Do dividends signal more earnings?

Aloisio Araujo

Instituto Nacional de Matemática Pura e Aplicada (IMPA)

and Escola de Pós-Graduação em Economia

Fundação Getulio Vargas (EPGE/FGV)

Humberto Moreira

Escola de Pós-Graduação em Economia

Fundação Getulio Vargas (EPGE/FGV)

Marcos H. Tsuchida

Escola de Economia de São Paulo

Fundação Getulio Vargas (EESP/FGV)

Address correspondence and requests for reprints to Humberto Moreira, Escola de Pós-Graduação em Economia, Fundação Getulio Vargas, Praia de Botafogo, 190, 11º andar, Rio de Janeiro, RJ, Brazil, CEP 22250-900; telephone number: +55-21-2559-5837; e-mail address: humberto@fgv.br

Do dividends signal more earnings?

Abstract

Signaling models contributed to the corporate finance literature by formalizing "the informational content of dividends" hypothesis. However, these models are under criticism as the empirical literature found weak evidences supporting a central prediction: the positive relationship between changes in dividends and changes in earnings. We claim that the failure to verify this prediction does not invalidate the signaling approach. The models developed up to now assume or derive utility functions with the single-crossing property. We show that, in the absence of this property, signaling is possible, and changes in dividends and changes in earnings can be positively or negatively related.

The information content of dividends is a controversial issue in corporate finance. The research started when Miller and Modigliani (1961) suggested that managers use dividend policy to convey their expectations of future prospects of the firm. With this hypothesis they proposed to explain the effect of dividend changes on the prices of shares. Since then, theoretical and empirical research advanced. Signaling models were the main tool that formalized the original intuition. Bhattacharya (1979), Miller and Rock (1985), and John and Williams (1985) were the initiators of a long list of signaling models.[1] The basic idea is that firm managers possess private information about future earnings and they like to convey it to the market. However, they cannot simply announce their expectations of future earnings publicly because every firm could imitate them. The information is conveyed by a costly signal. In the cited models, the respective costs are: financing of a committed level of dividend, suboptimal investment and tax on dividends.

On the empirical side, researchers tries to verify the testable implications derived from the models. In all these models the single-crossing property holds. As a consequence, the models predict that dividends, market price and future or current earnings are positively related. The correlation between dividend and returns was a strongly established result even before the signaling models have appeared. Aharony and Swary (1980) show that announcements of dividend increase or decrease result in, respectively, positive or negative abnormal returns.

The controversy lays on the relationship between dividends and subsequent earnings. Watts (1973) analysis found the positive relation, however, the effect was very small and not conclusive. Healy and Palepu (1988) found a significant relation, but they focused in the particular situation of initiation and omission of dividend payment. Exploring a larger data set Benartzi, Michaely and Thaler (1997) found no significant relation between dividends and future earnings and concluded that dividends are more related to past and present earnings. More recently, Nissim and Ziv (2001) using an improved measure of future earnings concluded that dividends matters for earnings prediction. Many other works contributed to this debate and a definitive conclusion seems far to be reached.

We claim that the lack of a clear relation between dividends and earnings is not incompatible with information content of dividends. Common to all the previous models of signaling is the existence of single-crossing in the objective function of manager, i.e., the marginal cost of signaling is monotonic in the type of firms. This property generates the monotonic relationship between dividends and earnings. In this work, we drop this assumption and develop a model employing the techniques presented in Araujo and Moreira (2001). In non-single-crossing signaling, a new kind of equilibrium may exist. In this equilibrium the relation between firm current earnings and dividends is U-shaped, that is, it is decreasing for low-earning firms and increasing for high-earning firms. Consequently, two types of firms signal with the same level of dividend, and, as they are indistinguishable by the market, this situation can be interpreted as low types pretending to be high types. The market value of shares is an average of the values of two types and increases with dividend. Finally, we found that, compared to the traditional separating equilibrium, dividends in U-shaped schedule are a weaker signal for future earnings, which can explain why the statistical testing has not succeeded in detecting signaling in dividends. The results suggest that the empirical research testing for information asymmetry should investigate the distribution of the signal variable and the relationship between signal and variance of the types, in order to detect U-shaped signaling.

The model is presented in section 1 and it is similar to the model of Miller and Rock (1985). For concreteness, we assume a quadratic production function, in section 2. Computations of equilibrium are performed and results are presented in this section. We comment the connections with the empirical literature in section 3. The conclusions are presented in section 4.

1 The model

This model builds on Miller and Rock (1985). There is a firm with production function $F(\cdot)$. The usual properties, $F'(\cdot) > 0$ and $F''(\cdot) < 0$, are assumed. Let X and Y be the earnings, respectively, at period 1 and 2. At the beginning of period 1, managers know X and announce dividends, D. Then shareholders may sell their shares, dividends are distributed and $X - D$ are invested. At the end of period 1, the firm production is subject to a multiplicative shock, $\delta > 0$, so that

$$Y = \delta F(X - D),$$

where $0 \leq D \leq X$. At period 2 the earnings are distributed and the firm is disassembled. We assume the firm cannot issue debt and investments have to be financed exclusively by internal resources. The information asymmetry is on the knowledge of X, which we assume randomly distributed on $[X_1, X_2]$, with density $p(X)$. Consistent with the terminology used in theory of contracts, we refer to X as the *type* of the firm. At period 1, managers know X before the dividend announcement, but the market does not, and managers cannot credibly convey their private information to the market. The shock δ is unknown to both manager and market, but may be correlated with X.

Both X and δ affect the value of firm. Now suppose the manager has an estimate of δ based on private information in period 1. If managers are interested on market value of firm, they would like to sign X and δ when the implied value is high. The signaling of X is clear. If a firm pays more dividends, it incurs in increasing costs due to underinvestment. Since higher type firms have higher earnings, their sacrifice in output is lower for the same level of dividend. The decreasing cost of dividend allows signaling in a way that higher types signal with higher dividends. Also, they would like to reveal δ to the market but, in this case, the signaling is not possible. They cannot signal high productivity distributing more dividends, as high δ firms has higher marginal cost of dividend because optimal investment level is higher and sacrifice in output is higher. On the other side, a high δ firm cannot signal higher investment opportunity paying less dividends because low δ firms could imitate paying low dividends. However, we will show that if X and δ are correlated, the shock in productivity may produce a signaling in which, for a subset of firms, dividend choice is decreasing with respect of type.

Assumption 1 *The shocks are correlated,*

$$\mathrm{E}[\delta|X] = \varepsilon(X) > 0.$$

1.1 The value of the firm

At period 1, managers estimate the fundamental cum-dividend value of the firm as the present value of dividend flow:

$$\begin{aligned} V(X, D) &= D + \frac{1}{1+i}\,\mathrm{E}[\delta F(X-D)|X] \\ &= D + \frac{1}{1+i}\varepsilon(X)F(X-D). \end{aligned} \tag{1}$$

Under symmetric information, this would be the value of shares and the manager would choose the investment in order to maximize V. The first-best dividend level, D^*, would be given by the Kuhn-Tucker conditions:

$$F'(X-D^*) - \frac{1+i}{\varepsilon(X)} \begin{cases} = 0, & \text{if } D^* > 0, \\ \geq 0, & \text{if } D^* = 0. \end{cases} \tag{2}$$

But under asymmetric information, the market value may not coincide with V. Let V^m denote the market value. We will assume that V^m is determined as a signaling equilibrium, that is, firms signal to the market by the choice of dividend level and market estimates the value observing the dividend choice. The firms will choose dividend above the optimal level, paying an underinvestment cost for signaling.

Shareholders want to maximize V if they keep the share with them until period 2. The ones who intend to sell at period 1 prefer the maximization of the market value, V^m. As in Miller and Rock (1985), we assume the firm's managers are maximizing a welfare function that aggregates the interests of shareholders that desire to sell the shares and the ones who do not. Let $k \in (0, 1]$ be the fraction of shareholders that sell at period 1. This fraction is exogenous and can be motivated by necessity of liquidity by shareholders. The welfare function is

$$W(X, D, V^m) = kV^m + (1 - k)V(X, D).$$

For the purpose of signaling analysis, we are interested on marginal rate of substitution between V^m and D. Since W is quasi-linear with respect to V^m, all the properties are found in marginal welfare of D,[2]

$$W_D(X, D) = (1 - k)\left(1 - \frac{1}{1+i}\varepsilon(X)F'(X - D)\right).$$

The dependence of productivity shock on type may make W_D non-monotone on type. More precisely, higher X increases ε, but reduces F', for the same level of dividend.

In terms of the cross-derivative of W,

$$W_{XD}(X, D) = \frac{1-k}{1+i}\left[-\varepsilon(X)F''(X - D) - \varepsilon'(X)F'(X - D)\right], \tag{3}$$

may change its sign. We can define two regions in $X \times D$ plane, according to this sign.

Definition 1 *The CS+ region (resp. CS− region) is the set of points in $X \times D$ plane such that $W_{XD} > 0$ (resp. $W_{XD} < 0$).*

In equation (3), the first term in the brackets is the investment effect. Firms with higher earnings invest more and have lower marginal product. Consequently, the marginal cost of dividend is lower for higher types. The second term is the productivity effect. Firm earnings provide information about the future productivity, which affect the expected marginal cost of the signal.

Negative correlation between earnings and productivity

When $\varepsilon'(X) \leq 0$, higher earnings reduce the expected productivity and the cost of signaling is lower. Welfare function has the single-crossing property since both productivity and investment effects collaborate on $W_{XD} > 0$. The results are, therefore, similar to the ones found by Miller and Rock (1985).

Positive correlation between earnings and productivity

If $\varepsilon'(X) > 0$, higher earnings correspond to a higher optimal investment level and dividend becomes costlier since more earnings should be retained for investment. If, for some types and dividend level, productivity effect dominates, then $W_{XD} < 0$ and higher types will be more reluctant to pay dividends because of lost of investment opportunities. Conversely, $W_{XD} > 0$ holds when investment effect dominates. In this case, lower types are more reluctant to pay dividends because they have lesser investment resources. Note that from (2) the first-best dividend as a function of types, $D^*(X)$, is increasing for $W_{XD} > 0$ and decreasing for $W_{XD} < 0$. If investment effect dominates, firms with higher earnings may pay more dividends and, if productivity effect dominates, they should invest more paying less dividends. When $\varepsilon'(X) > 0$ is such that the signal of W_{XD} is ambiguous, the single-crossing propriety does not hold. The assumption on the constancy of sign (for instance in Riley (1979)) for cross-derivative of objective function is then violated and we need another approach developed in Araujo and Moreira (2001, 2003).[3]

1.2 The signaling equilibrium

As usual, the signaling equilibrium is a perfect Bayesian one (the formal definition is provided in appendix A.1). The basic description remains. The market generates a value function, $V^m(\cdot)$, and each type of firm, X, chooses a dividend level, D, that maximizes W. We have an equilibrium if zero expected profit condition holds, that is,

$$V^m(D) = \mathrm{E}_\mu[V(X, D)|D], \tag{4}$$

where E_μ denotes the expectation taken on the Bayesian updated distribution on X. The market value V^m should be the expected value of the firm with respect to the probability distribution of X, resulting from the Bayesian update given the choice of D by the firm.

Formally, the signaling problem consists in finding functions $\mathcal{V}^m(X)$ and $\mathcal{D}(X)$ such that the type

X firm chooses a dividend level $\mathcal{D}(X)$ and is evaluated as $\mathcal{V}^m(X)$ by the market. Since dividends and market value are linked by $V^m(\cdot)$, these functions are related by $\mathcal{V}^m(X) = V^m(\mathcal{D}(X))$.

Define the welfare of type X firm that declares to be type $\hat{X}$ as

$$\begin{aligned}\mathcal{W}(\hat{X}, X) &= W(X, \mathcal{D}(\hat{X}), V^m(\mathcal{D}(\hat{X}))) \\ &= kV^m(\mathcal{D}(\hat{X})) + (1-k)V(X, \mathcal{D}(\hat{X})).\end{aligned}$$

In order to be incentive compatible, each firm should prefer to tell the truth, that is

$$\mathcal{W}(X, X) \geq \mathcal{W}(\hat{X}, X), \tag{5}$$

for all $X, \hat{X} \in [X_1, X_2]$. A differential equation for $\mathcal{D}$ is derived from the first order condition

$$\frac{\partial \mathcal{W}}{\partial \hat{X}}(X, X) = 0. \tag{6}$$

It should be noted that the first order condition is not sufficient condition for implementability when the single-crossing property does not hold. Incentive compatibility should be checked globally after a candidate for equilibrium is obtained.

Additionally, the second order condition constrains $\mathcal{D}'(X)$:

Proposition 1 *In signaling equilibrium, $\mathcal{D}(X)$ is non-decreasing in $CS+$ region and non-increasing in $CS-$ region.*

Proof: See the appendix.

When single-crossing property is present, $CS+$ and $CS-$ do not show up simultaneously and contracts should be monotone. As a consequence, types are separated when $\mathcal{D}'(X) \neq 0$, or a interval of types is bunched when $\mathcal{D}'(X) = 0$. When single-crossing property does not hold, monotonicity is not assured and the relationship between type and signal may be, for example, U-shaped and a disconnected set of types may signal with the same dividend level.

1.3 Equilibria diversity

In a equilibrium, the same signal, D, may be chosen by many types. We are interested in classifying the equilibrium according to its degree of separability. The following definition will be useful:

Definition 2 *The pooling set, $\Theta(D)$, is the set of types whose signal is D, that is, $\Theta(D) = \{X \in [X_1, X_2] | \mathcal{D}(X) = D\}$.*

In particular, in a separating equilibrium, $\Theta(D)$ is singleton for every D that is chosen by a firm.

Definition 3 *The type X is separated if $\Theta(\mathcal{D}(X)) = \{X\}$. A separating equilibrium is a signaling equilibrium such that every X is separated.*

When X is separated, market correctly infers the type by the observation of D. So $V^m(D) = V(X, D)$, where X is the type that choose dividend level D.

Proposition 2 *In an interval of separated types, $\mathcal{D}(X)$ follows the differential equation*

$$\mathcal{D}'(X) = \frac{-kV_X(X, \mathcal{D}(X))}{V_D(X, \mathcal{D}(X))}. \tag{7}$$

Proof: See the appendix.

As in the single-crossing case, a pooling equilibrium may be characterized by a continuum of types that chooses the same signal level.

Definition 4 *The type X is continuously pooled, if there is a non-degenerate closed interval I, such that, $X \in I \subset \Theta(\mathcal{D}(X))$. A continuous pooling equilibrium is a signaling equilibrium such that for every X, $\Theta(\mathcal{D}(X)) = [X_1, X_2]$.*

In signaling games without the single-crossing condition, a new kind of pooling arises. As in the continuous pooling, some values of D will be chosen by more than one type of firm. However, the number of pooling types may be finite.

Definition 5 *The type X is discretely pooled, if X is an isolated point of $\Theta(\mathcal{D}(X))$ and $\Theta(\mathcal{D}(X)) \neq X$.*

The property aggregating the discretely pooled types is that they must have the same marginal welfare W_D.

Proposition 3 *If X_a and X_b are discretely pooled, $D = \mathcal{D}(X_a) = \mathcal{D}(X_b)$, $\mathcal{D}'(X_a) \neq 0$, and $\mathcal{D}'(X_b) \neq 0$, then*

$$W_D(X_a, D) = W_D(X_b, D). \tag{8}$$

Proof: See the appendix.

Equation (8) gives $\varepsilon(X_a)F'(X_a - D) = \varepsilon(X_b)F'(X_b - D)$. So different types can choose the same level of dividend when, for higher types, the higher productivity shock compensates the reduction in marginal productivity resulted from higher investment. In the discrete pooling, dividend choice does not fully reveal the type of the firm. The market knows the set of possible types but it cannot distinguish one type from the other. This fact is taken into account when the market estimates the value, so $\mathrm{E}_\mu[V(X, D)|D]$ is the average value of types in the pool.

Assumption 2 *The type X is uniformly distributed on the interval $[X_1, X_2]$.*

With assumption 2, each type has the same probability. In particular, when there are only two types in the pool, the expected value of firms is

$$\mathrm{E}_\mu[V(X, D)|D] = \frac{1}{2}V(X_a, D) + \frac{1}{2}V(X_b, D),$$

where X_a and X_b are the types that choose D.

Proposition 4 *Under assumption 2, in a interval with discretely pooled types, if exactly two types chooses the same dividend, $\mathcal{D}(X)$ follows the differential equation*

$$\mathcal{D}' = \frac{-k\left[V_X(X, \mathcal{D}) + V_X(\overline{X}(X, \mathcal{D}), \mathcal{D})\overline{X}_X(X, \mathcal{D})\right]}{kV_X(\overline{X}(X, \mathcal{D}), \mathcal{D})\overline{X}_D(X, \mathcal{D}) + 2V_D(X, \mathcal{D})}, \tag{9}$$

where $\overline{X}(X, D)$*, derived from (8), is the type pooled together with type* X*, when dividend* D *is chosen.*

Proof: See the appendix.

1.4 Equilibrium refinement

The disturbing fact in any signaling model is the existence of many equilibria. For the same parameters, different kinds of equilibrium may exist, and the choice of initial conditions may generate a continuum of equilibria. At this point a selection criterion is needed. The pro-separation criterion, defined below, chooses, among different kinds of equilibria, the one that minimizes pooling and maximizes efficiency.

Assumption 3 *Separability degree of a continuous pooled type, a discretely pooled type, and a separated type are, respectively, 1, 2, and 3.*

Definition 6 *Let* $\Pi(d) = \{X \in [X_1, X_2] | X \text{ has separability degree } d\}$*.*

Therefore $\Pi(1)$ is the set of continuously pooled types, $\Pi(2)$ is the set of discretely pooled types and $\Pi(3)$ is the set of separated types.

Definition 7 *The separation floor of a signaling equilibrium is the lowest separability degree associated to a type in* $[X_1, X_2]$*.*

Definition 8 *A pro-separation equilibrium is a signaling equilibrium with separating floor* φ*, such that (a) there is no other equilibrium with higher separation floor; (b) among equilibria with same separation floor, there is no other with lower probability of* $\Pi(\varphi)$*, according to density* $p(\cdot)$*; and (c) among equilibria with same separation floor and same probability of* $\Pi(\varphi)$*, there is no other equilibria with higher expected value, according to density* $p(\cdot)$*.*

Therefore, pro-separation equilibrium criterion chooses an equilibrium eliminating poorly separated equilibria and taking the most efficient among the surviving equilibria.

2 The quadratic case

For the computations we consider a quadratic production function

$$F(I) = aI(b - I), \tag{10}$$

where $0 \leq I \leq b/2$, $a > 0$, and $b > 0$. We assume a linear expected productivity shock

$$\varepsilon(X) = g + hX, \tag{11}$$

where $h > 0$.

The cross-derivative of welfare function of the firms is

$$W_{XD} = \frac{4ah(1-k)}{1+i}\left(X - \frac{D}{2} - \frac{b}{4} + \frac{g}{2h}\right).$$

The $CS+/CS-$ frontier is a straight line and is described by

$$X = \frac{D}{2} + \frac{b}{4} - \frac{g}{2h}, \tag{12}$$

and $CS+$ region is observed for X above or D below the frontier. Only two types can be present in the discrete pooling equilibrium, and types are associated by the function $\overline{X}$, derived from marginal condition (8), i.e.,

$$\overline{X}(X, D) = -\frac{g}{h} + \frac{b}{2} + D - X.$$

For a given level of dividend, higher type in $CS+$ region pools with lower type in $CS-$ region. The marginal cost of signaling is decreasing for high types, due to high investment, and is increasing for low types, due to productivity shock.

The differential equation for separating equilibrium is derived from (7),

$$\mathcal{D}' = \frac{-ka\left[-3hX^2 + 2(2h\mathcal{D} + hb - g)X - h\mathcal{D}^2 + (2g - hb)\mathcal{D} + bg\right]}{-2a(hX + g)\mathcal{D} + a(2hX^2 + (2g - hb)X - bg) + 1 + i}, \tag{13}$$

and, from (9), the differential equation for discrete pooling equilibrium is

$$\mathcal{D}' = \frac{-k\left(\mathcal{D} + \frac{b}{2} + \frac{g}{h}\right)\left(2X - \left(\mathcal{D} + \frac{b}{2} - \frac{g}{h}\right)\right)}{R(X, \mathcal{D})}, \tag{14}$$

where

$$R(X,D) = (4-3k)X^2 - \left[(2-k)(b+2D) - 4(1-k)\frac{g}{h}\right] X$$
$$- (2-k)(b+2D)\frac{g}{h} + k\left(\frac{b^2}{4} - \frac{g^2}{h^2}\right) + \frac{2(1+i)}{ah}.$$

2.1 An example

As a first example, we choose parameters values that generate a representative equilibrium with discretely pooled types. Equilibria for other parameter values are shown in section 2.3. In this example, we assume $a = 1$, $b = 4$, $i = 0$, $k = \frac{1}{2}$, $g = 0$ and $h = 1$, that is,

$$F(I) = I(4 - I),$$
$$\varepsilon(X) = X,$$

where $0 \leq I \leq 2$, and X uniformly distributed over $[X_1, X_2] = [0.8, 2]$.

In this context we have

$$W_{XD} = 2X - D - 2$$

and the $CS+$ region is defined by $X > \frac{D}{2} + 1$. The thick line in figure 1 divides the $X \times D$ plane in $CS+$ and $CS-$ regions. Note that we cannot have equilibrium in the region above $D = X$, since dividend must be lesser than earnings in period 1.

The distribution of probabilities of types imposes more restrictions on the possible contracts. The grayed area in figure 1 is the region where the discrete pooling is possible for our choice of support. For each level of dividend, every type, X, in grayed area has a correspondent type, $\overline{X}(X, D) = D + 2 - X$, inside the support of the distribution function. Note that for low dividend, high types cannot discretely pool because the correspondent type is out of the interval of possible types.

From (13), the differential equation for separating equilibria is

$$\mathcal{D}' = \frac{3X^2 - 4(\mathcal{D} + 2)X + \mathcal{D}(\mathcal{D} + 4)}{2 - 4X(2 - X + \mathcal{D})}, \tag{15}$$

and, from (14), the differential equation for discrete pooling is

$$\mathcal{D}' = \frac{-2(\mathcal{D}+2)X + (\mathcal{D}+2)^2}{5X^2 - 6(\mathcal{D}+2)X + 8}. \tag{16}$$

Figure 2 shows some possible paths derived from equation (16), which has a singularity point at $D = -2 + \frac{4}{7}\sqrt{14} \approx 0.1381$. Only paths on or above the saddle point are implementable, as they satisfy second order condition (proposition 1). The dashed curve is the first-best optimal dividend level given by equation (2). That is, for $X \in [1 - \frac{\sqrt{2}}{2}, 1 + \frac{\sqrt{2}}{2}]$ the firm should not pay dividend and invest all the earnings. For all other types $D^* = X - 2 + \frac{1}{2X}$.

An initial condition must be provided for differential equations (15) and (16). For each initial condition we have a different solution. We follow the pro-separation criterion and choose the signaling schedule with the lowest level of dividend. This is an efficiency criterion, since signaling is costly and dividend implies low investments. The efficient signaling is the path crossing the singularity point because this is the path closer to the optimal dividend level, if the market knows types.

The efficient discrete pooling path does not cover all types. Denote as X_s the type on which the saddle path crosses the boundary of grayed area. For types higher than X_s the matching types are out of the support of distribution. These types must be covered by a separating equilibrium that is a solution of the differential equation (15). The initial dividend level for the separating schedule should not break incentive compatibility. This dividend level should be on the indifference curve of X_s type that passes through the last point of discrete pooling schedule and satisfy zero-profit condition for separating equilibrium.

The optimal contract is plotted on $D \times V^m$ plane of figure 3. Observe that there is a jump in dividend and in value at type X_s assigned contract. The intermediate values can be filled by any contract that is not preferred for any type. In particular, the dotted curve is the lower envelope of the indifference curves that crosses the pair (D, V^m) assigned to each type by the signaling schedule.

On the $X \times D$ plane there is a jump when equilibrium changes from pooling to separating. This jump does not exist if the distribution were continuous and equal to zero at X_1. Separating equilibrium is full informative and market knows that only higher types are selected. Then, dividends must be higher in separating equilibrium than in discrete pooling to satisfy zero profit.

In discrete pooling equilibrium, lower types pool with higher types because, in $CS-$, benefits from being treated equally as higher types compensates the cost of higher dividend distribution. The market knows the disguising behavior and adjust the evaluation, but signaling is preserved since lower types do not reduce the expect valuation to the point that incentives are broken.

Whenever the single-crossing property holds, the first and second order conditions assure that incentive compatibility is valid globally. However, without single-crossing, additional checking is needed. In general, every type should correctly choose the signal level assigned to him. In the quadratic case, sometimes a type in the saddle path prefers a contract in the separating region (this situation does not occur in the example illustrated in figure 3). If incentive compatibility does not hold for the saddle path, there exist paths above it that satisfy global incentive compatibility. The lowest of these paths is chosen as the equilibrium of our model.

2.2 Other cases

The combination of discrete pooling with separating equilibrium is not the only equilibrium case for this model. Depending on the support of distribution and parameter values, other kinds of equilibrium may arise. Figure 4 shows other three cases of equilibrium. In graph (a) the $CS-$ set is small and a separating equilibrium schedule exists in $CS+$ region. In graphs (b) and (c) the saddle path crosses the upper limit of discrete pooling region and the contract is not implementable because dividends exceed earnings or because correspondent type is above support. In this case a combination of discrete pooling with continuous pooling may be a bunching equilibrium. Graph (b) shows a discrete pooling with bunching. A continuum of low types pools with a continuum of higher types so that the value of the pool coincides with the value of discrete pool, for the same level of dividend. In graph (c), the saddle path requires a higher level of dividend, such that discrete pooling is not possible even with bunching. The equilibrium is a continuous pooling of all types at efficient dividend level. Graph (d) shows the three cases on $D \times V^m$ plane.

2.3 Changing the parameter values

Figure 5 shows equilibria for some values of the parameter k and the remaining parameter values equal to the example above. The parameter k represents the weight of market price in the objective of the manager. For higher k, the manager is more inclined to sacrifice future earnings paying dividends in exchange for an increase in V^m, or, equivalently, a small increase in V^m is sufficient to induce the manager to pay given level of dividends. Consequently, indifference curves and the signaling equilibrium at $D \times V^m$ plane is flatter when k is higher. Comparing the equilibrium for $k = 0.5$ and $k = 0.7$, it is noticeable that the saddle path for higher k is at a higher dividend level. The flatter is the indifference curve, the higher must be the dividend in order to obtain local incentive compatibility, and, as the lowest types cannot pay the dividends required for discrete pooling, the equilibrium with bunching occurs. On the other side, for $k = 0.1$ and $k = 0.3$ the saddle path is not equilibrium because global incentive compatibility condition does not hold. For lower k, the relatively high level of V^m increases the probability of violation of global incentive compatibility condition. Consequently, the signaling equilibrium is higher and the set of separated types is shorter.

The change of parameter g is examined in figures 6 and 7. They repeat equilibria of figure 5 for $g = -0.2$ and $g = 0.2$. For lower g, the relative role of productivity shock on the value of the firm is greater, and, as lower types has more opportunity to pool with higher types, pooling is prevalent in figure 6. For higher g, the importance of productivity shock is diminished and separation is more likely as is shown in figure 7. For parameter h, the rationale is similar. The greater is the relative role of productivity shock, the stronger is the productivity effect, and pooling is more likely.

Conversely, for weaker productivity effect, separation is favored as the model is closer to Miller and Rock (1985).

Geometrically speaking, the influence of parameters g, h and b on equilibrium is determined by the placement of $CS+/CS-$ frontier. By equation (12), higher b, lower g, higher h for positive g or lower h for negative g moves the frontier to right and $CS-$ region enlarges. Consequently, separating equilibria is less frequent and continuous pooling more common. On the other side, with changes of parameters in opposite direction, $CS+$ region enlarges and separating equilibrium is more frequent. Discrete pooling is an intermediate case and will be present when $CS+$ and $CS-$ coexist in a favorable way.

3 Connections to the empirical literature

3.1 Future earnings

In the equilibria computed in section 2, we analyzed the relationship between dividends (D) and current earnings (X). As empirical studies are concerned to the future earnings (Y), we plot in figure **??** the relationship between dividends and expected future earnings ($\mathrm{E}[Y]$), corresponding to the base example developed in subsection 2.1. The essential point is that the segment of the curve for discretely pooled types is flatter compared to the separated type segment. This means that discrete pooling weakens the ability of dividends to signal. For the low dividend firms, large changes in dividends are related to relatively small changes in future earnings. On the other hand, high dividends firm are in the separated subset and maintain the ability to signal.

3.2 Testable implications

From the equilibria we found, three testable implications arise. First, in discrete pooling set, the variance of earnings is higher for higher dividend levels. For low earnings firms in discrete pooling, the productivity effect dominates. As earnings and productivity shocks are positively correlated, higher earning firms expect higher productivity in future and are more reluctant to sacrifice investment. Consequently, dividends are decreasing with respect to earnings. On the other hand, if earnings are high enough, decreasing returns of investments dominates over productivity shock, and the original Miller and Rock (1985) signaling reappear, that is, higher earnings firms pay more dividends because the signaling cost is lower for them. The resulting signaling schedule is U-shaped, and high dividends are used as a signal by more distant types. Conversely, firms that signal with low dividend have similar earnings. This result suggests that the variance of earnings in the subset of firms that signal with the same level of dividends is increasing with respect to the dividends. Such relationship can be used as a test for our model.

The second implication is that, when separation and discrete pooling coexist, signal levels are sharply distinct for these two kinds of signaling, as illustrated in figure 3. As very low earnings firms are in discrete pooling set, the average earning and, consequently, the dividend level is substantially lower in this class. A statistical analysis may detect a bimodal distribution of the signal, such that higher levels correspond to the separated types and lower levels, to the discretely pooled types. Furthermore, for discretely pooled types the signaling is weaker, that is, the corresponding earnings must be confined to a smaller interval.

The third implication derives from the assumption that firms cannot issue debt to finance investments. If firms can issue debt, investment is always at optimum level and dividends have no informational content because there is no underinvestment cost supporting signaling. This suggests that evidence for signaling should be more likely in a subset of firms subject to more credit constraints.

3.3 Other remarks

The correspondence between empirical results and the implications from signaling models is imprecise. The main point is that the signaling approach characterizes single-period dividends. Since the empirical analysis, as in Benartzi, Michaely and Thaler (1997), focus on changes in dividends and changes in earnings, the correspondence should be made to variables of a dynamic model, which is out of the scope of this work. The empirical literature implicitly assumes that the signaling profile is stable, thus, changes in dividends and changes in earnings are positively related in separating equilibrium.

When signal is non-monotonic, the same conclusion does not hold. The relationship between changes in dividends and the average change in earnings depends on the distribution of the changes, about which we did not made any assumption. As signal decreases with respect to earnings for some types, and increases for others, the resulting relationship is ambiguous. The equilibrium we found suggests that the empirical research should also examine the relationship among the levels of the variables.

4 Conclusion

Traditional models of signaling satisfy the single-crossing property for the objective function of the firm. This property leads to a monotone relationship between types and signals. We believe that this assumption is not essential and the results derived from non-single-crossing signaling are also plausible. For instance, in our dividend-signaling model, low-earning firms may pay high dividends in order to be considered as high-earning firms. The general result is that the monotone relationship between types (current earnings) and signals (dividends) is not assured and the signaling power of dividends is weakened.

This result has impact on empirical works that test for the presence of asymmetric information assuming single-crossing.[4] Critics on the informational contents of dividends claim that statistical evidences of positive relation between dividends and earnings are weak. But, as we have shown, this is a possible result when discrete pooling equilibrium occurs. Our results suggest that future empirical research should identify separated and pooled types and investigate the relation between variance of earnings and dividends.

We believe that the absence of the single-crossing property is not an uncommon situation and empirical research must take this possibility into account. The presence of discrete pooling enriches the interrelation among variables in signaling models and tests based simply on monotonicity should not be used to reject asymmetric information.

Figure legends

Figure 1: CS split and discrete pooling region.

Figure 2: Discrete pooling paths.

Figure 3: Signaling Equilibrium.

Figure 4: Signaling Equilibria.

Figure 5: Equilibria for different values for k.

Figure 6: Equilibria for $g = -0.2$.

Figure 7: Equilibria for $g = 0.2$.

Figure 8: Future earnings.

A Appendix

A.1 The perfect Bayesian equilibrium

A perfect Bayesian equilibrium (PBE) for dividend signaling model is a profile of strategies $\{c(X) = (\mathcal{D}(X), \mathcal{V}^m(X))\}_{X\in[X_1,X_2]}$ and ex-post beliefs $\mu(\cdot|c)$ such that the following conditions are satisfied:

1. *Zero expected profit constraint*:

$$\mathcal{V}^m(X) = \int V(\tilde{X}, \mathcal{D}(X)) d\mu(\tilde{X}|c(X))$$

2. *Maximization of firm welfare*:

$$X \in \arg \max_{\hat{X}\in[X_1,X_2]} k\mathcal{V}^m(\hat{X}) + (1-k)V(X, \mathcal{D}(\hat{X}))$$

3. *Consistence of beliefs*: $\mu(X|c)$ is the Bayesian updating given 1 and 2, i.e., it is the probability a posteriori of X given c.

A.2 Proof of propositions

A.2.1 Proof of proposition 1

The first order condition, (6) can be written as

$$W_D(X, \mathcal{D}(X))\mathcal{D}'(X) + kV^{m\prime}(\mathcal{D}(X))\mathcal{D}'(X) = 0,$$

whose derivative is

$$\begin{aligned} W_{XD}(X, \mathcal{D}(X))\mathcal{D}'(X) + \left[W_{DD}(X, \mathcal{D}(\hat{X})) + kV^{m\prime\prime}(\mathcal{D}(\hat{X}))\right](\mathcal{D}'(\hat{X}))^2 \\ + \left[W_D(X, \mathcal{D}(\hat{X}))) + kV^{m\prime}(\mathcal{D}(\hat{X}))\right]\mathcal{D}''(\hat{X}) = 0. \end{aligned} \tag{17}$$

From definition of $\mathcal{W}$,

$$\begin{aligned} \frac{\partial^2 \mathcal{W}}{\partial \hat{X}^2}(\hat{X}, X) = \left[W_{DD}(X, \mathcal{D}(\hat{X})) + kV^{m\prime\prime}(\mathcal{D}(\hat{X}))\right](\mathcal{D}'(\hat{X}))^2 \\ + \left[W_D(X, \mathcal{D}(\hat{X}))) + kV^{m\prime}(\mathcal{D}(\hat{X}))\right]\mathcal{D}''(\hat{X}). \end{aligned} \tag{18}$$

Equations (17) and (18) simplify second order condition to

$$\frac{\partial^2 \mathcal{W}}{\partial \hat{X}^2}(X, X) = -W_{XD}(X, \mathcal{D}(X))\mathcal{D}'(X) \leq 0,$$

which proves the proposition.

A.2.2 Proof of proposition 2

Since type is perfectly revealed by observation of D, the market correctly evaluate the firm, that is, $V^m(D) = V(X, D)$, where $X = \mathcal{D}^{-1}(D)$. The welfare is

$$\mathcal{W}(\hat{X}, X) = kV(\hat{X}, \mathcal{D}(\hat{X})) + (1-k)V(X, \mathcal{D}(\hat{X})),$$

and the first order condition (6) results (7).

A.2.3 Proof of proposition 3

The first order condition is

$$W_D(X, \mathcal{D}(X))\mathcal{D}'(X) + kV^{m\prime}(\mathcal{D}(X))\mathcal{D}'(X) = 0,$$

and, since $\mathcal{D}'(X) \neq 0$,

$$-kV^{m\prime}(\mathcal{D}(X)) = W_D(X, \mathcal{D}(X)).$$

The left hand side is independent of X, if types are drawn from the same pool, so $W_D(X, \mathcal{D}(X))$ is the same for pooled types.

A.2.4 Proof of proposition 4

Let X be the lowest type that chooses D and $\overline{X}(X, D)$ the highest. The market value is an average of the value of the types pooled in the same dividend level,

$$V^m(\mathcal{D}(X)) = \frac{1}{2}V(X, \mathcal{D}(X)) + \frac{1}{2}V(\overline{X}(X, \mathcal{D}(X)), \mathcal{D}(X)),$$

and the welfare as a function of declaration is

$$\mathcal{W}(\hat{X}, X) = \frac{k}{2} V(\hat{X}, \mathcal{D}(\hat{X})) + \frac{k}{2} V(\overline{X}(\hat{X}, \mathcal{D}(\hat{X})), \mathcal{D}(\hat{X})) + (1-k) V(X, \mathcal{D}(\hat{X})).$$

Differentiating in $\hat{X}$ and taking into account that (8) implies $V_D(X, \mathcal{D}(X)) = V_D(\overline{X}(X, \mathcal{D}(X)), \mathcal{D}(X))$, condition (6) results (9).

References

[1] Aharony, J., and I. Swary, 1980, "Quarterly dividend and earnings announcements and stockholders' returns: an empirical analysis," *Journal of Finance*, 35, 1–12.

[2] Allen, F., and R. Michaely, 1995, "Dividend Policy," in R. Jarrow et al. (eds.), *Handbooks in OR & MS*, vol. 9, Elsevier, New York, 793–837.

[3] Araujo, A., and H. Moreira, 2001, "Adverse selection problems without the Spence-Mirrlees condition," working paper, Escola de Pós-Graduação em Economia, Fundação Getulio Vargas.

[4] Araujo, A., and H. Moreira, 2003, "Non-monotone insurance contracts and their empirical consequences," working paper, Escola de Pós-Graduação em Economia, Fundação Getulio Vargas.

[5] Benartzi, S., R. Michaely, and R. H. Thaler, 1997, "Do changes in dividends signal the future or the past?," *Journal of Finance*, 52, 1007–1034.

[6] Bhattacharya, S., 1979, "Imperfect information, dividend policy, and 'the bird in the hand' fallacy," *Bell Journal of Economics*, 10, 259–270.

[7] Healy, P. M., and K. G. Palepu, 1988, "Earnings information conveyed by dividend initiations and omissions," *Journal of Financial Economics*, 21, 149–175.

[8] John, K., and J. Williams, 1985, "Dividends, dilution, and taxes: a signalling equilibrium," *Journal of Finance*, 40, 1053–1070.

[9] Miller, M. H., and F. Modigliani, 1961, "Dividend policy, growth, and the valuation of shares," *Journal of Business*, 34, 411–433.

[10] Miller, M. H., and K. Rock, 1985, "Dividend policy under asymmetric information," *Journal of Finance*, 40, 1031–1051.

[11] Nissim, D., and A. Ziv, 2001, "Dividend changes and future profitability," *Journal of Finance*, 56, 2111–2133.

[12] Riley, J. G., 1979, "Informational equilibrium," *Econometrica*, 47, 331–359.

[13] Watts, R., 1973, "The informational content of dividends," *Journal of Business*, 46, 191–211.

Footnotes

1. We thank Richard Kihlstrom, Daniel Ferreira, Luís Braido, Ricardo Cavalcanti, Walter Novaes, Fabio Kanczuk, Heitor Almeida, Boyan Jovanovic, Sudipto Bhattacharya, Jean-Charles Rochet and David Martimort for their useful comments, and thank CNPq and CAPES for financial support.

2. See Allen and Michaely (1995) for a survey on theoretical and empirical issues on dividend policy.

3. We use subscripts to denote partial derivative with respect to the subscripted variable.

4. An alternative setting, without the multiplicative shock, would be to assume that the production function has increasing returns for low levels of investment and decreasing returns otherwise. In this case, $W_{XD} = -\frac{1-k}{1+i}F''(X-D)$ and, consequently, marginal cost of dividend increases with earnings only for firms in increasing returns.

5. For instance, see Araujo and Moreira (2003) for the insurance case.

Figures

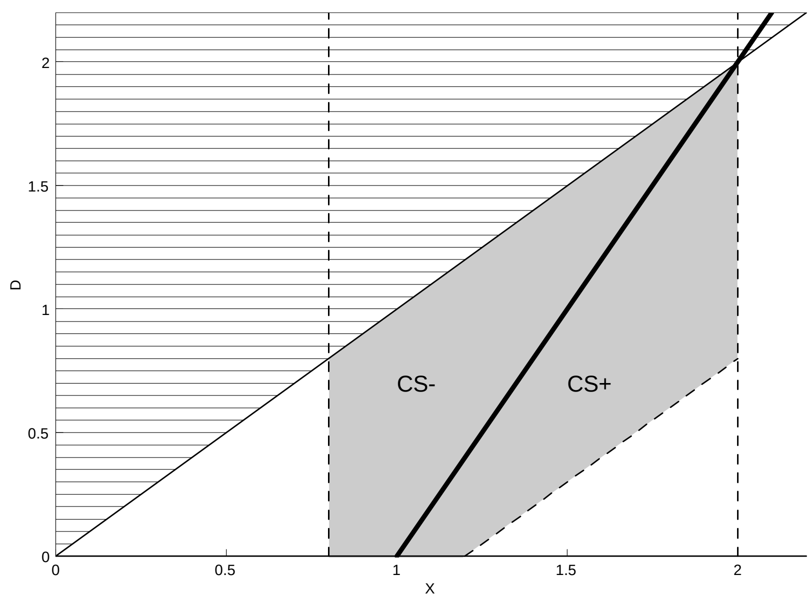

Figure 1:

Figure 8

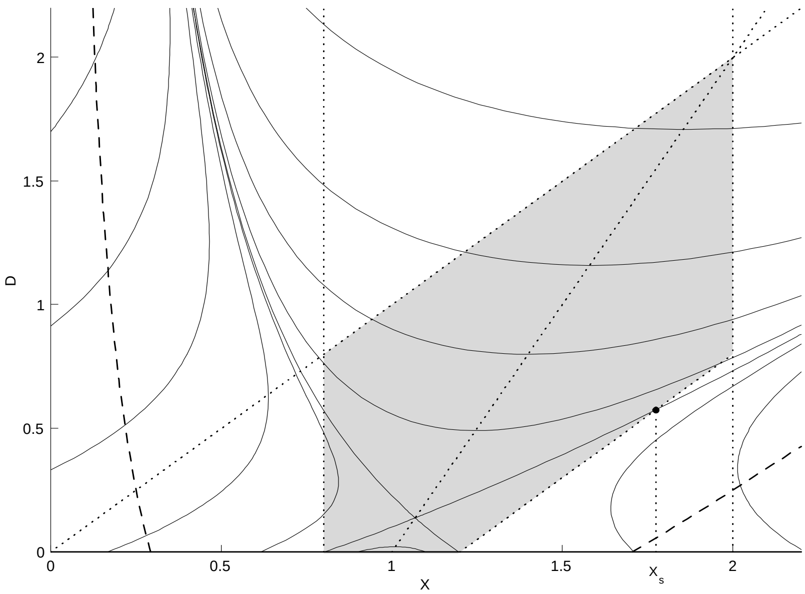

Figure 2:

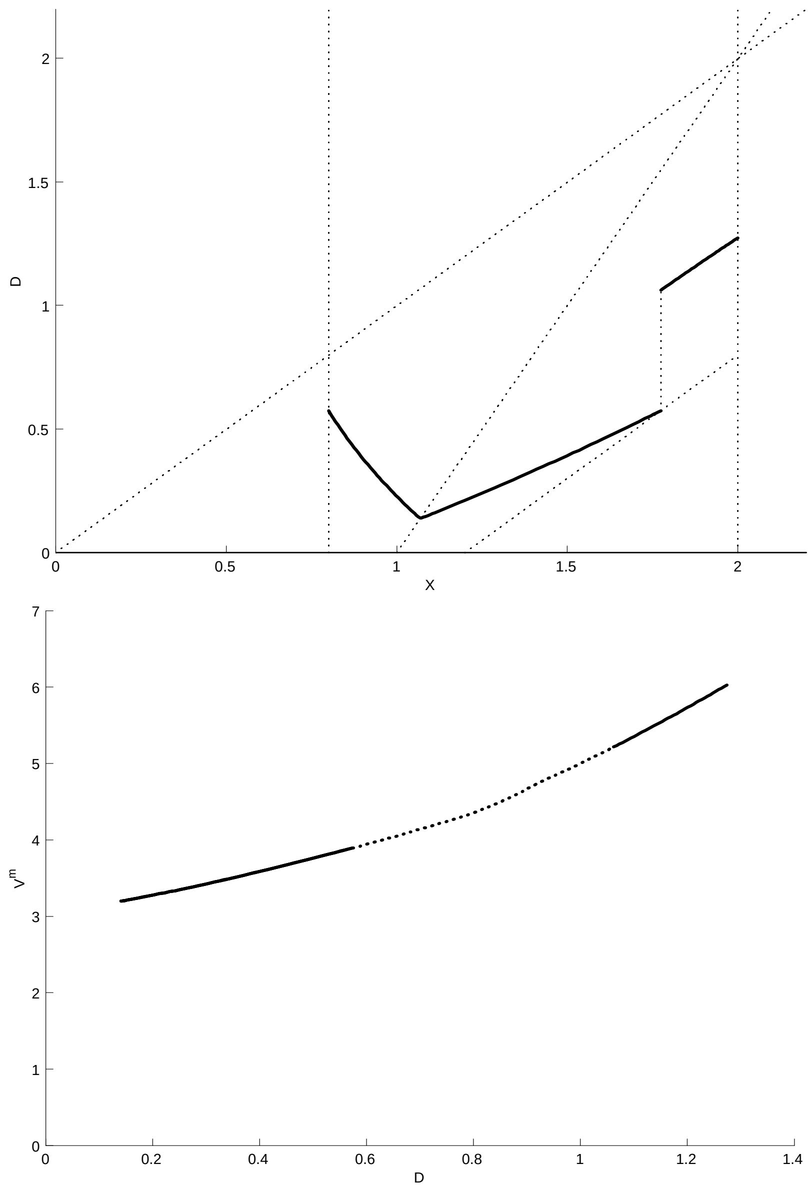

Figure 3:

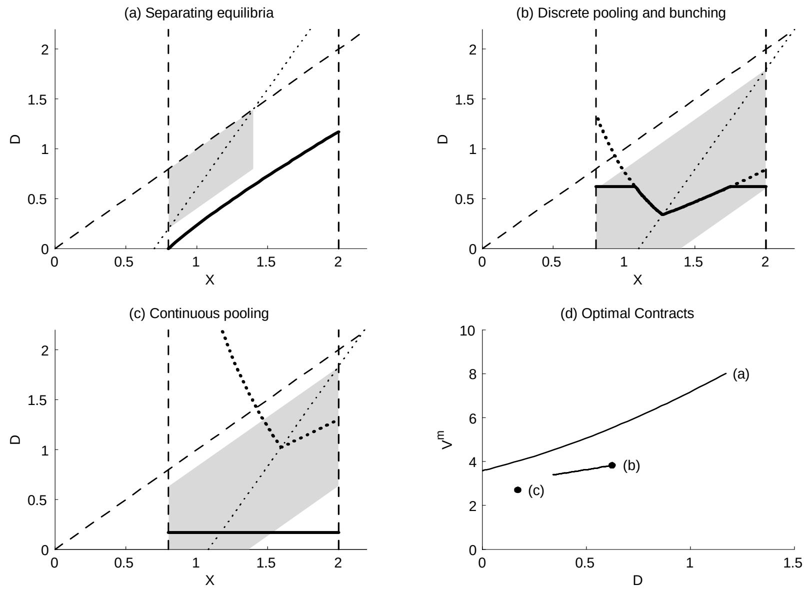

Figure 4:

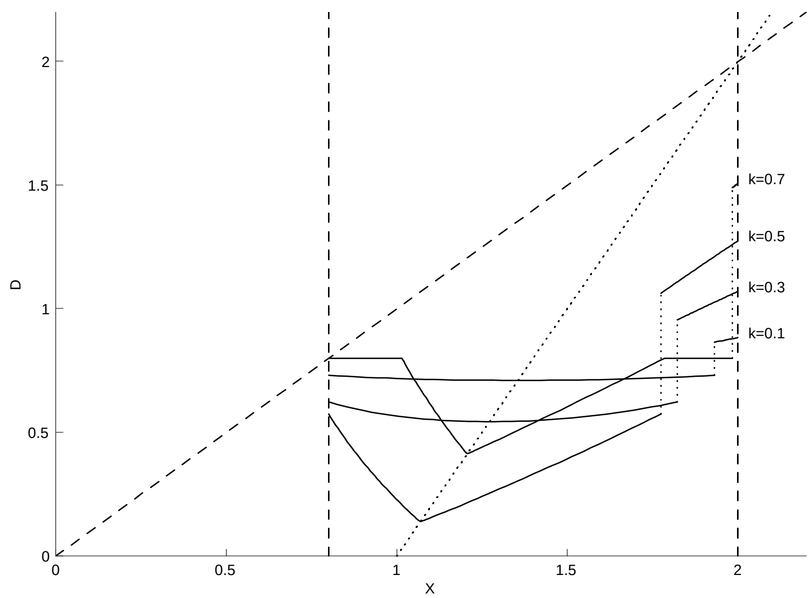

Figure 5:

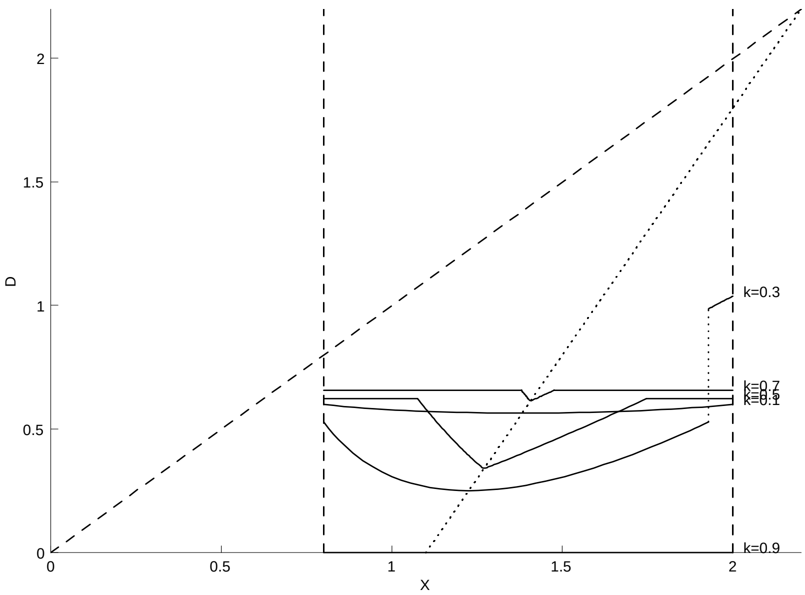

Figure 6:

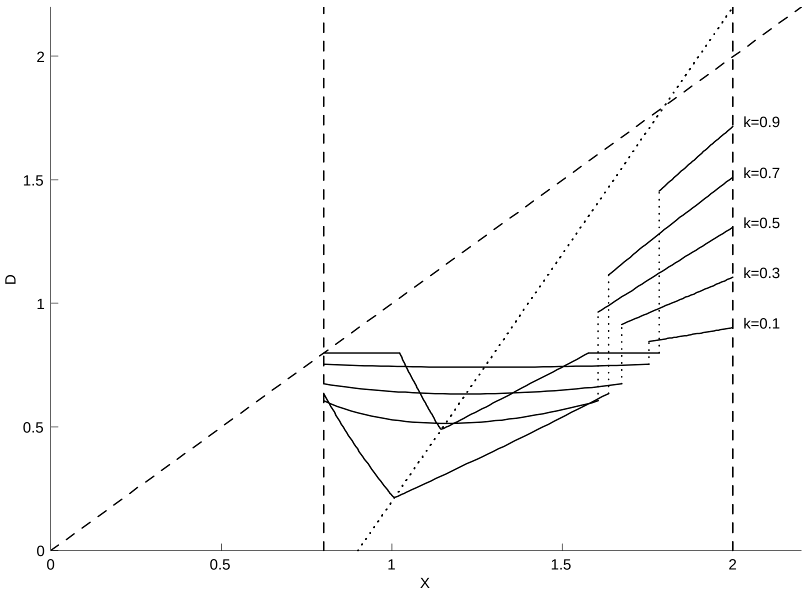

Figure 7:

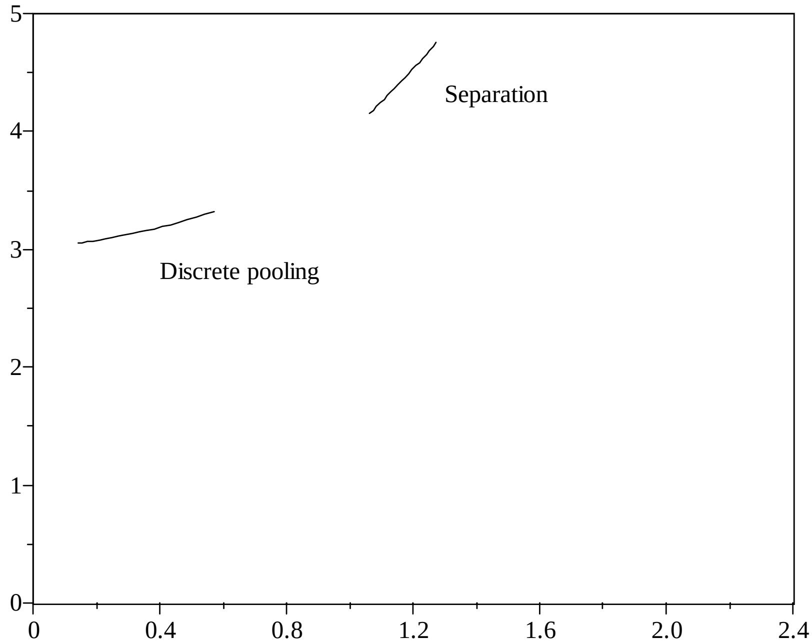

Figure 8: