manuscript No.
(will be inserted by the editor)

Optimal decisions in two stage bundling

Xeni Dassiou[1], **Dionysius Glycopantis**[2]

[1] Department of Economics, City University e-mail: x.dassiou@city.ac.uk
[2] Department of Economics, City University email: d.glycopantis@city.ac.uk

The date of receipt and acceptance will be inserted by the editor

Abstract We develop a generalised framework for pure bundling where buyer tastes for two goods are assumed to follow a normal distribution. In the literature optimal bundling decisions have been considered under the assumption that the weights of the two goods are fixed and equal. The only consideration is then to choose the profit maximising optimal price. Our approach is different and much more realistic. The monopolist first decides on the optimal weights of the two goods and in the second stage derives the profit maximising bundle price. Welfare and policy implications of our approach are derived and comparisons are made with those of the fixed weights approach.

Correspondence to: Dr. Xeni Dassiou, Department of Economics, City University, Northampton Square, London EC1V 0HB, United Kingdom, tel. +44 (0)20 70400206, fax. +44 (0)20 70408580. We are very grateful for all the incisive comments received on an earlier version of this paper. We also wish to thank our 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.

1 Introduction

Bundling has been discussed as an instrument of second degree price discrimination with distinct original contributions by a number of authors. Among them are Adams and Yellen (1976), McAfee, McMillan and Whinston (1989) and Schmalensee (1984). There is also an original contribution by Stigler (1968). Typically a uniform distribution was used to describe the valuation of consumers for two goods. On the other hand, Schmalensee used the bivariate normal distribution.

In Dassiou and Glycopantis (2006, 2008), using the uniform distribution we show that mixed bundling (where the consumers self-select among buying neither good, only one good, or both goods bundled together) leads to an increase in profits and if practiced by a monopsonist to an increase in trade for its trading partners.

The discussions in the literature have produced concrete results. We refer to these and where appropriate we explain the contribution in the present paper. We assume that the buyers' valuations of the two goods are distributed according to a bivariate normal distribution. The monopolist firm uses a pure bundling approach through the construction of a composite good. We will explain this approach in detail below.

Bundling increases the proportion of valuations around the mean by reducing the dispersion among the buyers. While this is good news in the case of low marginal costs, in the opposite case the seller will want to increase rather than decrease the dispersion of valuations. This is because if the marginal costs are greater than the mean valuation, bundling will decrease profits. By reducing the size of the lucrative fraction of buyers whose mean valuation exceeds the marginal cost of the bundle.

This suggests that the manipulation of the mean valuation is as important. Yet this is largely neglected in the bundling literature; instead the focus has been on the fact that bundling reduces dispersion and this allows the monopolistic firm to extract more of the consumers' surplus. In our paper we study a form of pure bundling which is not mean preserving. We return to this point below.

A second point regards the manipulation of dispersion. The standard deviations of the valuations for the goods in a bundle are sub-additive unless the goods are perfectly correlated. In other words, bundling reduces the effective dispersion of the reservation prices. However bundling is based on the conventional approach of assigning equal weights for the two goods in a bundle. Schmalensee notes that if symmetry (i.e. $\sigma_1 \simeq \sigma_2$) does not hold, pure bundling is less likely to be profit or welfare enhancing. He therefore argues that in this case the mixed bundling approach is preferable.

However, if the weights of pure bundling can be optimized by the firm this may no longer be the case. The ability to set weights may effectively 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. Moreover, it might also lead to an increase in the weight of a good which despite its relatively higher dispersion has a substantially higher mean valuation.

There are also a number of results in the literature that highlight the importance of the correlation in tastes. Crawford (2008) notes that the heterogeneity reduction in tastes achieved by bundling networks is greater the more negatively correlated the tasted for the goods included are. Using bundle-sized pricing, Chu *et al.* (2010) show that if consumers can select themselves which goods (theatrical plays in their example) to include in a bundle of pre determined size and price, they will include goods for which tastes are positively correlated. While this is still profit enhancing, it is less so than that of combining goods whose valuations are negatively correlated as in the Crawford example.

Issues connected explicitly with risk aversion have been considered by De Graba (2005) who introduces the idea of price discrimination without bundling. He finds that the firm will be willing to sacrifice some profit in order to increase the probability of making a sale to a large purchaser. Dana (1998, 1999a, 1999b) argues that price dispersion occurs because of stochasticity, and that firms set different prices, including discounts, to smooth out demand and reduce uncertainty. The work of Dana is discussed in Gaggero and Piga (2009) within the context of pricing strategies pursued by airlines.

Dana and De Graba show that the existence of price differentials is a "defensive mechanism" for responding to uncertainty, rather than an "aggressive mechanism" for extracting surplus through price discrimination as it is typically stressed in the bundling literature. In our model risk aversion intensifies the desire to reduce dispersion. The reduction in dispersion may enable the firm to better capture consumer surplus and this will lead to an increase in profits. The desire to reduce dispersion may be either tempered or strengthened by the desire to increase the average (mean) demand for the bundle.

In the literature the weights of the two goods in the bundle are taken as fixed. For example, Schmalensee assumes that the two goods participate in the bundle with a ratio 1:1. The firm determines the optimal bundle price through the maximisation of its profit function. Our paper breaks away from this tradition. We first formulate a utility function of the firm based on the composition of the bundle means, variances, correlation in tastes and the firm's degree of risk aversion. The weights of the two goods are chosen by the firm using a portfolio maximisation approach. It is then that the monopolist proceeds to determine through profit maximisation the bundle price. Comparing our optimal pure bundling to Schmalensee's, we show that in our framework the optimal relative weights are equal only when the difference in the net (of costs) means of the valuations for the two goods is equal to the degree of absolute risk aversion times the difference in the variances.

Our analysis has important policy implications. The first important point is that while portfolio optimisation attempts by a company to reduce dispersion through bundling may be detrimental to consumer surplus, this may be counteracted by the fact that at the same time an increase in the bundle mean may be achieved, which is both profit as well as consumer surplus enhancing. The second, more striking point is that the desire to increase the bundle mean may ultimately lead a choice of weights such that the result is an increase rather than a decrease in dispersion. In this case, bundling will lead to a further increase in consumer surplus and hence it be an unambiguously welfare enhancing practice.

Section 2 sets a preliminary background to the main investigation. It considers the case of unbundled sales and derives comparative static results for profits, consumer surplus and social welfare with respect to the mean and dispersion of the normal distribution which describes the consumers' valuations for a single good. The results are of relevance for the next section, as the bundle is in effect a composite good. Section 3 discusses in detail our approach to pure bundling where the consumer valuations of the two goods are given by a bivariate normal distribution. It analyses how the optimal weights and then the bundle price are obtained. It derives comparative static results and makes comparisons with the conventional fixed weights approach. Section 4 concludes the discussion.

2 Results based on separate sales of goods

Our model is based on Schmalensee's specification. We examine briefly the case of one good and derive *comparative static results* that we will use in our analysis of pure bundling. We are interested in the effect of changes in the mean and dispersion on profits, the consumer surplus and total welfare.

Assuming a density function of buyers for a single good sold separately $g(x)$, the demand for this good can be written as:

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

$g(x)$ is the underlying distribution of the buyers' valuations for the good and $Q(P)$ is the cumulative distribution. Each consumer who has a valuation for the good greater than its price buys one unit of the good. This valuation is described by $g(x)$ which is assumed to follow the normal distribution with mean and standard deviation μ and σ.

Let $f(t)$ be the standard normal density function, and define:

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

$F(x)$, the cumulative distribution function, is everywhere strictly increasing and it is straightforward to show that:

$$F'(x) = f(x), F''(x) = \frac{df(x)}{dx} = -xf(x), \int_x^\infty tf(t) = f(x), \frac{f(x)}{1-F(x)} > x.$$

The demand for a single good is thus given by $Q(P) = 1 - F(\frac{P-\mu}{\sigma})$. Below we will examine the impact on profits and on consumer surplus of changes in the mean and the dispersion.

Calculating the derivatives $Q'(P)$ and $Q''(P)$ we see that the demand function is decreasing and strictly convex for $P > \mu$, and strictly concave for $P < \mu$. We can think of the curvature of the demand function as being measured by $A = -\frac{Q''(P)}{Q'(P)} = \frac{P-\mu}{\sigma^2}$. A has the same sign as $Q''(P)$, which is negative for $P < \mu$; $A < 0$ is similar to the absolute risk aversion parameter applied to a utility function. For a concave function of utility of demand, its absolute risk aversion measure α ($\alpha > 0$) is enhanced by A in this case (e.g. $\alpha >\mid A \mid$). We will return to this discussion in the next section.

The profit function of the supplying firm also follows the normal distribution as it is linear in demand:

$$\Pi = (P - C)(1 - F(\frac{P-\mu}{\sigma})). \tag{3}$$

Profit maximisation results to the following FOC and SOC respectively:

$$(1 - F(\frac{P^*-\mu}{\sigma})) = \frac{P^*-C}{\sigma} f(\frac{P^*-\mu}{\sigma}), \tag{4}$$

$$(P^* - C)(P^* - \mu) < 2\sigma^2. \tag{5}$$

Combining the price cost margin condition $\frac{P^*-C}{P^*} = \frac{1}{\eta}$, where η is the absolute value of the elasticity of demand, with $\frac{f(x)}{1-F(x)} > x$ we obtain:

$$(P^* - C)(P^* - \mu) < \sigma^2. \tag{6}$$

The above means that the SOC is always satisfied. It can be re-written as:

$$\frac{\mu + C - \sqrt{(\mu - C)^2 + 4\sigma^2}}{2} < P^* < \frac{\mu + C + \sqrt{(\mu - C)^2 + 4\sigma^2}}{2} \tag{7}$$

The inclusion of $P^* = \mu$ in the above interval means that the demand, and through it the profit function, are not globally concave.

We need to obtain the effect of a change in the parameters of the demand on the profit maximising optimal price. From implicit differentiation of the FOC with respect to μ and invoking relation (6) is is easy to show that:

$$0 < \frac{dP^*}{d\mu} < 1. \tag{8}$$

Similarly,

$$\frac{dP^*}{d\sigma} = \frac{(P^* - C)(P^* - \mu)^2 - [2P^* - (\mu + C)]\sigma^2}{(P^* - C)(P^* - \mu) - 2\sigma^2} \frac{1}{\sigma}. \tag{9}$$

It is important to establish conditions for the sign of the change in optimum price with respect to the dispersion. The above two results are used repeatedly below in calculating the various comparative static results. Using straightforward calculations we establish the theorem below.

Theorem 1 *For $P^* > \mu$ we have that $\frac{dP^*}{d\sigma} > 0$. For $P^* < \mu$ we have that $\frac{dP^*}{d\sigma} < 0$ unless $\frac{\mu+C+\sqrt{(\mu-C)^2+4\sigma^2}}{2} > P^* > \frac{\mu+C.}{2}$*

This means that a higher dispersion raises the price when the latter is above the mean. A higher dispersion will continue to do so even if the price is a little below the mean (i.e. if P^* is larger than $\frac{\mu+C}{2}$ without violating the SOC upper boundary for P^*). Once the price is significantly below the mean a higher dispersion reduces the price.

Next, we wish to consider the impact mean and dispersion changes on the maximum profit. Taking $\frac{d\Pi^*}{d\mu}$ and using the envelope theorem we obtain:

$$\frac{d\Pi^*}{d\mu} = (P^* - C)\left(f(\frac{P^* - \mu}{\sigma})\right)\frac{1}{\sigma} = (1 - F(\frac{P^* - \mu}{\sigma})). \tag{10}$$

Invoking the properties of the cumulative distribution function we obtain that $0 < \frac{d\Pi^*}{d\mu} < 1$. Moreover we have $\frac{d^2\Pi^*}{d\mu^2} > 0$. Hence Π^* is an increasing, convex function of μ.

We also obtain:

$$\frac{d\Pi^*}{d\sigma} = (\frac{P^* - C}{\sigma})(\frac{P^* - \mu}{\sigma})f(\frac{P^* - \mu}{\sigma}) = (\frac{P^* - \mu}{\sigma})(1 - F(\frac{P^* - \mu}{\sigma})). \tag{11}$$

It therefore follows that Π^* is increasing in σ for $P^* > \mu$ and decreasing for $P^* < \mu$.

We next examine the consumer surplus, CS, given by the sum of the valuations which exceed the price[1]:

$$CS = \int_P^\infty (x - P)f(x)dx = CS = \sigma f(\tfrac{P-\mu}{\sigma}) - (P - \mu)\left(1 - F(\tfrac{P-\mu}{\sigma})\right)$$

This implies that, $CS\mid_{P=P^*} = \sigma f(\frac{P^*-\mu}{\sigma})\left(1 - \frac{(P^*-\mu)(P^*-C)}{\sigma^2}\right)$. This means that the consumer surplus is strictly positive unless $\sigma = 0$. Hence, the existence of dispersion in demand gives the consumers the opportunity to capture some of the welfare in society, while in the case where $\sigma = 0$ all consumers' valuations are the same and nobody extracts a surplus. Below we examine under what circumstances $CS\mid_{P=P^*}$ is an increasing function of dispersion.

Note that $\frac{\partial CS}{\partial P} = -\left(1 - F(\frac{P-\mu}{\sigma})\right)$. Therefore the consumer surplus is a strictly decreasing function of price. It is instructive to split between the direct and indirect effect:

$$\frac{dCS}{d\sigma}\mid_{P=P^*} = f(\frac{P^* - \mu}{\sigma}) + \frac{\partial P^*}{\partial\sigma}\frac{\partial CS}{\partial P^*}\mid_{P=P^*}, \tag{12}$$

[1] For a detailed analysis see Dassiou & Glycopantis (2011) where derivation of formulae and proofs are obtained in the text and in a mathematical appendix.

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

For $P^* < \mu$, the consumer surplus is an increasing function of σ always, as both effects are positive unless $P^* > \frac{\mu+C.}{2}$. This implies that the share of the consumer surplus in total surplus is larger the smaller dispersion is. As we show in the next section, one of the reasons (albeit not the only one) that may drive the firm into using pure bundling with flexible weights is to manipulate dispersion.

On the other hand, if $P^* > \mu$, then the indirect effect is now negative and hence the overall impact depends on whether the direct or indirect effect dominates.

Finally by adding together the producer and consumer surplus we have that:

$$W = \sigma f(\frac{P-\mu}{\sigma}) + (\mu - C)\left(1 - F(\frac{P-\mu}{\sigma})\right). \tag{13}$$

Straightforward calculations show that:

$$\frac{\partial W}{\partial P}\mid_{P=P^*} = -\frac{P^*-C}{\sigma} f(\frac{P-\mu}{\sigma}) < 0, \tag{14}$$

$$\frac{dW}{d\mu}\mid_{P=P^*} = [1 - F(\frac{P^*-\mu}{\sigma})](2 - \frac{\partial P^*}{\partial \mu})]. \tag{15}$$

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

$$\frac{dCS^*}{d\mu} = [1 - F(\frac{P^*-\mu}{\sigma})](1 - \frac{\partial P^*}{\partial \mu}). \tag{16}$$

Given (8), the above two relations imply that both W^* *and* CS^* *are increasing functions of the mean valuation.*

Furthermore we can obtain:

$$\frac{dW}{d\sigma}\mid_{P=P^*} = f\left(\frac{P^*-\mu}{\sigma}\right)\left(\frac{(P^*-C)^2 - 2\sigma^2}{(P^*-C)(P^*-\mu) - 2\sigma^2}\right). \tag{17}$$

The above means that if the profit maximising price is within $\left(C - \sqrt{2}\sigma, C + \sqrt{2}\sigma\right)$, welfare is an increasing function of dispersion. Using the normal distribution frequency tables we can conclude that for $P^* \in (\mu - 0.147\sigma, \mu + 0.147\sigma)$ welfare is a decreasing function of dispersion. We focus on the range of values for which $P^* < \mu$.

For P^* which is is less than μ and greater than the critical value $\mu - 0.147\sigma$, we have that $\frac{d\Pi^*}{d\sigma} < 0$ and $\frac{dW^*}{d\sigma} < 0$. As both profits and welfare are decreasing functions of dispersion, its reduction is beneficial to society as a whole as it also increases welfare. This means that a reduction in dispersion will, in this case, be beneficial to society as a whole. However this range of values for P^* corresponds to just 7% of the total range of values that the optimal price can take.

If P^* falls below this critical value 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 can damage both consumer surplus as well as overall welfare.[2]

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

3 Two stage pure bundling and comparisons with previous formulations

3.1 The framework of the analysis

The setting of optimal weights by a risk averse firm in constructing a pure bundle, and the willingness to sacrifice profits in the process of reducing risk can be better understood in the context of two articles that have inspired our approach.

Eckel and Smith (1992) focus their analysis on the cost side of things. The firm manipulates the dispersion of a convex cost function using a portfolio optimisation approach. By setting prices for the different demand groups that determine their contribution in total demand, expected outputs are affected and through them expected costs.

Anam and Chiang (2006) look at a risk averse monopolist. The firm faces two different markets with stochastic and correlated demands. They note that the ability to price discriminate if combined with risk aversion may lead to unconventional results. In its drive to reduce dispersion the firm may decide to price the good with the more elastic demand but also more risk a higher price. Hence the conventional direction of third price discrimination may be reversed if the sacrifice in expected profit is more than compensated for by the corresponding decrease in profit risk exposure.

We employ a portfolio optimisation approach; however unlike Eckel and Smith, this is not mean preserving. We show the existence of risk aversion may mean that the bundle weights chosen by the monopolist may not only lead to a change in the bundle mean, but also to an increase in the bundle dispersion. In this case, the conventional result of an equal weights bundling firm where pure bundling is used as an instrument of reduction in the dispersion of valuations is reversed. This is analogous to the Anam and Chiang's possibility of a reversion of the conventional result in third degree price discrimination.

As we discuss below the motives of altering the mean and the dispersion can be both in the same direction or in opposite directions. The consumer and welfare implications in the latter case will depend on which of the two effects dominates over the other. The existence of risk aversion from the side of the firm may ameliorate or even reverse competition policy concerns that

[2] Of course when demand is strictly convex ($P^* > \mu$), the dispersion increasing behaviour of a profit maximinsing firm will be beneficial to society as well for as long as $P^* > \mu + 0.147\sigma$.

bundling may be damaging to the interests of consumers; in fact in some instances the bundle dispersion in the valuations of the consumers may be end up to be super-additive (e.g. larger than that of the separate sales, $\sigma_1 + \sigma_2$) rather than sub-additive.

We analyse briefly below 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.

3.2 Step one: Optimal weights in fixed bundling

We start our analysis with a packaged good, a unit of which is defined. In a way analogous to the previous section, each consumer if he/she purchases the package has a one unit demand for it. A unit of the composite good consists of a given total number of individual units of Good 1 and Good 2. These goods are measured in similar units, for example their weights. Although throughout we are concerned with one unit of the packaged good, its composition may vary. Costs, means and standard deviations are all 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. Normalising to an overall sum of 2 we obtain $k = 2\lambda = 0.67$ and $2 - k = 2(1 - \lambda) = 1.33$ respectively[3]. The decision that each consumer then has to make is whether to buy this package, i.e. the 900 minutes with the weights and price offered by the monopolist.

A number of examples one can think of this type of bundling belong to the family of goods that can be digitized, i.e. information goods, where marginal costs are low and constant. We draw our inspiration from the paper by Crampes and Hollander (2005) and that by Crawford, which explain that TV bundles may be composed by combining movies and sport channels.

Optimisation is a two stage approach. In the first step when determining the optimal composition in the bundle we assume risk aversion from the point of view of the monopolist. Once the optimal bundle composition has

[3] It must be stressed that the use of k is only for computational convenience in the calculations that follow below, and in no way implies 2 units of the composite good.

been set we then proceed to the second step of the calculation of optimal profits. Hence risk aversion is incorporated 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.

The firm's objective before it sets the bundle price P_B, is to maximise its expected utility, given the bivariate normal distribution $f(x, y)$[4] of the buyers' reservation prices (valuations). The firm decides whether it is worthwhile to bundle and if so at what relative weights. In the second stage the firm calculates the bundle price[5]. As we have mentioned, by manipulating the weights we can affect dispersion and the level of demand.

The marginal densities are given by

$$\begin{aligned} f_X(x) &= \tfrac{1}{\sqrt{2\pi}\sigma_x}\exp\left(-\tfrac{(x-\mu_x)^2}{2\sigma_x^2}\right), \\ f_Y(y) &= \tfrac{1}{\sqrt{2\pi}\sigma_y}\exp\left(-\tfrac{(y-\mu_y)^2}{2\sigma_y^2}\right) \end{aligned}$$

The valuations x and y for Goods 1 and 2 can be in money form, and so will be the valuation of the bundle $xk + y(2 - k)$. The firm will first set weights to the two goods to maximise its utility function. The utility function 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. Hence,

[4] The underlying distribution for the buyers' valuations for the two goods x (Good 1 valuation) and y (Good 2 valuation) is

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

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

[5] If P_B is the bundle price then the demand function can be written as:

$$\int_{-\infty}^{\infty}\int_{P_B-y}^{\infty} f(x, y)dxdy$$

$$maxJ = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} -\exp[-\alpha\,(xk + y(2-k))]f(x,y)dxdy^{6\ ,7}$$

This means that we wish to maximise J, the utility function of the firm through a choice of the weights in the bundle. Using the moment generating function, $M_{X,Y}(t_1,t_2) = \exp[(\mu_x t_1 + \mu_y(2-t_2) + \frac{\sigma_x^2 t_1^2 + 2\rho\sigma_x\sigma_y t_1 t_2 + \sigma_y^2 t_2^2}{2}]$, with $t_1 = -\alpha k$ and $t_2 = -\alpha(2-k)$ we can re-write the optimisation function as:

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

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

$$maxK = \mu_1 k + \mu_2(2-k) - \alpha\frac{\sigma_1^2 k^2 + 2\rho\sigma_1\sigma_2 k(2-k) + \sigma_2^2(2-k)^2}{2}.$$

In other words the firm wishes to create an optimal bundle, where the weights of the two participating goods are such that the expected utility from the bundle is maximised given the valuations by the consumers and the production costs. In determining the bundle the firm will also take into account the costs of producing the good. We therefore adjust the means so that they are set as net from their corresponding costs. The firm maximises with respect to λ the following revenue certainty equivalent function which accounts for the loss of utility that the firm experiences given its risk aversion:

$$(\mu_1 - C_1)\lambda + (\mu_2 - C_2)(1-\lambda) - \alpha\lambda^2\sigma_1^2 - 2\alpha\rho\sigma_1\sigma_2\lambda(1-\lambda) - \alpha(1-\lambda)^2\sigma_2^2. \quad (18)$$

Theorem 2 *For $\mu_i - C_i > \mu_j - C_j$, a positive weight for good j requires that $\mu_i - \mu_j - (C_i - C_j) < 2\alpha\sigma_i(\sigma i - \rho\sigma_j)$. Hence the decision of whether to pure bundle or not, as well as how to balance the goods within the bundle,* ***are both endogenous decisions.***

[6] The 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

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

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 weights will be either $(0,2)$ or $(2,0)$.

[7] We show in Dassiou & Glycopantis (2011) that if the utility function is multiplicative and the correlation coefficient between the valuations of the two goods is equal to zero, the combination $k^* = 2 - k^* = 1$ is optimal. This corresponds to Schmalensee's type of conventional bundling.

Proof By first order and second order differentiation of the above expression with respect to λ, we derive the F.O.C. and the S.O.C. The S.O.C. is satisfied for $\alpha > 0$.[8] From the FOC we obtain the optimal split of the two unit bundle between the two goods as set out below

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

and

$$2 - k^* = 2(1 - \lambda) = \frac{(\mu_1 - \mu_2) - (C_1 - C_2) - 2\alpha\sigma_1(\sigma_1 - \rho\sigma_2)}{-\alpha(\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2)}. \tag{19}$$

The denominator in the above ratios is the S.O.C. , and hence negative. Footnote 6 derived previously explains why corner solutions will apply in the case where $\alpha = 0$. It follows that 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. □

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 that 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 inspection of the equalities in (19) and assuming that the weights are strictly positive 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 one unit of each good pure bundling 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 ($k^* < 1$).

(i) and (iii) set that the good with the higher net mean in consumer valuations will have a higher (lower) weight if the difference in the net means is more (less) than the difference in the variances multiplied by the risk aversion parameter.

In case (ii) the good with the higher net mean valuation has also a higher dispersion and the difference in the net means is exactly offset by the differences in the variances multiplied by the risk aversion parameter. As a result each good will receive an equal weight. This means that equal relative weights is a special case of our approach.

[8] This translates into $P < \mu$. If $A > 0$ ($\alpha < 0$) then there can only be corner solutions where goods are offered separately. This optimality of no bundling when the demand function is convex is also confirmed by Schmalensee who concludes that pure bundling is less profitable than unbundled sales for $P > \mu$.

Clearly case (i) will always be satisfied if $\sigma_i < \sigma_j$, as then the LHS of the inequality is negative. 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. As shown in (i) if the difference in the net means is sufficiently large 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.

Combining Theorem 2 and (i) we obtain the following result. For an optimal bundling decision such that the good with the higher net mean is given the higher, but strictly less than one 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 more easily (always) satisfied in the case of a low (negative) 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.

From the expression for k^* we calculate the derivatives with respect to these various parameters and we obtain the comparative static results.

As $\frac{dk^*}{d\alpha} = \frac{(\mu_2-\mu_1)-(C_2-C_1)}{\alpha^2(\sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2)}$, the weight of the good with the higher net mean (e.g. $k^* > 1$), is a decreasing function of α, the degree of absolute risk aversion. As implied by the derivative, the higher the degree of risk aversion, the lower will be the absolute value of its impact on that weight. In other words, risk aversion encourages bundling by enhancing the contribution of the weaker good (the one with the lower net mean) into the bundle all other things being equal.

Moreover, it is straightforward to show that if $k^* > 1$, $\frac{dk^*}{d\rho} > 0$, while if $k^* < 1$, $\frac{dk^*}{d\rho} < 0$. In other words, the share of the good with the higher weight in the bundle is an increasing function of the correlation coefficient, ρ, while the share of the other good is a decreasing function. A lower ρ boosts the share of the good with the lower weight in the bundle and improves its odds of having a non zero participation in the bundle. Therefore low positive correlation values promote bundling, and even more so negative values of ρ.

As $\frac{dk^*}{d(\sigma_1)} = \frac{-\rho\sigma_2(2-k^*)-2k^*\sigma_1}{(\sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2)} < 0$ and $\frac{dk^*}{d(\sigma_2)} = \frac{(2\sigma_2-\rho\sigma_1)(2-k^*)}{(\sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2)} > 0$, the weight of each good is inversely (directly) related to its own (the other good's dispersion). If the dispersion in the valuations for the other good is so small that $2\sigma_2 - \rho\sigma_1 < 0$, then this implies that $\sigma_2 - \rho\sigma_1 < 0$. It follows that this may imply that the share of good 1 in the bundle could be less than zero if the mean of good 2 is greater than the mean of good 1. In such a case $k^* = 0$.

Regarding the mean $\frac{dk^*}{d(\mu_1)} = \frac{1}{\alpha(\sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2)} > 0$. Hence, the optimal weight of each good in the bundle is directly related to its mean and, as can be easily shown, inversely related to the mean of the other good in the bundle. Also $\frac{dk^*}{d(C_1)} = \frac{1}{-\alpha(\sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2)} < 0$, i.e. 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.

3.3 Step two: Optimal bundle 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 good at the specific price as a take it or leave it option (pure bundling). As this composite good consists of Goods 1 and 2, its valuation will be the weighted sum of the valuations x and y and it will itself be normally distributed.

We define the profit function of the composite good, using our profit function as defined in relation (3) of Section 2. As Π^*_{B,k^*} is the optimal profit when the weights are set as $k = k^*$ and $2 - k = 2 - k^*$, i.e.

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

Equivalently $\Pi^*_{B,k=2-k}$ is optimal profit when weights are set as $k = 2 - k = 1$, e.g. in the Schmalensee model. Π^*_{B,k^*} is different to $\Pi^*_{B,k=2-k}$ unless $\mu_i - \mu_j - (C_i - C_j) = \alpha(\sigma_1^2 - \sigma_2^2)$.

The mean, cost and dispersion of the composite good are defined as:

$$\begin{aligned} \mu_{B,k^*} &= k^*\mu_1 + (2-k^*)\mu_2, \\ C_{B,k^*} &= k^*C_1 + (2-k^*)C_2, \\ \sigma_{B,k^*} &= \sqrt{(k^*)^2\sigma_1^2 + (2-k^*)^2\sigma_2^2 + 2k^*(2-k^*)\theta_{k^*}(1-\theta_{k^*})} \Leftrightarrow \\ \sigma_{B,k^*} &= (k^*\sigma_1 + (2-k^*)\sigma_2)\sqrt{1 - 2(1-\rho)\theta_{k^*}(1-\theta_{k^*})} \end{aligned}$$

where ρ is the correlation coefficient of the joint reservation distribution and $\theta_{k^*} = \frac{k^*\sigma_1}{k^*\sigma_1+(2-k^*)\sigma_2}$.

We note the case where the good i whose net mean and dispersion are both relatively larger also receives a higher weight in the bundle if $(\sigma_i^2 - \sigma_j^2) < \mu_i - \mu_j - (C_i - C_j)$. Setting as $\delta_{k^*} = \sqrt{1 - 2(1-\rho)\theta_{k^*}(1-\theta_{k^*})}$ $(0 \leq \delta_{k^*} \leq 1)$, this means that in this case $k^*\sigma_1 + (2-k^*)\sigma_2 > \sigma_1 + \sigma_2$ and $\delta_{k^*} \geq \delta_{k=2-k}$. Consequently, $\sigma_{B,k^*} > \sigma_{k=2-k}$. This case is further discussed as Case 3 in Section 3.4 and it is of particular interest as it is possible that σ_{B,k^*} 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.

More generally, the above results imply that optimisation alters the bundle dispersion in two ways: by affecting δ_{k^*} as well as by affecting the weighted sum of the two dispersions. θ_{k^*} is in our model a function of ρ; by differentiating with respect to ρ we find that θ_{k^*} is an 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 $\theta_{k^*} = \frac{\sigma_2 - \rho\sigma_1}{(1-\rho)(\sigma_1+\sigma_2)}$, which for as long as $\sigma_2 > \rho\sigma_1$ guarantees a positive weight for good 1.

Using relation (10) in Section 2, it is easy to calculate in an analogous manner that $\frac{\partial \Pi_B^*}{\partial(\mu_{B,k^*} - C_{B,k^*})} = 2\left[1 - F(\frac{P^*_{B,k^*} - \mu_{B,k^*}}{\sigma_{B,k^*}})\right]$, which means that:

$$0 < \frac{\partial \Pi^*_{B,k^*}}{\partial(\mu_{B,k^*} - C_{B,k^*})} < 2. \tag{20}$$

The derivative of the optimal profit with respect to the variance of the bundle is:

$$\frac{\partial \Pi^*_{B,k^*}}{\partial \sigma^2_{B,k^*}} = \frac{1}{2\sigma_{B,k^*}} \frac{\partial \Pi^*_{B,k^*}}{\partial \sigma_{B,k^*}} = \left(\frac{P^*_{B,k^*} - \mu_{B,k^*}}{2\left(\sigma_{B,k^*}\right)^2}\right)\left[1 - F(\frac{P^*_{B,k^*} - \mu_{B,k^*}}{\sigma_{B,k^*}})\right]. \tag{21}$$

Corresponding range of values and expressions can be derived for $\frac{\partial W^*_{B,k^*}}{\partial(\mu_{B,k^*} - C_{B,k^*})}$ $(0 < \frac{\partial W^*_{B,k^*}}{\partial(\mu_{B,k^*} - C_{B,k^*})} < 3)$ and $\frac{\partial W^*_{B,k^*}}{\partial \sigma^2_{B,k^*}}$ respectively.[9] These will be of relevance in the discussion below.

3.4 Welfare and policy implications of two stage bundling in relation to conventional pure bundling

In order to analyse the consumer and welfare implications of our pure bundling and make comparisons with the conventional equal weights bundling found in Schmalensee's model we consider three cases.

Using the small increments formula we can define an approximate equality for the difference between the welfare in our model minus the welfare in the conventional equal weights bundling as $\Delta W^*_{B,k^*}$:

$$\Delta W^*_{B,k^*} \cong \frac{\partial W^*_{B,k^*}}{\partial(\mu_{B,k^*} - C_{B,k^*})}\Delta_{nm} + \frac{\partial W^*_{B,k^*}}{\partial \sigma^2_{B,k^*}}\Delta_{var}. \tag{22}$$

We set as $\Delta_{nm} = \mu_{B,k^*} - C_{B,k^*} - (\mu_{B,k=2-k} - C_{B,k=2-k})$ the difference in the net bundle means of our model minus that of Schmalensee's pure bundling. We also define $\Delta_{var} = \sigma^2_{B,k^*} - \sigma^2_{B,k=2-k}$, which is the difference in the variance of the bundle in our model minus the bundle variance in the 1:1 bundling model.

Case 1 $\mu_i - C_i > \mu_j - C_j$ and $\sigma_i < \sigma_j$.

[9] The derivations of these can be found in the mathematical appendix of the Dassiou & Glycopantis 2011 paper as relations (50) and (51) respectively.

Then $k^* > (2-k^*)$, $\Delta_{nm} > 0$ and $\Delta_{var} < 0$. Since profit is an increasing function of the net mean and a decreasing function of the dispersion, we have that $\Pi^*_{B,k^*} \geqq \Pi^*_{B,k=2-k}$. The impact of using optimal bundling on consumer surplus is ambiguous. According to the findings from Section 2, consumer surplus is an increasing function of both the net mean as well as dispersion (as $P^*_{B,k^*} < \mu_{B,k^*}$).

Inspecting the impact on welfare, if $(P^*_{B,k^*} - C_{B,k^*})^2 - 2\sigma^2_{B,k^*} > 0$ then $\frac{\partial W^*_{B,k^*}}{\partial \sigma^2_{B,k^*}} < 0$ as derived from relation (17) in Section 2. A smaller bundle dispersion increases welfare, and this is further reinforced by a larger 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 $(P^*_{B,k^*} - C_{B,k^*})^2 - 2\sigma^2_{B,k^*} < 0$ [10], then $\frac{\partial W^*_{B,k^*}}{\partial \sigma^2_{B,k^*}} > 0$. A larger net mean would still increase welfare, but this will have to be compared to the reduction in welfare from a smaller bundle dispersion. Hence, in this case it is unclear whether the reduction in welfare is smaller or larger than the decrease in welfare induced in the pure bundling of Schmalensee. In other words, while dispersion is further reduced by our brand of pure bundling, the bundle net mean is increased. The latter can more or less than offset the negative impact on welfare of a larger reduction in the former. We attempt to bring some clarity to this ambiguity below.

It can be shown that:

$$\Delta_{var} = \frac{\Delta_{nm}}{\alpha} + \Delta_{nm} \frac{\sigma_2^2 - \sigma_1^2}{(\mu_2 - \mu_1) - (C_2 - C_1)} \frac{\sigma_2^2 + \sigma_1^2 + 2\rho\sigma_1\sigma_2(1-2\alpha)}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}. \tag{23}$$

As in Case 1 $\Delta_{var} < 0$ and $\Delta_{nm} > 0$, we have that $\Delta_{nm} < |\Delta_{var}|$ if and only if

$$\frac{(\sigma_2^2 - \sigma_1^2)}{(\mu_1 - \mu_2) - (C_1 - C_2)} \frac{\sigma_2^2 + \sigma_1^2 + 2\rho\sigma_1\sigma_2(1-2a)}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2} > 1 + \frac{1}{\alpha}. \tag{24}$$

Comparing the size of the rate of change of welfare with respect to the variance to that of the rate of change with respect to the net mean we establish (see Dassiou & Glycopantis 2011) that the absolute size of the former is smaller than the size of the latter, i.e. $\left|\frac{\partial W^*_{B,k^*}}{\partial \sigma^2_{B,k^*}}\right| < \frac{\partial W^*_{B,k^*}}{\partial(\mu_{B,k^*} - C_{B,k^*})}$. Hence, $\Delta W^*_{B,k^*} \geq 0$ if $\frac{\partial W^*_{B,k^*}}{\partial \sigma^2_{B,k^*}} \leq 0$, or if $\frac{\partial W^*_{B,k^*}}{\partial \sigma^2_{B,k^*}} > 0$, and $\Delta_{nm} \geq |\Delta_{var}|$. The change in welfare $\Delta W^*_{B,k^*}$ will still be non-negative even if $\frac{\partial W^*_{B,k^*}}{\partial \sigma^2_{B,k^*}} > 0$, and $\Delta_{nm} < |\Delta_{var}|$, unless the following proposition (proved in Dassiou & Glycopantis, 2011) holds:

[10] As $\frac{P^*_{B,k^*} - C_{B,k^*}}{\sigma_{B,k^*}} < \sqrt{2}$ implies that P^*_{B,k^*} is below the critical value $\mu_{B,k^*} - 0.147\sigma_{B,k^*}$.

Proposition 1 $\Delta W^*_{B,k^*} < 0$ *if and only if* $\frac{\partial W^*_{B,k^*}}{\partial \sigma^2_{B,k^*}} > 0, \Delta_{nm} < |\Delta_{var}|$ *and*
$$\frac{|\Delta_{var}|}{\Delta_{nm}} = \frac{\sigma^2_{B,k=2-k} - \sigma^2_{B,k^*}}{\mu_{B,k^*} - C_{B,k^*} - (\mu_{B,k=2-k} - C_{B,k=2-k})} > \frac{\frac{\partial W^*_{B,k^*}}{\partial(\mu_{B,k^*} - C_{B,k^*})}}{\frac{\partial W^*_{B,k^*}}{\partial \sigma^2_{B,k^*}}}.$$

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

Below we discuss Cases 2 and 3 in which there is a trade off as far as the firm is concerned. It faces a dilemma as the good with the larger net mean also has a larger dispersion. Therefore it has to choose between its desire to increase the former and decrease the latter in order to enhance its profits.

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

Then $k^* < (2 - k^*)$, $\Delta_{nm} < 0$ and $\Delta_{var} < 0$. Hence both the mean as well as the dispersion are smaller than that in conventional pure bundling. The inequality $\mu_i - \mu_j - (C_i - C_j) < \alpha(\sigma_i^2 - \sigma_j^2)$ means that in this case the overriding criterion for the company in setting the bundle weights is a decrease in dispersion rather than an increase in the bundle net mean. Hence, the optimal weights chosen by the firm will not only produce a smaller dispersion relative to that of conventional bundling, but also a smaller bundle mean. As a result, consumer surplus will be smaller than that in the equal weights bundling as it is an increasing function of both the mean as well as the dispersion. This means that Case 2 is a case where a competition authority is justified to intervene if its primary objective is the protection of consumer surplus.

We now turn our attention to the overall welfare implications. Given that both Δ_{var} and Δ_{nm} are negative, if $\frac{\partial W^*_{B,k^*}}{\partial \sigma^2_{B,k^*}} > 0$ then both terms in the small increments formula will be negative and welfare will unambiguously decrease. Only in the case where $\frac{\partial W^*_{B,k^*}}{\partial \sigma^2_{B,k^*}} < 0$ 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 that (Dassiou & Glycopantis, 2011):

Proposition 2 $\Delta W^*_{B,k^*} > 0$ *if and only if* $\frac{\partial W^*_{B,k^*}}{\partial \sigma^2_{B,k^*}} < 0$, $|\Delta_{nm}| < |\Delta_{var}|$ *and*
$$\frac{|\Delta_{var}|}{|\Delta_{nm}|} = \frac{\sigma^2_{B,k=2-k} - \sigma^2_{B,k^*}}{(\mu_{B,k=2-k} - C_{B,k=2-k}) - \left(\mu_{B,k^*} - C_{B,k^*}\right)} > \frac{\frac{\partial W^*_{B,k^*}}{\partial(\mu_{B,k^*} - C_{B,k^*})}}{\left|\frac{\partial W^*_{B,k^*}}{\partial \sigma^2_{B,k^*}}\right|}.$$

Case 3 $\mu_i - C_i > \mu_j - C_j$, $\sigma_i > \sigma_j$ and $\mu_i - \mu_j - (C_i - C_j) > \alpha(\sigma_i^2 - \sigma_j^2)$.

Then $k^* > (2-k^*)$, $\Delta_{nm} > 0$ and $\Delta_{var} > 0$. Here the overriding criterion for the company in setting the bundle weights is an increase in the bundle mean rather than a decrease in dispersion. Clearly, the assignment of a larger relative weight to the good with the larger dispersion in valuations will mean that consumer surplus will definitely be larger than that of the conventional case of equal relative weights pure bundling as the company constructs a composite good whose mean and dispersion are both larger. This will lead to a larger consumer surplus on both counts. This is one case where the policy makers should view bundling in variable proportions favourably as such a strategy followed by the company is very likely to increase consumer and welfare surplus as compared to the case where pure bundling is restricted to fixed proportions.[11] This stance is further strengthened if the weight assigned to the good with the relatively larger mean and variance is such that σ_{B,k^*} is even be larger than $\sigma_1 + \sigma_2$.

3.5 Concluding remarks

In this paper we have analysed a pure bundling model which reflects economic reality and decisions better that the conventional assumption made in the literature. In our discussion we break away from the usual approach which assumes that the two goods form a bundle of a fixed 1:1 ratio. Instead we obtain the optimal participation of the goods in the bundle through the maximisation of the utility function. This captures through an absolute risk aversion parameter the attitude of the firm as regards demand valuation dispersion.

While the corresponding bundle price is obtained through maximising profits, our model has the distinct characteristic that the firm is not simply interested in profit, but also in taking only an acceptable level of risk as defined through the optimal bundle composition.

We proceed to consider the consumer surplus and welfare implication of our bundling approach. We distinguish between three different cases characterized by the relation between the net means and dispersions of the implied marginal distributions for the two goods.

In Case 1, the inequalities in the differences in the net means and dispersions are in opposite directions and the company will end up with a smaller bundle dispersion that the one obtained in conventional bundling, while achieving a large profit through a larger bundle mean. If welfare depends

[11] The case of $\Delta W^*_{B,k^*} < 0$ is quite remote: First, we need P^*_{B,k^*} to exceed the critical value $\mu_{B,k^*} - 0.147\sigma_{B,k^*}$ for $\frac{\partial W^*_{B,k^*}}{\partial \sigma^2_{B,k^*}} < 0$ - by inspection of the normal distribution tables there is only a 7% probability of this happening. Additionally, we need $\Delta_{nm} < \Delta_{var}$ and $\frac{\Delta_{var}}{\Delta_{nm}} = \frac{\sigma^2_{B,k^*} - \sigma^2_{B,k=2-k}}{\mu_{B,k^*} - C_{B,k^*} - (\mu_{B,k=2-k} - C_{B,k=2-k})} > \frac{\frac{\partial W^*_{B,k^*}}{\partial(\mu_{B,k^*} - C_{B,k^*})}}{\left|\frac{\partial W^*_{B,k^*}}{\partial \sigma^2_{B,k^*}}\right|}$ for $\Delta W^*_{B,k^*} < 0$.

positively on dispersion it is possible that optimal bundling will have an adverse effect on welfare, which is however tempered by the fact that impact of a larger bundle mean. Unless the dispersion of our bundling method is so substantially smaller than that of the conventional bundling that the impact on the net means is not sufficient to offset the damage this will do, our flexible weights bundling will cause less welfare damage than conventional bundling does.

Case 2 is characterized by the fact that (i) the differences in the net means and dispersions are in the same direction and (ii) the difference in the net means is smaller than that in the variances multiplied by the parameter of risk aversion. The firm will have bundle net mean and variance smaller than the ones derived in conventional bundling as it will attach a lower weight to the good with the higher dispersion and larger net mean. This affects adversely both the consumer surplus and welfare to a larger extent than conventional fixed weights bundling does.

Case 3 is characterized by the fact that (i) the differences in the net means and dispersions are in the same direction and (ii) the difference in the net means is larger than that in the variances multiplied by the parameter of risk aversion. The firm will construct a package with a larger variance in tastes by attaching a larger weight to the good with the relatively larger dispersion. The composite good with also have a larger net mean. This will have a beneficial effect in both consumer surplus and welfare relative to that of fixed equal weights.

Finally we note that the existence of risk aversion makes it possible to have cases of optimal pure bundling where there is no conflict between the bundling decisions of the firm and its effects on society. If the optimal weights are such that the reduction in bundle dispersion is substantial we may end up with $\sigma_{B,k^*} > \sigma_1 + \sigma_2$. This means that optimal pure bundling will in this case lead to an increase rather than a decrease in the dispersion not only relative to conventional bundling, but also in relation to the case of separate selling of the two goods.

References

Adams, W.J.and J.L. Yellen (1976), 'What makes advertising profitable?', *Economic Journal,* **87**, 427-449.

Anam, M. and S.-H. Chiang (2006), 'Price discrimination and social welfare with correlated demand', *Journal of Economic Behavior & Organization,* **61**, 110-122.

Chu, S., P. Leslie and A. Sorensen (2011), 'Bundle-size pricing as an approximation to mixed bundling', *American Economic Review*, forthcoming.

Crampes C. and A. Hollander (2005), 'Product specification, multi-product screening and bundling: the case of pay TV', *Information Economics and Policy,* **17** (1)**,** 35-59.

Crawford, G.S. (2008) 'The discriminatory incentives to bundle in the cable television industry', *Quantitative Marketing and Economics,* **6** (1), 41-78.

Dana, J.D. (1999a), 'Equilibrium price dispersion under demand uncertainty', *Rand Journal of Economics,* **30** (4), 632-660.

Dana, J.D. (1999b), 'Using yield management to shift demand when the peak time is unknown', *Rand Journal of Economics,* **30** (3), 456-474.

Dana, J.D. (1998), 'Advance-purchase discounts and price discrimination in competitive markets', *Journal of Political Economy,* **106** (2), 395-422.

DeGraba, P. (2005) 'Quantity discounts from risk averse sellers', Federal Trade Commission - Bureau of Economics, Working Paper 276.

Dassiou, X. and D. Glycopantis (2011) 'Optimal pure bundling under Gaussian demand', mimeo.

Dassiou, X. and D. Glycopantis (2008), 'Price discrimination through transactions bundling: The case of monopsony', *Journal of Mathematical Economics,* **44** (7-8)**,** 672-681.

Dassiou, X. and D. Glycopantis (2006), 'The economic theory of price discrimination via transactions bundling: An assessment of the policy implications', *Review of Law & Economics*, Berkeley Electronic Press **2** (2).

Eckel, C.E.and W.T. Smith (1992), 'Price discrimination with correlated demands', *Southern Economic Journal,* **59** (1), 58-65.

Gaggero, A.A. and C.A. Piga, (2009), 'Airline market power and intertemporal price discrimination', Discussion Paper Series, Department of Economics, Loughborough University.

Kreps, D.M. (1990), *A Course in Microeconomic Theory*, Harvester-Wheatsheaf, UK.

McAfee, R.P., J. McMillan and M. Whinston (1989), 'Multiproduct monopoly, commodity bundling and correlation of values', *Quarterly Journal of Economics,* **103**, 371-383.

Schmalensee, R. (1984), 'Gaussian demand and commodity bundling', *The Journal of Business,* **57** (1), part 2, S211-S230.

Stigler, G. (1968), 'A note on block booking', in Stigler, G. (ed.), *The Organisation of Industry*, Richard D. Irwin.