How to Pick a Winner*

Andrea Mattozzi† Fabio Michelucci‡
EUI CERGE-EI

August 28, 2019

Abstract

We study a two-period model in which two agents of different and privately-known abilities compete for a single prize by exerting costly effort. A risk-neutral principal can affect the outcome of the contest by allocating resources to agents in each period, and her net payoff depends on the relative share allocated to the winner. We analyze how the principal optimal strategy depends on the severity of moral hazard. The results we derive are consistent with stylized facts regarding the dynamics of political campaign contributions and of third-party intervention in military conflicts.

Keywords: Dynamic Games, Contests, Experimentation, Lobbies, Campaign Contributions
JEL Classification: D72 D78 C72 C73

*An earlier version of this paper has circulated under the title "Electoral Contests with Dynamic Campaign Contributions". The study was supported by the Grant Agency of the Czech Republic (GACR 13-35452S). All opinions expressed are those of the authors and have not been endorsed by GACR. We would like to thank Marco Battaglini, Alessandra Casella, Massimo Morelli, Dana Sisak, and Jan Zápal for very helpful suggestions and Veronika Vraná for excellent research assistance.

†Department of Economics, European Economic Institute, Florence, Italy, Email: andrea.mattozzi@eui.eu.

‡Corresponding author. CERGE-EI, Charles University and the Academy of Sciences, Prague. CERGE-EI, a joint workplace of Charles University and the Economics Institute of the Czech Academy of Sciences, Politickych veznu 7, P.O. Box 882, 111 21 Prague 1, Czech Republic. Telephone (+420) 224-005-119; Fax: (+420) 224-005-117; E-mail: fabio.michelucci@cerge-ei.cz.

"I am right now in the learning phase and I'm trying to learn about the candidates, learn about their thinking, their capabilities of being president." Ken Abramowitz, venture capitalist contributing to 6 Republican candidates in the 2016 U.S. Presidential primary. Time Magazine, July 2015.

1 Introduction

Contests rarely take place in isolation. When interested third-parties have a chance of influencing the outcome, they often take it and their intervention assume a variety of forms depending on the type of contest. A large literature in economics and political science has studied under various angles the competition among two or multiple contestants for a prize but its focus has always been on the agents directly involved in the contest.[1] The novel contribution of this paper is to provide a tractable framework to analyze the impact of third parties intervention in contests.

Applications abound. In the case of elections, for example, interest groups and donors actively contribute to affect the electoral outcome by supporting one or more candidates and by switching support between them and across time as new information arrives. Likewise, the outcome of a military conflict is affected by third parties strategically taking and changing sides over the course of the conflict (Regan (2000)). In a labor market environment where workers compete for promotions, the performance measure determining the selection might be affected by how much time and resources the employer has invested in any given worker. In all these examples, agents are by and large motivated by winning a contest, they rely on external resources, and the reason and intensity of a third party's intervention is related to the benefits that can be seized from having identified and supported the ultimate winner.

We focus on dynamic contests characterized by a high initial uncertainty regarding the contestants' abilities, and such that the payoff of a third party is increasing in the total amount of resources allocated to the winner. Specifically, one risk-neutral principal decides how to allocate a fixed budget between two risk-neutral agents. Agents compete in a two-period contest whose outcome depends on the amount of

[1] See, e.g., Corchon (2007) for a survey.

resources received in each period, agents' ability, and their costly effort. We assume that effort is observable but non-contractible, and that agents rely exclusively on the principal's budget. While the winner is determined at the end of the contest, intermediate outcomes convey information about agents' abilities.

Under these assumptions, a natural trade-off between experimentation and selection arises. On the one hand, conditional on selecting the winner the third party has an incentive to concentrate resources on a single agent in the last period. On the other hand, the fact that agents are heterogenous in their ability and that intermediate outcomes are informative about them pushes towards "testing" the contestants in the first period. Furthermore, if the principal decides to side with one agent in the first period, she faces a *locked-in effect*, i.e. the payoff from switching to the other agent in the second period diminishes the larger the first period budget allocation is. How is the trade-off between experimentation and selection resolved?

We show that the principal optimal strategy depends on the severity of moral hazard. In the absence of moral hazard, the optimal strategy is to allocate a fraction of the budget to one agent in the first period and then switch resources to the other upon observing unfavorable information. The presence of moral hazard, however, constraints the set of first period allocations of budget inducing a given effort. To see why, consider a strategy profile in which both agents exert effort and the principal allocates budget to one agent only in the first period. If this agent is a low type, i.e. with a low chance of producing favorable information, the ex-ante most likely outcome will be that the principal, upon receiving unfavorable information, will allocate the remaining budget to the agent that did not receive any initially. But then, given the low chances of being allocated the residual budget, the low type prefers not to exert effort in the first place.

In equilibrium, if the cost of exerting effort is low, the principal allocates resources to *both* agents such that they pool in exerting effort in the first period. In the second period she allocates the remaining resources to one agent only based on her posterior. If instead the cost of effort is intermediate, separating agents is optimal but it is still achieved by allocating resources to both agents in the first period. Finally, if the cost of effort is high enough, the principal allocates all the budget to one agent in the last period giving up on experimentation.

In summary, the key features of our dynamic model are that in equilibrium, and despite risk neutrality, the principal may either allocate resources to both agents in the first period, which would resemble "hedging" or allocate a share of the budget to one agent and, depending on first period outcomes, switch to the other agent in the last period. As we discuss in the next section it is not rare that in the context of US Presidential primary elections or in military conflicts third parties choose to support more contestants early on and switch sides across time. Such strategies would hardly be distinguishable if one were to interpret the data using a static model, and would require ad-hoc assumptions on risk preferences.[2] In this sense our paper naturally complements earlier literature on political campaign contributions that has focused on static settings (see, for example, Prat (2002a)).

The rest of the paper is organized as follows. Section 2 presents our motivating evidence. Section 3 introduces the model. Section 4 solves for the second period allocation. Section 5 characterizes the optimal first period allocation in a benchmark model without moral hazard, and Section 6 is devoted to the moral hazard case. In Section 7 we discuss our key assumptions and Section 8 relates our work to the existing literature. Section 9 concludes.

2 Motivating Evidence

In this section we present stylized facts on third-party intervention in political campaigns and in military conflict. The intervention strategies of third-parties in these contests are quite sophisticated and appear to be consistent with equilibrium features of a dynamic model with endogenous learning.

Political Campaign Contributions. A defining feature of the process of campaign giving is that it is inherently dynamic. Interest groups and donors strategically

[2]Notice that risk diversification alone would not be enough to account for such behaviors. Since third parties often modify their allocation of resources over time, assuming specific risk preferences within a dynamic model would not be enough to reconcile theory and evidence unless new information arrives over time. But if new information arrives exogenously, the principal should optimally allocate her resources only in the last period. Our theory, based on endogenous learning, does not require risk-aversion.

choose to whom and when to contribute during the course of a political campaign.[3] Furthermore, it is by no means an exception that donors contribute to multiple candidates and sometimes even opposing candidates at different stages of the electoral cycle. For example, in the 2012 U.S. Presidential election Deloitte LLP was among the top ten contributors to both Barack Obama (43% of Deloitte total contributions) and Mitt Romney (57%).[4]

A richer picture suggesting the systematic adoption of non-trivial strategies in campaign giving emerges if we look at donations from all Political Action Committees (PACs) in U.S. Presidential Elections in the last thirty-five years as reported from the Federal Election Commission.[5] In particular, focusing on relatively large PACs, four key facts stand out.[6] First, the timing of political contributions displays a large variation: While a number of PACs start contributing as early as twenty months before the elections, contributions usually peak in the three months just preceding the elections. Second, if we split the campaign cycle into two roughly equal subperiods - we will henceforth refer to them as early and late campaign- the majority of PACs contribute in both periods.[7] Third, early contributions are particularly common in primaries. Fourth, the share of PACs switching support between candidates and the share of PACs "experimenting" - contributing to multiple candidates in the early campaign and then supporting a unique candidate in the late campaign - is substantial and much higher in primaries than in general elections.[8]

[3]While we abstract from potential competition among multiple interest groups, we believe that our approach represents a useful starting point especially since different interest groups may specialize on particular political issues. We further discuss this point in Section 3 and Section 9.

[4]Similarly, Citygroup Inc. was among the top ten contributors to both Barack Obama (70%) and John McCain (30%) in 2008. Data are taken from www.opensecrets.org.

[5]One of the first papers providing empirical evidence of PACs' strategic behavior in campaign giving is Stratmann (1992).

[6]For all PACs contributing within a specific electoral cycle, we compute the total number of donations, order them, and focus on PACs with total contribution in the top half of the distribution, which amounts to more than two-hundred PACs for every electoral cycle.

[7]We describe in the Appendix how we compute these breaks, which are reported in Table 1.

[8]For example, averaging over the electoral cycles 1980-2016, 78.2% of PACs contributing in both periods supported only one candidate in the early phase of primaries. Of those, 17.7% switched their contributions to a different candidate later. The corresponding figures for general elections are 81.7% and 6.5%, respectively. Furthermore, 80.3% of PACs contributing in both periods supported only one candidate in the late phase of primary campaign. Of those 11.2% of PACs started contributing to multiple candidates in the early phase. The corresponding figures for general elections are 71.8%

In line with the earlier quote from Ken Abramovitz, data suggest that donors are willing to experiment in primaries. Furthermore, the use of strategies consistent with our theory becomes more prevalent as the size of PACs increases.

Third-Party Intervention in Military Conflicts. In military conflicts third parties can and do influence the outcome of the conflict by strategically taking and changing sides. Powell (2017) reports that at least one state intervened in an ongoing war in 27% of wars between 1816 and 2007. Furthermore, as documented in Regan (2000), third parties prefer to be on the winning side and are equally likely to intervene in a civil conflict on behalf of government or opposition. As for the propensity of many third parties to change sides during the conflict, anecdotal evidence abounds. In the Afghan history, warring leaders were repeatedly switching support between the Taliban and government (Christia (2012) and Powell (2017)). During the Congo crisis in the sixties, Belgium initially supported the opposition forces of the Katanga district fighting for independence from Zaire and subsequently switched its support in favor of the government. Similarly, Cuba and the Soviet Union switched their support from opposition to government in the thirty years long war of independence that Eritrea fought against Ethiopia (Regan (2000)).

3 The Model

Players and Actions

One principal P (she) and two agents A_i (he) $i \in \{1, 2\}$ compete for a prize in a contest that lasts two periods, $t = 1, 2$. All players are risk neutral. The principal chooses how much of her total budget $B > 1$ to allocate to each agent across the two periods.[9] We call $B_i^t \in [0, B]$ the observable amount allocated to agent i in period t, and we define $B_i \equiv B_i^1 + B_i^2 \leq B$ to be the total allocation to agent i, and by $B^t \equiv \{B_1^t, B_2^t\}$ the vector of allocations at time t.

and 5.8%, respectively. For details see the Appendix.

[9] Assuming $B > 1$ is needed for experimentation to occur in equilibrium. If the total budget is small, the principal will never invest in experimentation.

Agents choose whether to exert effort or not in each period, where $e_i^t = \{0, 1\}$ denotes agent i's effort in period t, and $e^t \equiv \{e_1^t, e_2^t\}$ is the vector of chosen efforts at time t. Effort costs C to each agent, and we assume that is ex-post observable but it is not contractible. Conditional on exerting effort, in each period each agent may produce either a low output or a high output so that $O_i^t \in \{0, 1\}$ denotes the output of agent i in period t. We define $O^t \equiv \{O_1^t, O_2^t\}$ the vector of realized outputs at time t. The probability that A_i produces a high output in period t depends on his private ability or type $\theta_i = \{\theta_L, \theta_H\}$, on his effort decision e_i^t, and on the principal's allocation B_i^t.[10] Specifically, we assume that when A_i receives an amount of resources B_i^t, he produces high output with probability $e_i^t \theta_i f(B_i^t)$, where $\theta_i f(B) < 1$. The function $f(\cdot)$ is strictly increasing and strictly concave, and such that $f(0) = 0$ and $f'(0) \to \infty$.[11] The winner of the contest is the agent that produces high output in the second period.[12] In case of a tie, each agent has an equal probability of winning.

We further assume that one agent has a high type, one a low type, and that the principal has a symmetric prior probability that each agent is the high type (types are private information). Finally, for simplicity we assume that $\theta_H = \theta > 0$ and $\theta_L = 0$. This case provides all the intuition for our main mechanism, simplifies the notation, and it generalizes to the case where θ_L is small and positive.

Timing

The timing of the game is as follows:

1. Agents observe their private type

[10] In case of elections, one can interpret outputes as polls conducted before the election and on the election day, which generate binary signals. The cost of exerting effort could represent the political cost associated to the endorsement of a policy put forward by a particular donor. Along these lines, agents' abilities can be interpreted as the capacity of the candidate to successfully and credibly embrace the policy issue championed by a particular donor.

[11] Concavity of f is needed for an interior solution. Our results would not change if $f(0)$ was positive but small enough. If $f(0)$ is large, instead, we conjecture that a "waiting strategy" in which the principal free rides on agents in the first period may be optimal.

[12] This assumption implies that first period output may only affect the principal's beliefs. While relaxing it does not qualitatively affect our results, it seems a reasonable assumption. For example, in electoral campaigns a poll provides information but the winner is determined solely by the actual election results.

2. Principal chooses first-period allocation of resources
3. Agents choose simultaneously whether to exert effort or not
4. Outputs and efforts are observed
5. Principal chooses second-period allocation of resources
6. Agents choose simultaneously whether to exert effort or not
7. Outputs and efforts are observed
8. Payoffs are collected

Payoffs

The utility of the principal depends on her allocation of resources to the winning agent relative to the losing one. Specifically, the payoff for the principal is

$$\Pi(B_i, O_i^2) \equiv \begin{cases} g(B_i)W & \text{if } O_i^2 = 1 \text{ and } B_i \geq B/2 \\ 0 & \text{otherwise,} \end{cases}$$

where $g(B_i) : B_i \to [0, 1]$ is a continuous function with $g(B_i) = 0$ for $B_i < B/2$ and, for $B_i \geq B/2$, $g(B_i)$ is concave, strictly increasing and such that $g(B) = 1$. The interpretation is that the principal can receive a prize from the winning agent, and this is worth $W > 0$ to her. However, the probability of receiving the prize (or, equivalently, its size) depends on the principal relative budget allocation to agents, and it is described by the $g(\cdot)$ function.[13] In particular this probability is positive only if the principal is actually "taking sides" and it increases the larger is the budget allocated to the winning side.[14] We assume for simplicity and without any real loss of generality that in case of a tie in the contest the principal does not receive the prize.

[13] The fact that $g(\cdot)$ is not contractible seems reasonable for the applications we are interested in. As for concavity, the ability of a politician in office to please its donors is naturally constraint by a number of factors, e.g. political costs.

[14] This assumption rules out somewhat uninteresting cases in which the possibility that the principal chooses to "hedge her bets" would be a straightforward consequence of the payoff function. All our analysis would go through, if the $g(\cdot)$ was instead continuous but close to zero for $B_i < B/2$ and with a sharp increase at $B/2$.

The utility of an agent depends on whether he wins the contest or not and on the effort exerted. Specifically, the payoff for a winning agent i is

$$V - (I_i^1 + I_i^2)C,$$

where V is the payoff of winning the contest and $I_i^t = 1$ if A_i exerted effort in period t and zero otherwise. On the other hand, the losing agent $(-i)$ gets $-(I_{-i}^1 + I_{-i}^2)C$.

Finally, we denote by $\lambda_i(O^1, B^1, e^1) \in [0, 1]$ the principal's posterior belief at the end of period $t = 1$ that A_i is type θ, i.e. the high type. Note that if A_1 is type θ, the set of first period outcomes that may realize is $\{\{0, 0\}, \{1, 0\}\}$, while if it is A_2 the set is $\{\{0, 0\}, \{0, 1\}\}$. Hence, the only case in which the principal does not have deterministic beliefs about the agents is when agents pool on exerting effort and the outcome is $\{0, 0\}$. Our equilibrium concept is PBE.[15]

4 A Model of Experimentation and Selection

In our model the inter-temporal decision of the principal entails a fundamental trade-off. An early allocation of resources to agents generates valuable information since intermediate outcomes are more informative the larger the budget allocated early is. The flip side of the coin is that the more resources are allocated early, the less is available to win the contest in the second period. Furthermore, if the principal decides to side with one agent in the first period, she needs to consider that the payoff from switching to the other agent in the second period diminishes in the amount already allocated. We refer to this strategic effect as *locked-in* effect. In short the principal faces a trade-off between experimentation and selection.

Notice that all action takes place in the first period of our model, the one in which experimentation may take place. Indeed, the optimal strategy in the second period does not entail any trade-off.[16]

[15] In our model both the relative likelihood of reaching the two possible outcomes and the posterior conditioning on not observing any success change smoothly with a change of the budget allocation. Similarly, the expected prize for the principal changes smoothly with a change of the budget allocation. This smoothness allows a simple (per-period) binary-outcome, binary-effort model to capture the key strategic trade-off between experimentation and selection.

[16] All omitted proofs can be found in the Appendix.

Lemma 1. *The optimal strategy in the second period is to allocate all unspent budget to one agent only. In particular, the principal allocates to A_i if $\lambda_i(O^1, B^1, e^1) > g(B - B_i^1)/(g(B - B_i^1) + g(B - B_{-i}^1))$. When indifferent she allocates to either agent with equal probability.*

The result follows from the second period choice being equivalent to selecting which agent receives the largest fraction of total budget, and from noticing that once the selection is made, the principal payoff increases in the amount allocated to that agent.

In what follows we refer to a *lopsided* (allocation) strategy if the principal allocates a strictly positive amount only to one agent in the first period. We refer to a *hedging* strategy if the principal allocates a strictly positive amount to both agents in the first period. Within hedging strategies we further distinguish between *symmetric hedging* when the principal allocates a strictly positive and equal amount to both agents, and *asymmetric hedging* otherwise. Finally, we define a *betting* strategy as a degenerate lopsided strategy such that the principal allocates the entire budget to one agent in the second period. We also distinguish between *separating* and *pooling* equilibria. In a pooling equilibrium both agents exert effort in the first period, while in a separating equilibrium only the high type agent does. Recall that in a pooling equilibrium the principal posterior is based on the bayesian updating given first period information, whereas in a separating equilibrium the principal can perfectly infer agents private information from observed effort levels. Since the principal can control via her budget allocation whether to induce separation or not, we need to characterize which of the two alternatives way of learning is better for her.

We are now ready to consider the first period. We start from a benchmark case in which moral hazard is absent. Then, we consider the case in which the principal must provide incentives for the agents to exert effort.

5 Experimentation in the Benchmark Case ($e_i^t = 1$)

In the benchmark model if the principal decides to give an amount B_i^t to A_i in period t, she has a probability of observing A_i producing a high output - equal to $\theta f(B_i^t)$ - if and only if A_i is the high type. In short, agents effort levels are exogenously set equal to 1. We already solved for the second-period optimal decision conditional on any

first-period allocation of budget and on the intermediate outcome. Next, in Lemma 2 and Lemma 3 we prove a number of properties of the optimal first-period decision. Proposition 1 characterizes the equilibrium.

Lemma 2. *It is never optimal for the principal to allocate budget to both agents in the first period.*

While intuitive, Lemma 2 is not an obvious result. In fact, a countervailing effect of using of a lopsided strategy in the first period (as opposed to hedging) is that, for a fixed total amount allocated in the first period, it yields a lower probability of observing a high output.[17] In the proof of the lemma we show that for any hedging strategy there exists a lopsided one that allocates the same amount to one of the agents in the first period and nothing to the other agent. This strategy dominates the hedging strategy. In fact, whenever the hedging strategy allocates all remaining budget to a given agent in the second period so does the alternative lopsided strategy. However, the lopsided strategy allocates to such agent the same total budget in some continuation play and strictly more in others. Intuitively, the key aspect is that the amount that the lopsided "saves" from the first period is allocated under superior information in the second period.

Now we turn to the determination of the optimal lopsided allocation by first establishing the existence of an upper bound on how much to allocate in the first period.

Lemma 3. *A necessary condition for a lopsided strategy B_i^1 to be optimal is $B_i^1 < \hat{B}_i^1$, where $\hat{B}_i^1 \in (0, B/2)$ is unique and increasing in θ and B.*

If the principal sides with one agent in the first period, it must be that following a $(0, 0)$ outcome in the first period she switches into allocating the remaining budget to the other agent in the second period. This is because if the principal does not switch after $(0, 0)$ she would never, which would be dominated by flipping a coin and giving

[17] To see why, fix the total amount of first period budget allocation to B^*, and assume w.l.o.g that an hedging strategy allocates a fraction x of it to A_1 and $1 - x$ to A_2. The probability of observing high output is then $\theta/2(f(xB^*) + f((1 - x)B^*)$ under the hedging strategy and $\theta/2(f(B^*))$ under the lopsided strategy, with $\theta/2(f(xB^*) + f((1 - x)B^*) > \theta/2(f(B^*))$ by concavity of $f(\cdot)$.

the entire budget to one agent in the second period. Essentially, $B_i^1 < \hat{B}_i^1$ must hold because the principal anticipates she would be locked-in otherwise.

Consider a lopsided allocation of B_i^1 to agent A_i in the first period. The expected payoff of the principal after a $(0,0)$ if she keeps allocating to A_i in the second period is equal to[18]

$$\lambda_i \left((0,0),(0,B_i^1)\right) \theta f(B - B_i^1)W$$

where

$$\lambda_i = \frac{1-\theta f(B_i^1)}{2-\theta f(B_i^1)} < \frac{1}{2}$$

is the principal's posterior belief that A_i is the productive agent conditional on the observed outcome and given the allocation $(0, B_i^1)$. On the other hand, the expected payoff of the principal after $(0,0)$ if she switches and allocates to A_{-i} in the second period is equal to

$$(1-\lambda_i)\theta f(B - B_i^1) g(B - B_i^1)W$$

It is immediate to see that when B_i^1 approaches $B/2$ switching is never optimal. On the other hand, when B_i^1 approaches zero, both payoffs become identical. The condition for the principal to switch after $(0,0)$ is

$$\lambda_i = \frac{1-\theta f(B_i^1)}{2-\theta f(B_i^1)} < \frac{g(B - B_i^1)}{1+g(B - B_i^1)}$$

which can be rewritten as

$$1-\theta f(B_i^1) < g(B - B_i^1)$$

and, as it is shown in the proof of Lemma 3, it holds if and only if $B_i^1 < \hat{B}_i^1 < B/2$. Notice that the LHS of the inequality above is the probability of committing a type II error (i.e., switching to the unproductive agent in the last period), and the RHS is the benefit of switching which is inversely related to the locked-in effect.

A large $\hat{B}_i^1$ implies that the range of first period allocations for which, following $(0,0)$ it is optimal to switch sides is also large. The comparative statics with respect to θ and B have a clear interpretation. If B is larger, more resources can be allocated

[18] We omit effort levels as they are exogenously set to 1.

following $(0,0)$, which increases the value of switching sides. Similarly, a larger θ makes switching sides more attractive because it implies both a better posterior after observing $(0,0)$ as well as a higher chance of generating high output conditional on switching to the high type. Combining the previous results we can characterize the equilibrium of the benchmark model in the next proposition.

Proposition 1. *There exists a $\tilde{B}_i^1 \in (0, \hat{B}_i^1)$ such that, in all equilibria of the game, the principal chooses a lopsided strategy. In the first period the principal allocates $\tilde{B}_i^1$ randomly to A_i and then keeps allocating the remaining budget to A_i if and only if she observes high output.*

The optimal strategy is the solution to the following problem

$$\max_{B_i^1 \in [0, \hat{B}_i^1]} \frac{\theta W}{2} \left(\theta f(B_i^1) f\left(B - B_i^1\right) + f\left(B - B_i^1\right) g\left(B - B_i^1\right) \right)$$

which yields the unique $\tilde{B}_i^1$ that resolves optimally the trade-off between experimentation and selection. The equilibrium is unique up to the initial randomization (due to the equal prior assumption). Most of the results stated in Proposition 1 follow as a straightforward corollary of the results shown in the previous lemmas. In addition, $f'(0) = \infty$ guarantees that $\tilde{B}_1^i > 0$, which establishes that a lopsided strategy dominates a betting strategy ($B_1^i = 0$). Finally, the comparative statics of the optimal level of first period budget allocation with respect to θ follow from IFT with $\tilde{B}_1^i$ being increasing in θ.[19]

Notice that the optimal allocation of the benchmark model with exogenous efforts can be implemented if efforts were endogenous *and* contractible. In particular, the following mechanism would implement the optimal allocation. The principal picks one of the agents at random and makes him the following take it or leave it offer. She offers the lopsided amount that is optimal in the benchmark case and a plan of action that is contingent on the effort observed: if low effort is observed, the agent does not receive any budget in the second period (which is then allocated to the competing agent); if high effort is observed, the agent is given a monetary transfer that perfectly

[19] Uniqueness of $\tilde{B}_1^i$ is intuitive. Any lopsided strategy yields the same payoff to the principal if the high type is selected. When the low type is selected in the first period, the strategy with the lowest first period budget leaves more budget to be allocated in the second period to the high type.

offsets his effort cost and he is allocated all remaining budget if and only if he produces a high output in the first period. If the offer is rejected, the principal allocates all budget to the opponent in the second period (betting in favor of the opponent). In this mechanism the low type is made indifferent between accepting and rejecting the offer and the high type strictly prefers to accept. Since both types will exert effort for any lopsided allocation, everything becomes identical to the benchmark model. Clearly, the implementation comes at the cost of paying a transfer, which is optimally set equal to the cost of effort. If this cost is low - which we would argue is the case in the applications we consider - then the mechanism described is optimal when efforts are contractible and thus can be viewed as the first best.

6 Experimentation with Moral Hazard

In this section we characterize the optimal strategy of the principal depending on the value of the effort cost C. The key difference with the benchmark model is that the principal needs to incentivize agents to exert effort. Furthermore, the principal may use the first period to screen the agents. The interaction of moral hazard and screening is a defining feature of the model.

To characterize the equilibrium, we first prove a series of lemmata that narrow down the type of strategies that are incentive compatible (i.e., such that agents' effort choices can be supported in equilibrium) and clarify the strategic forces at play in this setting. First, the possibility of being screened creates a strong incentive for the low type either to mimic the high type or to save the effort cost. A consequence of this is stated in the next lemma.

Lemma 4. *In any equilibrium of the game each agent effort decision must be independent of whether he received a larger or smaller first period budget than his opponent.*

This lemma implies that in any equilibrium and regardless of which agent gets a larger first period allocation (if any), either both agents exert effort (pooling equilibrium) or only one does (separating equilibrium). The case in which no agent exerts effort when a positive amount of budget is allocated in the first period is ruled out as

the principal would strictly prefer a betting strategy. We henceforth refer to pooling to indicate that both agents exert effort.

Our next result, Lemma 5, has two parts. The first shows that the principal cannot allocate the budget deterministically in the second period if in the first agents are pooling and the outcome is $(0, 0)$. In such event, which may occur on the equilibrium path, types are not revealed with certainty. If the principal allocates the budget deterministically to A_i and he is the high type, the low type would deviate from exerting effort in the first period anticipating that he would not receive the second period budget. In a separating equilibrium, instead, the above event may occur out of equilibrium when both agents shirk. If the out of equilibrium second period allocation were deterministic, the high type would prefer not to exert effort when expecting to get the second period budget anyway. Such deviation shows that the locked-in effect has an additional bite when moral hazard is present and it implies that separation can be achieved if and only if out of equilibrium the principal is indifferent between agents, which is the second result stated in the lemma.

Lemma 5.

a. A pooling strategy profile is incentive compatible only if the Principal, following a $(0, 0)$ *outcome in the first period, is indifferent between allocating to* A_1 *or* A_2 *in the second period.*

b. A separating strategy profile is incentive compatible only if out of equilibrium the Principal, following a $(0, 0)$ *outcome in the first period, is indifferent between allocating to* A_1 *or* A_2 *in the second period.*

Part a. of Lemma 5 has an important implication: With a symmetric prior, there exists no pooling equilibrium in asymmetric first period contribution strategies. To see why this is the case consider any asymmetric hedging strategy profile that makes the principal indifferent between agents after a $(0, 0)$ outcome. For any such profile consider a symmetric hedging with same total first period contribution. Such alternative strategy preserves the principal indifference and relaxes the incentive compatibility constraint of the low type. If the relative likelihood of observing a $(0, 0)$ versus observing either $(1, 0)$ or $(0, 1)$ was the same under the two alternative strategies, one

could easily show that the symmetric strategy dominates because of the concavity of $g(.)$. In fact, the relative likelihood of observing a $(0,0)$ is actually lower under the symmetric strategy. Hence the dominance of the symmetric hedging strategy profile strategy holds a fortiori. This combined with part b. of Lemma 5 implies the following corollary result.

Corollary. *For generic out of equilibrium beliefs and a symmetric prior, the only candidate equilibrium strategies are either a symmetric hedging strategy profile with agents pooling or a betting strategy.*

Notice that in a pooling equilibrium the first period allocation is such that conditional on observing $(0,0)$ the principal does not improve her knowledge over her prior. This is precisely the negative impact of moral hazard on the ability of the principal to learn with respect to the benchmark case.

Before characterizing a symmetric hedging strategy profile a few remarks are in order. First, regardless of the prior assumed, the optimal first period allocation of the benchmark model is no longer an equilibrium in the presence of moral hazard. In fact, for a lopsided strategy not to be dominated by a betting strategy it must be that in the second period following a $(0,0)$ and both agents exerting effort, the principal allocates to the agent who initially did not receive any budget. But then, following the intuition of Lemma 5, when the low type gets allocated the first period budget, he anticipates that a $(0,0)$ outcome would follow with certainty and the principal would never allocate to him in the second period. Hence, he has no incentive to exert effort.

Second, while we use a symmetric prior as a tractable benchmark to represent situations with high uncertainty about agents abilities, Lemma 5 holds for generic priors and -away from the equal prior - requires that the first period allocation should be asymmetric in a way to keep the principal indifferent between allocating to A_1 or A_2 following a $(0,0)$ outcome. Hence, first period allocations that are roughly symmetric should be expected only when uncertainty is relatively large.

Third, as it is transparent from the proof of Lemma 5, the assumption that the low type cannot produce a high output is what triggers a deviation when he anticipates he will not be allocated any budget in the second period after a $(0,0)$ outcome. This conclusion still holds if the low agent was able to produce a high output with a probability small enough compared to the cost of exerting effort.

6.1 Characterization of Pooling Equilibria

We are now ready to characterize the pooling equilibrium with symmetric hedging in the first period. Since agents are pooling in effort, the characterization is the result of a maximization problem, which resembles the case of the benchmark model. The description of a pooling PBE, however, requires specifying out of equilibrium beliefs. Out of equilibrium either no agent exert effort or only one agent exerts effort. The first case cannot be reached by a unilateral deviation from the equilibrium, and thus does not call for any restriction. If only one agent is observed to exert effort, we assume that the principal believes that the deviator is the low type (i.e., we assume punitive beliefs). Any less stringent out of equilibrium belief, which assumes that the agent that did not exert effort is more likely to be the low type is consistent with the proposed equilibrium as it also makes it optimal for the principal to allocate all remaining budget to the agent who exerted effort.

Let B^{Pool} be the total budget allocated in the first period in a pooling symmetric hedging strategy profile. The next lemma provides a necessary and sufficient condition for incentive compatibility, and characterizes the optimal $\tilde{B}^{Pool}$.

Lemma 6. *Consider a pooling symmetric hedging strategy profile and assume punitive out of equilibrium beliefs. Such a strategy profile is incentive compatible if and only if $f(B)\theta V/4 > C$. Furthermore, there exists a unique optimal total budget allocated in the first period $\tilde{B}^{Pool}$ and an upper bound $\bar{B}^{Pool}$ such that $\tilde{B}^{Pool} \leq \bar{B}^{Pool} < B$.*

Incentive compatibility requires that the first period strategy is such that the low type (and hence the high type) prefers to exert effort rather than not. The proof of the lemma shows that this requires that the total budget allocated to the first period is bounded above by $\bar{B}^{Pool}$, with the bound being strictly positive if and only if $f(B)\theta V/4 > C$. By inspecting the following condition that determines the upper bound $\bar{B}^{Pool}$

$$\frac{V}{4}\left(1 - f\left(\frac{\bar{B}^{Pool}}{2}\right)\theta\right) f\left(B - \bar{B}^{Pool}\right)\theta = C$$

it is immediate to see that it becomes tighter when either C is larger or V is smaller. Both changes correspond to an aggravated moral hazard. Indeed, to induce the low type to exert effort, it must be the case that the high type can produce a high

output with a relatively low probability. Hence $\bar{B}^{Pool}$ has to decrease. Finally, $\bar{B}^{Pool}$ increases with θ. The reason is that the benefit for the low type from following the equilibrium strategy comes from decreasing the probability the high type gets the remaining budget in the second period. Such benefit is larger the smaller θ is, which corresponds to a less severe moral hazard problem. Hence as θ increases more resources can be allocated to first period learning. In the proof of Lemma 6 we characterize the optimal $\tilde{B}^{Pool}$, which is increasing in θ, and show that when C is sufficiently small the upper bound $\bar{B}^{Pool}$ is tight.

We are now ready to state our main result, which follows directly from the lemmata.

Proposition 2. *If $C \leq f(B)\theta V/4$, the following strategies and set of beliefs form an equilibrium:*

- *The principal chooses a symmetric hedging strategy allocating a total budget equal to $\tilde{B}^{Pool}$ in the first period and agents pool in effort.*
- *If one agent produces a high output in the first period, he gets allocated the remaining budget in the second period.*
- $\lambda_i\Big((0,0),(\tilde{B}^{Pool}/2,\tilde{B}^{Pool}/2),(1,1)\Big) = 1/2$ *and the Principal randomizes with equal probability between agents in the second period.*
- $\lambda_1\Big((0,0),(\tilde{B}^{Pool}/2,\tilde{B}^{Pool}/2),(0,1)\Big) = \lambda_2\Big((0,0),(\tilde{B}^{Pool}/2,\tilde{B}^{Pool}/2),(1,0)\Big) < 1/2$.

If instead $C > f(B)\theta V/4$, in the unique equilibrium the principal adopts a betting strategy.

As long as we restrict attention to non generic out-of-equilibrium beliefs, the equilibrium described in Proposition 2 is unique. If not, separation can be achieved. We discuss this case in the next subsection, and show under what conditions the strategy profiles characterized in Proposition 2 remain optimal.

6.2 Optimal Strategy for Any Value of C

Consider first symmetric hedging strategies. A separating equilibrium with symmetric hedging requires that the first period allocation induces only the high type to exert effort. Since the payoff maximizing allocation for the principal must minimize the total first period budget, the characterization of the optimal strategy is direct: The low type must be indifferent between exerting effort or not.[20]

The lemma below ensures that the symmetric strategy identified above maximizes the payoff of the principal among all possible allocation of budget that are feasible for some separating equilibrium.

Lemma 7. *For any strategy with asymmetric allocation of first period budget and agents separating in effort that is incentive compatible, there exists an incentive compatible symmetric strategy that allocates the same total first period budget and achieves a higher expected payoff for the principal.*

The dominance established in the result above follows from the concavity of $g(\cdot)$. The intuition for the first part of the lemma comes from the fact that fixing the same total budget, asymmetric allocations tighten the incentive constraint in one state of the world and soften it in the another. Since separation must hold in both states symmetric allocations are easier to sustain.

Some remarks are in order. The description of a separating equilibrium must include a specification of out of equilibrium beliefs following a $(0, 0)$ and no agent exerting effort. In view of Lemma 5 such out of equilibrium beliefs, along with the first period allocation, must guarantee that the principal would allocate the second period budget with equal probability among the two agents following the deviation. Hence, out of equilibrium beliefs must be unique. In particular, for symmetric hedging and under our symmetric prior assumption, the principal's beliefs must be that each type is equally likely to be the high type. While this seems a natural condition given the perfectly symmetric continuation game, it is in fact rather arbitrary.[21]

[20] This contrasts with the determination of the pooling equilibrium in the previous subsection, which was the outcome of a maximization problem. Furthermore, notice that the determination of the first period allocation that guarantees the highest payoff for the principal does not depend on the prior belief assumed.

[21] For example, if the prior is asymmetric and both agents shirk, it might be natural to assume

The following lemma illustrates how the payoff of the principal varies with C when adopting the payoff maximizing separating strategy.

Lemma 8. *The payoff of the principal in a symmetric hedging strategy profile with agents separating in their effort decisions is strictly increasing in C.*

This result might at first seem surprising as it states that as moral hazard becomes more severe (C increases) the principal is better off. To see why it holds, notice that as C increases the principal is able to separate agents allocating less budget in the first period. Hence, more budget is available for the high type in the last period. Alternatively, when C is low, the principal has to increase the chance that the high type succeeds in the first period in order to discourage the low type from exerting effort. This is achieved increasing the total budget allocated in the first period.

We are now ready to describe the equilibrium of the model with moral hazard for any level of C in the next proposition. The message of the proposition below is that for low values of C - arguably a better fit for the type of applications we have described - even if we consider the somewhat more fragile separating strategies, a pooling symmetric hedging remains optimal for the principal.

Proposition 3. *There exists a unique value of the cost of effort $\tilde{C} < f(B)\theta V/4$ such that the optimal first period equilibrium strategy is:*
i) symmetric hedging with agents pooling in effort, if $C < \tilde{C}$;
ii) symmetric hedging with agents separating in effort, if $C \in (\tilde{C}, f(B)\theta V/4)$;
iii) betting, if $C \geq f(B)\theta V/4$.

Each of the candidates for an optimal strategy as previously determined is an equilibrium strategy for some space of the parameter C. A betting strategy is optimal provided that is the only incentive compatible strategy. This is the case when moral hazard is the most severe. Instead, there is a range of C such that both separating and pooling symmetric hedging strategy are incentive compatible. In this case, given that separation perfectly identifies types, it is not obvious that pooling may be optimal. However, the optimal amount of first period budget allocation needed to make separation incentive compatible approaches the total budget B as C approaches

that out of equilibrium the principal would deterministically allocate to one agent.

zero. This implies that while separation is feasible even for arbitrarily small C, in that case the payoff of the principal is minimized (there is no budget left to generate a high output in the second period). Furthermore, while the payoff of the principal in a pooling equilibrium remains constant or decreases as C increases (it decreases if the incentive constraint becomes binding), the payoff of the principal under separation increases in C. As a result, for small values of C pooling is optimal and for larger values of C separation is optimal provided that it is incentive compatible. Interestingly, an implication of the proposition is that the amount of learning in equilibrium is non monotonic in C.

A final observation is that the characterization in Proposition 3 holds even if the low type could generate a positive output with small probability. In this case, however, a lopsided strategy similar to the benchmark case is optimal for a range of C close to zero.

7 Discussion

In our model agents know their opponent's type. Assuming that each agent does not know the type of the opponent does not affect the analysis substantially. In that case each agent plays against the average type, this being equivalent to compressing the value of θ. A different situation would be to consider a model in which agents do not know their own type. This alternative model is closer in spirit to the classic models of experimentation.

Another simplification that we adopt is that the low type cannot generate a high output. If this was not the case a third outcome (high output, high output) could be observed. As long as the ability of the low type is low enough, i.e. the less able agent cannot do significantly better than the more able one even if allocated more budget, our results are not affected.[22]

We also assumed that agents cannot roll over their budget allocation across periods. In other words, each agent uses all available B_i^t. Relaxing this assumption would imply that agents could potentially undo the principal's strategy. In particular, the

[22]Otherwise, if agents were too similar, betting would dominate and there would not be any experimentation in equilibrium.

high type might have an incentive to use less resources in the first period and use the remaining budget in the last period. As a result, the principal would give only an amount that she anticipates will be fully spent. Similarly, we assumed that the principal commits to exhaust the overall budget over the two periods. By relaxing this assumption, the moral hazard would be exacerbated and the principal might need a higher first period allocation for incentive compatibility to hold.

In some applications a regulator might want to affect the principal inter-temporal allocation of resources imposing constraints on total or per-period spending. While our model is silent about welfare, certain welfare implications could be analyzed assuming that the high type is also the more able agent and produces a higher surplus. Under this interpretation, one possible way to think about regulation within our model would be to investigate the effect of imposing caps on the budget allocation, and the most interesting case is when caps are imposed in the last period. The direct effect of a cap to late period allocations in a model with a fixed budget and two periods is that the first period budget allocation must increase whenever the cap is binding. Because of the additional strategic effects, the consequences of a cap may depend on how tight the cap is and on the type of equilibrium played.

In the lopsided equilibrium of the benchmark model we have that the effect of a binding second period cap depends on its level. If the cap is not too tight, the optimal second period strategy conditional on the first period outcome observed will not be affected. If the cap is tighter, however, the principal would not switch between agents even after receiving unfavorable information (locked-in effect). This implies that while a mildly binding cap improves learning, which may be welfare improving, a tighter one is unambiguously detrimental as in equilibrium the principal entirely foregoes the informational content coming from the intermediate outcome. On the other hand, in a pooling equilibrium, a binding cap induces trade-off: It increases the likelihood of identifying the more able agent, but also decreases the chances that this agent, if allocated the remaining budget, produces the high output in the second period. Finally in the betting equilibrium a cap can only improve learning.[23]

[23] In the separating equilibrium, a binding cap makes separation more costly and reduces resources for the second period. Hence, the effect of a cap is unambiguously bad.

8 Related Literature

In terms of theoretical underpinning, our work relates to several strands of literature including dynamic contracts, experimentation and contests. While related, our framework differs substantially from all of them. It differs from the literature on contests because, to the best of our knowledge, the influence of a third external player who is strategic has not been investigated. Similarly, it differs from the literature on experimentation because we focus on the strategic allocation of a fixed total budget over time. Finally, since in all our applications committing ex-ante to a formal contract may not be feasible, we choose not to focus on the contracting dimension.

From the more applied perspective, our paper is related to a vast political economy literature studying political campaigns. Since the seminal contribution of Tullock (1980) on rent-seeking, the impact of money in political elections has rightly become a central theme of the political economy literature. The effect of campaign contributions on political elections has been studied under various assumptions on both the source and the use of campaign funds in election, from informative advertising to vote buying (e.g., Coate (2004), Prat (2002a) and Prat (2002b)). Recently, Bouton et al. (2018) have focused on small campaign contributions driven by an electoral motive. In particular, they analyze separately a "demand side" model where candidates must exert effort to obtain contributions and a "supply side" model where donors are the only strategic actors. All these models are essentially static in nature. Our model not only features a dynamic setting but also combines elements of supply (the third party chooses the allocation) and demand (agents choose effort levels).[24]

A dynamic component is present in some papers studying social learning and momentum building in sequential elections. To this end Knight and Schiff (2010) empirically document a substantial momentum effect focusing on 2004 Democratic primaries. Klumpp and Polborn (2006) show theoretically that victory in an early primary increases the likelihood of victory in the general election and creates an important asymmetry in campaign spending. Denter and Sisak (2015) study the effect

[24]The fact that the action of two sides is needed for some outcome to result connects indirectly our work to the communication framework of Dewatripont and Tirole (2005). Our model could be reinterpreted as having two receivers competing for the attention of a sender where the competition element is absent in Dewatripont and Tirole (2005).

of polls in amplifying ex-ante asymmetries between candidates in a given election.[25] Acharya et al. (2019) study the strategic inter-temporal allocation of a stock of resources by two candidates that wish to raise their popularity in order to win the electoral contest. Differently from our theoretical framework, none of these papers models the role of third parties such as interest groups affecting the outcome of the election. Finally, Schwabe (2015) studies how the dynamics of sequential elections influence campaign financing. While our focus is on the donor strategic inter-temporal allocation of a budget, Schwabe (2015) focuses on the optimal design of the calendar of voting in order to stimulate donors contributions.[26]

Our paper is also related to existing works on outside intervention in conflicts. In particular, Powell (2017) studies a war of attrition between two conflicting parties and it is the only other theoretical contribution we are aware of that studies the role of a third party in affecting a contest. In his model the conflicting parties are aware that a third party will side with one of them at an exogenously given future date T. Expectations about future intervention affect the set of types that will be active at T and, due to adverse selection, the side that is not expected to receive support becomes relatively stronger inducing a "boomerang effect".[27] This effect forces the third party to randomize in equilibrium over the decision of which side to support. Powell (2017) extends the model introducing multiple intervention dates and shows that switching sides can occur in equilibrium if a preference for contestants is introduced. However, differently from our model, switching sides follows from the necessity of preserving uncertainty at any intervention date, which means that the third party is completely indifferent between parties when switching sides. Furthermore, and differently from Powell (2017), our model focuses on the trade off caused by the scarcity of the third party resources.

Among the theoretical contributions, one that like ours lie at the intersections of the literature on contests and experimentation is Halac et al. (2017). Their illustrating example also studies a setting with one principal, two agents and two periods. They

[25] See also de Roos and Sarafidis (2018).

[26] In particular, the model shows that under some conditions a super Tuesday calendar may be optimal.

[27] Some types of the side expecting support stay active longer, and some types of the opponent quit earlier. Hence, conditional on reaching T the opponent becomes relatively stronger.

show that under some conditions a principal that aims at maximizing the chances of observing a success should opt for a contest such that: i) the prize is shared equally among the two agents if both produce a success in one of the two periods and ii) the first period outcome is not revealed. Their model, however, differs in a number of dimensions from ours. First, agents do not know their probabilities of success. Second, the principal, unlike in our model, has control over the disclosure policy of the contest. Lastly and even more fundamentally, the principal cannot affect the chances of success of an agent.

Finally, among the literature on experimentation, Hörner and Samuelson (2013) studies the incentives for experimentation in a dynamic agency problem. In their model a principal provides the necessary funding for a dynamic experiment carried over by an agent exerting costly effort over time. Since the model considers one agent only, the strategic considerations that arise in our setting are absent.

9 Conclusions

Our framework allows to investigate the inter-temporal allocation of resources by a principal deciding which of two agents to support. The model can be applied directly to model campaign contributions and conflict scenarios in which an external player has to choose which side of the conflict to finance and when. Our model could be extended in a number of potentially interesting directions.

First, the third party might have preferences regarding the identity of the winner. Having the principal being biased towards one of the two agents can be accommodated without affecting the key features of the equilibrium. To see why notice that the introduction of a bias has a similar effect to the one of moving away from an equal prior. At the polar case in which the principal knows the identity of the productive agent, the outcome of the model is trivial and a betting strategy is the only optimal strategy. However, as long there is enough uncertainty, the strategies we have described are unchanged.[28]

Second, agents could have preferences regarding when they receive units of budget.

[28]Clearly the principal is not any longer indifferent about to whom giving first in a lopsided strategy.

This could be modeled assuming that the units of budget are scaled by weights that depend on the time period. A preference over early budget allocation would obviously imply an equilibrium increase in first period allocations.

Third, agents may have budgets of their own. Our model is a first order approximation for the more realistic case of agents having small budgets ("war chests") relative to the one of the principal.

Fourth, there are important applications where the principal and the agents may belong to the same organization. This is the case of workers competing for promotions mentioned earlier or, for example, of R&D companies deciding how to optimally allocate funds across competing research teams. A natural assumption for these settings is that a single surplus is shared among the principal and the winning agent. This departure from the current model might alter the strategic forces at play and offer new insights. We leave this for future research.

Finally, and more fundamentally one may wish to add an additional principal to allow for competition among them. In this case the amount of experimentation might be affected by strategic free-riding. In particular, if principals have heterogenous budgets, it is likely that those with smaller budgets would free ride on the one(s) with larger budget. While we expect the strategic effect we highlight to carry over for principals with large enough budgets, we leave this extension for future research. Note, however, that whether or not the principal faces effective competition depends on the specific application. For instance, in the case of political campaigns, examples abound of donors championing rather specific issues where they enjoy a monopolistic position.

References

Acharya, Avidit, Edoardo Grillo, Takuo Sugaya, and Eray Turkel, "Dynamic Campaign Spending," 2019. mimeo, Stanford University, https://stanford.edu/ avidit/campaigns.pdf.

Bouton, Laurent, Micael Castanheira, and Allan Drazen, "A Theory of Small Campaign Contributions," 2018.

Christia, Fotini, *Alliance Formation in Civil Wars*, New York, NY: Cambridge University Press, 2012.

Coate, Stephen, "Pareto-improving campaign finance policy," *American Economic Review*, Jan 2004.

Corchon, Luis, "The Theory of Contests: A Survey," *Review of Economic Design*, 2007, *11*, 69–100.

de Roos, Nicolas and Yianis Sarafidis, "Momentum in dynamic contests," *Economic Modelling*, 2018, *70*, 401 – 416.

Denter, Philipp and Dana Sisak, "Do Polls Create Momentum in Political Competition?," *Journal of Public Economics*, 2015, *130*, 1–14.

Dewatripont, Mathias and Jean Tirole, "Modes of Communication," *Journal of Political Economy*, 2005, *113* (6), 1217–1238.

Halac, Marina, Navin Kartik, and Qingmin Liu, "Contests for Experimentation," *Journal of Political Economy*, 2017, *125* (5), 1523–1569.

Hörner, Johannes and Larry Samuelson, "Incentives for experimenting agents," *The RAND Journal of Economics*, 2013, *44* (4), 632–663.

Klumpp, Tilman and Mattias K. Polborn, "Primaries and the New Hampshire Effect," *Journal of Public Economics*, 2006, *90*, 1073–1114.

Knight, Brian and Nathan Schiff, "Momentum and Social Learning in Presidential Primaries," *Journal of Political Economy*, 2010, *118*, 1110–1150.

Powell, Robert, "Taking Sides in Wars of Attrition," *American Political Science Review*, 2017, *111*, 219–236.

Prat, Andrea, "Campaign Advertising and Voter Welfare," *Review of Economic Studies*, Oct 2002, *69*, 999–1017.

_ , "Campaign spending with office-seeking politicians, rational voters, and multiple lobbies," *Journal of Economic Theory*, Jan 2002.

Regan, Patrick M., *Civil Wars and Foreign Powers: Outside Intervention in Interstate Conflict* 2000.

Schwabe, Rainer, "Super Tuesday: campaign finance and the dynamics of sequential elections," *Social Choice and Welfare*, 2015, *44* (4), 927–951.

Stratmann, Thomas, "Are Contributors Rational? Untangling Strategies of Political Action Committees," *Journal of Political Economy*, 1992, *100*, 647–664.

Tullock, G., "Efficient Rent Seeking," in J. Buchanan, R. Tollison, and G. Tullock, eds., *Toward a Theory of the Rent-seeking Society*, 1980, pp. 97–112.

Appendix

Contributions to U.S. Presidential campaign

Data are taken from http://classic.fec.gov/finance/disclosure/ftpdet.shtml. We consider all contributions to candidates from political action committees (PACs) for all U.S. Presidential election cycles (primaries and general elections) in the period 1980-2016.

Elections and timing We analyze separately primary elections and general elections. To identify early and late campaign periods we proceed in two steps: If no structural break in the frequency of donations during the cycle is observed we divide the electoral cycle into two periods of equal length. For example, this is the case of all primary elections in the period 1980-2016 with the only exception of 2004. Otherwise, in the case in which there appears to be a clear structural break in the frequency of donations we split the period accordingly. For general elections, an increase in the frequency of donations occurred in one of the weeks of September for most of the years in our sample - in 2008 and 2012 the break seems to occur already in August. The dates are listed in Table 1.

The resulting amount of observations in the early and late periods of each election cycle is reported in Table 2 and (in dollar values) in Table 3. Observe that the dynamics of donations between the two periods follow a similar pattern.

Primary Elections Focusing on primary elections, Table 4 shows for each election year the number of PACs following one of three different strategies: contribute in both periods (early and late), contribute only in the early period, or only in the late period of the campaign. In the top half of the table, the left part shows the total number of PACs using a given contribution strategy in the election period, whereas the right part concentrates only on above median contributors. In the bottom half of Table 4 we report the corresponding numbers for PACs in the top quartile and top deciles, respectively. The share of PACs contributing only early and the share contributing only late oscillate between one third and one half with the only exception being the

2004 election cycle where most of the action took place late in the campaign.[29] The share of PACs contributing in both periods starts at the highest level in 1980 (23%), it decreases in 1996-2004 and it increases again later. The share of PACs contributing in both periods is significantly higher among the above-median donors and further increases with the size of total donations.

Table 5 provides information on the number of candidates supported by three types of PACs: those that were active only in the early period of campaign, only in the late period or in both periods, respectively. It is worth noting that close to 90% of PACs choosing to give only in the early or only in the late phase of the campaign give only to a single candidate. None of the PACs in any election period give to 7 or more candidates, rarely they give to more than three. A similar picture holds for both-period contributors. In this case the share of PACs contributing only to a single candidate, however, is slightly lower (78%).

Table 6 shows the distribution of six distinct strategies for PACs contributing in both periods. PACs that contributed to a single candidate in the early campaign, can either **stay** with the same (single) candidate in the late campaign, **switch** and contribute to a different single candidate, or **increase** the number of candidates (from one). Those that contibuted to more than one candidate in the early campaign are divided into those who **increase**, **decrease**, or give to the same number of candidates, but **more** than one. Averaging over the electoral cycles 1980-2016, 78.2% of PACs contributing in both periods supported only one candidate in the early phase of primary campaign. Of those, 17.7% switched their contributions to a different candidate in the second half of the primary campaign. As for giving to multiple candidates, 80.3% of PACs contributing in both periods supported only one candidate in the late phase of primary campaign. Of those 11.2% of PACs started contributing to multiple candidates in the early primary campaign.

General Elections Focusing on general elections, the 1980-2016 average share of PACs contributing only early or in both periods is about 28% and 19%, respectively. Differently from primaries, the average share of PACs contributing only late is more

[29]Recall that 2004 was the only case for primary elections where there appears to be a clear structural break in the frequency of donations. Since in this case we split the period accordingly, the 2004 late campaign period is much longer.

than half. As in the case of primaries, the share of PACs contributing in both periods is higher among the above-median donors and further increases with the size of total donations. Furthermore, and similarly to primaries, nearly all PACs that are active only in the early or late phase of the campaign give only to a single candidate. The share of PACs active in both periods and contributing only to a single is slightly higher than in primaries (81.7%). Of those, only 6.47% switched their contributions to a different candidate in the second half of the primary campaign. As for giving to multiple candidates, 71.8% of PACs contributing in both periods supported only one candidate in the late phase of primary campaign. Of those 5.8% of PACs started contributing to multiple candidates in the early primary campaign (see Table 7).

In summary, the main differences between the picture emerging from focusing on primaries or general elections are: i) in general elections there is much more action in the late phase of the campaign; ii) the share of PACs switching support between candidates and the share of PACs "experimenting" in the early campaign and then supporting a unique candidate eventually is much lower in general elections than in primaries. From this perspective, our theoretical framework seems better suited to analyze primary elections.

Table 1: Dates of primary and general election periods

	begin primary	primary break	end primary	general break	elect. date
80	1980-01-21	1980-03-27	1980-06-03	1980-09-26	1980-11-04
84	1984-02-20	1984-04-16	1984-06-12	1984-09-12	1984-11-06
88	1988-01-14	1988-03-29	1988-06-14	1988-09-20	1988-11-08
92	1992-02-10	1992-04-10	1992-06-09	1992-09-12	1992-11-03
96	1996-01-29	1996-04-01	1996-06-04	1996-09-26	1996-11-05
00	2000-01-24	2000-03-31	2000-06-06	2000-09-26	2000-11-07
04	2004-01-14	2004-02-10	2004-06-08	2004-09-12	2004-11-02
08	2008-01-03	2008-03-18	2008-06-03	2008-08-05	2008-11-06
12	2012-01-03	2012-03-19	2012-06-05	2012-08-20	2012-11-06
16	2016-02-01	2016-04-04	2016-06-07	2016-09-12	2016-11-08

Table 2: Total number of contributions and share (%) in early and late campaign

	Primaries			General Elections		
	early %	late %	total	early %	late %	total
80	43	57	3297	46	54	8576
84	48	52	2032	43	57	5763
88	56	44	2104	40	60	2943
92	48	52	680	30	70	1967
96	53	47	661	45	55	1371
00	66	34	684	24	76	2831
04	80	20	4053	25	75	6620
08	48	52	2839	9	91	14306
12	61	39	4347	4	96	49593
16	58	42	7559	41	59	102451
average	58	42	2826	29	61	19642

Table 3: Total value of contributions and share (%) in early and late campaign

	Primaries			General Elections		
	early %	late %	total	early %	late %	total
80	43	57	5781954	33	67	39666434
84	57	43	6083079	39	61	24640947
88	50	50	2754952	32	68	29385496
92	49	51	2476282	37	63	27041642
96	31	69	1235025	39	61	20299506
00	50	50	4444372	33	67	44326587
04	46	54	6209931	31	69	246371910
08	46	54	36565634	6	94	208039320
12	70	30	90565067	24	76	546621631
16	81	19	136083331	27	73	420218807
average	70	30	29219963	25	75	160661228

Table 4: Distribution of strategies in Primaries

	All PACs			PACs>median*		
	Only early %	Only late %	Both %	Only early %	Only late %	Both %
80	35.14	41.43	23.43	30.11	36.80	33.09
84	47.89	35.00	17.11	48.98	29.80	21.22
88	56.87	26.91	16.22	53.12	26.56	20.33
92	45.05	41.58	13.37	45.07	40.14	14.79
96	52.44	39.02	8.54	50.96	39.42	9.62
00	51.93	41.40	6.67	46.85	41.96	11.19
04	19.50	73.90	6.60	19.64	69.64	10.71
08	42.25	42.25	15.50	38.78	41.99	19.23
12	39.57	42.81	17.63	35.25	38.13	26.62
16	44.73	33.42	21.85	36.41	26.67	36.92
average	43.54	41.76	14.70	40.52	39.11	20.37
	PACs>75%*			PACs>90%*		
	Only early %	Only late %	Both %	Only early %	Only late %	Both %
80	21.74	26.96	51.30	21.74	17.39	60.87
84	42.11	22.11	35.79	28.95	18.42	52.63
88	44.27	22.14	33.59	31.48	16.67	51.85
92	47.17	26.42	26.42	31.82	18.18	50.00
96	44.44	33.33	22.22	37.50	40.62	21.88
00	42.86	45.45	11.69	50.00	20.00	30.00
04	16.47	67.06	16.47	18.75	50.00	31.25
08	36.28	38.60	25.12	18.18	21.82	60.00
12	24.66	30.14	45.21	21.43	17.86	60.71
16	31.63	17.35	51.02	33.33	10.26	56.41
average	35.16	32.95	31.88	29.32	23.12	47.56

Note*: Only PACs with total contributions above median/top quartile/top decile.

Table 5: Distribution of number of candidates supported in Primaries

PAC contribution strategy / No. of Candidates	1	2	3	4	5 or more
Only in early campaign	73.8	8	2.3	1.1	0.3
Only in late campaign	71.4	6.2	0.4	0.1	0.0
Both periods	34.3	7	2.2	0.8	0.8

Averages 1980-2016, PACS with total contribution above the median.

Table 6: Contribution strategies in Primaries

	stay with same	late switch	late increase	early give to 1
Average 1980-2016	51.42	17.66	9.11	78.20
	early to 1		early to multiple	late give to 1
Average 1980-2016	69.09		11.18	80.27

Note: Share of PACs contributing in both periods with total contribution above the median divided by early-period and late-period number of candidates.

Table 7: Contribution strategies in General Elections

	stay with same	late switch	late increase	early give to 1
Average 1980-2016	59.57	6.47	15.64	81.68
	early to 1		early to multiple	late give to 1
Average 1980-2016	66.04		5.77	71.81

Note: Share of PACs contributing in both periods with total contribution above the median divided by early-period and late-period number of candidates.

Omitted Proofs

Proof. **of Lemma 1** Assume first that either $B_1^1 > B/2$ or $B_2^1 > B/2$. Then it is optimal for the principal to allocate all remaining budget to the agent that had more than $B/2$ in period 1. In fact, giving to the other agent yields zero for sure, while the proposed strategy yields a strictly positive expected payoff regardless of the history. Now assume that $B_2^1 \leq B/2$ and $B_1^1 \leq B/2$. In this case, the principal choice determines which of the two agents should end up with a total budget greater than $B/2$. How much the other agent receives is irrelevant for the principal payoff. But then given that the principal payoff is increasing in the share allocated to the winner, we can conclude that the agent who received a total budget above $B/2$ should receive all budget not allocated in period 1. To determine the identity of who gets the remaining budget, the condition in the lemma simply compares the expected payoff of giving to A_i and A_{-i} given the information available at the end of period 1. ■

Proof. **of Lemma 2**

To prove the lemma, we argue that any hedging strategy in the first period, that is any (B_1^1, B_2^1) with $B_1^1 > 0$, $B_2^1 > 0$, is dominated by a lopsided strategy. To do so assume (w.l.o.g) that $B_2^1 \geq B_1^1$. Consider first the case in which $B_2^1 > B/2 > B_1^1 > 0$. This type of allocation is such that, regardless of the history that unfolds, the principal does not get any payoff from having allocated $B_1^1 > 0$ to A_1. Hence, she is strictly better off by setting $B_1^1 = 0$ and allocating $B - B_2^1$ to A_2 in period 2 as this increases her payoff conditional on A_2 winning and does not affect it conditional on A_1 winning.

Consider next any hedging strategy in the first period such that $B/2 > B_2^1 \geq B_1^1 > 0$. We want to show that for any such strategy we can find a lopsided strategy that dominates it. In particular the lopsided strategy we propose depends on whether the considered hedging strategy allocates to A_1 (case a) or to A_2 (case b) in the second period after observing a $(0, 0)$. These are the only relevant cases since after observing a high output the principal must optimally allocate all remaining budget to the agent who succeeded.

Case a) We compare the hedging strategy (B_1^1, B_2^1) with a lopsided strategy $(0, B_2^1)$ in which the principal allocates to A_1 in the second period after observing a $(0, 0)$.

Since the principal cannot distinguish ex-ante between the two agents, we compare the payoffs under the two strategies depending on whether A_2 is the high type. First, consider that A_2 is the high type. Suppose that with $(0, B_2^1)$ a high output is observed. Then the second period decision is to give all remaining budget to A_2 under both type of strategies. However, $B_1^1 > 0$ has two negative effects: it reduces the probability of getting a prize because $g(B) > g(B - B_1^1)$, and it reduces the probability of generating a high output in the second period. Suppose instead that a low output is observed. Then, both strategies allocate all remaining budget to A_1, which yield a zero payoff for both.

Second, consider that A_1 is the high type. In this case, the second period decision under both strategies is to give all remaining budget to A_1, so that the total budget share allocated to A_1 is the same under the two strategies, which means that the resulting value for $g(\cdot)$ is the same. Furthermore, the lopsided strategy has a higher chance of generating a high output in the last period because the second period allocation is now $B - B_2^1 > B - B_2^1 - B_1^1$. Hence the lopsided strategy yields a higher payoff.

Thus, we can conclude that any hedging strategy of the type considered in case a) is dominated by the proposed lopsided strategy.

Case b) We compare the hedging strategy (B_1^1, B_2^1) with a lopsided strategy $(B_1^1, 0)$ in which the principal allocates to A_2 in the second period after observing a $(0, 0)$. The proof is identical to case a), it only requires to invert the roles of A_1 and A_2. ■

Proof. **of Lemma 3**

We now show that there exists a bound on B_i^1 smaller then $B/2$ such that the first period budget allocation of an optimal lopsided strategy will be always smaller than this bound. Recall that the condition for the principal to switch is

$$\lambda_i = \frac{1 - \theta f(B_2^1)}{2 - \theta f(B_2^1)} < \frac{g(B - B_2^1)}{1 + g(B - B_2^1)}$$

which can be rewritten as

$$1 - \theta f(B_2^1) < g(B - B_2^1)$$

Consider then the function

$$Q(B_i^1) \equiv 1 - \theta f(B_i^1) - g(B - B_i^1)$$

Notice that $Q(0) = 0$ and $Q(B/2) > 0$. Furthermore,

$$Q'(B_i^1) = -\theta f'(B_i^1) + g'(B - B_i^1),$$

$$Q'(0) = -\theta f'(0) + g'(B) < 0$$

and

$$Q''(B_i^1) = -\theta f''(B_i^1) - g''(B - B_i^1) > 0.$$

Hence, there exists a unique positive $\hat{B}_i^1 < B/2$ such that $Q(B_i^1) < 0$ if and only if $B_i^1 < \hat{B}_i^1$. It follows that the optimal lopsided strategy is to give some $B_i^1 < \hat{B}_i^1$ (since any $B_i^1 > \hat{B}_i^1$ implies that the principal will keep allocating to the same agent no matter what the outcome is at the end of the first period).
The comparative static results follow from straightforward application of the implicit function theorem to $Q(\hat{B}_2^1) = 0$. Indeed, existence of $\hat{B}_2^1$ implies that $Q'(\hat{B}_i^1) > 0$. Hence

$$\frac{dQ(\hat{B}_i^1)}{dB} = -\frac{-g'(B - \hat{B}_i^1)}{Q'(\hat{B}_i^1)} > 0$$

and

$$\frac{dQ(\hat{B}_i^1)}{d\theta} = -\frac{-f(\hat{B}_i^1)}{Q'(\hat{B}_i^1)} > 0$$

Proof. **of Proposition 1**

Conditional on a lopsided budget allocation in the first period, the optimal strategy is the solution to the following problem

$$\max_{B_i^1 \in [0, \hat{B}_i^1]} \frac{\theta W}{2} \left(\theta f(B_i^1) f\left(B - B_i^1\right) + f\left(B - B_i^1\right) g\left(B - B_i^1\right) \right)$$

Taking the first order conditions we get

$$\theta(f'(B_i^1)f(B-B_i^1)-f(B_i^1)f'(B-B_i^1))-f'(B-B_i^1)g(B-B_i^1)-f(B-B_i^1)g'(B-B_i^1)=0,$$

and notice that the FOC evaluated at $\hat{B}_i^1$ is equal to

$$f(B-\hat{B}_i^1)\left(\theta f'(\hat{B}_i^1)-g'(B-\hat{B}_i^1)\right)-f'(B-\hat{B}_i^1)<0$$

since by the uniqueness of $\hat{B}_i^1$ it must be that

$$Q'(\hat{B}_i^1)=-\theta f'(\hat{B}_i^1)+g'(B-\hat{B}_i^1)>0.$$

Furthermore, by evaluating the FOC at $B_i^1=0$ yields

$$\theta f'(0)f(B)-f'(B)-f(B)g'(B)>0$$

Hence, an interior solution exists. However, taking the second order condition yields

$$\theta\left(f''(B_i^1)f(B-B_i^1)-2f'(B_i^1)f'(B-B_i^1)+f(B_i^1)f''(B-B_i^1)\right)+$$

$$+f''(B-B_i^1)g(B-B_i^1)+2f'(B-B_i^1)g'(B-B_i^1)+f(B-B_i^1)g''(B-B_i^1),$$

which is not necessarily negative since $2f'(B-B_i^1)g'(B-B_i^1)$ is positive.

To show uniqueness of an optimal $\tilde{B}_i^1$ we proceed by contradiction. Suppose that the solution is not unique and without loss of generality assume that there exists another lopsided allocation strategy $\tilde{B}_i^{1'}\in(\tilde{B}_i^1,\hat{B}_i^1)$ that gives the same utility to the principal. Notice that, conditional on observing a high output in the first period, the two strategies give the same payoff to the principal. Furthermore, conditional on observing a low output in the first period the two strategies give the same payoff to the principal if the principal has chosen the high type in the first period. If not, since $\tilde{B}_i^1<\tilde{B}_i^1$, the principal can allocate more resources to the high type in the second period under $\tilde{B}_i^1$, which yields a higher expected payoff. But this is a contradiction to the assumption that the two strategies are optimal. Notice also that, although the objective function is not necessarily globally concave, given uniqueness and the fact

that the FOC evaluated at $B_i^1 = 0$ is strictly positive, the second order condition must be negative at the optimum. ■

Proof. **of Lemma 4**

The lemma refers to situations in which first period contributions are asymmetric. Asymmetric allocations imply the existence of two states of the world depending on whether the agent who received the larger budget is the high type or the low type. Assume, by way of contradiction, that one agent exerts effort in one state only, say w.l.o.g. Agent 1. We have to consider three cases:
a) Agent 2 never exerts effort: Then agent 1 is always identified as high type and he is therefore allocated all remaining budget in the second period. But then in the state in which no agent should exert effort, agent 2 is better off deviating by exerting effort as the principal would wrongly believe to be along the equilibrium path and would identify the deviator as high type.
b) Agent 2 exerts effort in one state only: Similarly to case a), from observed efforts the principal would infer the state and allocate in the second period to the high type. But then low type deviates to save the effort cost.
c) Agent 2 exerts effort in all states: For the same reason as above, the low type deviates to save the effort cost. ■

Proof. **of Lemma 5**

We argue by contradiction. Assume that there exist a pooling equilibrium in which first period contributions are such that after observing a $(0, 0)$ outcome the principal strictly prefers to allocate the remaining budget to one agent in the second period. When such agent is the high type, the low type anticipates that he will never be allocated any budget in the second period and he is therefore better off deviating to save the effort cost. Hence for a pooling strategy profile to be incentive compatible, it must be that the principal, following a $(0, 0)$, allocates with positive probability to either agent (and allocating with equal probability makes the relevant constraint less binding). Let us now consider a separating equilibrium. Since only one agent must exert effort, it must be the high type. Suppose that he deviates, which implies

that a $(0,0)$ realizes. If the principal strictly prefers allocating to one agent in the second period, the high type deviation is profitable when he anticipates that out of equilibrium he will be allocated the second period budget. ■

Proof. **of Lemma 6**

We characterize the optimal total value of budget allocated in the first period in a pooling symmetric hedging strategy profile B^{Pool}. In order to do so, first we need to ensure that the first period strategy is such that the low type (and hence the high type) prefers to exert effort rather than not. In doing so, we assume punitive out-of-equilibrium beliefs, that is beliefs such that if an agent shirks in the first period, he is considered a low type with probability one. Hence, we need

$$\frac{V}{2}\left(1-f\left(\frac{B^{Pool}}{2}\right)\theta\right)\left[\frac{1}{2}+\frac{1}{2}\left(1-f\left(B-B^{Pool}\right)\theta\right)\right]+$$
$$\frac{V}{2}f\left(\frac{B^{Pool}}{2}\right)\theta\left(1-f\left(B-B^{Pool}\right)\theta\right)-C\geq\frac{V}{2}\left(1-f\left(B-B^{Pool}\right)\theta\right),$$

which simplifies to

$$\frac{V}{4}\left(1-f\left(\frac{B^{Pool}}{2}\right)\theta\right)f\left(B-B^{Pool}\right)\theta\geq C$$

The condition is satisfied when evaluated at $B^{Pool}=0$ (it simplifies to $f(B)\theta V/4\geq C$, which is a necessary condition for a pooling symmetric hedging equilibrium to exist). Furthermore, its left hand side is strictly decreasing for all range of feasible B^{Pool}, and the condition is clearly violated at $B^{Pool}=B$. Hence, a unique solution to the inequality above exists for each positive C, and it yields an upper bound on B^{Pool}. We call such upper bound $\bar{B}^{Pool}$. Next, we compute the optimal level $\tilde{B}^{Pool}$. The principal maximizes

$$\max_{B^{Pool}\leq\bar{B}^{Pool}} W\left(1-f\left(\frac{B^{Pool}}{2}\right)\theta\right)\frac{1}{2}f\left(B-B^{Pool}\right)\theta g\left(B-\frac{B^{Pool}}{2}\right)+$$
$$Wf\left(\frac{B^{Pool}}{2}\right)\theta f\left(B-B^{Pool}\right)\theta g\left(B-\frac{B^{Pool}}{2}\right)$$

which simplifies to

$$\max_{B^{Pool} \leq \bar{B}^{Pool}} \frac{\theta W}{2} f(B - B^{Pool}) g\left(B - \frac{B^{Pool}}{2}\right)\left(1 + f\left(\frac{B^{Pool}}{2}\right)\theta\right)$$

Taking first order conditions yields

$$-f'(B - B^{Pool}) g\left(B - \frac{B^{Pool}}{2}\right)\left(1 + f\left(\frac{B^{Pool}}{2}\right)\theta\right) +$$

$$-\frac{1}{2} f(B - B^{Pool}) g'\left(B - \frac{B^{Pool}}{2}\right)\left(1 + f\left(\frac{B^{Pool}}{2}\right)\theta\right) +$$

$$+\frac{1}{2} f(B - B^{Pool}) g\left(B - \frac{B^{Pool}}{2}\right) f'\left(\frac{B^{Pool}}{2}\right)\theta = 0$$

Evaluating the first order conditions at $B^{Pool} = 0$, we have that a sufficient condition for a strictly positive $\tilde{B}^{Pool}$ is

$$\frac{f'(0)\theta - g'(B)}{2} > \frac{f'(B)}{f(B)},$$

which always holds given that $f'(0) \to \infty$. Furthermore, when C is relatively small, $\bar{B}^{Pool}$ approaches B and for $\bar{B}^{Pool}$ close to B we have that

$$f'(B - \bar{B}^{Pool})\left(1 + f\left(\frac{\bar{B}^{Pool}}{2}\right)\theta\right) > \frac{1}{2} f(B - \bar{B}^{Pool}) f'\left(\frac{\bar{B}^{Pool}}{2}\right)\theta$$

so that the left hand side of the first order condition must be negative. Hence, when C is small, the optimal $\tilde{B}^{Pool}$ is strictly smaller than $\bar{B}^{Pool}$. To show uniqueness of $\tilde{B}^{Pool}$ notice that since in a pooling symmetric hedging strategy profile both agents exert effort, we can apply a similar argument as the one used in Proposition 1. ■

Proof. **of Lemma 7**

Fix the total level of first period budget at B^S, and consider the strategy (B_1^1, B_2^1) such that $B_1^1 + B_2^1 = B^S$, and let $B_1^1 > B_2^1 > 0$ without loss of generality. We want to consider all possible strategies and hence allow for any out of equilibrium beliefs.

First we consider out of equilibrium beliefs such that following both agents exerting efforts and a $(0,0)$ the principal allocates the remaining budget to the agent that got B_2^1 in the first period. In this case the conditions for the low type to find optimal to shirk are:

$$\left(1-f(B-B^S)\theta\right)\frac{V}{2} \geq \left(1-f\left(B_1^1\right)\theta\right)\frac{V}{2}+$$
$$+f\left(B_1^1\right)\theta(1-f(B-B^S)\theta)\frac{V}{2}-C$$

and

$$\left(1-f(B-B^S)\theta\right)\frac{V}{2} \geq$$
$$\left(1-f\left(B_2^1\right)\theta\right)\left(1-f(B-B^S)\theta\right)\frac{V}{2}+f\left(B_2^1\right)\theta(1-f(B-B^S)\theta)\frac{V}{2}-C$$

.

The conditions above can be compared with the one for the symmetric hedging:

$$\left(1-f(B-B^S)\theta\right)\frac{V}{2} \geq \left(1-f\left(\frac{B^S}{2}\right)\theta\right)\frac{1}{2}\frac{V}{2}+$$
$$\left(1-f\left(\frac{B^S}{2}\right)\theta\right)\frac{1}{2}\left(1-f(B-B^S)\theta\right)\frac{V}{2}+f\left(\frac{B^S}{2}\right)\theta(1-f(B-B^S)\theta)\frac{V}{2}-C$$

Clearly, the first condition implies that the low type also shirks in the symmetric hedging.

Similarly, we can consider out of equilibrium beliefs such that following both agents exerting efforts and a $(0,0)$ the principal allocates the remaining budget to the agent that got B_1^1 in the first period. In this case the conditions for the low type to find optimal to shirk are:

$$\left(1-f(B-B^S)\theta\right)\frac{V}{2} \geq$$
$$\left(1-f\left(B_1^1\right)\theta\right)\left(1-f(B-B^S)\theta\right)\frac{V}{2}+f\left(B_1^1\right)\theta(1-f(B-B^S)\theta)\frac{V}{2}-C$$

and

$$\left(1-f(B-B^S)\theta\right)\frac{V}{2} \geq \left(1-f\left(B_2^1\right)\theta\right)\frac{V}{2}+$$

$$+f\left(B_2^1\right)\theta(1-f(B-B^S)\theta)\frac{V}{2}-C$$

Clearly, the first condition implies that the low type also shirks in the symmetric hedging.

Finally, notice that the conditions for the high type to exert effort are the same both in the symmetric and asymmetric hedging. Hence, we can conclude that whenever the asymmetric separating hedging is a feasible strategy (for some out of equilibrium beliefs) so it is a symmetric separating hedging with the same total level of first period budget (for some out of equilibrium beliefs).

To establish that the constructed symmetric hedging dominates the asymmetric hedging we need to compare the principal expected payoff under the two strategies. The comparison gives:

$$\frac{1}{2}g(B-B_1^1)+\frac{1}{2}g(B-B_2^1)<g\left(B-\frac{B^S}{2}\right),$$

where the left hand side is the payoff under the asymmetric strategy and the right hand side the payoff under the symmetric strategy. The inequality stated holds because of the assumed concavity of $g(\cdot)$. ■

Proof. **of Lemma 8**

To see why this is the case, notice that the condition that determines the optimal budget allocation in the first period (within this class of strategies)

$$\frac{4C}{\theta V}=\left(1-f\left(\frac{B^S}{2}\right)\theta\right)f(B-B^S)$$

implies that as C increases the principal can separate agents with a lower level of first period allocation B^S.

Proof. **of Proposition 3**

In order to compare the principal payoff under the best pooling symmetric hedging, separating symmetric hedging, and betting strategies we first need to characterize the best separating symmetric hedging. We do so below.

If we define B^S to be the total level of first period budget allocated in a symmetric hedging equilibrium, we have that separation can occur via a symmetric hedging strategy if:

$$\left(1 - f(B - B^S)\theta\right)\frac{V}{2} \geq \left(1 - f\left(\frac{B^S}{2}\right)\theta\right)\frac{1}{2}\frac{V}{2} +$$

$$\left(1 - f\left(\frac{B^S}{2}\right)\theta\right)\frac{1}{2}\left(1 - f(B - B^S)\theta\right)\frac{V}{2} + f\left(\frac{B^S}{2}\right)\theta(1 - f(B - B^S)\theta)\frac{V}{2} - C$$

The left hand side is the equilibrium payoff of the low type agent if he follows the equilibrium strategy. The right hand side is the payoff of the low type agent if he deviates to exerting effort. In this case if the principal observes a $(0, 0)$ she has no way of distinguishing the two types and the out of equilibrium beliefs coincide with the prior. The above inequality simplifies to

$$C \geq \frac{\left(1 - f\left(\frac{B^S}{2}\right)\theta\right) f(B - B^S)\theta V}{4}$$

Furthermore, the high type exerts effort in the first period if and only if

$$f(B - B^S)\theta V + \left(1 - f(B - B^S)\theta\right)\frac{V}{2} - 2C \geq \frac{V}{2}\left(f(B - B^S)\theta + (1 - f(B - B^S)\theta)\frac{1}{2} + \frac{1}{2}\right) - C$$

which yields

$$f(B - B^S) \geq \frac{4C}{\theta V} \text{ or } C \leq \frac{\theta V}{4} f(B - B^S)$$

Hence, for $C \in \left(\frac{\left(1 - f\left(\frac{B^S}{2}\right)\theta\right) f(B - B^S)\theta V}{4}, \frac{f(B - B^S)\theta V}{4}\right)$ a symmetric hedging strategy can be supported in equilibrium. Clearly, the optimal symmetric hedging minimizes the total level of budget B^S allocated in the first period, and hence equals the unique solution to

$$\frac{4C}{\theta V} = \left(1 - f\left(\frac{B^S}{2}\right)\theta\right) f(B - B^S),$$

which is strictly positive if and only if $C < \frac{f(B)\theta V}{4}$. Note that, at the optimal B^S, it is always the case that

$$f(B - B^S) \geq \frac{4C}{\theta V}$$

since $1 - f\left(\frac{B^S}{2}\right) < 1$.

We can now move to the comparison between the three alternative strategies for the principal.

First notice that for small C, a symmetric hedging strategy with agents separating cannot be optimal. To see this notice that the optimal B^S characterized above approaches B as C approaches zero. Hence the principal payoffs in a symmetric hedging strategy with agents separating approaches zero for small C. Next, a symmetric hedging strategy with agents pooling on exerting effort - whenever it is incentive compatible - dominates a betting strategy since the optimal B^{Pool} is strictly positive (this is proved in Lemma 6). This establishes the optimality of a symmetric hedging strategy with agents pooling for small C. Notice that the optimal B^{Pool} does not depend on C as long as $B^{Pool} < \bar{B}^{Pool}$, i.e., the optimal B^{Pool} is strictly interior, which also implies that the principal's payoff does not depend on C (and if B^{Pool} hits the upper bound the profit can only decrease). On the other hand, the principal's payoff in a symmetric hedging strategy with agents separating increases with C, since the principal can attain separation using less resources in the first period. Since for C close to $f(B)\theta V/4$ the principal attains the highest possible profit, namely she can achieve screening spending almost no resources in the first period, there must exist a level of $C^* < f(B)\theta V/4$ such that a symmetric hedging strategy with agents separating dominates any other possible strategy for $C \in (C^*, f(B)\theta V/4)$. Finally, when $C \geq f(B)\theta V/4$ the only (trivially) incentive compatible strategy is a betting strategy.

■