



City Research Online

City St George's, University of London

Citation: Dassiou, X. & Glycopantis, D. (2013). Welfare analysis and policy implications of bundling decisions by firms. *Advances in Economics and Business*, 1(1), pp. 6-21. doi: 10.13189/aeb.2013.010102

This is the unspecified version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: <https://city-ref-test.eprints-hosting.org/id/eprint/3418/>

Link to published version: <https://doi.org/10.13189/aeb.2013.010102>

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to [City Research Online policy](#).

Welfare Analysis and Policy Implications of Bundling Decisions by Firms

Xeni Dassiou*, Dionysius Glycopantis

Department of Economics, City University London, London EC1V 0HB, U.K

*Corresponding Author: x.dassiou@city.ac.uk

Copyright ©2013 Horizon Research Publishing All rights reserved.

Abstract This paper models an original approach where the monopolist firm no longer uses pure bundling with weights 1:1. The decisions of the firm take place in a two-step optimisation process. In the first stage, using portfolio optimisation it decides whether to bundle and also sets the optimal weight for each good. In the second stage, it sets the profit maximising bundle price on the basis of the chosen weights. We compare the profit, consumer and welfare implications of our pure bundling model to the usual 1:1 pure bundling model, and comment on the competition policy making implications.

Keywords Pure bundling, Bundle weights, Risk aversion, Portfolio optimisation, Bundle dispersion, Bundle mean

1 Introduction

The use of bundling as an instrument of second degree price discrimination is well grounded in the theoretical literature of industrial organisation (Adams and Yellen [1]; McAfee, McMillan and Whinston [2]; Stigler [3]). Pure bundling, in contrast to mixed bundling, only allows the sale of two goods together. In both the original contribution of Adams and Yellen and that of McAfee et al. the consumer surplus and profit implications of using both pure and mixed bundling are analysed within a model of two goods. This assumes a uniform distribution of reservation prices describing the valuations of consumers for these goods. On the other hand, Schmalensee [4] analyses pure bundling assuming that the reservation values of consumers are described by the bivariate normal distribution.

Bundling can be used to extract consumers' surplus. It is profit enhancing because it reduces the effective dispersion of the reservation prices.¹ In Schmalensee's paper bundling is based on the conventional approach of assigning equal weights to the goods. However, if symmetry (i.e. $\sigma_1 \simeq \sigma_2$) does not hold, pure bundling is less likely to be profit or welfare enhancing. This is used as an argument for mixed bundling. But as our paper shows, lack of symmetry may not be a problem if the weights of pure bundling can be optimized by the firm.

Moreover, it should be stressed that the manipulation of the mean valuation is as important as that of the dispersion. This is largely neglected in the literature, which mainly focuses on dispersion reduction alone, as an instrument for extracting consumers' surplus. In our paper we study a change in the spread of valuations which is *not mean preserving*.

This paper is a continuation of our research on the theory of bundling in the context of alternative forms of transactions by firms (see Dassiou *et al.* [5,6], Dassiou and Glycopantis [7-9]). In Dassiou and Glycopantis [7,8], we have already examined the use of mixed bundling as a mechanism for price discrimination leading to an enhancement of profits. It can also lead to an increase in trade if practiced by a price setting monopsonist. In common with Adams and Yellen, we used the assumption of a uniform distribution to describe the valuations of consumers for the goods.

In this present paper we consider pure bundling. We assume that the consumers' valuations follow the Gaussian distribution. We generalise Schmalensee's approach, and show that the weights of the two goods in the bundle can be chosen optimally. This can resolve the problem of a lack of symmetry in standard deviations without having to resort to a mixed bundling approach. Weight manipulation may ameliorate the desire to reduce dispersion. It may

¹This is based on the simple mathematical fact that unless the correlation of the reservation values between the demand for the two goods is equal to 1 ($\rho = 1$), we shall have $\sigma_B < \sigma_1 + \sigma_2$.

also lead to an increase in the weight of a good which despite its relatively higher dispersion has a substantially higher mean valuation.

An additional motive behind bundling is the desire by the firm to deal with risk aversion. This has been considered by De Graba [10] who introduces the idea of price discrimination but without bundling. If the seller can not observe the buyers' valuations, then a larger, "group" purchaser is a more risky source of profits than a small, "single unit" buyer. Hence the firm will be willing to sacrifice some profit in order to increase the probability of making a sale to a large purchaser.

Additionally, Dana [11-13] draws from the theory of yield (revenue) management. He argues that price dispersion occurs because of stochasticity, and that firms set different prices to smooth out demand and reduce uncertainty. The work of Dana is discussed in some length in Gaggero and Piga [14] within the context of pricing strategies pursued by airlines.

From the work of Schmalensee we know that dispersion reduction leads to higher profits through better extraction from the consumer surplus. In contrast, in the works of Dana and De Graba we see an alternative interpretation to the existence of price differentials, as a "defensive mechanism" for responding to uncertainty, rather than an "aggressive mechanism" for extracting consumers' surplus. The desire to reduce dispersion for profit enhancement may be either tempered or strengthened by the desire to increase the average demand for the bundle.

What is missing in the bundling literature is an integrative framework which combines both mechanisms to endogenously determine an optimal bundling strategy. To achieve this we set up a two-step decision process. First we formulate a utility function of the firm which depends on absolute risk aversion and then calculate the weights. This deals with the "defensive" aspect of bundling by using a portfolio optimisation approach to determine the relative weights of the component goods. Then, in the second step, by inserting these optimal weights into the profit function we calculate the optimal bundle price. Our analysis has distinct merits as we consider explicitly risk aversion, dispersions and means of valuation of goods. We also define the conditions under which in our framework the optimal weights are equal (this case is referred to as 1:1 bundling henceforth).

We then compare our optimal pure bundling to Schmalensee's equal weights (1:1) case in terms of the profit surplus, consumer and welfare implications. In terms of policy implications, we also show that the desire to increase the bundle mean and/or deal with risk aversion may ultimately lead to the choice of such weights that the result is an increase rather than a decrease in dispersion. Hence this will also lead to a further increase in consumer surplus when this an increasing function of dispersion.

Section 2 looks at the unbundled sales paradigm and derives the comparative statics results for profits, consumer surplus and welfare with respect to the mean and the dispersion of the normal distribution which describes the tastes of consumers for each good. In Section 3 we set out a two step process of optimal pure bundling. The firm determines the optimal weights in the bundle and then the bundle price. Comparisons are made with Schmalensee's approach of fixed weights. In Section 4 we discuss our conclusions. In the mathematical appendix we provide mathematical proofs of results stated in the text.

2 Unbundled sales

We derive here *comparative static results* for the case of no bundling for each good separately. Schmalensee also looks at the non bundling case briefly. However he does not separate the effects of the mean, dispersion and cost, which in our case is essential for our later analysis of pure bundling.

The two separate demand functions when the two goods are sold unbundled can be written as $Q^1(P^1)$ and $Q^2(P^2)$, i.e. their demands are independent of the price of the other good. Either demand function can be written as:

$$Q(P) = \int_P^{\infty} g(x)dx \quad (1)$$

where $g(x)$ is the underlying distribution of the buyers' valuations and $Q(P)$ is the cumulative distribution. $g(x)$ is assumed to be a univariate normal distribution with mean μ and standard deviation σ ; it gives the density function of buyers. As it is standard in the literature we assume that each consumer, if he/she purchases a good, has a unit demand.

Let $f(t)$ be the standard normal density function, where $f(t) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{t^2}{2})$, and define:

$$1 - F(x) = 1 - \int_{-\infty}^x f(t)dt \quad (2)$$

where F is the cumulative distribution function. F is everywhere strictly increasing and it can be shown (see Schmalensee, p. S214) that:

$$F'(x) = f(x) \quad (3)$$

$$F''(x) = \frac{df(x)}{dx} = -xf(x), \quad (4)$$

$$\int_x^\infty tf(t)dt = f(x), \quad (5)$$

$$H(x) = \frac{f(x)}{1 - F(x)} > x. \quad (6)$$

The proofs of (3), (4) and (5) are straightforward. $H(\cdot)$ is the hazard function; the proof for (6) is obtained by integrating (5) by parts.

The demand for a single good² under unbundled sales is thus given by:

$$Q(P) = 1 - F\left(\frac{P - \mu}{\sigma}\right).$$

Strictly speaking since the normal distribution allows negative values, and since valuations are typically non negative, some parameter restrictions should be applied when using the normal distribution to represent these valuations. It is simple to achieve this by assuming that $\frac{\mu}{\sigma} > 3$, in which case only positive values occur 99.93% of the time and negative valuations are thus virtually eliminated. Of course this inequality has implications both for the left as well as for the right hand side tail; valuations will fall within the region $0 < P < 2\mu$, 99.86% of the time.

While we do follow Schmalensee's analysis, we *do not* re-define variables by standardizing the mark up or the mean by subtracting costs and dividing by the standard deviation. This is because we wish to examine the impact on profits and on consumer surplus of a change in mean and a change in deviation separately.

For $P > \mu$ ($\mu > 0$) the demand curve is *strictly convex* and for $P < \mu$ it is *strictly concave* (given that $Q'(P) = -\frac{1}{\sigma}f(\frac{P-\mu}{\sigma}) < 0$, while $Q''(P) = \frac{1}{\sigma^2}\frac{P-\mu}{\sigma}f(\frac{P-\mu}{\sigma})$ is positive if $P > \mu$ and negative if $P < \mu$). In order to understand the relationship between the demand concavity or convexity and that of risk averse or risk loving behaviour respectively, we can think of the curvature of the demand curve as being measured by $A = -\frac{Q''(P)}{Q'(P)} = \frac{P-\mu}{\sigma^2}$. $A < 0$ is similar to the absolute risk aversion parameter applied to a utility function. Since the demand curve is downward sloping, A has the same sign as $Q''(P)$, which is negative for $P < \mu$.

Cowan [15] refers to the analogous nature of a curvature of the demand function and the concept of risk aversion. Kimball [16] further shows that the curvature of the slope of demand is analogous to the concept of relative prudence applied to utility functions. Moreover, the existence of risk aversion in our model would provide an additional source of concavity to that of $Q''(P) < 0$, as the utility function would be a monotonically increasing concave function of demand, which in turn is monotonically decreasing and concave. Therefore an absolute risk aversion measure α ($\alpha > 0$) would exceed the absolute value of A in the former case. We will return to this discussion in Section 3.

The profit function also follows the normal distribution as it is a linear function of the demand function. It is given by

$$\Pi = (P - C)(1 - F(\frac{P - \mu}{\sigma})). \quad (7)$$

where C is the constant unit cost of producing the good considered. The FOC for profit maximisation for the choice of P can be expressed as:

$$H(\frac{P^* - \mu}{\sigma}) = \frac{\sigma}{P^* - C}, \quad (8)$$

while the SOC is:

$$(P^* - C)(P^* - \mu) < 2\sigma^2. \quad (9)$$

Substituting (6) into (8) we obtain:

²Essentially this gives you the percentage of the people buying one unit of this good.

$$(P^* - C)(P^* - \mu) < \sigma^2. \quad (10)$$

The above means that the SOC is always satisfied. The above inequality implies a range of values:

$$\frac{\mu + C - \sqrt{(\mu - C)^2 + 4\sigma^2}}{2} < P^* < \frac{\mu + C + \sqrt{(\mu - C)^2 + 4\sigma^2}}{2}. \quad (11)$$

The range of relevant values for a non negative profit can be reduced³ to:

$$C \leq P^* < \frac{\mu + C + \sqrt{(\mu - C)^2 + 4\sigma^2}}{2}. \quad (12)$$

For any pair of values of μ and σ the above condition is satisfied for P^* in this interval. The inclusion of $P^* = \mu$ in particular, means that the interval does include the origin of the standard normal distribution F . Hence the demand, and through it the profit function, are not globally concave.

From implicit differentiation of the FOC with respect to μ it is easy to show that

$$\frac{dP^*}{d\mu} = \frac{(P^* - C)(P^* - \mu) - \sigma^2}{(P^* - C)(P^* - \mu) - 2\sigma^2}. \quad (13)$$

Invoking relation (10) we obtain that $0 < \frac{dP^*}{d\mu} < 1$. It is also fairly easy to show that $\frac{dP^*}{dC} = \frac{-\sigma^2}{(P^* - C)(P^* - \mu) - 2\sigma^2} >$

0. For $P^* < (>)\mu$, $\frac{dP^*}{dC} < (>)\frac{dP^*}{d\mu}$.

Similarly,

$$\frac{dP^*}{d\sigma} = \frac{(P^* - C)(P^* - \mu)^2 - [2P^* - (\mu + C)]\sigma^2}{(P^* - C)(P^* - \mu) - 2\sigma^2} \frac{1}{\sigma}. \quad (14)$$

It is important to establish conditions for the sign of the change in optimum price with respect to the dispersion. This, together with the sign of relation (13), is used repeatedly below in calculating the various comparative static results. Using straightforward but tedious calculations we derive the theorem below.

For $P^* < \frac{\mu+C}{2}$, $\frac{dP^*}{d\sigma} < 0$. For $P^* \geq \mu$ we have that $\frac{dP^*}{d\sigma} > 0$. For $\mu > P^* > \frac{\mu+C}{2}$ (that requires $\mu > C$), as P^* increases $\frac{dP^*}{d\sigma}$ is first negative and then turns positive.

. Given (10), the denominator in (14) is negative. For $(\frac{\mu+C-\sqrt{(\mu-C)^2+4\sigma^2}}{2}) < P^* < \frac{\mu+C}{2}$ (a necessary condition for this to hold is that $P^* < \mu$ since $P^* > C$), the nominator in (14) is positive. This means that a higher dispersion lowers the profit maximising price since $\frac{dP^*}{d\sigma} < 0$. On the other hand, if $(\frac{\mu+C+\sqrt{(\mu-C)^2+4\sigma^2}}{2}) > P^* \geq \mu$, then it follows from (10) that $(P^* - C)(P^* - \mu)^2 \leq \sigma^2(P^* - \mu) < \sigma^2[2P^* - (\mu + C)]$. Hence the nominator in (14) is negative and given the negative denominator, $\frac{dP^*}{d\sigma} > 0$. Hence a higher dispersion raises the optimal price when the latter is above the mean. Finally, for some value of P^* in the range $\frac{\mu+C}{2} < P^* < \mu$, (a necessary condition for this inequality to hold is that $\mu > C$; else $P^* > \mu$) the nominator in (14) is for this range of values a decreasing function of the optimal price and will turn from positive to negative.⁴ Hence $\frac{dP^*}{d\sigma}$ increases from negative to positive values as the optimal price increases. If the price is significantly below the mean a higher dispersion reduces the optimal price.

We can now calculate the impact of μ and σ on the optimal profit function.

$$\frac{d\Pi^*}{d\mu} = \left[\left(1 - F\left(\frac{P^* - \mu}{\sigma}\right)\right) - \frac{P^* - C}{\sigma} f\left(\frac{P^* - \mu}{\sigma}\right) \right] \frac{\partial P^*}{\partial \mu} + (P^* - C) \left(f\left(\frac{P^* - \mu}{\sigma}\right) \right) \frac{1}{\sigma}. \quad (15)$$

Using the envelope theorem we obtain:

$$\frac{d\Pi^*}{d\mu} = (P^* - C) \left(f\left(\frac{P^* - \mu}{\sigma}\right) \right) \frac{1}{\sigma} = (1 - F\left(\frac{P^* - \mu}{\sigma}\right)). \quad (16)$$

³As $\frac{\mu+C-\sqrt{(\mu-C)^2+4\sigma^2}}{2} < C$.

⁴The value of $(P^* - C)(P^* - \mu)$ for this range of values is negative and an increasing function of P^* (e.g. approaching zero as P^* approaches μ) while the expression $\frac{[2P^* - (\mu + C)]\sigma^2}{(P^* - \mu)}$ in this same range of values is negative and a decreasing function of P^* . Hence overall $(P^* - C)(P^* - \mu) - \frac{[2P^* - (\mu + C)]\sigma^2}{(P^* - \mu)}$ will at some point turn positive as P^* increases. This implies that $((P^* - C)(P^* - \mu) - \frac{[2P^* - (\mu + C)]\sigma^2}{(P^* - \mu)})(P^* - \mu)$ will turn from positive to negative as P^* approaches μ in this range of values.

Invoking the properties of the cumulative distribution function we obtain that $0 < \frac{d\Pi^*}{d\mu} < 1$. Moreover,

$$\frac{d^2\Pi^*}{d\mu^2} = f\left(\frac{P^* - \mu}{\sigma}\right) \frac{\sigma}{2\sigma^2 - (P^* - C)(P^* - \mu)} > 0. \quad (17)$$

Relations (16) and (17) imply that Π^* is an increasing, convex function of μ .

$$\frac{d\Pi^*}{d\sigma} = \left[\left(1 - F\left(\frac{P - \mu}{\sigma}\right)\right) - \frac{P - C}{\sigma} f\left(\frac{P - \mu}{\sigma}\right) \right] \frac{\partial P}{\partial \sigma} + (P - C)(P - \mu) \frac{1}{\sigma^2} \left(f\left(\frac{P - \mu}{\sigma}\right) \right). \quad (18)$$

Similarly, employing the envelope theorem we obtain:

$$\frac{d\Pi^*}{d\sigma} = \left(\frac{P^* - C}{\sigma}\right) \left(\frac{P^* - \mu}{\sigma}\right) f\left(\frac{P^* - \mu}{\sigma}\right) = \left(\frac{P^* - \mu}{\sigma}\right) \left(1 - F\left(\frac{P^* - \mu}{\sigma}\right)\right). \quad (19)$$

It therefore follows that Π^* is an increasing function of σ for $P^* > \mu$ and a decreasing function for $P^* < \mu$. Hence *decreasing dispersion in valuations is not always profit enhancing as commonly assumed*.

We next examine the consumer surplus, $CS = \sigma \int_{\frac{P^* - \mu}{\sigma}}^{\infty} \left(\frac{x - \mu - P}{\sigma}\right) g\left(\frac{x - \mu}{\sigma}\right) d\left(\frac{x - \mu}{\sigma}\right)$. It is straightforward to show that for $P = P^*$, using (5) and the FOC, can be expressed as:

$$CS|_{P=P^*} = \sigma f\left(\frac{P^* - \mu}{\sigma}\right) \left(1 - \frac{(P^* - \mu)(P^* - C)}{\sigma^2}\right). \quad (20)$$

Given (10), the consumer surplus is strictly positive for $\sigma \neq 0$. In other words, the existence of dispersion in demand gives the consumers the opportunity to capture some of the total surplus from the transactions in this good. We examine immediately below under what circumstances $CS|_{P=P^*}$ is an increasing function of dispersion. It follows that for $\sigma \neq 0$, the value of P that maximises the welfare function differs from that of profit maximisation.

Note that through routine calculations we obtain:

$$\frac{\partial CS}{\partial P} = - \left(1 - F\left(\frac{P - \mu}{\sigma}\right)\right). \quad (21)$$

Therefore the consumer surplus is a strictly decreasing function of price.

Splitting between a direct and an indirect effect, the impact of dispersion on consumer surplus at $P = P^*$ is equal to the additive expression:

$$\frac{dCS}{d\sigma} |_{P=P^*} = f\left(\frac{P^* - \mu}{\sigma}\right) + \frac{\partial P^*}{\partial \sigma} \frac{\partial CS}{\partial P^*} |_{P=P^*}. \quad (22)$$

This means that the direct effect of dispersion on consumer surplus is always positive.

For P^* sufficiently smaller than μ , (in the way detailed in Theorem 1) the consumer surplus is an increasing function of σ always, as both effects are positive since $\frac{\partial P^*}{\partial \sigma} < 0$. Hence the indirect effect is positive and reinforces the direct effect. Denoting $CS|_{P=P^*}$ by CS^* , this means that CS^* is an increasing function of σ for $P^* < \mu$.

The above property of CS^* combined with the fact that Π^* is a decreasing function of σ for $P^* < \mu$, means that the firm dislikes risk in the form of dispersion because the concavity of the demand function implies that capturing consumer surplus becomes more difficult as the dispersion in valuations increases. Therefore for $A < 0$ the concavity of the demand function elicits risk averse behaviour. A acts as a measure of the degree of this dispersion sensitivity. As we show below, the desire to manipulate dispersion may drive the firm into using bundling.

If, on the other hand, $P^* > \mu$, then the indirect effect is now negative as $\frac{dP^*}{d\sigma} > 0$ and hence the overall impact depends on which effect dominates over the other.

We now obtain comparative static results for the welfare function as a way of consolidating the overall impact of mean and dispersion on both the producer and the consumers:

$$W = \sigma f\left(\frac{P - \mu}{\sigma}\right) + (\mu - C) \left(1 - F\left(\frac{P - \mu}{\sigma}\right)\right). \quad (23)$$

So,

$$\frac{dW}{dP} = -\frac{P - \mu}{\sigma} f\left(\frac{P - \mu}{\sigma}\right) - \frac{\mu - C}{\sigma} f\left(\frac{P - \mu}{\sigma}\right) = -\frac{P - C}{\sigma} f\left(\frac{P - \mu}{\sigma}\right) < 0. \quad (24)$$

It is easy to check for the above that for $P = P^*$, the terms that correspond to the profit component become zero which implies:

$$\frac{dW}{dP} |_{P=P^*} = -\frac{P^* - C}{\sigma} f\left(\frac{P^* - \mu}{\sigma}\right) < 0. \quad (25)$$

In other words, a lower price than P^* will give a higher welfare. Moreover:

$$\frac{dW}{d\mu} = \frac{\partial W}{\partial P} \frac{\partial P}{\partial \mu} + [1 - F(\frac{P-\mu}{\sigma})] + \frac{(\mu-C)}{\sigma} f(\frac{P-\mu}{\sigma}) + \frac{(P-\mu)}{\sigma} f(\frac{P-\mu}{\sigma}),$$

and

$$\frac{dW}{dC} = \frac{\partial W}{\partial P} \frac{\partial P}{\partial C} - [1 - F(\frac{P-\mu}{\sigma})].$$

Using the FOC the above two expressions become:

$$\frac{dW}{d\mu} |_{P=P^*} = [1 - F(\frac{P^*-\mu}{\sigma})](2 - \frac{dP^*}{d\mu}) = [1 - F(\frac{P^*-\mu}{\sigma})] \left(\frac{(P^*-C)(P^*-\mu) - 3\sigma^2}{(P^*-C)(P^*-\mu) - 2\sigma^2} \right) > 0, \quad (26)$$

and $\frac{dW}{dC} |_{P=P^*} = -[1 - F(\frac{P^*-\mu}{\sigma})](1 + \frac{\partial P^*}{\partial C}) = -[1 - F(\frac{P^*-\mu}{\sigma})] \frac{(P^*-C)(P^*-\mu) - 3\sigma^2}{(P^*-C)(P^*-\mu) - 2\sigma^2} = -\frac{dW}{d\mu} |_{P=P^*} < 0$.

As $\frac{dW}{d\mu} |_{P=P^*} = \frac{dCS^*}{d\mu} + \frac{d\Pi^*}{d\mu}$, and since $\frac{d\Pi^*}{d\mu} = \frac{(P^*-C)}{\sigma} f(\frac{P^*-\mu}{\sigma})$ we also have that:

$$\frac{dCS^*}{d\mu} = [1 - F(\frac{P^*-\mu}{\sigma})](1 - \frac{\partial P^*}{\partial \mu}). \quad (27)$$

So both W^* and CS^* are increasing functions of the mean valuation. Correspondingly, as $\frac{dW}{dC} |_{P=P^*} = \frac{dCS^*}{dC} + \frac{d\Pi^*}{dC}$, and since $\frac{d\Pi^*}{d\mu} = -\frac{d\Pi^*}{dC}$ and $\frac{dW}{d\mu} |_{P=P^*} = -\frac{dW}{dC} |_{P=P^*}$ we also have that $\frac{dCS^*}{d\mu} = -\frac{dCS^*}{dC}$. Also at $P = P^*$ using the FOC we have:

$$\frac{dW}{d\sigma} |_{P=P^*} = \frac{\partial W}{\partial P} \frac{\partial P^*}{\partial \sigma} + f(\frac{P^*-\mu}{\sigma})(1 + \frac{(P^*-\mu)(P^*-C)}{\sigma^2}). \quad (28)$$

Using (19) and (14), this becomes:

$$\frac{dW}{d\sigma} |_{P=P^*} = \frac{d\Pi^*}{d\sigma} + \frac{dCS}{d\sigma} |_{P=P^*} = f\left(\frac{P^*-\mu}{\sigma}\right) \left(\frac{(P^*-C)^2 - 2\sigma^2}{(P^*-C)(P^*-\mu) - 2\sigma^2} \right). \quad (29)$$

The derivation of the above expression is given in the mathematical appendix. It shows that for $-\sqrt{2} < \frac{P^*-C}{\sigma} < \sqrt{2}$ welfare is an increasing function of dispersion and a decreasing function otherwise (Schmalensee, p. S217).

Hence if the profit maximising price is within $(C - \sqrt{2}\sigma, C + \sqrt{2}\sigma)$, welfare is an increasing function of dispersion. Setting $\frac{P^*-C}{\sigma} = \sqrt{2}$ in the FOC and using the normal distribution frequency tables we derive that $\frac{P^*-\mu}{\sigma} = -0.147$. This results to $\frac{\mu-C}{\sigma} = 1.561$. Hence we can conclude that for $P^* \in (\mu - 0.147\sigma, \mu + 0.147\sigma)$ welfare is a decreasing function of dispersion. This corresponds to a range of values which covers almost 14% of the distribution.

In the range $\mu - 0.147\sigma < P^* < \mu$, $\frac{d\Pi^*}{d\sigma} < 0$ ($\frac{\mu-C}{\sigma} > 1.561$) and as we have that $\frac{dW^*}{d\sigma} < 0$ ($\frac{P^*-C}{\sigma} > \sqrt{2}$), both profits and welfare are decreasing functions of dispersion. This means that a reduction in dispersion will, in this case, be beneficial to society as a whole, not just the company.

If $P^* < \mu - 0.147\sigma$, then $\frac{d\Pi^*}{d\sigma} < 0$ while $\frac{dCS^*}{d\sigma}, \frac{dW^*}{d\sigma} > 0$. In other words, the desire of a profit maximising firm to decrease dispersion will damage both consumer surplus as well as overall welfare.⁵

The above comparative static results are important for the bundling analysis and conclusions below.

3 Pure Optimal Bundling

We analyse here how the firm determines its portfolio of goods in the bundle. We derive comparative static results for the optimal composition of goods in the bundle. On the basis of this choice the firm then chooses the optimal bundle price. The comparative statics results derived in Section 2 are used in comparing the consumer and welfare implications of our results to those of the fixed weights approach by Schmalensee.

There is only limited literature on the topic of bundle composition manipulation by a monopoly. A recent exemption is the paper by Crampes and Hollander [17] referring to pay-TV distributors deciding on the composition of programme content, such as how many channels should be devoted to each type of programming (sports, documentaries, movies, news and current affairs, etc.).

An original idea in their paper is the introduction of a weights determination process. For example, how many type 1 (Good 1) and how many type 2 (Good 2) channels should be in the bundle, as well as the offering of alternative bundles where some have full coverage and some partial coverage. However the focus of their paper is different. There is no optimal calculation of the optimal bundle; instead they study when it is optimal to offer two alternative bundles as opposed to a single bundle.

⁵Of course when demand is strictly convex ($P^* > \mu$), the dispersion increasing behaviour of a profit maximising firm will be beneficial to society as well for as long as $P^* > \mu + 0.147\sigma$. If $\mu < P^* < \mu + 0.147\sigma$ then $\frac{d\Pi^*}{d\sigma} > 0$, while $\frac{dW^*}{d\sigma} < 0$; this implies that a negative $\frac{dCS^*}{d\sigma}$ more than offsets the impact of $\frac{d\Pi^*}{d\sigma}$.

Anam and Chiang [18] show that third degree price discrimination ameliorates the problem of risk by allowing the risk averse company to diversify its risk between different goods. They show that the conventional direction of third degree price discrimination may be reversed because a good with more elastic demand, but more risk, may be charged a higher price. Expected profit is sacrificed if this more than compensated for by the corresponding decrease in profit risk exposure.

Eckel and Smith [19] in a model of price discrimination use a convex cost function of expected quantities. By manipulating the contribution of each demand group on aggregate demand (through setting different prices for different goods) the firm can influence the mean and variance, and through it, the expected cost.

We also employ a portfolio optimisation approach. Unlike Eckel and Smith, we do not restrict our portfolio to mean preserving combinations. Portfolio manipulation may lead to a bundle combination of the two goods in a way that increases the bundle mean ($\mu_B > \mu_1 + \mu_2$). It has already been shown in Section 2 that a higher mean is both profit and consumer enhancing. In the subsection below we show that if weights are determined through portfolio optimisation, it is possible for pure bundling to take a form such that both the mean as well as dispersion are increased relative to conventional pure bundling. This ameliorates (or even eliminates) policy concerns that bundling may be damaging to the interests of consumers because in this case the dispersion in the valuations of the consumers may in some instances end up to be larger than that of separate sales $\sigma_1 + \sigma_2$.

In our analysis we adopt the following approach. With respect to determining the optimal composition in the bundle we assume risk aversion from the point of view of the firm. The unique optimal allocation supports the observable result that the firm offers one bundle, rather than two bundles, for example each half of the time. Having calculated the optimal bundle composition we proceed to the calculation of the optimal profits. The focus is now the size of the profits as such. This is what is reported in the annual profits of companies and this is the criterion by which comparison of the performance of the firms is made. *Namely having incorporated risk aversion in the construction of the optimal bundle, which is an internal calculation, the firm is then risk neutral in the calculation of the bundle price which will maximise profits.*

3.1 Constructing the bundle

We start our analysis with a packaged good for which we assume, in a way analogous to the previous section, that each consumer if he/she purchases the package has a one unit demand for it. A unit of the packaged good consists of a given total number of individual units of Good 1 and Good 2. These goods have a similar dimension, for example the time involved in consuming them. Although throughout the one unit of the composite good is fixed, the number of units of each constituent good in the bundle can vary. The relative weights of the two goods in the package can be anything between 0 and 1 and the cost of such participation in the bundle is equal to the relative weight of the good times the cost of producing a package that consists only of this good. Means and standard deviations are also perfectly divisible into fractions of less than one.

For example imagine a package in mobile phone services that both offers text messages and phone calls (all measured in minutes) for a given monthly fee (bundle price). If the bundle consists of say, 300 phone calls and 600 text messages, this means that one third of the bundle is phone calls and the other two thirds is text messages. Hence 1 package has 900 minutes.

The decision that the monopolist has to make is first to optimally divide this package into texts and phone calls, and second to price it. We set as λ the relative weight of Good 1 (phone calls, $\lambda = \frac{1}{3}$) and as $1 - \lambda$ the relative weight of Good 2 (texts) in the package. In terms of the Crampes and Hollander paper, one may think in similar terms in setting a TV bundle composition between movies and sport channels. The decision that each consumer then has to make is whether to buy this package of one unit of the composite good, i.e. the 900 minutes with the relative weights and price offered by the monopolist.

As another example one can think of Cadbury's Roses chocolates package which contains mini chocolates of various types (black, milk, white, with nuts, etc.). 1 box has a content of 400g. The decision of how many chocolates (in grams) of each type to include in the package is made by the company. The purchaser of the assortment will eat some types of chocolate, but can discard some others for which he does not have a positive valuation. Apart from chocolates, most of the examples one can think of this type of bundling belong to the family of information goods (e.g., goods that can be digitized) where the marginal costs are low and constant.

As we have already mentioned, by manipulating the relative weights we can affect dispersion and the level of demand. Given bivariate normality, any linear combination of the buyers' pairs of reservation prices will also follow the normal distribution. The underlying distribution for the buyers' valuations for the two goods x (valuation of Good) and y (valuation of Good 2) is

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\left(-\frac{z}{2(1-\rho^2)}\right),$$

where $z = \frac{(x-\mu_x)^2}{\sigma_x^2} - \frac{2\rho(x-\mu_x)((y-\mu_y))}{\sigma_x\sigma_y} + \frac{(y-\mu_y)^2}{\sigma_y^2}$ and ρ is the correlation coefficient of the reservation valuations.

The valuations x and y are in money form, and so is the valuation of the bundle $x\lambda + y(1 - \lambda)$. We now assume $\mu_x, \mu_y > 0$ and that the firm will assign weights in order to maximise a utility function which depends on the

consumers' valuations of the bundle. The firm will then calculate the bundle price⁶. This is clarified below.

Given the underlying (bivariate normal) distribution of the valuations, the moment generating function for $f(x, y)$ is given by $M_{X,Y}(t_1, t_2)$:

$$\begin{aligned} M_{X,Y}(t_1, t_2) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp(t_1 x + t_2 y) f(x, y) dx dy = \\ &= \exp[(\mu_x t_1 + \mu_y(1 - t_2) + \frac{\sigma_x^2 t_1^2 + 2\rho\sigma_x\sigma_y t_1 t_2 + \sigma_y^2 t_2^2}{2})]. \end{aligned}$$

The firm's objective before it sets the bundle price P_B , is to maximise its expected utility, given the bivariate normal distribution $f(x, y)$ of the buyers' reservation prices (valuations). Optimisation is split into two stages, as the firm first decides whether it is worthwhile to bundle in the first place (hence endogenizes the bundling decision as set out in Theorem 2 below) and if so at what relative weights. This is determined by the valuations distribution characteristics (mean, variance and correlation) as well as the degree of risk aversion by the firm. We will use the term α to express risk aversion; this will reflect the firm's sensitivity to dispersion and other risk aversion factors. We assume that α is constant and positive, as in the case of constant absolute risk aversion. The coefficient α corresponds to the normalization $k = 2\lambda$, which we adopt as this will convenience our calculations. Hence,

$$\max J = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} -\exp[-\alpha(xk + y(2 - k))] f(x, y) dx dy. \quad ^7$$

This utility function of the firm depends on the valuations of the bundle by all consumers. Obviously, in determining the bundle the firm will also take into account the costs of producing the good in the bundle and this is done below. Using the moment generating function, with $t_1 = -\alpha k$ and $t_2 = -\alpha(2 - k)$ we can re-write the optimisation function as

$$\max J = -\exp[-\alpha k \mu_1 - \alpha(2 - k) \mu_2 + (-\alpha)^2 \frac{\sigma_1^2 k^2 + 2\rho\sigma_1\sigma_2 k(2 - k) + \sigma_2^2 (2 - k)^2}{2}].$$

Applying a monotonic transformation to the above the problem simplifies into:

$$\max K = \mu_1 \lambda + \mu_2 (2 - \lambda) - \alpha \frac{\sigma_1^2 \lambda^2 + 2\rho\sigma_1\sigma_2 \lambda(1 - \lambda) + \sigma_2^2 (1 - \lambda)^2}{2}.$$

The combination of the parameters above is such that K is positive. The firm wishes to create an optimal bundle, where the weights of the two participating goods are such that the net return to (the expected utility from) the bundle is maximised given the valuations by the consumers and the production costs. Adjusting the means so that they are set as net from their corresponding costs, the firm maximises with respect to relative weight λ the following revenue certainty equivalent function:

⁶If P_B is the bundle price then the demand function can be written as:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy$$

⁷The firm's utility function has a constant absolute risk aversion functional form $-\exp[-\alpha(xk + y(2 - k))]$. If $\alpha = 0$ then the utility function is of the linear form $xk + y(2 - k)$ (Kreps, 1990, p.85). Hence, expected value maximisation will be of the form

$$\max J = \max \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (xk + y(2 - k)) f(x, y) dx dy = \max[k\mu_1 + (2 - k)\mu_2]$$

The above expression indicates that in this case the firm has zero dispersion sensitivity, corresponding to the case of a risk neutral firm. So for maximising the firm will need to assign all the weight to either good 1 or good 2 depending on which of the two goods has the highest mean (or net mean). Hence, the relative weights $(\lambda, 1 - \lambda)$ will be either $(0, 1)$ or $(1, 0)$ and there will be no bundling for different means.

⁸There is a further interesting case to consider which will result in an equal weights outcome: the firm's utility function has a functional form $k(2 - k)xy$ with $\rho = 0$, which denotes an expected value maximisation for:

$$\max J = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \alpha k(2 - k) xy f(x, y) dx dy = \max[\alpha k(2 - k) \mu_1 \mu_2]$$

Maximising the above expression with respect to λ since $\mu_1, \mu_2 > 0$ gives relative weights $(\lambda, 1 - \lambda) = (\frac{1}{2}, \frac{1}{2})$ which corresponds to Schmalensee's type of conventional bundling. Hence this combination is optimal when $\rho = 0$ and the utility function of the firm is multiplicative. In the analysis that follows, we will derive the specific circumstances under which this combination is optimal within our framework.

$$(\mu_1 - C_1)\lambda + (\mu_2 - C_2)(1 - \lambda) - \alpha\lambda^2\sigma_1^2 - 2\alpha\rho\sigma_1\sigma_2\lambda(1 - \lambda) - \alpha(1 - \lambda)^2\sigma_2^2. \quad (30)$$

For $\mu_i - C_i > \mu_j - C_j$, a positive weight for good j requires that $\mu_i - \mu_j - (C_i - C_j) < 2\alpha\sigma_i(\sigma_i - \rho\sigma_j)$.

. Differentiating (30) with respect to λ , gives the F.O.C.:

$$-2\lambda\alpha(\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2) = (\mu_2 - \mu_1) - (C_2 - C_1) - 2\alpha\sigma_2(\sigma_2 - \rho\sigma_1) \quad (31)$$

The S.O.C. is satisfied as $(\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2) > 0$, and $\alpha > 0$.⁹

Normalising the previously defined weights to an overall sum of 2 we obtain $k = 2\lambda = 0.67$ and $2 - k = 2(1 - \lambda) = 1.33$ respectively in our mobile phone services example¹⁰. From the FOC above we can derive the optimal split between the two goods:

$$k^* = 2\lambda^* = \frac{(\mu_2 - \mu_1) - (C_2 - C_1) - 2\alpha\sigma_2(\sigma_2 - \rho\sigma_1)}{-\alpha(\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2)},$$

and

$$2 - k^* = 2(1 - \lambda^*) = \frac{(\mu_1 - \mu_2) - (C_1 - C_2) - 2\alpha\sigma_1(\sigma_1 - \rho\sigma_2)}{-\alpha(\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2)}. \quad (32)$$

The denominator in the above ratios is the S.O.C. derived above, and hence negative.

Therefore unless $(\mu_i - \mu_j) - (C_i - C_j) < 2\alpha\sigma_i(\sigma_i - \rho\sigma_j)$, good j will receive a zero weight within the bundle. This implies either the deletion of an entire product line from the bundle or, less drastically, that the firm needs to consider the use of mixed bundling. *Hence the decision whether to bundle or not, as well as how to balance the goods within the bundle, are both endogenous decisions.* $\square$

Obviously for $\rho > 0$, a necessary (though not sufficient) requirement for a strictly positive share of the good with the lower mean net of cost, say j , is that the dispersion of tastes in good j multiplied with the correlation coefficient is smaller than the dispersion for other good, i.e. $\sigma_i > \rho\sigma_j$. This is obviously increasingly binding as ρ increases. On the other hand, if $\rho < 0$ this is no longer a requirement.

From Theorem 2 it follows that:

- (i) For $\mu_i - C_i > \mu_j - C_j$, if $\alpha(\sigma_i^2 - \sigma_j^2) < \mu_i - \mu_j - (C_i - C_j)$, good i has a greater share in the bundle than good j ($k^* > 1$).
- (ii) The Schmalensee format of equal weights becomes optimal when $(\mu_i - \mu_j) - (C_i - C_j) = \alpha(\sigma_i^2 - \sigma_j^2)$.
- (iii) For $\mu_i - C_i > \mu_j - C_j$, if $\alpha(\sigma_i^2 - \sigma_j^2) > \mu_i - \mu_j - (C_i - C_j)$, good j has a greater share in the bundle than good i ($k^* < 1$).

We use relations (32) describing the optimal split of the two unit bundle. Suppose the weights of both goods are strictly positive. Then, given the common negative denominator, if $\mu_i - C_i > \mu_j - C_j$, in order for i to be given a greater weight than j we require that:

$$\alpha(\sigma_i^2 - \sigma_j^2) < \mu_i - \mu_j - (C_i - C_j).$$

This completes the proof of (i).

If the good which is superior (higher) in the net mean valuation is also inferior (higher) in the terms of its dispersion, and the difference in the net means is exactly offset by the differences in the variances multiplied by the risk aversion parameter, then each good will receive an equal weight. This proves part (ii), which shows that the Schmalensee format is a special case of our approach.

Part (iii) of the corollary means that the good with the higher net mean in consumer valuations will be assigned a lower weight if the difference in the net means is less than the differences in the variances multiplied by the risk aversion parameter. The proof of this follows from reversing the last inequalities above. $\square$

Part (i) in Corollary 1 will always be satisfied if $\sigma_i < \sigma_j$, as then the LHS of the above inequality is negative. Clearly that is so because in this case good i is superior to good j both in terms of the net of cost mean, as well as in terms of the variance criterion.

However good i may still be given a greater weight than j in the bundle even when $\sigma_i > \sigma_j$. Hence pure bundling in our model is not always dispersion reducing, as the latter is not always optimal. If the difference between the net means on the RHS of the inequality discussed in Corollary 1 is sufficiently large to exceed the LHS, then good i will have a greater weight in the bundle than j despite the fact that $\sigma_i > \sigma_j$. This may cause the dispersion of the bundle to exceed the sum of the dispersion of the two stand alone goods. We shall return to this point latter.

⁹This translates into $P < \mu$. If $A > 0$ ($\alpha < 0$) and utility is positive then there can only be corner solutions where goods are offered separately. This optimality of non bundling when the demand function is convex is also confirmed by Schmalensee who concludes that pure bundling is less profitable than unbundled sales for $P > \mu$.

¹⁰This is for convenience purposes only and in no way implies 2 units.

Hence combining Theorem 2 and part (i) of Corollary 1 we obtain the following result. For an optimal bundling decision such that the good with the higher net mean is also given a larger but strictly less than one ($\lambda < 1$) relative weight, the condition is

$$2\alpha\sigma_i(\sigma_i - \rho\sigma_j) > \mu_i - \mu_j - (C_i - C_j) > \alpha(\sigma_i^2 - \sigma_j^2).$$

Going back to the case where $\sigma_i < \sigma_j$, it is still possible for good j to feature in the bundle as long as $\sigma_i > \rho\sigma_j$ which will be always (more easily) satisfied in the case of a negative (low positive) value of ρ . Hence low correlation between the two goods makes bundling more desirable. While this is a result shared with the Schmalensee paper, given that pure bundling lowers profits the closer ρ is to 1, the weight of each good here is determined, among other things, by the correlation coefficient. We show this more formally below, by first differentiating k^* with respect to ρ .

As already mentioned, the decision whether to bundle in the first place is endogenous, while the weights of the goods in the bundle are also endogenously determined by their means, dispersion, the degree of correlation, and finally the degree of risk aversion as set out in relations (35). We derive below the *comparative static results* on the impact of each of these parameters.

Implicit differentiation of (31), keeping means, costs, variances and ρ constant gives:

$$\frac{dk^*}{d\alpha} = \frac{(\mu_2 - \mu_1) - (C_2 - C_1)}{\alpha^2(\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2)}. \quad (33)$$

The relation immediately above implies that the weight of the good with the higher net mean, $\mu_i - C_i$, is a decreasing function of risk aversion, α . Also, the higher the degree of risk aversion, the lower will be the absolute value of its impact on that weight.

This means that the existence of risk aversion is encouraging bundling by enhancing the contribution in the bundle of the good with the lower net mean all other things being equal.

Implicitly differentiating (31) with respect to ρ for given means, costs, variances and α gives:

$$\frac{dk^*}{d\rho} = \frac{2\sigma_2\sigma_1(k^* - 1)}{(\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2)}. \quad (34)$$

Hence for $k^* > 1$, $\frac{dk^*}{d\rho} > 0$, while for $k^* < 1$, $\frac{dk^*}{d\rho} < 0$. Hence a lower ρ boosts the share of the good with the lower relative weight in the bundle and improves its odds of having a non zero participation in the bundle. Consequently, low positive correlation values promote bundling, and even more so negative values of ρ .

By implicitly differentiating (31) with respect to the dispersion, for given means, costs, α , ρ , $d\sigma_2 = 0$ we obtain:

$$\frac{d(k^*)}{d\sigma_1} = \frac{2k^*(\rho\sigma_2 - \sigma_1) - 2\rho\sigma_2}{(\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2)} < 0^{11}. \quad (35)$$

This means that the share of a good in the bundle is inversely related to its own dispersion. It follows that:

$$\frac{d(2 - k^*)}{d\sigma_1} = -\frac{dk^*}{d\sigma_1} > 0. \quad (36)$$

This means that the share of a good in the bundle is directly related to the dispersion of the other good.

Implicitly differentiating (31) for given costs and variances, α and ρ , setting $d\mu_2 = 0$ and dividing by $d\mu_1$ gives $\frac{dk^*}{d\mu_1} = \frac{1}{\alpha(\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2)} > 0$. Hence, the optimal weight of each good is directly related to its mean and, as can be easily shown, inversely related to the mean of the other good.

Finally, implicitly differentiating (31) for given means, variances, α and ρ , setting $dC_2 = 0$ and dividing by dC_1 we obtain:

$$\frac{d(k^*)}{dC_1} = \frac{1}{-\alpha(\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2)} < 0. \quad (37)$$

It follows that the optimal weight of each good in the bundle is inversely related to its cost, and, as can easily be shown, directly related to the cost of the other good in the bundle.

¹¹The proof is as follows. Assume that the net mean of good 1 is greater than the net mean of good 2. Distinguish two cases:

(i) For $k^* \geq 1$, $\sigma_1 \geq \rho\sigma_2$ (else $2 - k^* < 0$).

(ii) For $k^* < 1$, $\sigma_1 \geq \sigma_2$.

Hence in both cases $\rho\sigma_2 - \sigma_1 < 0$.

In (i) if $\rho \geq 0$, $\frac{dk^*}{d\sigma_1} < 0$ as $2k^*(\rho\sigma_2 - \sigma_1) - 2\rho\sigma_2 < 0$. If $\rho < 0$, since $2k^*\rho\sigma_2 < 2\rho\sigma_2 < 0$, it also follows that $\frac{dk^*}{d\sigma_1} < 0$.

Taking (ii), if $\rho \geq 0$, $(\rho\sigma_2 - \sigma_1) < 0$ and $-2\rho\sigma_2 < 0$. Hence $\frac{dk^*}{d\sigma_1} < 0$. Finally, if $\rho < 0$ then as $\sigma_1 > \sigma_2$ and $|k^* - 1| < 1$, it follows that $0 < 2(k^* - 1)\rho\sigma_2 < 2k^*\sigma_1$. It follows that the nominator is negative and this means that $\frac{dk^*}{d\sigma_1}$ is negative.

3.2 Derivation of the profit maximising price

Having derived the optimal weights for the two goods, we now proceed to derive the profit maximising bundle price. The firm will offer the bundle at the specific price as a take it or leave it option (pure bundling).

The mean and the cost of the bundle are respectively defined as¹² $\mu_B = k\mu_1 + (2-k)\mu_2$, $C_B = kC_1 + (2-k)C_2$. The standard deviation of the bundle can be written as follows:

$$\sigma_B = (k\sigma_1 + (2-k)\sigma_2) \sqrt{1 - 2(1-\rho)\theta(1-\theta)}, \quad (38)$$

where ρ is the correlation coefficient of the joint reservation distribution, and

$$\theta = \frac{k\sigma_1}{k\sigma_1 + (2-k)\sigma_2}, \quad (39)$$

is the share of Good 1 in the weighted sum of dispersions. Hence, $\sigma_B = \delta(k\sigma_1 + (2-k)\sigma_2)$, where:

$$\delta = \sqrt{1 - 2(1-\rho)\theta(1-\theta)}. \quad (40)$$

Since $0 \leq \delta \leq 1$, it follows that $\sigma_B < k\sigma_1 + (2-k)\sigma_2$. The value of δ is minimised for any given ρ if $\theta = \frac{1}{2}$ so that the proportion of the two goods in the bundle is set as $\frac{\sigma_1}{\sigma_2} = \frac{2-k}{k}$. Such weight-setting is clearly not possible in the Schmalensee paper. A value of $\theta = \frac{1}{2}$ requires that $k\sigma_1 + (2-k)\sigma_2 < \sigma_1 + \sigma_2$. In other words, the good with the higher dispersion will receive the lower weight in the pack so that the weights of each good's contribution to the weighted sum of the dispersions are equal.

Relations (32) imply that *a good will receive a higher relative weight in the bundle if it has a net mean that more than makes up for a relatively higher dispersion*, as seen in part (i) of Corollary 1. This means that $k^*\sigma_1 + (2-k^*)\sigma_2 > \sigma_1 + \sigma_2$ and $\delta_{k^*} > \delta_{k=2-k}$. Hence in this case σ_B , as defined in relation (38), is for $k = k^*$ larger than what it is for $k = 2 - k = 1$. (This is further discussed as Case 3 in the next subsection.) In fact, it is possible that σ_B may even be larger than $\sigma_1 + \sigma_2$, if $k^*\sigma_1 + (2-k^*)\sigma_2$ and δ_{k^*} are sufficiently larger than $\sigma_1 + \sigma_2$ and $\delta_{k=2-k}$ respectively.

Weights optimisation alters the bundle mean, μ_B , in relation to the fixed weights case. It also alters σ_B in two ways: by affecting δ as well as the weighted sum of the two dispersions. Using (39) and inserting k^* from (32) we obtain the optimal share of Good 1 in dispersions:

$$\theta^* = \frac{\frac{(\mu_2 - \mu_1) - (C_2 - C_1)}{\sigma_2} - 2\alpha(\sigma_2 - \rho\sigma_1)}{-2\alpha(1-\rho)(\sigma_1 + \sigma_2) + \frac{[(\mu_1 - \mu_2) - (C_1 - C_2)](\sigma_2 - \sigma_1)}{\sigma_1\sigma_2}}. \quad (41)$$

The above implies that θ^* in our framework is a function of ρ ; differentiating with respect to ρ we find that θ^* is an increasing (decreasing) function of ρ when $\sigma_2 > \sigma_1$ ($\sigma_2 < \sigma_1$). If the net means of the two goods are equal, then we have $\theta^* = \frac{\sigma_2 - \rho\sigma_1}{(1-\rho)(\sigma_1 + \sigma_2)}$, which for as long as $\sigma_2 > \rho\sigma_1$ guarantees a positive weight for good 1.

We now proceed into defining the profit functions for the composite good (bundle), using the definition in relation (7). Π_{B,k^*}^* is the optimal profit when the weights are k^* and $2 - k^*$, i.e.:

$$\Pi_{B,k^*}^* = (P_{B,k^*}^* - C_{B,k^*}) \left[1 - F\left(\frac{P_{B,k^*}^* - \mu_{B,k^*}}{\sigma_{B,k^*}}\right) \right].$$

Equivalently $\Pi_{B,k=2-k}^*$ is the optimal profit when weights are set at 1:1, as for example in the Schmalensee model. Π_{B,k^*}^* is different to $\Pi_{B,k=2-k}^*$ unless part (ii) in Corollary 1 is satisfied, i.e. $\mu_i - \mu_j - (C_i - C_j) = \alpha(\sigma_1^2 - \sigma_2^2)$.

Using relation (16) in Section 2, it is easy to show that the derivative of the optimal profit with respect to the mean is:

$$\frac{\partial \Pi_{B,k^*}^*}{\partial \mu_{B,k^*}} = 1 - F\left(\frac{P_{B,k^*}^* - \mu_{B,k^*}}{\sigma_{B,k^*}}\right),$$

$$\text{while } \frac{\partial \Pi_{B,k^*}^*}{\partial C_{B,k^*}} = - \left(1 - F\left(\frac{P_{B,k^*}^* - \mu_{B,k^*}}{\sigma_{B,k^*}}\right) \right) = - \frac{\partial \Pi_{B,k^*}^*}{\partial \mu_{B,k^*}}.$$

The analysis of the changes in the profit and welfare functions is markedly different and rather more satisfactory than the analysis of the calculations when constraints are imposed on the mean valuation and cost of the bundle to move in the same proportion and in opposite directions (as in Dassiou and Glycopantis [2012]). By allowing separate and independent changes of the bundle mean and costs we capture the full effect of changes in profit and welfare. The derivative with respect to the bundle variance is:

$$\frac{\partial \Pi_{B,k^*}^*}{\partial \sigma_{B,k^*}^2} = \frac{1}{2\sigma_{B,k^*}} \frac{\partial \Pi_{B,k^*}^*}{\partial \sigma_{B,k^*}} = \left(\frac{P_{B,k^*}^* - \mu_{B,k^*}}{2(\sigma_{B,k^*})^2} \right) \left[1 - F\left(\frac{P_{B,k^*}^* - \mu_{B,k^*}}{\sigma_{B,k^*}}\right) \right]. \quad (42)$$

¹²As k , $2 - k$ sum up to 2, the bundle mean and cost are exactly half that as $k = 2\lambda$. As this normalization does not affect our conclusions we continue our discussion in terms of k 's rather than λ 's.

3.3 Consumer and welfare implications of optimal bundling

For the welfare function, $W_{B,k^*}^* = \sigma_{B,k^*} f\left(\frac{P_{B,k^*}^* - \mu_{B,k^*}}{\sigma_{B,k^*}}\right) + (\mu_{B,k^*} - C_{B,k^*}) \left[1 - F\left(\frac{P_{B,k^*}^* - \mu_{B,k^*}}{\sigma_{B,k^*}}\right)\right]$, the corresponding relations are:

$$\frac{\partial W_{B,k^*}^*}{\partial \mu_{B,k^*}} = \frac{P_{B,k^*}^* - C_{B,k^*}}{\sigma_{B,k^*}} f\left(\frac{P_{B,k^*}^* - \mu_{B,k^*}}{\sigma_{B,k^*}}\right) \frac{(P_{B,k^*}^* - C_{B,k^*})(P_{B,k^*}^* - \mu_{B,k^*}) - 3\sigma_{B,k^*}^2}{(P_{B,k^*}^* - C_{B,k^*})(P_{B,k^*}^* - \mu_{B,k^*}) - 2\sigma_{B,k^*}^2} = -\frac{\partial W_{B,k^*}^*}{\partial C_{B,k^*}},$$

while $\frac{\partial W_{B,k^*}^*}{\partial \sigma_{B,k^*}^2}$ (derived as relation (46) in the mathematical appendix) is shown to be such that $\left|\frac{\partial W_{B,k^*}^*}{\partial \sigma_{B,k^*}^2}\right| < \frac{\partial W_{B,k^*}^*}{\partial \mu_{B,k^*}}$. In other words, *the absolute value of the rate of change in welfare with respect to the variance is smaller to the rate of change in welfare with respect to the mean.*

In order to analyse the consumer and welfare implications of our pure bundling approach and make comparisons with that of Schmalensee approach we consider three cases.

. The firm faces no dilemma as the good with the larger mean also has the lower dispersion. Hence it assigns a larger weight to this good.

In this case it is relatively rare that our optimal packaging approach is welfare inferior to that of the conventional approach. It occurs only when the ratio of the absolute value of the change in the bundle variance to the change in the net bundle mean exceeds the fraction of the marginal rate of change of welfare with respect to the mean divided by the marginal rate of change with respect to the variance, provided that the latter is positive.

. This means that for $\mu_i - C_i > \mu_j - C_j$ and $\sigma_i < \sigma_j$, we have that $\mu_{B,k^*} - C_{B,k^*} = k^*(\mu_i - C_i) + (2 - k^*)(\mu_j - C_j) > \mu_i - C_i + \mu_j - C_j$ and $\sigma_{B,k^*} \leq \sigma_{B,k=2-k}$ as $k^* > (2 - k^*)$. Clearly in this case as profit is an increasing function of the net mean and a decreasing function of the dispersion, we have that $\Pi_{B,k^*}^* \geq \Pi_{B,k=2-k}^*$. Using the small increments formula we can define an approximate equality for the difference between the welfare in our model minus the welfare in Schmalensee's, as $\Delta W_{B,k^*}$ that will allows us to determine the sign of the change in welfare of using our model instead of that of equal weights:¹³

$$\Delta W_{B,k^*} \cong \frac{\partial W_{B,k^*}^*}{\partial \mu_{B,k^*}} \nu \Delta_m + \frac{\partial W_{B,k^*}^*}{\partial C_{B,k^*}} \nu \Delta_c + \frac{\partial W_{B,k^*}^*}{\partial \sigma_{B,k^*}^2} \nu \Delta_{var} = \frac{\partial W_{B,k^*}^*}{\partial \mu_{B,k^*}} \nu \Delta_{nm} + \frac{\partial W_{B,k^*}^*}{\partial \sigma_{B,k^*}^2} \nu \Delta_{var} \quad (43)$$

since $\frac{\partial W_{B,k^*}^*}{\partial \mu_{B,k^*}} = -\frac{\partial W_{B,k^*}^*}{\partial C_{B,k^*}}$. We identify $\Delta_{nm} = \Delta_m - \Delta_c = \mu_{B,k^*} - \mu_{B,k=2-k} - (C_{B,k^*} - C_{B,k=2-k})$ as the difference in the net bundle mean of our model minus that of Schmalensee's pure bundle net mean. We also identify $\Delta_{var} = \sigma_{B,k^*}^2 - \sigma_{B,k=2-k}^2$ as the difference in the variance of the bundle in our model minus the bundle variance in Schmalensee's model. To ensure the accuracy of the approximation in the case of large changes in net means and variances, we multiply by ν , $0 \leq \nu \leq \epsilon$, where ϵ is arbitrarily small.

Considering the impact on welfare using (29) if $(P_{B,k^*}^* - C_{B,k^*})^2 - 2\sigma_{B,k^*}^2 > 0$, we derive that $\frac{\partial W_{B,k^*}^*}{\partial \sigma_{B,k^*}^2} < 0$. This means that a decrease in the bundle dispersion increases welfare, which will be further reinforced by the increase in the bundle mean. Hence in this case our bundling is superior to Schmalensee's both in terms of the profits as well as the total welfare.

On the other hand, if $(P_{B,k^*}^* - C_{B,k^*})^2 - 2\sigma_{B,k^*}^2 < 0$ ¹⁴, then $\frac{\partial W_{B,k^*}^*}{\partial \sigma_{B,k^*}^2} > 0$. The rise in the net mean would still increase welfare, but this will have to be compared with the reduction that a decrease in the bundle dispersion will cause. Hence, in this case it is ambiguous whether the reduction in welfare is smaller or larger than in the Schmalensee case. In other words, while the dispersion is further reduced, the bundle mean is increased. The latter can more or less than offset the negative impact on welfare of a reduction in dispersion. We attempt to clarify this ambiguity below.

We have that,

$$\Delta_{nm} = \frac{(\mu_2 - \mu_1) - (C_2 - C_1) - \alpha(\sigma_2^2 - \sigma_1^2)}{\alpha(\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2)} (\mu_2 - \mu_1 - (C_2 - C_1)), \quad (44)$$

and

$$\Delta_{var} = \frac{\Delta_{nm}}{\alpha} + \Delta_{nm} \frac{\sigma_2^2 - \sigma_1^2}{\mu_2 - \mu_1 - (C_2 - C_1)} \frac{\sigma_2^2 + \sigma_1^2 + 2\rho\sigma_1\sigma_2(1 - 2\alpha)}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}. \quad (45)$$

The derivation of expressions (44) and (45) is rather involved and it is given in the appendix. Given the conditions within Case 1, $\Delta_{var} < 0$, $\Delta_{nm} > 0$ and $k^* > 2 - k^*$. Using (45) it is easy to show Lemma 1.

¹³The small increments formula in (43) gives in effect:

$$-\Delta W_{B,k^*}^* = -(W_{B,k^*}^* - W_{B,k=2-k}).$$

We call the left-hand side $\Delta W_{B,k^*}$ since we are interested in changes in welfare resulting from using bundling as determined in our case relative to welfare in Schmalensee's case, and for this calculation we want to use the specific derivatives we have calculated for the optimal welfare function.

¹⁴Which means that $\frac{P_{B,k^*}^* - C_{B,k^*}}{\sigma_{B,k^*}} < \sqrt{2}$, $\frac{\mu_{B,k^*} - C_{B,k^*}}{\sigma_{B,k^*}} < 1.561$ and $\frac{P_{B,k^*}^* - \mu_{B,k^*}}{\sigma_{B,k^*}} < -0.147$.

$\Delta_{nm} < |\Delta_{var}|$ if and only if $\frac{(\sigma_2^2 - \sigma_1^2)}{(\mu_1 - \mu_2) - (C_1 - C_2)} \frac{\sigma_2^2 + \sigma_1^2 + 2\rho\sigma_1\sigma_2(1-2a)}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2} > 1 + \frac{1}{\alpha}$.

Lemma 1 will be used in assessing whether welfare has been increased or decreased relative to the Schmalensee case. As $\left| \frac{\partial W_{B,k^*}^*}{\partial \sigma_{B,k^*}^2} \right| < \frac{\partial W_{B,k^*}^*}{\partial \mu_{B,k^*}}, \Delta W_{B,k^*} \geq 0$ if either $\frac{\partial W_{B,k^*}^*}{\partial \sigma_{B,k^*}^2} \leq 0$, or if $\frac{\partial W_{B,k^*}^*}{\partial \sigma_{B,k^*}^2} > 0$ and $\Delta_{nm} \geq |\Delta_{var}|$. The change in welfare $\Delta W_{B,k^*}$ will still be non-negative even if $\frac{\partial W_{B,k^*}^*}{\partial \sigma_{B,k^*}^2} > 0$, and $\Delta_{nm} < |\Delta_{var}|$, $\frac{\partial W_{B,k^*}^*}{\partial \sigma_{B,k^*}^2} > 0$, unless the following proposition holds:

$\Delta W_{B,k^*} < 0$ if and only if $\frac{\partial W_{B,k^*}^*}{\partial \sigma_{B,k^*}^2} > 0$, Lemma 1 is satisfied ($\Delta_{nm} < |\Delta_{var}|$) and $\frac{|\Delta_{var}|}{\Delta_m} > \frac{\frac{\partial W_{B,k^*}^*}{\partial \mu_{B,k^*}}}{\frac{\partial W_{B,k^*}^*}{\partial \sigma_{B,k^*}^2}}$.

The above follows from inspection of (43).

We can then summarise below the findings relating to Case 1:

- (i) If $\frac{\partial W_{B,k^*}^*}{\partial \sigma_{B,k^*}^2} \leq 0$, then $\Delta W_{B,k^*} \geq 0$.
- (ii) If $\frac{\partial W_{B,k^*}^*}{\partial \sigma_{B,k^*}^2} > 0$, and $\Delta_{nm} \geq |\Delta_{var}|$, then $\Delta W_{B,k^*} \geq 0$.
- (iii) For $\frac{\partial W_{B,k^*}^*}{\partial \sigma_{B,k^*}^2} > 0$, and Lemma 1 satisfied ($\Delta_{nm} < |\Delta_{var}|$), to determine $\Delta W_{B,k^*}$ we will have to directly compare $\frac{\partial W_{B,k^*}^*}{\partial (\mu_{B,k^*})} \Delta_m$ to $\frac{\partial W_{B,k^*}^*}{\partial \sigma_{B,k^*}^2} |\Delta_{var}|$. Unless Proposition 1 holds, in Case 1 our optimal pure bundling will be more welfare enhancing relative to the outcome of the Schmalensee case of pure bundling.

Below we discuss Cases 2 and 3 where the firm faces a dilemma as the good with the larger net mean also has a larger dispersion. It has to choose between its desire to increase the bundle mean and decrease the dispersion in order to enhance its profits.

The firm faces a dilemma; it chooses to assign a larger weight to the good with the lower dispersion, e.g. $k^* < (2 - k^*)$.

Here both the mean as well as the dispersion are decreased. $\mu_i - C_i > \mu_j - C_j$, $\sigma_i > \sigma_j$ and $\mu_i - \mu_j - (C_i - C_j) < \alpha(\sigma_i^2 - \sigma_j^2)$, we have that $\mu_{B,k^*} - C_{B,k^*} = k^*(\mu_i - C_i) + (2 - k^*)(\mu_j - C_j) < \mu_i - C_i + \mu_j - C_j$ and $\sigma_{B,k^*} < \sigma_{B,k=2-k}$ as $k^* < (2 - k^*)$. When the difference in the variance of the two goods multiplied by the degree of risk aversion is greater than the difference in the net means (e.g. $\mu_i - \mu_j - (C_i - C_j) < \alpha(\sigma_i^2 - \sigma_j^2)$), the overriding criterion for the company in to bundle in such a way as to decrease the dispersion rather than an increase in the net mean.

The consumer surplus will be decreased unambiguously, given that it is an increasing function of both the mean as well as the dispersion. The optimal weights chosen by the firm will not only decrease dispersion further but will also reduce the bundle mean. Hence, Case 2 is a case where a competition authority is justified to intervene if its primary objective is the protection of the consumer surplus.

We now turn our attention to the overall welfare implications. We note that both Δ_{var} and Δ_{nm} are negative. So if $\frac{\partial W_{B,k^*}^*}{\partial \sigma_{B,k^*}^2} > 0$ then both terms in the small increments formula will be negative and welfare will unambiguously decrease. Only in the case where $\frac{\partial W_{B,k^*}^*}{\partial \sigma_{B,k^*}^2} < 0$ and $|\Delta_{nm}| < |\Delta_{var}|$ we will have to weight the overall positive impact of the change in the variance against the overall negative impact of the change in the net mean in small increments formula. Then it follows:

$\Delta W_{B,k^*} > 0$ if and only if $\frac{\partial W_{B,k^*}^*}{\partial \sigma_{B,k^*}^2} < 0$, and $\frac{|\Delta_{var}|}{|\Delta_{nm}|} > \frac{\frac{\partial W_{B,k^*}^*}{\partial (\mu_{B,k^*})}}{\left| \frac{\partial W_{B,k^*}^*}{\partial \sigma_{B,k^*}^2} \right|} > 1$.

Again this follows from using (43).

The firm faces a dilemma; it chooses to assign a larger weight to the good with the higher dispersion, e.g. $k^* > (2 - k^*)$.

Hence in Case 3 the overriding criterion for the company when setting the weights is an increase in the bundle mean rather than a decrease in dispersion. For $\mu_i - C_i > \mu_j - C_j$, $\sigma_i > \sigma_j$ and $\mu_i - \mu_j - (C_i - C_j) > \alpha(\sigma_i^2 - \sigma_j^2)$ we have that $\mu_{B,k^*} - C_{B,k^*} = k^*(\mu_i - C_i) + (2 - k^*)(\mu_j - C_j) > \mu_i - C_i + \mu_j - C_j$ and $\sigma_{B,k^*} > \sigma_{B,k=2-k}$ as $k^* > (2 - k^*)$.

The assignment of a larger weight to the good with the higher dispersion in valuations will lead to an unambiguous substantial increase in consumer surplus relative to the Schmalensee case. The company will not only increase the bundle mean, but also increase, rather than reduce, its bundle dispersion. This will lead to an increase in consumer surplus on both counts. Here the policy makers should view bundling in variable proportions favourably. Such a strategy followed by the company will lead to a definite increase in consumer surplus and an almost definite increase in increase in welfare surplus¹⁵ as compared to the case where pure bundling is restricted to fixed proportions.

¹⁵The case of $\Delta W_{B,k^*} < 0$ is quite remote for the following reasons. First, as we discussed in Case 1, given relation (29) we need that $(P_{B,k^*}^* - C_{B,k^*})^2 - 2\sigma_{B,k^*}^2 > 0$ for $\frac{\partial W_{B,k^*}^*}{\partial \sigma_{B,k^*}^2} < 0$. There is only a 7% probability of this happening if $P_{B,k^*}^* \in (\mu_{B,k^*} - 0.147\sigma_{B,k^*}, \mu_{B,k^*})$.

Second, if we are in this case we need to additionally have that $\Delta_{nm} < \Delta_{var}$ and that $\frac{\Delta_{var}}{\Delta_{nm}} = \frac{\sigma_{B,k^*}^2 - \sigma_{B,k=2-k}^2}{\mu_{B,k^*} - C_{B,k^*} - (\mu_{B,k=2-k} - C_{B,k=2-k})} > \frac{\frac{\partial W_{B,k^*}^*}{\partial (\mu_{B,k^*} - C_{B,k^*})}}{\left| \frac{\partial W_{B,k^*}^*}{\partial \sigma_{B,k^*}^2} \right|}$ for $\Delta W_{B,k^*} < 0$.

4 Conclusions

This paper discusses a model of pure bundling where a monopolistic firm does not restrict itself to assuming that the weights of the two goods are fixed. Therefore it is a far more realistic reflection of bundling in practice, compared with the conventional approach. We contrast the implications of an optimal choice of weights in terms of profit, consumer and welfare outcomes with those of a conventional model of pure bundling which assumes equal weights 1:1.

The decision whether to bundle or not is endogenous, and so are the weights chosen. By choosing the weights through an optimisation process, bundling is no longer mean preserving. Hence, both the means (net of costs) as well as the dispersions of the consumers' valuations of the two goods are important in determining the weights within the bundle. Additionally, this optimisation introduces the impact of risk aversion into the bundling decision.

We consider three different cases of the profit and welfare implications of our bundling approach. These depend on the values of the net means and dispersions, and the degree of risk aversion.

In Case 1 the inequalities in the net means and the dispersions of the two goods are in opposite directions. Our findings show that the firm is more sharply focused in reducing dispersion through bundling, which may reduce welfare. The adverse impact on the latter is however tempered by the fact that it also increases the net bundle mean. Hence typically in this case our optimal bundling will inflict a lower damage on welfare than the one using the Schmalensee approach.

In Case 2 the inequalities in the net means and the dispersions are in the same direction and the difference in the net means is smaller than the difference in the variances times the degree of risk aversion. This will induce the firm to offer a bundle such that the good that has both a higher net mean as well as a higher dispersion will be assigned a lower weight. This means that the firm will bundle in a way that is both mean and variance reducing. This leads to adverse effects in terms of consumer surplus as well as welfare.

In Case 3 the inequalities in the net means and the dispersions are in the same direction; however this time the difference in the net means is larger than the difference in the variances times the degree of risk aversion. The firm will then bundle in a way that is both mean and variance increasing, leading to beneficial effects both in terms of the consumer surplus as well as welfare. The choice of optimal weights in the bundle is such that, the standard deviation $\sigma_B |_{k^* \neq 2-k^*} > \sigma_B |_{k=2-k=1}$, as the good with the higher dispersion is assigned a larger weight. Such a weight may even lead to a case where $\sigma_B |_{k^* \neq 2-k^*} > \sigma_1 + \sigma_2$. This means that pure optimal bundling will in this case lead to an increase rather than a decrease in the dispersion not only in relation to conventional bundling, but also in relation to the case of no bundling. Obviously such a typically welfare enhancing case was not a possibility of pure bundling in the Schmalensee paper.

Mathematical Appendix

Below we provide proofs of certain expressions and show that $\left| \frac{\partial W_{B,k^*}^*}{\partial \sigma_{B,k^*}^2} \right| < \frac{\partial W_{B,k^*}^*}{\partial \mu_{B,k^*}}$ in the sequence in which they appear in the text.

Explanation of Expression (29):

Substituting relations (14) and (21) into the second term on the right hand side expression in (22), and adding (19) to (22) we get:

$$\begin{aligned} \frac{d\Pi^*}{d\sigma} + \frac{dCS}{d\sigma} |_{P=P^*} &= f\left(\frac{P^*-\mu}{\sigma}\right) + \left(\frac{1-F\left(\frac{P^*-\mu}{\sigma}\right)}{\sigma}\right) \left(P^* - \mu - \frac{(P^*-C)(P^*-\mu)^2 - [2P^*-(\mu+C)]\sigma^2}{(P^*-C)(P^*-\mu)-2\sigma^2}\right) \Leftrightarrow \\ &= \frac{d\Pi^*}{d\sigma} + \frac{dCS^*}{d\sigma} = f\left(\frac{P^*-\mu}{\sigma}\right) \left(1 + \frac{(P^*-C)(\mu-C)}{(P^*-C)(P^*-\mu)-2\sigma^2}\right) \end{aligned}$$

Explanation of Expression (44):

$$\begin{aligned} \Delta_{nm} = \Delta_m - \Delta_c &= \mu_{B,k^*} - \mu_{B,k=2-k} - (C_{B,k^*} - C_{B,k=2-k}) = \mu_{B,k^*} - C_{B,k^*} - (\mu_{B,k=2-k} - C_{B,k=2-k}) = \\ &= (k^*-1)(\mu_1 - C_1) + (2-k^*-1)(\mu_2 - C_2) = \\ &= (k^*-1)[(\mu_1 - \mu_2) - (C_1 - C_2)] \end{aligned}$$

Replacing k^* by its value we obtain:

$$\Delta_{nm} = \Delta_c - \frac{(\mu_2 - \mu_1) - (C_2 - C_1) - \alpha(\sigma_2^2 - \sigma_1^2)}{\alpha(\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2)} [\mu_2 - \mu_1 - (C_2 - C_1)]$$

Explanation of Expression (45):

$$\Delta_{var} = \sigma_{B,k^*}^2 - \sigma_{B,k=2-k}^2 = ((k^*)^2 - 1)\sigma_1^2 + ((2-k^*)^2 - 1)\sigma_2^2 + 2\rho\sigma_1\sigma_2(k^*(2-k^*) - 1)$$

Replacing k^* by its value we obtain:

$$\begin{aligned}\Delta_{var} &= \frac{[(\mu_2 - \mu_1) - (C_2 - C_1)]^2 - \alpha(\sigma_2^2 - \sigma_1^2)[(\mu_2 - \mu_1) - (C_2 - C_1)]}{\alpha^2(\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2)^2} (\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2) \\ &+ \frac{\alpha(\sigma_2^2 - \sigma_1^2)[\sigma_2^2 + \sigma_1^2 - 4\alpha\rho\sigma_1\sigma_2][(\mu_2 - \mu_1) - (C_2 - C_1)] - \alpha(\sigma_2^2 - \sigma_1^2)}{\alpha^2(\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2)^2} + 2\rho\sigma_1\sigma_2 \frac{\alpha(\sigma_2^2 - \sigma_1^2)[(\mu_2 - \mu_1) - (C_2 - C_1)]}{\alpha^2(\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2)^2} \\ &+ 2\rho\sigma_1\sigma_2 \frac{4\sigma_1\sigma_2(\sigma_1 - \rho\sigma_2)(\sigma_2 - \rho\sigma_1) - (\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2)^2}{(\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2)^2} \Leftrightarrow \\ &\Delta_{var} = \frac{[(\mu_2 - \mu_1) - (C_2 - C_1)]^2 - \alpha(\sigma_2^2 - \sigma_1^2)[(\mu_2 - \mu_1) - (C_2 - C_1)]}{\alpha^2(\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2)^2} (\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2) + \frac{\alpha(\sigma_2^2 - \sigma_1^2)[\sigma_2^2 + \sigma_1^2 + 2\rho\sigma_1\sigma_2(1 - 2\alpha)][(\mu_2 - \mu_1) - (C_2 - C_1)]}{\alpha^2(\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2)^2} \\ &- \frac{(\sigma_2^2 - \sigma_1^2)^2(\sigma_2^2 + \sigma_1^2 - 4\alpha\rho\sigma_1\sigma_2 + 2\rho\sigma_1\sigma_2)}{(\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2)^2} \Leftrightarrow \\ \Delta_{var} &= \frac{\Delta_{nm}}{\alpha} + \Delta_{nm} \frac{\sigma_2^2 - \sigma_1^2}{[\mu_2 - \mu_1 - (C_2 - C_1)]} \frac{\sigma_2^2 + \sigma_1^2 + 2\rho\sigma_1\sigma_2(1 - 2\alpha)}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}.\end{aligned}$$

Proof that $\left| \frac{\partial W_{B,k^}^*}{\partial \sigma_{B,k^*}^2} \right| < \frac{\partial W_{B,k^*}^*}{\partial \mu_{B,k^*}}$:*

We can easily show that the rate of change in welfare with respect to the variance is:

$$\begin{aligned}\frac{\partial W_{B,k^*}^*}{\partial \sigma_{B,k^*}^2} &= \frac{\partial W_{B,k^*}^*}{\partial \sigma_{B,k^*}} \frac{1}{2\sigma_{B,k^*}} \Leftrightarrow \\ \frac{\partial W_{B,k^*}^*}{\partial \sigma_{B,k^*}^2} &= \frac{f\left(\frac{P_{B,k^*}^* - \mu_{B,k^*}}{\sigma_{B,k^*}}\right)}{2\sigma_{B,k^*}} \left[\frac{(P_{B,k^*}^* - C_{B,k^*})^2 - 2\sigma_{B,k^*}^2}{(P_{B,k^*}^* - C_{B,k^*})(P_{B,k^*}^* - \mu_{B,k^*}) - 2\sigma_{B,k^*}^2} \right].\end{aligned}$$

Hence,

$$\frac{\partial W_{B,k^*}^*}{\partial \sigma_{B,k^*}^2} = \frac{\partial W_{B,k^*}^*}{\partial \mu_{B,k^*}} \frac{2\sigma_{B,k^*}^2 - (P_{B,k^*}^* - C_{B,k^*})^2}{2(P_{B,k^*}^* - C_{B,k^*}) [3\sigma_{B,k^*}^2 - (P_{B,k^*}^* - C_{B,k^*})(P_{B,k^*}^* - \mu_{B,k^*})]} \quad (46)$$

For $\mu_{B,k^*} > P_{B,k^*}^* > C_{B,k^*}$ and $2\sigma_{B,k^*}^2 - (P_{B,k^*}^* - C_{B,k^*})^2 > 0$:

$$0 < \frac{2\sigma_{B,k^*}^2 - (P_{B,k^*}^* - C_{B,k^*})^2}{2(P_{B,k^*}^* - C_{B,k^*}) [3\sigma_{B,k^*}^2 - (P_{B,k^*}^* - C_{B,k^*})(P_{B,k^*}^* - \mu_{B,k^*})]} < 1$$

as $(P_{B,k^*}^* - C_{B,k^*})(P_{B,k^*}^* - \mu_{B,k^*}) < 0$. Hence $0 < \frac{\partial W_{B,k^*}^*}{\partial \sigma_{B,k^*}^2} < \frac{\partial W_{B,k^*}^*}{\partial \mu_{B,k^*}}$.

Analogously, for $\mu_{B,k^*} > P_{B,k^*}^* > C_{B,k^*}$ and $2\sigma_{B,k^*}^2 - (P_{B,k^*}^* - C_{B,k^*})^2 < 0$ we obtain $0 < \left| \frac{\partial W_{B,k^*}^*}{\partial \sigma_{B,k^*}^2} \right| < \frac{\partial W_{B,k^*}^*}{\partial \mu_{B,k^*}}$.

Acknowledgements

We are grateful for all the incisive comments received on an earlier draft to this paper. We also wish to thank the discussant Professor K. Serfes and other participants in the July 2011 CRESSE conference in Rhodes, Greece. Responsibility for the final version stays of course with the authors.

REFERENCES

- [1] Adams, W.J., Yellen, J.L. What makes advertising profitable?, *Economic Journal*, Vol.87, 427-449, 1976.
- [2] McAfee, R.P., McMillan, J., Whinston, M. Multiproduct monopoly, commodity bundling and correlation of values, *Quarterly Journal of Economics*, Vol.103, 371-383, 1989.
- [3] Stigler, G. A note on block booking. In: Stigler, G. (Ed.) *The Organisation of Industry*, Richard D. Irwin, 1968.
- [4] Schmalensee, R. Gaussian demand and commodity bundling, *The Journal of Business* Vol.57, No.1, Part 2, S211-S230, 1984.
- [5] Dassiou, X., Choi, C.J. and Maldoom, D. Trade and linked exchange; Price discrimination through transaction bundling, *Topics in Theoretical Economics*, Berkeley Electronic Press, Vol.4, No.1textbf, 2004.

- [6] Dassiou, X., Glycopantis, D. and Stavropoulou, C. Bundling in general markets and in health care systems, *The Journal of Economic Asymmetries*, Vol.6, No.2, 47-68, 2009.
- [7] Dassiou, X., Glycopantis, D. The economic theory of price discrimination via transactions bundling: An assessment of the policy implications, *Review of Law Economics*, Berkeley Electronic Press, Vo.2, No.2textbf, 2006.
- [8] Dassiou, X., Glycopantis, D. Price discrimination through transactions bundling: The case of monopsony, *Journal of Mathematical Economics*, Vol.44, No.7-8textbf, 672-681, 2008.
- [9] Dassiou, X., Glycopantis, D. Optimal decisions in two-stage bundling. In: Joseph E. Harrington Jr. and Yannis Katsoulacos (Eds.) *Recent Advances in the Analysis of Competition Policy and Regulation*, Edward Elgar, 95-115, 2012.
- [10] DeGraba, P. Quantity discounts from risk averse sellers, *Federal Trade Commission - Bureau of Economics*, Working Paper 276, 2005.10
- [11] Dana, J.D. Advance-purchase discounts and price discrimination in competitive markets, *Journal of Political Economy*, Vol.106, No. 2, 395-422, 1998.
- [12] Dana, J.D. Equilibrium price dispersion under demand uncertainty, *Rand Journal of Economics*, Vol.30, No.4, 632-660, 1999.
- [13] Dana, J.D. Using yield management to shift demand when the peak time is unknown, *Rand Journal of Economics*, Vol.30, No.3, 456-474, 1999.
- [14] Gaggero, A.A., Piga, C.A. Airline market power and intertemporal price discrimination, *Discussion Paper Series*, Department of Economics, Loughborough University, 2009.
- [15] Cowan, S. The welfare effects of third degree price discrimination with non-linear demand functions, *Rand Journal of Economics*, Vol. 38, No.2, 419-428, 2007.
- [16] Kimball, M. Precautionary saving in the small and in the large, *Econometrica*, Vol.58, 53-73, 1990.
- [17] Crampes C., Hollander, A. Product specification, multi-product screening and bundling: the case of pay TV, *Information Economics and Policy*, Vol.17, No.1textbf, 35-59, 2005.
- [18] Anam, M., Chiang, S.-H. Price discrimination and social welfare with correlated demand, *Journal of Economic Behavior Organization*, Vol.61, 110-122, 2006.
- [19] Eckel, C.E., Smith, W.T. Price discrimination with correlated demands, *Southern Economic Journal*, Vol.59, No.1, 58-65, 1992.