Optimal Dynamic Resource Allocation to Prevent Defaults

U. Ayesta[a,b,c], M. Erausquin[c], E. Ferreira[c], P. Jacko[d]

[a] LAAS-CNRS, Univ. de Toulouse, CNRS, Toulouse, France
[b] Univ. de Toulouse, LAAS, Toulouse, France
[c] UPV/EHU, University of the Basque Country, Bilbao, Spain
[d] Lancaster University, Department of Management Science, Lancaster, LA1 4YX, UK

Abstract

We consider a resource allocation problem, where a rational agent has to decide how to share a limited amount of resources among different companies that might be facing financial difficulties. The objective is to minimize the total long term cost incurred by the economy due to default events. Using the framework of multi-armed restless bandits and, assuming a two-state evolution of the default risk, the optimal dynamic resource sharing policy is determined. This policy assigns an index value to each company, which orders its priority to be funded. We obtain an analytical expression for this index, which generalizes the return-on-investment (ROI) index under the static setting, and we analyse the influence of the future events on the optimal dynamic policy. A discussion about the structure of the optimal dynamic policy is provided, as well as some extensions of the model.

Keywords: Markov Decision Processes, Multi-Armed Bandit Problem, Default Risk Management, Dynamic Resource Allocation Policies.

1 Introduction

There are many situations where a rational agent (e.g. a public institution) spends a certain amount of resources in order to obtain social rewards (i.e. to prevent social losses). How to manage those resources in an optimal way is a problem that has been widely analysed in the literature. For instance, [2] developed models where the control instruments are income tax rate, government deficit size or debt-money ratio. These instruments are used to maximize functions defined in terms of certain social criteria (e.g. growth rate). In recent years, and more intensely due to economical and financial situation, economic agents have been facing the resource allocation problem of minimizing the social cost incurred by the economy when a firm defaults. Many countries have been sharing public resources among companies which have been close to default, in order to avoid social costs (e.g. sudden increase in unemployment due to collective redundancies) derived from the default event. Although there is a lot of literature devoted to the problem of how to deal with bank failures (see e.g. [1] and references there), to the best of our knowledge, there is not much literature modelling how to solve the resource allocation optimization problem to prevent defaults in a dynamic manner.

This paper represents a first attempt to obtain an optimal dynamic resource allocation policy, with the only objective of preventing possible future default events. Usually, the evolution of the default probability of a company has been modelled as a Markov chain, where each state represents a different rating, the last state being the default event. This Markovian structure of the evolution of ratings has been used for different purposes. For instance, [8] provided a new framework for quantifying credit risk in portfolios, based on historical transition probabilities for the modelling of the rating migrations. In [10], the authors refine the model presented in [5], and proposed a new discrete-time Markovian model to estimate changes in bonds and loans prices. This model is still considered one of the most important approaches when it comes to pricing bonds and credit derivatives. In general, most of those applications focused on management of portfolios with assets that involve credit risk. In this paper we use the Markovian setting to dynamically describe and analyse the resource allocation problem. Our model takes advantage of the Markovian structure of the evolution of default probabilities. It is formulated as an extension of the Multi Armed Restless Bandit (MARB) framework described in [20], which is a special type of constrained Markov Decision Processes (MDP).

MARB problems are a class of sequential resource allocation problems concerned with allocating resources among several alternative arms in a competitive manner. Depending on how the resources are allocated, the dynamics of the states of the arms might change. Moreover, the rewards earned by the decision maker depend on the state of the arms. And, therefore, the decision maker has to find an optimal policy taking into account both the influence of the actions chosen in the dynamics of the states, and the reward structure of each arm. There is a large recent literature where similar models have been shown to be useful in a variety of application fields (see, for instance, [14, 9, 6], and the references therein). In some cases, the solution of MARB problems is easy to implement and analyse. The optimal policy is given in terms of the (Whittle) index function, which assigns an index value to each arm. The larger the index value of an arm is, the higher is its priority when resources are allocated. Therefore, if this index function is obtained analytically, it also gives some insights about the structure of the optimal policy.

The structure of the MARB framework fits to several problems in economics, where rational agents must allocate resources. For instance, [17] approached the problem in which a single firm is facing a market with unknown demand. [18] analyzed the possibility of mispricing in a two-armed bandit problem when the frequency of change is small. If several firms were to experiment independently in the same market, they might offer different prices in the long run. Optimal experimentation may therefore lead to price dispersion in the long run, as shown formally in [13]. [4] used both the MAB and MARB to identify the best strategy for managing obsolescence in such instances wherein organizations have to deal with continuous technological evolution under uncertainty. Another situation appears when we must choose between various research projects. [19] analyses this problem where each arm represents a different research project with an associated random. In [16] a richer model of choice between R & D processes is considered. The MARB problem has also been used for defining a canonical model of experimentation in teams. In [3] and [11] a set of players choose independently between the different arms.

In this paper, we propose the use of MARB to analyse a decision problem in a default risk framework. That is, a rational agent (referred as the decision maker) must face the problem of selecting how to share funds among different companies that might be facing risk of default. Those companies are competing to obtain those resources, in order to improve their financial situation and avoid the default. The goal of the rational agent is to minimize the long run average costs incurred by the default events.

Although we focus on default events related with firms in financial difficulties, the formulation presented in this paper could be adapted for approaching other resource allocation problems, where the objective is to keep as many individuals as possible in the system. For instance, similar models have been used in problems where resources are shared among patients [12], or human resources to maintain machinery [7]. The rest of the paper is organized as follows: the model, and the related optimization problem, are stated in Section 2. In Section 3, the dynamic optimal policy, characterized by an index function, is provided. In Section 4, we discuss the role of the parameters of the problem in the index function. Proofs are included in the Appendix.

2 Problem and model description

We consider a set $\mathcal{K} = \{1, \cdots, K\}$ of companies asking for funding. We assume that, every time the decision of allocating the available resources is taken, the companies are divided in two groups: those which have a positive probability of defaulting before the next decision period, and those which do not. For simplicity, we say that the state of the companies that belong to the first group is *Bad*, and the state of the companies that are in the second group is *Good*.

Let us denote by c_k the cost incurred to the economy if company $k \in \mathcal{K}$ defaults. For any period $t \geq 0$, we consider that, if company k defaults at time t, the value of the costs derived from the default at $t = 0$ is given by $\beta^t c_k$, with $0 \leq \beta < 1$. In other words, β^t is a discount factor that measures the present value of the cost of default events at time t. Therefore, if $\beta < 1$, the later a company defaults, the smaller the cost derived from the default is, since this cost will be multiplied by a discount factor. Thus, the discount factor represents the importance of the future given by the decision maker, and the decision maker will select the best value of β according to its planning horizon. Clearly, β might include information related with inflation and, possibly, other future effects.

If the decision is not to give resources to company k while being in *Bad* state, then, this company will default with probability d_k. However, if the decision is to fund it, then the default probability is decreased to $\mu_k d_k$, with $0 < \mu_k < 1$. Both d_k and $\mu_k d_k$ can be interpreted as "expected probabilities".

The size of each company plays an important role too, since the amount of resources needed in order to improve the financial situation depends on it. We denote by w_k the resources that the decision maker has to give

to company k, in order to decrease its default probability. The evolution of the state of each company is modelled with a Markov chain, where the transition probabilities depend on the action taken by the decision maker (see (1) and (2)).

The objective is to find a policy which prescribes how to share the resources in order to minimize the expected total discounted costs derived from the defaults of the companies, given an expected budget per period. There is a clear trade-off related to the described optimization problem, between funding companies in the bad state versus funding companies in the good state. By giving funding to companies in the bad state, the decision maker obtains a reward in the short term. On the other hand, by funding companies in the good state, the decision maker prevents those companies from moving to the bad state, keeping them alive as long as possible.

We represent this optimization problem using a discrete-time MDP formulation, which fits in the framework of the MARB problem [20, 14]. Consider the time slotted in periods, $t \in \mathcal{T} = \{0, 1, \cdots\}$. At each time period, some of the companies in $\mathcal{K}$ are selected to receive funding from the decision maker, according to a budget constraint. We denote by $\mathcal{A} = \{0, 1\}$ the action space for each company, where 0 represents not selecting a company for funding, whereas 1 means selecting it. Each company is characterized independently from other companies, by the tuple $\left(\mathcal{N}, (W_k^a)_{a\in\mathcal{A}}, (R_k^a)_{a\in\mathcal{A}}, (P_k^a)_{a\in\mathcal{A}}\right)$, where:

- $\mathcal{N} = \{G, B, D\}$ represents the state space of the company. State G represents the *Good* state, B represents the *Bad* state, and D represents the company is default.
- $W_k^a = \left(W_{k,n}^a\right)_{n\in\mathcal{N}}$, where $W_{k,n}^a$ represents the one period budget consumption of company k, being in state n, if action a is taken by the decision maker. In particular, we have, for any $n \in \mathcal{N}$,

$$W_{k,n}^1 = w_k \qquad\qquad W_{k,n}^0 = 0$$

  In other words, the amount of resources given is w_k if company k is selected and 0 otherwise.
- $R_k^a = \left(R_{k,n}^a\right)_{n\in\mathcal{N}}$, where $R_{k,n}^a$ represents the one period expected reward earned by the system due to company k at state $n \in \mathcal{N}$ if action a is decided in the beginning of the period. In particular, in our model we have

$$\begin{aligned} R_{k,G}^1 &= 0 & R_{k,G}^0 &= 0, \\ R_{k,B}^1 &= -\mu_k d_k c_k & R_{k,B}^0 &= -d_k c_k, \\ R_{k,D}^1 &= 0 & R_{k,D}^0 &= 0. \end{aligned}$$

- $P_k^a = \left(p_{k,m,n}^a\right)_{m,n\in\mathcal{N}}$ represents the one-period probability of company k of migrating from state m to state n if action a is taken by the decision maker. The following transition matrices describe the transition probabilities under different actions:

$$\boldsymbol{P}_k^1 := \begin{array}{c} \\ G \\ B \\ D \end{array} \begin{array}{c} \begin{array}{ccc} G & B & D \end{array} \\ \begin{pmatrix} p_{k,G,G}^1 & p_{k,G,B}^1 & 0 \\ p_{k,B,G}^1 & p_{k,B,B}^1 & \mu_k d_k \\ 0 & 0 & 1 \end{pmatrix} \end{array} \tag{1}$$

$$\boldsymbol{P}_k^0 := \begin{array}{c} \\ G \\ B \\ D \end{array} \begin{array}{c} \begin{array}{ccc} G & B & D \end{array} \\ \begin{pmatrix} p_{k,G,G}^0 & p_{k,G,B}^0 & 0 \\ p_{k,B,G}^0 & p_{k,B,B}^0 & d_k \\ 0 & 0 & 1 \end{pmatrix} \end{array} \tag{2}$$

Note that D is an absorbing state.

The dynamics of company k is then captured by the state process $(X_k(t))_{t\in\mathcal{T}}$, and the action process $(a_k(t))_{t\in\mathcal{T}}$. As a result of deciding action $a_k(t)$ in state $X_k(t)$, the company consumes the allocated funding, provides a reward, and evolves its state for the next time period.

2.1 Optimization problem

Let $\Pi_{X,a}$ be the set of randomized and non-anticipative policies that at the beginning of period t decide an action $a(t) = (a_1(t), ..., a_K(t)) \in \mathcal{A}^K$ based only on the evolution of the state space, $X(0), X(1), \cdots, X(t)$, where $X(t) = (X_1(t), ..., X_K(t)$ and the history of the action process, $a(0), a(1), \cdots, a(t-1)$. Let $\mathbb{E}_\tau^\pi$ denote the expectation over the state process $X(\cdot)$ and over the action process $a(\cdot)$, conditioned on the state-process history $X(0), X(1), \ldots, X(\tau)$ and on policy π.

We formulate now the optimization problem we want to solve. For any $0 \leq \beta \leq 1$, the goal is to find a policy $\pi \in \Pi_{X,a}$ that maximizes the expected total discounted reward starting from the initial time period 0, subject to the family of sample path allocation constraints, i.e.,

$$\max_{\boldsymbol{\pi} \in \Pi_{\boldsymbol{X},\boldsymbol{a}}} \quad \mathbb{E}_0^{\boldsymbol{\pi}} \left[{}_{t=0}^{\infty} \sum_{k \in \mathcal{K}} \beta^t R_{k,X_k(t)}^{a_k(t)} \right] \tag{3}$$

$$\text{s. t.} \quad \mathbb{E}_0^{\boldsymbol{\pi}} \left[\sum_{t=0}^{\infty} \sum_{k \in \mathcal{K}} \beta^t W_{k,X_k(t)}^{a_k(t)} \right] \leq W \tag{4}$$

Note that W represents the total budget that the decision maker can spend on average. It seems natural to consider this quantity as the present value of the discounted budgets available at each period. For instance, if we assume that the budget at each period is given by a fixed value $\tilde{W}$, W could be interpreted as $W = \sum_{t=0}^{\infty} \beta^t \tilde{W} = \frac{\tilde{W}}{1-\beta}$, for any $0 \leq \beta < 1$. For $\beta = 1$, (3) becomes a total cost minimization problem. In Section 4.3.2 we argue why in this case any feasible policy is optimal.

3 Dynamic optimal policy

Following the method given in [20], this optimization problem described in (3)-(4) can be approached using Lagrangian multipliers, and decomposing it into K single-company subproblems.

Notice that any joint policy $\boldsymbol{\pi} \in \Pi_{\boldsymbol{X},\boldsymbol{a}}$ defines K single-company policies, $\widetilde{\pi}_k \in \Pi_{\boldsymbol{X},a_k}$, which depend on the joint state-space process and action process. We will therefore analyse the single-company subproblem

$$\max_{\widetilde{\pi}_k \in \Pi_{\boldsymbol{X},a_k}} \mathbb{E}_0^{\widetilde{\pi}_k} \left[\sum_{t=0}^{\infty} \beta^t \left(R_{k,X_k(t)}^{a_k(t)} - \nu W_{k,X_k(t)}^{a_k(t)} \right) \right], \tag{5}$$

being ν the Lagrangian multiplier.

Thus, the main idea of our approach is to identify a set of optimal policies $\widetilde{\pi}_k^*$ for each company k, and using them to construct a joint policy $\widetilde{\pi}^*$, which will be optimal for the maximization problem (3)-(4). The optimal policy is obtained assigning an index value, $\nu_{k,n}$, to each state $n \in \mathcal{N}$ under the so-called *indexability condition* [14].

The optimization problem described in (5) is a standard MDP problem for any value of ν. It is known that there exists an optimal policy in $\Pi_{\boldsymbol{X},a_k}$, which is deterministic, Markovian, and independent of the initial state (see, for instance, [15, Chapter 6]). In particular, this implies the existence of an optimal policy which depends on company k-s state-process $X_k(\cdot)$. Therefore, in order to find an optimal policy, it is enough to focus on policies that are described in terms of a subset of states, $\mathcal{S} \subseteq \mathcal{N}$, which prescribes to allocate resources to company k whenever the company is in state $n \in \mathcal{S}$. Thus, an optimal policy can be obtained by solving

$$\max_{\mathcal{S} \subseteq \mathcal{N}} \mathbb{E}_0^{\mathcal{S}} \left[\sum_{t=0}^{\infty} \beta^t \left(R_{k,X_k(t)}^{a_k(t)} - \nu W_{k,X_k(t)}^{a_k(t)} \right) \right], \tag{6}$$

Let us define $\Delta_{k,G,G} = p_{k,G,G}^1 - p_{k,G,G}^0$ and $\Delta_{k,B,G} = p_{k,B,G}^1 - p_{k,B,G}^0$. To derive our main results, we assume the following:

$$\frac{\Delta_{k,B,G}}{\Delta_{k,G,G}} \geq 1 \text{ or } \Delta_{k,G,G} = 0. \tag{7}$$

Remark 1. *Condition (7) implies that both differences, $\Delta_{k,B,G}$ and $\Delta_{k,G,G}$, have the same sign. Clearly, the natural situation comes from companies where this sign is positive; that is, acting increases the probability of being*

in state G. Moreover, a negative sign leads to a trivial situation, where not funding is always optimal. In general situations, where $\Delta_{k,B,G} > 0$ and $\Delta_{k,G,G} > 0$, and $\Delta_{k,G,G} > 0$, condition (7) implies that the effect of funding is greater in companies in the Bad state than in companies being in state G, which seems to be a realistic assumption.

Let f_k denote the total discounted probability of hitting state B if starting from state G and not funding company k in state G. It can easily be seen that

$$f_k = \frac{\beta p^0_{k,G,B}}{1 - \beta p^0_{k,G,G}}. \tag{8}$$

Note that f_k is decreasing in $p^0_{k,G,G}$.

Next we define the index values for company k

1.
$$\nu_{k,D} := 0. \tag{9}$$

2.
$$\nu_{k,G} := \frac{\mu_k c_k d_k}{w_k} \frac{\phi_k}{1 - \phi_k}, \tag{10}$$

where

$$\phi_k = -\beta \frac{\left(p^1_{k,G,G} - p^0_{k,G,G}\right) f_k + \left(p^1_{k,G,B} - p^0_{k,G,B}\right)}{1 - \beta \left(p^1_{k,B,G} f_k + p^1_{k,B,B}\right)}. \tag{11}$$

3.
$$\nu_{k,B} = \frac{c_k d_k}{w_k} \left(1 - \mu_k - \theta_k\right), \tag{12}$$

where

$$\theta_k = \beta \frac{\left(p^1_{k,B,G} - p^0_{k,B,G}\right) f_k + \left(p^1_{k,B,B} - p^0_{k,B,B}\right)}{1 - \beta \left(p^0_{k,B,G} f_k + p^0_{k,B,B}\right)}. \tag{13}$$

We describe now the main theoretical results of this paper. First, for each company k, we establish an ordering between the states with respect to the optimal policy.

Proposition 3.1. *Assume (7).*

1. *Suppose that $\nu > 0$ holds. Then, if it is optimal for Problem (5) to fund company k in state G, it is optimal to fund it in state B as well. Moreover, it is never optimal to fund company k in state D.*

2. *Suppose that $\nu \leq 0$. Then, it is optimal for Problem (5) to fund company k in any state $n \in \mathcal{N}$.*

Proof. See Appendix A. □

Remark 2. *Proposition 3.1 establishes a natural order between the states. Since the optimal policy is a threshold-type solution, the states can be ordered following the priority given by the optimal policy. Throughout the rest of the paper, we refer to this order using the following notation: $D < G < B$. In view of this, we define by $S_{n:\mathcal{N}} = \{m \in \mathcal{N} : m > n\}$ the subset of $\mathcal{N}$ with states* bigger *than a given state n.*

Theorem 3.2 (Whittle's indexability)**.** *Assume (7). The following holds for Problem (5)*

1. *If $\nu \leq \nu_{k,n}$, then it is optimal to fund company k if it is in state $n \in \mathcal{N}$.*

2. *If $\nu \geq \nu_{k,n}$, then it is optimal not to fund company k if it is in state $n \in \mathcal{N}$,*

where $\nu_{k,n}$ are given in (9), (10) and (12).

Proof. See Appendix B.

□

Theorem 3.2 characterizes the optimal policy for the single-company problem. The optimal policy for the joint problem (3) is built, employing at any period t, the policy given by Theorem 3.2 to all the companies in $\mathcal{K}$. Thus, if the Lagrangian multiplier ν is known by the decision maker, the optimal policy for the joint problem becomes very simple to implement. At the beginning of any period t, the decision maker has to proceed as follows:

1. First, observe the current state of each company k.
2. Second, depending on the current state, compute the index value of each company k.
3. Third, assign w_k units of resources to each company k with index value higher than ν. Companies with index value smaller or equal to ν are not funded within this decision period.

As usual in threshold type solutions, the value of the threshold parameter ν can be computed numerically, using the duality theory (see [14]).

4 Discussion

Theorem 3.2 gives the optimal solution to Problem (3) in terms of the index values given in Equations (9), (10) and (12).

Although in practice the implementation of the optimal policy requires a parameter estimation procedure, the theoretical analysis of the structure of the indices gives some insights of how the optimal policy behaves. Since these indices establish a priority to each company, it is worth looking at the influence of the different parameters in the indices, to have an idea of the role that each one plays. First, it is important to point out that the index values are nonnegative. Due to the indexability property, as pointed out in Remark 2, it is known that for any company $k \in \mathcal{K}$, $\nu_{k,D} \leq \nu_{k,G} \leq \nu_{k,B}$. Since $\nu_{k,D} = 0$, we have that the index is greater than 0 in the other two states.

4.1 Dependence of the index on c_k,d_k and w_k

Since the index values are positive both for state B and G, it is straightforward to see in (10) and (12) that the indices are increasing in c_k and d_k, and decreasing in w_k.

These properties coincide with what we could expect from the optimal policy. If we consider two companies being in the same state, and with equal parameters except for the costs derived from a default, then it makes sense to give priority to the company with the highest cost. Similarly, if there is a big difference between the amount of resources needed by two companies, but the rest of parameters are of the same order, then it seems also natural to give resources to the company that needs less resources, because the other resources can be saved for the future. Thus, properties that were expected to be followed by the optimal resource sharing policy hold in our priority index policy. Moreover, the expressions of the indices provide further information. For instance, the priority given to a company is not only increasing on c_k and decreasing on w_k. The index shows that the priority is proportional to the ratio between the two, c_k/w_k. Thus, the higher this ratio is, the more priority should be given by the decision maker.

4.2 Dependence of the index on transition probabilities

We observe that if $p^1_{k,B,G} - p^0_{k,B,G} \geq 0$ (a realistic assumption, as noted in Remark 1), then it is easy to show that θ_k is decreasing in $p^0_{k,G,G}$, and therefore $\nu_{k,B}$ is increasing in $p^0_{k,G,G}$. This means that a company with higher probability of maintaining by itself in state G without funding will have a higher index. Somehow, the intuition behind is that such companies would make use of the resource just in the current period and are not very likely to need funding in the near future.

Parameters	c	w	d	μ	$p^1_{G,G}$	$p^1_{G,B}$	$p^1_{B,G}$	$p^1_{B,B}$	$p^0_{G,G}$	$p^0_{G,B}$	$p^0_{B,G}$	$p^0_{B,B}$
Company 1	10000	1000	0.7	0.7857	0.6	0.4	0.25	0.2	0.5	0.5	0.1	0.2
Company 2	60000	1000	0.7	0.7857	0.6	0.4	0.25	0.2	0.5	0.5	0.1	0.2

Table 1: Parameters that represent companies 1 and 2.

4.3 Influence of the discount factor

The importance that the decision maker gives to the future plays a really important role in this resource allocation problem. Thus, analysing the influence of β on the indices, and therefore, on the optimal policy, gives useful information that can be used when making a decision.

4.3.1 The $\beta = 0$ case

If $\beta = 0$, the decision maker is said to be *myopic*. She only cares about the expected cost incurred by the economy at only one time slot, and forgets about minimizing possible future costs. Therefore, the optimal policy is given by that policy that minimizes the one-period expected cost. This is well captured by the indices given in (10) and (12).

If $\beta = 0$, from (10), we have that $b = 0$, and therefore, $\nu_{k,G} = 0$. That is, for those companies that are in state G, the probability of moving to state D in only one time slot is 0, and therefore, there is no reason to give them any resources. Thus, in this case companies in state G have the lowest priority.

On the other hand, from (12), we obtain that $a = 0$, and therefore, $\nu_{k,B} = \frac{c_k d_k}{w_k}(1 - \mu_k) = \frac{c_k}{w_k}(d_k - \mu_k d_k)$. Thus, among those companies with the same ratio $\frac{c_k}{w_k}$, those with the biggest difference $d_k - \mu_k d_k$ will have absolute priority, where $d_k - \mu_k d_k$ represents the difference between the probabilities of moving to state D in one step, not giving and giving resources to the company, respectively.

An interesting interpretation of the index $\nu_{k,B}$ when $\beta = 0$ is its coincidence with the *return-on-investment* (ROI) index, as explained in Section 4.4. Moreover, since the decision maker is myopic, no information about the rest of transition probabilities is used when assigning priority to the companies.

4.3.2 The $\beta = 1$ case

If $\beta = 1$, we are considering the total cost case, which in our case is always finite since all companies end up in finite time in the absorbing state D. This means that the cost incurred by the economy when a company defaults is the same, no matter when this default event happens, and therefore, there is no reason to keep companies in the economy as much time as possible. Since state D is absorbing, it is almost surely known that all the companies will reach state D in a finite time, which means that the cost incurred by the economy due to default events is independent of the policy, and is given by $\sum_{k=1}^{K} c_k$. Thus, all the policies are optimal for the optimization problem described in (3).

Setting $\beta = 1$ in the expression for the indices as obtained in Section 3 we obtain that $\nu_{k,G} = \nu_{k,B} = \nu_{k,D} = 0$. Thus, the indices are not able to establish an order based on priority, assigning the same priority to all the companies.

4.3.3 The $0 < \beta < 1$ case

For $0 < \beta < 1$, the indices depend on the transition probabilities. Not only the instantaneous transition probability to state D matters, but also the evolution of the states in the future. An interesting property of the indices when $0 < \beta < 1$ is that, in some cases, it could be optimal to give resources to companies in state G and, at the same time, not to give resources to some other companies in state B.

We illustrate this fact with the following numerical example. Consider two companies, described by the set of parameters given in Table 1. The statistics that describe their evolution are the same, being the only difference that the cost incurred by the economy if Company 2 defaults is 6 times bigger than the cost when Company 1 defaults.

In Figure 1, we show the evolution of the indices when we move β between 0 and 1. If $0 \leq \beta < 0.5$, the decision maker gives priority to the company in state B. However, if $0.5 \leq \beta \leq 1$, Company 2 has priority over Company 1, including the case where Company 2 is in state G and Company 1 is in state B.

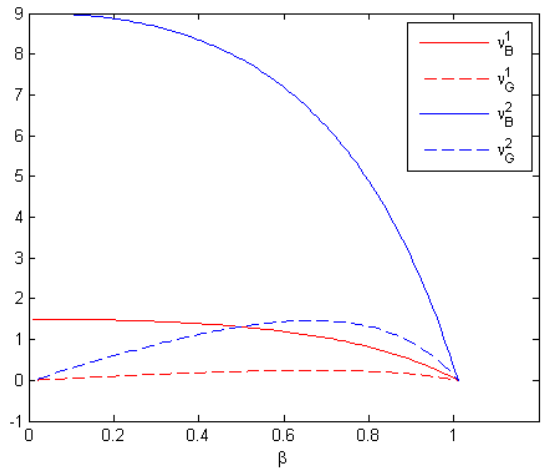

Figure 1: Comparison between indices

4.4 Comparison with ROI and Profitability Index

The *return-on-investment* (ROI) index is a fundamental and well-known performance measure often used in practice to evaluate investments and marketing expenditures. Broadly defined, ROI is the following measure:

$$\text{ROI} = \frac{\text{Gain from investment - Cost of investment}}{\text{Cost of investment}}. \tag{14}$$

In particular the single-period ROI for our model is

$$\text{ROI}_{k,G} + 1 = 0, \tag{15}$$

$$\text{ROI}_{k,B} + 1 = \frac{c_k d_k}{w_k}(1 - \mu_k). \tag{16}$$

It is interesting to note that the optimal myopic policy obtained from our proposed index in case $\beta = 0$ coincides with the single period ROI policy. Comparing $\text{ROI}_{k,B} + 1$ to the general index (12), we see that the latter contains an extra term θ_k, which can be interpreted as an *multi-period adjustment*.

Our index seems also related to the *profitability index* (PI), which is defined as the following measure:

$$\text{PI} = \frac{\text{Present value of future cashflows}}{\text{Initial investment}}. \tag{17}$$

It would be interesting to make an analogy and find out an interpretation for the terms in our proposed index.

Acknowledgements

The work of Urtzi Ayesta is partially supported by the French "Agence Nationale de la Recherche (ANR)" through the project ANR-15-CE25-0004 (ANR JCJC RACON). The work of Eva Ferreira, Martin Erausquin and Peter Jacko has been supported by MICINN (Ministerio de Ciencia e Innovación), Spain, through grant MTM2010-17405 regarding Peter Jacko and grants ECO2011-29268 and ECO2014-51914-P regarding Eva Ferreira and Martin Erausquin.

References

[1] V.V. Acharya and T. Yorulmazer. Cash-in-the-Market Pricing and Optimal Resolution of Bank Failures. *The Review of Financial Studies*, 21:2705–2742, 2008.

[2] K.J Arrow and M. Kurz. *Public Investment, The Rate of Return, and Optimal Fiscal Policy.* Johns Hopkins University Press, 1970.

[3] P. Bolton and C. Harris. Strategic Experimentation. *Econometrica*, 67:349–374, 1999.

[4] U. Dinesh Kumar and H. Saranga. Optimal Selection of Obsolescence Mitigation Strategies Using a Restless Bandit Model. *European Journal of Operational Research*, 200:170–180, 2010.

[5] J.S. Fons. Using Default Rates to Model the Term Structure of Credit Risk. *Financial Analysts Journal*, pages 25–33, Sep./Oct. 1994.

[6] J.C. Gittins, K. Glazebrook, and R. Weber. *Multi-armed Bandit Allocation Indices.* Wiley, 2011.

[7] K. Glazebrook, Mitchell H., and P. Ansell. Index Policies for the Maintenance of a Collection of Machines by a Set of Repairment. *European Journal of Operational Research*, 165:267284, 2005.

[8] G.M. Gupton, C.C. Finger, and M. Bhatia. Credit metrics, technical document. Technical report, JP Morgan, 1997.

[9] P. Jacko. *Dynamic Priority Allocation in Restless Bandit Models.* Lambert Academic Publishing, 2010. Invited book.

[10] R.A Jarrow, D. Lando, and S.M Turnbull. A Markov Model for the Term Structure of Credit Risk Spreads. *The Review of Financial Studies*, 10(2):481–523, 1997.

[11] G. Keller, S. Rady, and M. Cripps. Strategic Experimentation with Exponential Bandits. *Econometrica*, 73:39–68, 2005.

[12] D. Li and K. Glazebrook. A Bayesian approach to the Triage Problem with Imperfect Classification. *European Journal of Operational Research*, 215:169180, 2011.

[13] A. McLennan. Price Dispersion and Incomplete Learning in the Long-Run. *Journal of Economic Dynamics and Control*, 7:331–347, 1984.

[14] J. Niño-Mora. Dynamic priority allocation via restless bandit marginal productivity indices. *TOP*, 15(2):161–198, 2007.

[15] M. L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming.* John Wiley & Sons, 2005.

[16] Weitzman Roberts. Funding Criteria for Research, Development and Exploration of Projects. *Econometrica*, 49:1261–1288, 1981.

[17] M. Rothschild. A Two-Armed Bandit Theory of Market Princing. *Journal of Economic Theory*, 9:185–202, 1974.

[18] A. Rustichini and A. Wolinsky. Learning About Variable Demand in the Long Run. *Journal of Economic Dynamics and Control*, 19:1283–1292, 1995.

[19] M.L. Weitzman. Optimal Seach for the Best Alternative. *Econometrica*, 47:641–654, 1979.

[20] P. Whittle. Restless bandits: Activity allocation in a changing world. *Journal of Applied Probability*, 25:287–298, 1988.

Appendix: Notation

Let us define the following quantities:

- $\mathbb{E}_0^{\mathcal{S}}\left[\sum_{t=0}^{\infty}\beta^t R_{k,X_k(t)}^{a_k(t)}\right] = \mathbb{R}_n^{\mathcal{S}}$, being the initial state $X_k(0) = n \in \mathcal{N}$.
- $\mathbb{E}_0^{\mathcal{S}}\left[\sum_{t=0}^{\infty}\beta^t W_{k,X_k(t)}^{a_k(t)}\right] = \mathbb{W}_n^{\mathcal{S}}$, being the initial state $X_k(0) = n \in \mathcal{N}$.

The value function under policy $\mathcal{S}$ is given by

$$\mathbb{V}_n^{\mathcal{S}} = \mathbb{R}_n^{\mathcal{S}} - \nu\mathbb{W}_n^{\mathcal{S}}. \tag{18}$$

Remark 3. *Throughout the rest of the Appendices, we are solving a single-company subproblem that arises from the original optimization problem, say, for user k. Thus, in order to simplify the notation, we avoid writing the user subscript k.*

Appendix A: Proof of Proposition 3.1

In Proposition 3.1, we want to show that the optimal policy is of a threshold-type. Let us denote by $\mathbb{V}_n^*$ the value function following the optimal policy, under the state $n \in \mathcal{N}$. This value function satisfies the Bellman equation:

$$\mathbb{V}_n^* = \max_{a\in\mathcal{A}}\{R_n^a - \nu W_n^a + \beta\sum_{m\in\mathcal{N}} p_{n,m}^a \mathbb{V}_n^*\} \tag{19}$$

For the three different states, using (19), we have that the optimal value function in given by

$$\begin{aligned}\mathbb{V}_B^* = \max\{ &-\mu cd - \nu w + \beta\left(q_{B,G}^1\mathbb{V}_G^* + q_{B,B}^1\mathbb{V}_B^* + \mu d\mathbb{V}_D^*\right),\\ &-cd + \beta\left(q_{B,G}^0\mathbb{V}_G^* + q_{B,B}^0\mathbb{V}_B^* + d\mathbb{V}_D^*\right)\}\end{aligned} \tag{20}$$

$$\begin{aligned}\mathbb{V}_G^* = \max\{ &-\nu w + \beta\left(q_{G,G}^1\mathbb{V}_G^* + q_{G,B}^1\mathbb{V}_B^*\right),\\ &\beta\left(q_{G,G}^0\mathbb{V}_G^* + q_{G,B}^0\mathbb{V}_B^*\right)\}\end{aligned} \tag{21}$$

$$\mathbb{V}_D^* = \max\{-\nu w + \beta\mathbb{V}_D^*, \beta\mathbb{V}_D^*\}, \tag{22}$$

where the first option in the brackets represents the value earned by the economy if the action 1 is taken in the current period, and the second option in the brackets represents action 0.

1. First, assume that $\nu > 0$. Since $w > 0$, it is clear that it is optimal not to fund this company, and we have $\mathbb{V}_D^* = 0$. From (18), it is easy to check that, for every policy $\pi \in \Pi$, and for any initial state $n \in \mathcal{N}$, $\mathbb{V}_n^{\mathcal{S}} \leq 0$. Assume that it is optimal to fund a company in state G. Then, by (21),

$$\mathbb{V}_G^* = -\nu w + \beta\left(q_{G,G}^1\mathbb{V}_G^* + q_{G,B}^1\mathbb{V}_B^*\right). \tag{23}$$

   Now, assume that it is not optimal to fund in state B. Then, by (20)

$$\mathbb{V}_B^* = -cd + \beta\left(q_{B,G}^0\mathbb{V}_G^* + q_{B,B}^0\mathbb{V}_B^*\right). \tag{24}$$

   Our goal is to show that there is a contradiction in this case. (23) and (24) can be written as a system of two equations and two unknowns:

$$\begin{cases}-\beta q_{G,B}^1\mathbb{V}_B^* + \left(1-\beta q_{G,G}^1\right)\mathbb{V}_G^* = -\nu w\\ \left(1-\beta q_{B,B}^0\right)\mathbb{V}_B^* - \beta q_{B,G}^0\mathbb{V}_G^* = -cd\end{cases} \tag{25}$$

   Using the first equation in (25), we can write $\mathbb{V}_B^*$ in terms of $\mathbb{V}_G^*$ as follows:

$$\mathbb{V}_B^* = \frac{1-\beta q_{G,G}^1}{\beta q_{G,B}^1}\mathbb{V}_G^* + \frac{\nu w}{\beta q_{G,B}^1} \tag{26}$$

Since we are assuming that it is optimal to fund in state G, from (21), we have that $-\nu w+\beta\left(\left(q^1_{G,G}-q^0_{G,G}\right)\mathbb{V}^*_G+\left(q^1_{G,B}-q^0_{G,B}\right)\right.$ 0. Let us denote by $\Delta_{G,G}=q^1_{G,G}-q^0_{G,G}$ and $\Delta_{G,B}=q^1_{G,B}-q^0_{G,B}$. Then, we have $-\nu w+\beta\left(\Delta_{G,G}\mathbb{V}^*_G+\Delta_{G,B}\mathbb{V}^*_B\right)>$ 0. Since $q^1_{G,B}=1-q^1_{G,G}$ and $q^0_{G,B}=1-q^0_{G,G}$, $\Delta_{G,B}=q^1_{G,B}-q^0_{G,B}=1-q^1_{G,G}-1+q^0_{G,G}=-\Delta_{G,G}$, and therefore, we obtain the following inequality:

$$-\nu w+\beta\Delta_{G,G}\left(\mathbb{V}^*_G-\mathbb{V}^*_B\right)>0. \tag{27}$$

Using basic algebra, we obtain the simplified condition to be funded under state G,

$$-\frac{\Delta_{G,G}}{q^1_{G,B}}\left(\nu w+(1-\beta)\,\mathbb{V}^*_G\right)>\nu w. \tag{28}$$

Since we assume that it is not optimal to fund under state B, from Equation (20), $-\mu cd-\nu w+\beta\left(q^1_{B,G}\mathbb{V}^*_G+q^1_{B,B}\mathbb{V}^*_B\right)<$ $-cd+\beta\left(q^0_{B,G}\mathbb{V}^*_G+q^0_{B,B}\mathbb{V}^*_B\right)$. Following the same steps as we did for state G, and using basic algebra, we obtain the following inequality:

$$\begin{aligned}&\nu w+\tfrac{\Delta_{B,G}}{q^1_{G,B}}\left(\nu w+(1-\beta)\,\mathbb{V}^*_G\right)\\ -\quad & d\,(1-\mu)\left[c+\tfrac{\nu w}{q^1_{G,B}}-\tfrac{\left(1-\beta q^1_{G,G}\right)}{q^1_{G,B}}\mathbb{V}^*_G\right]\quad>0.\end{aligned} \tag{29}$$

Our goal is to prove that, if (28) holds, then there must be a contradiction in (29). If $\Delta_{G,G}=0$, then this implication is true. If $\Delta_{G,G}\neq 0$, then introducing $\Delta_{G,G}$ in previous expression, we obtain $\nu w+\frac{\Delta_{B,G}}{\Delta_{G,G}}\left(\frac{\Delta_{G,G}}{q^1_{G,B}}\left(\nu w+(1-\beta)\,\mathbb{V}^*_G\right)\right)-d\,(1-\mu)\left[c+\frac{\nu w}{q^1_{G,B}}-\frac{\left(1-\beta q^1_{G,G}\right)}{q^1_{G,B}}\mathbb{V}^*_G\right]>0$. However, it is easy to show that there must be a contradiction in last inequality, since (28) holds, and using the fact that $\Delta_{G,G}>0$ and $\Delta_{B,G}>0$, $\mathbb{V}^*_G\leq 0$ and $\frac{\Delta_{B,G}}{\Delta_{G,G}}>1$, we obtain $\nu w+\frac{\Delta_{B,G}}{\Delta_{G,G}}\left(\frac{\Delta_{G,G}}{q^1_{G,B}}\left(\nu w+(1-\beta)\,\mathbb{V}^*_G\right)\right)-d\,(1-\mu)\left[c+\frac{\nu w}{q^1_{G,B}}-\frac{\left(1-\beta q^1_{G,G}\right)}{q^1_{G,B}}\mathbb{V}^*_G\right]<0$, which leads to the contradiction. Thus, assuming it is optimal to fund a company under state G, then it is optimal to fund this company under state B as well.

2. If $\nu<0$, then, it is straightforward to see that the maximizing policy gives $a^\pi(t)=1$ for all t, since the trade-off related to the optimization problem disappears. Funding at the period maximizes $\mathbb{R}^{\mathcal{S}}_n$, because the expected default time for the company is maximized. On the other hand, $\mathbb{W}^{\mathcal{S}}_n$ is also maximized, since it adds w at any period. Together with $\nu<0$, it is obvious to see that the value function is maximized as well.

Appendix B: Proof of Theorem 3.2

From definitions (4.4) and (4.4) we have $\mathbb{R}^{\mathcal{S}}_n=R^{n\in\mathcal{S}}_n+\beta\sum_{m\in\mathcal{N}}p^{n\in\mathcal{S}}_{n,m}\mathbb{R}^{\mathcal{S}}_m$ and $\mathbb{W}^{\mathcal{S}}_n=W^{n\in\mathcal{S}}_n+\beta\sum_{m\in\mathcal{N}}p^{n\in\mathcal{S}}_{n,m}\mathbb{W}^{\mathcal{S}}_m$. To simplify the notation, the expression $n\in\mathcal{S}$ is used in the superindex instead of the indicator function $I_{\mathcal{S}}(n)$. That is, $n\in\mathcal{S}$ equals 1 if true and 0 otherwise. Substituting the values of $R^{n\in\mathcal{S}}_n$, $W^{n\in\mathcal{S}}_n$ and $p^{n\in\mathcal{S}}_{n,m}$ and simplifying, we get the following *balance equations:*

Lemma 4.1.

$$\mathbb{R}^{\mathcal{S}}_G=\begin{cases}\beta\left(p^1_{G,G}\mathbb{R}^{\mathcal{S}}_G+p^1_{G,B}\mathbb{R}^{\mathcal{S}}_B\right) & \text{if } G\in\mathcal{S},\\ \beta\left(p^0_{G,G}\mathbb{R}^{\mathcal{S}}_G+p^0_{G,B}\mathbb{R}^{\mathcal{S}}_B\right) & \text{if } G\notin\mathcal{S},\end{cases} \tag{30}$$

$$\mathbb{R}^{\mathcal{S}}_B=\begin{cases}-\mu cd+\beta\left(p^1_{B,G}\mathbb{R}^{\mathcal{S}}_G+p^1_{B,B}\mathbb{R}^{\mathcal{S}}_B\right) & \text{if } B\in\mathcal{S},\\ -cd+\beta\left(p^0_{B,G}\mathbb{R}^{\mathcal{S}}_G+p^0_{B,B}\mathbb{R}^{\mathcal{S}}_B\right) & \text{if } B\notin\mathcal{S},\end{cases} \tag{31}$$

$$\mathbb{W}^{\mathcal{S}}_G=\begin{cases}w+\beta\left(p^1_{G,G}\mathbb{W}^{\mathcal{S}}_G+p^1_{G,B}\mathbb{W}^{\mathcal{S}}_B\right) & \text{if } G\in\mathcal{S},\\ \beta\left(p^0_{G,G}\mathbb{W}^{\mathcal{S}}_G+p^0_{G,B}\mathbb{W}^{\mathcal{S}}_B\right) & \text{if } G\notin\mathcal{S},\end{cases} \tag{32}$$

$$\mathbb{W}^{\mathcal{S}}_B=\begin{cases}w+\beta\left(p^1_{B,G}\mathbb{W}^{\mathcal{S}}_G+p^1_{G,B}\mathbb{W}^{\mathcal{S}}_B\right) & \text{if } B\in\mathcal{S},\\ \beta\left(p^0_{B,G}\mathbb{W}^{\mathcal{S}}_G+p^0_{B,B}\mathbb{W}^{\mathcal{S}}_B\right) & \text{if } B\notin\mathcal{S},\end{cases} \tag{33}$$

Proof. Trivial from the definitions of $\mathbb{R}_n^{\mathcal{S}}$ and $\mathbb{W}_n^{\mathcal{S}}$, for any $n \in \mathcal{N}$. □

Following the methodology given in [14], we will study ν_n under all policies $\mathcal{S}$, defined as $\nu_n^{\mathcal{S}} := \frac{\mathbb{R}_n^{\langle 1,\mathcal{S}\rangle} - \mathbb{R}_n^{\langle 0,\mathcal{S}\rangle}}{\mathbb{W}_n^{\langle 1,\mathcal{S}\rangle} - \mathbb{W}_n^{\langle 0,\mathcal{S}\rangle}}$, where by $\langle 1, \mathcal{S}\rangle$ ($\langle 0, \mathcal{S}\rangle$) we refer to the policy that funds (does not fund) in the initial period and follows the strategy $\mathcal{S}$ from that moment on.

Our objective is to find the optimal index values ν_n in terms of Theorem 3.2. In other words, for each state $n \in \mathcal{N}$, we have to find the action set $\mathcal{S}$ such that $\nu_n = \nu_n^{\mathcal{S}}$. This action set will be given by the following lemma where we use the notation $S_{n:\mathcal{N}}$ to refer to the subset of $\mathcal{N}$ that are bigger than n.

Lemma 4.2. *For all $n \in \mathcal{N}$, $\nu_n = \nu_n^{S_{n:\mathcal{N}}}$, where $S_{n:\mathcal{N}} = \{m \in \mathcal{N} : m > n\}$.*

Proof. A sufficient condition to verify Lemma 4.2 is to show the LP-indexability, as described in [14], which in our problem can be simplified to the following:

Definition 1. *Problem (5) is LP-indexable if the following conditions hold:*

1. $\mathbb{W}_n^{\langle 1,\emptyset\rangle} - \mathbb{W}_n^{\langle 0,\emptyset\rangle} \geq 0$, *and* $\mathbb{W}_n^{\langle 1,\mathcal{N}\rangle} - \mathbb{W}_n^{\langle 0,\mathcal{N}\rangle} \geq 0$, *for all* $n \in \mathcal{N}$.
2. $\mathbb{W}_n^{\langle 1,S_{n:\mathcal{N}}\rangle} - \mathbb{W}_n^{\langle 0,S_{n:\mathcal{N}}\rangle} \geq 0$, *and* $\mathbb{W}_{n+1}^{\langle 1,S_{n:\mathcal{N}}\rangle} - \mathbb{W}_{n+1}^{\langle 0,S_{n:\mathcal{N}}\rangle} \geq 0$, *for all* $n \in \mathcal{N}$.
3. *For every real-valued ν there exists $n \in \mathcal{N}$ such that the policy defined by the set $S_{n:\mathcal{N}}$ is optimal.*

It is easy to check that the conditions given in Definition 1 hold in our case. The first two conditions are straightforward and the third is already proved in Lemma 4.2. □

We use the action sets given in Lemma 4.2 to obtain the desired indices in terms of Theorem 3.2. First, we calculate the index value of a company when its state is B. From Lemma 4.2, we know that $\nu_B = \nu_B^{S}$, with $S = \emptyset$. This policy does not allocate any resources, no matter which the state of the company is, and we have that

$$\nu_n^{\emptyset} := \frac{\mathbb{R}_n^{\langle 1,\emptyset\rangle} - \mathbb{R}_n^{\langle 0,\emptyset\rangle}}{\mathbb{W}_n^{\langle 1,\emptyset\rangle} - \mathbb{W}_n^{\langle 0,\emptyset\rangle}}. \tag{34}$$

First, it is clear that:

$$\mathbb{R}_B^{\langle 1,\emptyset\rangle} - \mathbb{R}_B^{\langle 0,\emptyset\rangle} = (1-\mu)cd + \beta\left(\left(p_{B,G}^1 - p_{B,G}^0\right)\mathbb{R}_G^{\emptyset} + (p_{B,B}^1 - p_{B,B}^0)\mathbb{R}_B^{\emptyset}\right). \tag{35}$$

On the other hand, we also have $\mathbb{W}_B^{\langle 1,\emptyset\rangle} - \mathbb{W}_B^{\langle 0,\emptyset\rangle} = w$. The index value will be obtained solving the systems obtained from the balance equations:

$$\begin{cases} \mathbb{R}_B^{\emptyset} = -cd + \beta\left(p_{B,G}^0\mathbb{R}_G^{\emptyset} + p_{B,B}^0\mathbb{R}_B^{\emptyset}\right) \\ \mathbb{R}_G^{\emptyset} = \beta\left(p_{G,G}^0\mathbb{R}_G^{\emptyset} + p_{G,B}^0\mathbb{R}_B^{\emptyset}\right) \end{cases} \tag{36}$$

and we obtain that $\mathbb{R}_B^{\langle 1,\emptyset\rangle} - \mathbb{R}_B^{\langle 0,\emptyset\rangle} = (1-\mu)cd + \beta\left(\frac{\left(p_{B,G}^1 - p_{B,G}^0\right)\beta p_{G,B}^0}{1-\beta p_{G,G}^0} + \left(p_{B,B}^1 - p_{B,B}^0\right)\right)\mathbb{R}_B^{\emptyset}$. From (34), we obtain the following index:

$$\nu_B = \frac{(1-\mu)\,cd}{w} - \frac{\beta cd\left(\frac{\left(p_{B,G}^1 - p_{B,G}^0\right)\beta p_{G,B}^0}{1-\beta p_{G,G}^0} + \left(p_{B,B}^1 - p_{B,B}^0\right)\right)}{\left(1 - \beta\left(\frac{\beta p_{B,G}^0 p_{G,B}^0}{1-\beta p_{G,G}^0} + p_{B,B}^0\right)\right)w}, \tag{37}$$

that can be written as $\nu_B = cd\,(1-\mu-\theta)\,/w$, where

$$\theta = \frac{\beta\left(\frac{\left(p_{B,G}^1 - p_{B,G}^0\right)\beta p_{G,B}^0}{1-\beta p_{G,G}^0} + \left(p_{B,B}^1 - p_{B,B}^0\right)\right)}{1 - \beta\left(\frac{\beta p_{B,G}^0 p_{G,B}^0}{1-\beta p_{G,G}^0} + p_{B,B}^0\right)}. \tag{38}$$

Second, we calculate the index value when the company is in state G. In this case, from Lemma (4.2), it is clear that

$$\nu_G = \nu_G^{\{B\}} = \frac{\mathbb{R}_G^{\langle 1,\{B\}\rangle} - \mathbb{R}_G^{\langle 0,\{B\}\rangle}}{\mathbb{W}_G^{\langle 1,\{B\}\rangle} - \mathbb{W}_G^{\langle 0,\{B\}\rangle}}. \tag{39}$$

First, we write (39) in terms of $\mathbb{R}_B^B$, $\mathbb{R}_G^B$, $\mathbb{W}_B^B$ and $\mathbb{W}_G^B$. Following similar steps as in the first case, we obtain,

$$\nu_G = \frac{\mathbb{R}_G^{\langle 1,B\rangle} - \mathbb{R}_G^{\langle 0,B\rangle}}{\mathbb{W}_G^{\langle 1,B\rangle} - \mathbb{W}_G^{\langle 0,B\rangle}} = \frac{\mu c d}{w}\frac{\phi}{1-\phi}, \tag{40}$$

where

$$\phi = -\frac{\beta\left(\left(p_{G,G}^1 - p_{G,G}^0\right)\frac{\beta p_{G,B}^0}{1-\beta p_{G,G}^0} + \left(p_{G,B}^1 - p_{G,B}^0\right)\right)}{1-\beta\left(\frac{\beta p_{B,G}^1 p_{G,B}^0}{1-\beta p_{G,G}^0} + p_{B,B}^1\right)}. \tag{41}$$