^{c_o^*}\left(\frac{c}{c^*}\right)^{-\frac{\alpha\beta\gamma}{\gamma(1+\alpha)-\beta}}\mathrm{d}G(c) + F_e},$$

which is equivalent to

$$G(c_o^*) + \frac{F_e}{F} = \int_0^{c_o^*}\left(\frac{c}{c_o^*}\right)^{-\frac{\alpha\beta\gamma}{\gamma(1+\alpha)-\beta}}\mathrm{d}G(c).$$

The latter equation determines c_o^* in the social optimum. As can be seen, this is exactly the same equation as that determining c^* in the market equilibrium. In other words, the cutoffs are the same: $c_o^* = c^*$. Since the cutoffs are the same and the resource constraint holds both in the market equilibrium and in the social planner outcome, it is straightforward to show using the derived findings that the number of entrants in the social optimum is the same as that in the market outcome: $N_o = N$.

To summarize, we show that in the market equilibrium, the total firm output, the number of entrants, and the production cost cutoff are "socially optimal", while the choice of qualities is not. Specifically, using (44) and (42), one can see that $Q_o^\alpha(c)q_o(c) = Q^\alpha(c)q(c)$ with $Q_o(c) > Q(c)$ and $q_o(c) < q(c)$ (see (40)). In the market equilibrium, firms invest too much in the horizontal quality and too little in the vertical one.

Appendix D. The Proof of Lemma 3

Note that combining (39) and (40) results in

$$Q_o(c) = \alpha^{\frac{1}{\rho}} \frac{1-\alpha\gamma}{\alpha\gamma} \frac{cL}{f} q_o(c).$$

Moreover, plugging (42) into (44) yields

$$Q_o^{\alpha}(c) q_o(c) = \left(\frac{(1-\alpha\gamma)\beta}{(1+\alpha)\gamma - \beta} \frac{F}{f} \right)^{\frac{1}{\gamma}} \left(\frac{c}{c^*} \right)^{-\frac{\alpha\beta}{\gamma(1+\alpha)-\beta}}.$$

Substituting the former into the latter, we obtain

$$\left(\alpha^{\frac{1}{\rho}} \frac{1-\alpha\gamma}{\alpha\gamma} \frac{cL}{f} q_o(c) \right)^{\alpha} q_o(c) = \left(\frac{(1-\alpha\gamma)\beta}{(1+\alpha)\gamma - \beta} \frac{F}{f} \right)^{\frac{1}{\gamma}} \left(\frac{c}{c^*} \right)^{-\frac{\alpha\beta}{\gamma(1+\alpha)-\beta}}.$$

After some simplifications, we have

$$q_o(c) = \left(\frac{L}{f} \right)^{-\frac{\alpha}{1+\alpha}} \left(\frac{1-\alpha\gamma}{\alpha\gamma} \alpha^{\frac{1}{\rho}} c^* \right)^{-\frac{\alpha}{1+\alpha}} \left(\frac{c}{c^*} \right)^{-\frac{\alpha\gamma}{\gamma(1+\alpha)-\beta}} \left(\frac{F}{f} D \right)^{\frac{1}{\gamma(1+\alpha)}}.$$