Repeated Auctions under Budget Constraints: Optimal bidding strategies and Equilibria

Ramakrishna GUMMADI, Stanford University
Peter KEY, Microsoft Research
Alexandre PROUTIERE, KTH Royal Institute of Technology

How should agents bid in repeated sequential auctions when they are budget constrained? A motivating example is that of sponsored search auctions, where advertisers bid in a sequence of generalized second price (GSP) auctions. These auctions, specifically in the context of sponsored search, have many idiosyncratic features that distinguish them from other models of sequential auctions: First, each bidder competes in a large number of auctions, where each auction is worth very little. Second, the total bidder population is often large, which means it is unrealistic to assume that the bidders could possibly optimize their strategy by modeling specific opponents. Third, the presence of a virtually unlimited supply of these auctions means bidders are *necessarily* expense constrained.

Motivated by these three factors, we first frame the generic problem as a discounted Markov Decision Process for which the environment is independent and identically distributed over time. We also allow the agents to receive income to augment their budget at a constant rate. We first provide a structural characterization of the associated value function and the optimal bidding strategy, which specifies the extent to which agents underbid from their true valuation due to long term budget constraints. We then provide an explicit characterization of the optimal bid shading factor in the limiting regime where the discount rate tends to zero, by identifying the limit of the value function in terms of the solution to a differential equation that can be solved efficiently. Finally, we proved the existence of Mean Field Equilibria for both the repeated second price and GSP auctions with a large number of bidders.

1. INTRODUCTION

In sponsored search auctions, such as those run by Google (AdWords) or Bing (adCenter), a query to a search engine triggers an auction for positions on the page returned by the query. Advertisers bid for positions or *slots*, and successful bids result in ads being displayed alongside "organic" results to the query. The slots are ordered, where higher slots are more valuable. Under a Generalized Second Price (GSP) auction mechansim for slots, an agent places a single bid, bids are ranked by weight and ranked bids determine the slot allocation. If an advert displayed in a slot is clicked, then the advertiser pays the price to be in the current slot (i.e. the weighted bid of the advertiser in the slot immediately below), a Pay-Per-Click payment model.

In addition, *budgets* constrain advertisers, who can specify a periodic budget limit, typically for a day, month, or for a campaign. Budgets are important to both advertisers and the auctioneer, but alter the dynamics of auctions: in consequence, how should an advertiser bid "optimally" in a sequence of repeated auctions when budget constrained? We shall address this question, and also the related question: what happens if all the advertisers bid optimally? Do they reach an equilibrium?

We shall use ad-auctions as a real-life exemplar of repeated auctions, where the same good or similar goods (slots or slot-keyword pairs) is repeatedly auctioned over time, and bidders (advertisers) repeatedly bid for these goods. In such auctions first, budgets are typically large compared to the price of an action, at least 100 times large than the price; for example bids may be in cents and budgets in dollars. Second, the auctions are rarely "homogeneous": apart from the disparity between the market prices for different keywords, there can be significant variability between successive auctions for the same keyword. The latter happens because the weights (CTR — click through rates) are variable (since they are estimated and depend on the searcher [Athey and Nekipelov 2010]), because of targeting, which affects the effective bid in an auction, and because of budget constraints, variable weights and searcher dependent effects that cause the number of bidders for a particular keyword to vary over time, [Pin and Key 2011].

Guided by this setting, we model a generic repeated auction where

— there are a large number of *potential* bidders (the market is thick),
— bidders are price takers,
— bidders are constrained by budgets that are large compared to the price paid in an individual auction,
— in any one auction, a bidder faces i.i.d (independent and identically distributed) bids, and where the bids faced are *as if* independent across auctions,
— bidders attempt to bid optimally, to optimize their net return over time given the auction design.

The assumption of i.i.d. bidders may seem strong, but it has some support from experimental evidence [Pin and Key 2011]. It can also be theoretically justified using Mean Field asymptotics. We use a model similar to that developed in [Iyer et al. 2011] for repeated auctions with learning where at each step of the auction, a small number of participants are randomly selected from a large population to compete. Under this model, the i.i.d. assumption can be formally justified. There are other ways to validate this assumption, see e.g. [Adlakha et al. 2011].

The fact that budgets are typically large compared to the price of an auction means that a bidder needs to participate to a large number of auctions to spend her budget. We shall exploit this observation to work with deterministic *fluid approximations*, which in turn will help us characterize optimal bidding strategies.

Our Contribution. The main contribution of our paper is twofold in nature: we characterize optimal bidding strategies for a general *Online Budgeting Problem*, and then prove existence of Mean Field Equilibria (MFE).

The simplest case of the our online Budgeting Problem is that of a repeated second price auction, where an agent competes against opponents whose highest bids at each time step are i.i.d.. We formulate this problem as a Markov Decision Process (MDP), and provide, for any arbitrary discount factor $e^{-\beta}$, the linkage between the value function and the optimal bidding strategy (Theorem 2.1), which in turn gives a fixed point equation for the value function.

The general Online Budgeting Problem is described in Section 3.1: we assume that a single agent plays repeatedly against a partially observable *environment* whose evolution is i.i.d. over time. The agent takes actions based on the observable part of the environment, whilst respecting her own resource or budget constraint, and seeks to maximize her discounted return, where time is discounted at rate $e^{-\beta}$. Actions result in a potential utility gain g to the agent, and potentially incur a cost c, which depletes the budget of the agent. Both g and c can be arbitrary bounded functions that depend upon both the actions of agents taken and the environment. We also allow the budget to be increased at a constant rate a. By scaling the value function and taking the limit as $\beta \to 0$, we are able to explicitly characterize both the value function, and the optimal control. The main result of the paper, stated in Theorem 3.1, gives a differential equation that the limiting value function satisfies, and describes the optimal control in terms of the derivative of limiting value function, V.

We demonstrate the power of this theorem by applying it to two examples: our repeated second price auction exemplar, and a repeated GSP. In both examples, the environment at each step or auction includes the valuation of the object sold in this auction to the agent, and the bids of the opponents. The valuation may be observed or not: in the case of second-price auctions, it makes sense to assume that the agent observes the valuation before choosing an appropriate bid, whereas in the case of sponsored search

GSP auctions, the valuation depends on whether the displayed advert is clicked, which cannot be observed before bidding. For repeated second-price auctions, we show that the optimal bid u^* when the observed valuation is v and scaled budget B is given by

$$u^*(v, B) = \frac{v}{1 + V'(B)}.$$

In other words, $1/(1 + V'(B))$ is a bid *shading* factor. This same factor is involved in the corresponding optimal bidding strategy for the GSP. The value function and hence the optimal bid can be explicitly computed. Both depend on the opponent's bid distribution. As a consequence, the knowledge of the opponent's bid distribution gives a shading factor that reacts to opponent's bids. This contrasts with the results derived in [Zhou et al. 2008] for repeated auctions with budget constraints in an adversarial environment. In such adversarial setting, the shading factor, proposed in [Zhou et al. 2008], is just a generic function, and does not take advantage of the knowledge of opponent's bid statistics, which potentially yields a much smaller utility.

Note also that we have a shading factor, whereas in the learning problem considered by [Iyer et al. 2011], exploration necessary to determine the agent's unknown valuation requires *overbidding*. Also notice that the overbidding progressively disappears as the agent learns her valuation with increased accuracy. On the contrary, in the case of budget constraints, the shading factor has a persistent effect. It actually increases with time, as the balance drops. Further, our results go a step beyond merely providing structural insight by giving an explicit characterization to compute the optimal bid shading factor.

The second feature of our work is the specification of a Mean Field game. In this game, we assume that the number of competing agents is large, and that for each auction, a finite number of agents is randomly selected (as in [Iyer et al. 2011]). Each agent has a random life-time, which is exponentially distributed with unit mean, and optimizes her utility over their life-time (As we shall explain in Section 4, the agent's optimal bidding strategy corresponds to the strategy maximizing her discounted utility with discount factor very close to 1). At the end of their life-time, agents are replaced by new agents whose initial budget, valuation distribution and income is sampled according to some probability measure. We prove that our Mean Field game always admits equilibria (Theorem 4.2), which are potentially easy to compute. We also give, in Section 5, corresponding theorems for both characterizations of the optimal control and Mean Field equilibrium existence for finite time-horizon without a discount factor.

Related Work. The imposition of financial constraints is known to alter the properties of even simple standard auctions. Che and Gale, [Che and Gale 1998] discuss some of the implications of constraints for first price and second price auctions. There is some literature on sequential auctions with budget constraints (see [Pitchik 2009] and references therein), however the three critical modeling issues we motivate in the abstract do not apply to existing body of work.

We have given a high level, simplified description of a GSP auction mechanism, and in practice, different providers adapt this basic model, through the use of reserve prices ([Ostrovsky and Schwarz 2011]), quality scores affecting the weights and so on. The application of GSP to sponsored search auctions is described in more detail in, for example, [Jansen and Mullen 2008], [Edelman et al. 2007], [Varian 2007]. There is also an intermediary step between the query and the auction: advertisers bid on a keywords, and the keywords in the query trigger the choice of advertisers that match the query. The recent work on budget constrained bidding in auctions, termed *bid optimization*, has looked at how to bid across *different* keywords (different goods, in our terminology). For example, Borgs et al. [Borgs et al. 2007] consider how to bid across keywords when users want to maximize their ROI (return on investment). [Feldman et al. 2007] show that a simple randomized algorithm has a good competitive ratio. [Even-Dar et al. 2009] discuss the complexity of this bid optimization, and the hardness associated with broadmatch queries.

The literature on learning or strategies in repeated games has links with our work. In the context of auctions [Cary et al. 2007] discuss a greedy (balanced bidding) strategy for playing against the same set of players, and [Vorobeychik and Reeves 2007] look at the evolution of best response dynamics. [Athey and Nekipelov 2010] and [Gomes and Sweeney 2009] consider Bayes-Nash equilibria in a single round: This has some interesting connections with our work. In particular, they show that a symmetric Nash Equilibrium may not exist under certain circumstances, whereas by introducing both budgets and asynchronous regeneration, we show that equilibria do exist. Our framework includes the online stochastic knapsack problem as a special case. The connection between the knapsack and sponsored search has been remarked on by several authors (e.g. [Zhou et al. 2008],[Zhou and Naroditskiy 2008] and references therein).

2. THE OPTIMAL BIDDING STRATEGY IN REPEATED SECOND PRICE AUCTIONS

In the absence of budget constraints, repeated second price auctions from a bidder's perspective are effectively independent, where truthful bidding is a dominant strategy. With budget constraints, truthful bidding may not be sustainable indefinitely, since the bidder could eventually run out of money. In optimizing her bidding strategy over a sequence of auctions, there are two obvious competing factors, whose balance determines the optimal bidding strategy:

— *O*ption Value: Winning the auction at any price below the true valuation clearly generates positive utility. However, by not winning the auction at a certain price, the bidder creates a future opportunity to win an equivalent auction at a lower price. This reduces the amount that a strategic bidder is willing to pay for winning at a value below the true valuation, which we call the *b*id shading effect.
— *L*imited Patience: We assume that bidders have a preference for present utility over future utility. However, bid shading postpones possible present utility to the future. Therefore, a limited patience on part of the bidder moderates the extent of underbidding that she finds to be optimal.

In order to formally characterize the above effects, we consider a budget constrained bidder's optimization problem over repeated second-price auctions with i.i.d. competing bid distribution, with a discounted utility. We formulate the problem as a Markov Decision Process (MDP), and derive an explicit expression for the optimal bid as a function of the balance.

2.1. Model

Time is discrete and indexed by $i = 0, 1, 2, \ldots$. The bidder's initial budget is b. At each time period, this budget is increased by a fixed income $a \geq 0$. At time period i, the bidder participates in a second-price auction. Denote by $b(i)$ her budget at the beginning of time period i. She observes the valuation $v(i)$ of the object sold in that auction, and bids accordingly. The highest bid of her opponents in auction i is $b'(i)$. The instantaneous reward or utility for auction i when bidding $u(i)$, is $(v(i) - b'(i))\mathbb{1}_{\{b(i) \geq u(i) > b'(i)\}}$, and the corresponding cost is $b'(i)\mathbb{1}_{\{b(i) \geq u(i) > b'(i)\}}$. Then the agent's budget evolves as follows:

$$b(i+1) = b(i) + a - b'(i)\mathbb{1}_{\{b(i) \geq u(i) > b'(i)\}}. \tag{1}$$

Assume that valuations and highest opponent bids are i.i.d. over time with distributions, $v(i) \sim v$ and $b'(i) \sim b'$[1] for some *generic* random variables v and b'. The agent chooses a bid $u(i)$ for auction i after observing the valuation $v(i)$, subject to the budget constraint at all times. Her objective is to maximize her long-term discounted utility defined as:

$$\sum_{i=0}^{\infty} e^{-\beta i} \mathbb{E}[(v(i) - b'(i)) \mathbb{1}_{\{b(i) \geq u(i) > b'(i)\}}].$$

This resulting MDP has a state at time i defined by $(b(i), v(i))$. The set of actions corresponds to the set of feasible bids, and the state transitions are dictated by Equation (1). We assume that an agent needs to bid strictly more than all opponents to win an auction, for simplicity. However, this can be relaxed to include any other tie breaking rule without loss of generality.

2.2. The value function and the optimal bidding strategy

Let $v_\beta(b)$ be the value function of the above MDP, i.e., it is the supremum, over all possible bidding strategies, of the expectation of the discounted utility starting from an initial budget b. The dynamic programming principle implies that[2]:

$$v_\beta(b) = \mathbb{E}_v \left[\max_{u \leq b} \mathbb{E}_{b'} \left[\mathbb{1}_{\{u > b'\}} T_1 + \mathbb{1}_{\{u \leq b'\}} T_2 \right] \right] \tag{2}$$

where $T_1 = v - b' + e^{-\beta} v_\beta(b + a - b')$ represents the expected utility obtained if the agent wins the auction with valuation v against a highest opponent bid of b', and $T_2 = e^{-\beta} v_\beta(b + a)$ is the expected utility when the auction is lost. Note that the optimization over the bid is performed inside the expectation operator $\mathbb{E}_v$, which captures the fact that the valuation is known to the bidder before bidding. Rewriting, we can express the optimal bid as a function of b, v, and a as:

$$u^\star_\beta(b, v, a) = \arg\max_{u \leq b} \mathbb{E}_{b'} \left[\mathbb{1}_{\{b \geq u > b'\}} \big(v - \lambda(b', b)\big) \right], \tag{3}$$

where the function λ is defined by:

$$\lambda(b', b) = b' + e^{-\beta}(v_\beta(b + a) - v_\beta(b + a - b')) \tag{4}$$

The following theorem gives explicit compact expressions for the optimal bid and the value function.

THEOREM 2.1. *For any v and b, let $\lambda^{-1}(v, b)$ be the unique x that satisfies $\lambda(x, b) = v$. λ^{-1} is well defined and the optimal bidding strategy is characterized as:*

$$u^\star_\beta(b, v, a) = \min(\lambda^{-1}(v, b), b).$$

Assume without loss of generality that the support for valuations and the maximum opponent bid is in $[0, 1]$. Then the value function satisfies the following fixed-point equation, for all $b \geq 1$:

$$v_\beta(b) = e^{-\beta} v_\beta(b + a) + \mathbb{E}\big[v - \lambda(b', b)\big]^+. \tag{5}$$

Remark 2.2. From Equation (4), $\lambda(., b)$ may be interpreted as the *magnification* function of the highest opponent's bid as seen by the agent at a current budget b.

[1] $a \sim b$ means that a and b are identically distributed.
[2] $\mathbb{E}_x$ denotes the expectation w.r.t. the distribution of r.v. x.

In other words, for each possible value of the opponent's bid b', the agent's optimal bid is obtained by mapping the current auction to a stand alone second price auction, where the opponent's bid is mapped to $\lambda(b', b)$. The two terms in $\lambda(b', b)$ represent the following two costs which are subtracted from the valuation v to optimize the agent's bid. (1) The first term, b' is the instantaneous cost of having to pay the price b', equal to the opponent's bid, if the agent were to win the auction. (2) The second term, $e^{-\beta}(v_\beta(b+a) - v_\beta(b+a-b'))$, represents the *opportunity cost* involved in winning the current auction at price b', with respect to extracting utility from bidding the same amount over the future auctions, if the agent were to lose the current auction.

In the next Section, the issue of computing the value function and the associated optimal bidding strategy is explored in further detail. A key result is the identification of value function in a limiting scenario of practical interest for ad auctions.

3. COMPUTATION OF THE VALUE FUNCTION AND THE OPTIMAL BID

Section 2 provides a structural characterization of the optimal bid. For practical applications, it is of interest to be able to explicitly identify the value function and the associated optimal bid as a function of the remaining balance for any given opponent bid distribution, which will be the focus of this section. Theorem 2.1 suggests the following iterative procedure to numerically estimate v_β. Define $v_\beta^0(b) = 0, \;\; \forall b \in \mathbb{R}_+$ and recursively compute v_β^{t+1}, for $t = 1, 2, \ldots$ as follows. Let $b_{\max}$ be an upper bound for the bid support. For all $b \geq b_{\max}$,

$$v_\beta^{t+1}(b) = e^{-\beta} v_\beta^t(b+a) + \mathbb{E}[v - (b' + e^{-\beta}(v_\beta^t(b+a) - v_\beta^t(b+a-b')))]^+.$$

In ad auctions, there are two important factors that motivate the scaling behavior considered in this section. These are: (1) bidders typically participate in a large number of auctions, i.e. the optimization horizon is large. (2) The transaction amounts in each auction are miniscule (cents or a few dollars) in comparison to the magnitudes of budgets and utilities generated over the long run (which could be several hundreds or thousands of dollars). Even though the factors (1) and (2) might seem unrelated, scaling by a single parameter, β turns out to be sufficient to characterize the limiting behavior motivated by considering both these constraints simultaneously. More precisely, (1) implies that it is of interest to model a discount rate that is close to zero, i.e. $\beta \approx 0$. (2) implies that the arguments of interest for the value function $v_\beta(.)$ involve ranges much larger than a, b'. Furthermore, the magnitude of the value function itself is much larger than a, b', which is consistent with $\beta \approx 0$, since the utility is effectively aggregated over a large time horizon for small discount rates. This regime makes it inefficient to use value iteration for estimating v_β, since both $v_\beta(.)$ and the argument range over which v_β has to be approximated need to be large. This is illustrated in the appendix, where the convergence of the value iteration procedure is graphically represented for various discount factors.

The following scaling of value function models both (1) and (2) simultaneously:

$$V_\beta(B) \triangleq \beta v_\beta(B/\beta)$$

Theorem 3.1 shows that $V_\beta(.)$ has a well defined limit when $\beta \to 0$, and explicitly identifies its limit, $V(.)$. In fact, this limiting characterization applies to a more general class of online budget problems, of which repeated auctions is one specific instance. Other instances of this framework in the context of broadcast scheduling may also be found in [Gummadi 2011].

Practical Implications: Besides the theoretical interest in characterizing the limit, Theorem 3.1 is also useful as a practical tool for approximately computing the optimal bidding strategy under scenarios where (1) and (2) discussed above apply. To find the optimal bidding strategy, the limiting value function is first computed using theorem 3.1 by modeling the opponent bid distribution and the income parameter. Since the scaling parameter $\beta \to 0$ has two complementary interpretations, either could be used in practice to fix its value. One method is to interpret it directly as the discount factor, by perhaps using an estimate for the probability that the bidder participates in a fixed large number of auctions. The other method is to estimate the asymptotic value function $v_\beta(b)$ as $b \to \infty$, which can be interpreted as the maximum utility that the bidder could obtain if she were to participate repeatedly without any budget constraints. β can then be fixed to be the ratio of $V(\infty)/v_\beta(\infty)$, from the scaling definition used. Having modeled β appropriately, at any current balance, b, the bidder can estimate her optimal bid shading factor using the derivative of the limiting value function, $V(.)$ at the scaled argument, βb. Note that under assumption (2), b could be large, but βb remains finite since β is small.

3.1. A General Online Budgeting Problem

We now describe a general model of stochastic online budget optimization that includes both the repeated second price auctions considered in Section 2, and the Generalized Second Price Auctions. The model consists of an agent taking actions repeatedly in a random environment, generating utilites that require payments subject to a budget constraint.

Random environment. A random environment that is i.i.d. over time affects the utilities and payments. For time period i, the environment is described by the random variable $\xi(i)$, taking values in $\mathbb{R}^N$. $\xi(i)$, $i = 0, 1, \ldots$ are independent and identically distributed, with $\xi(i) \sim \xi$, where ξ denotes a generic independent copy. Let $\mathcal{F}(i)$ be the σ-algebra generated by $\xi(i)$. ξ is partially observable, and $\mathcal{F}_0(i) \subset \mathcal{F}(i)$ denotes a σ-algebra that represents the observable part of the environment for time period i.

Actions. At the beginning of time period i, the agent chooses an action from a compact set U based on the observable part of the environment. Formally, this action is represented by a $\mathcal{F}_0(i)$-measurable random variable $u(i)$.

Utility. The utility obtained in a time period depends on the action and the random environment, and is given by a bounded function $g : U \times \mathbb{R}^N \to \mathbb{R}$. The agent's objective function is her expected infinite horizon discounted utility, $\sum_{i=0}^{\infty} e^{-\beta i}\mathbb{E}[g(u(i), \xi(i))]$.

Resource constraint and feasible strategy. The agent's resources are limited, and involve a consumption or payment at each time. The payment made by the agent in a time period depends on the action and the random environment, denoted by a positive and bounded function $c : U \times \mathbb{R}^N \to \mathbb{R}_+$. We assume that there exists an action which consumes no resource and provides no utility, i.e. $\exists 0 \in U$ such that $c(0, \xi) = 0 = g(0, \xi)$, for all ξ. A fixed income a is also added to the resource at each time period. Let $b(i)$ denote the amount of resources or budget available at the beginning of time period i. The agent's budget evolves as follows:

$$b(i + 1) = b(i) + a - c(u(i), \xi(i)), \quad b(0) = b, \tag{6}$$

where b is the initial budget. The agent is forbidden from taking any sequence of actions that could lead to a negative balance. More precisely, given an initial budget b, the set $\mathcal{U}_b$ of feasible policies is the set of all sequences of actions $u(i) \in \mathcal{F}_0(i)$ for all i such that almost surely, $b(i) \geq 0$ for all i.

Value Function. The resulting discrete time constrained MDP has the following value function:

$$v_\beta(b) = \sup_{u \in \mathcal{U}_b} \left(\mathbb{E} \sum_{i=0}^{\infty} e^{-\beta i} g(u(i), \xi(i)) \right). \tag{7}$$

3.2. Examples of Auctions

3.2.1. The second-price auction. Following section 2, the agent is a bidder participating in a sequence of second-price auctions who wishes to maximize the infinite horizon discounted surplus utility. The random environment in time period i is given by $\xi(i) = (v(i), b'(i))$, where $v(i)$ is the agent's valuation of the object sold in the corresponding auction, assumed to be observable, and $b'(i)$ is the highest opponent's bid. The agent's action, represented by u, denotes the bid. The utility and payment functions are: $g(u,\xi) = \mathbb{1}_{u>b'}(v - b')$ and $c(u,\xi) = \mathbb{1}_{u>b'}b'$, for all $u \in \mathbb{R}^+$ and all $\xi = (v, b')$.

3.2.2. The Generalized Second Price auction (GSP). In a GSP auction, the agent is an advertiser competing for one of L ad-slots, in decreasing order of preference, for every auction. Let $b'(i) = (b'_1(i), \ldots, b'_L(i))$ the random vector of bid thresholds required to win the corresponding slot. Upon being displayed, the advertiser may or may not receive a click, which has a valuation $v(i)$. Let $\chi_l(i)$ denote the indicator variable for a click in slot l, with $\chi(i) = (\chi_1(i), \ldots, \chi_L(i))$. The random environment for auction i is $\xi(i) = (v(i), b'(i), \chi(i))$. Let $u(i)$ denote the bid. The utility and payment are: $g(u,\xi) = \sum_{l=1}^{L} \mathbb{1}_{\{b'_{l-1}<u\le b'_l\}}\chi_l(v - b'_l)$ and $c(u,\xi) = \sum_{l=1}^{L} \mathbb{1}_{\{b'_{l-1}\le u<b'_l\}}\chi_l b'_{l-1}$, for all $u \in \mathbb{R}^+$ and all $\xi = (v, b', \chi) \in \mathbb{R}_+ \times \mathbb{R}_+^L \times \{0,1\}^L$.

3.3. Limiting Value Function and Optimal Control

The following theorem provides a method for approximating the value function and the associated optimal control when the discount factor $e^{-\beta}$ is close to 1. This represents a situation where the discount factor is effective only over a large number of time slots. An alternative view is that this describes a scenario where the magnitude of transactions in individual time slots is miniscule in comparison to the infinite horizon valuations. Consider the scaling described below:

$$V_\beta(B) = \beta v_\beta(b) = \beta v_\beta(B/\beta),$$

$v_\beta(b)$ is defined in (7) as the value function of the initial MDP. An easy way to remember the above notation is to note that a capital letter variable is related to its corresponding small letter version, scaled down by a factor of β.

Let ξ be a generic copy of $\xi(i)$, and $\mathcal{F}_0$ be the σ-algebra corresponding to its observable part. For $u \in \mathcal{F}_0$, the expectations of the utility and payments are denoted as: $\bar{g}(u) = \mathbb{E}[g(u,\xi)]$ and $\bar{c}(u) = \mathbb{E}[c(u,\xi)]$, respectively.

THEOREM 3.1. *Let $\phi : \mathbb{R}_+ \mapsto \mathbb{R}$ be defined as:*

$$\phi(x) = ax + \sup_{u\in\mathcal{F}_0} (\bar{g}(u) - \bar{c}(u)x), \tag{8}$$

then ϕ is a convex and Lipschitz function with a minimum denoted $\eta_\star = \min_{x\ge 0}\phi(x)$. Let $\phi^{-1} : \mathbb{R} \mapsto \mathbb{R}_+$ be an inverse to ϕ defined as: $\phi^{-1}(y) = \min\{x \ge 0 : \phi(x) = y\}$. Then, for all $B \ge 0$, $V(B) = \lim_{\beta\to 0} \beta v_\beta(B/\beta)$ is well defined, and satisfies the ODE:

$$\frac{dV}{dB} = \phi^{-1}(V), \quad V(0) = \eta_\star. \tag{9}$$

Furthermore, when the scaled budget is B, an optimal control solves $\sup_{u\in\mathcal{F}_0}(\bar{g}(u) - \bar{c}(u)V'(B))$.

The proof involves relating the discrete time problem being considered to that of a continuous time version, which can be solved using the HJB equation. To do this, a parametrized version of the discrete time problem is considered, where the agent receives a fixed terminal utility (the parameter) upon exhausting the balance. The value function of this variant is shown to converge to the value function of an analogous continuous time MDP using the limit theorems of [Kushner 1990]. The primary technical difficulty involves showing that the limit of V_β matches the limit of the value function of the parametrized variant, for an appropriate value of the terminal utility.

3.3.1. Application to the repeated second-price auction. Consider the repeated second-price auction, defined in Subsection 3.2.1. To apply Theorem 3.1, we first need to compute the function ϕ corresponding to this problem. We have:

$$\phi(x) = ax + \sup_{u\in\sigma(v)} \mathbb{E}\left[\mathbb{1}_{\{u>b'\}}(v - b'(1+x))\right] \tag{10}$$

$$= ax + (1+x)\sup_{u\in\sigma(v)} \mathbb{E}\left[\mathbb{1}_{\{u>b'\}}\left(\frac{v}{1+x} - b'\right)\right]. \tag{11}$$

Note that the above supremum is achieved using the smallest possible bid greater (or equal) than $v/(1+x)$. When the set of possible bids U is continuous, i.e., is $\mathbb{R}_+$, then the supremum is achieved for bid $u = v/(1+x)$. When the set of bids is finite, $U = \{u_1 = 0, \ldots, u_K\}$ (the set is ordered, $u_{k+1} > u_k$), then the supremum is achieved for $u = \min\{b \in U : b \geq v/(1+x)\}$. We have (in both cases, either continuous or finite possible bids):

$$\phi(x) = ax + \mathbb{E}\left[(v - b'(1+x))^+\right]. \tag{12}$$

$V(B)$ is now computed using Theorem 3.1. Finally, when the (scaled) budget is B, the optimal bid solves:

$$\sup_{u\in\sigma(v)} \mathbb{E}\left[\mathbb{1}_{\{u>b'\}}(v - b'(1+V'(B)))\right].$$

In the case of a continuous set of bids, the optimal bid is a function of the valuation v and the (scaled) budget B, and is:

$$u^\star(v,B) = \frac{v}{1+V'(B)}. \tag{13}$$

This bid may be interpreted as the optimal bid in a second-price auction in which the bidder has a *shaded valuation* $\frac{v}{1+V'(B)}$.

Similarly, in the case of a finite set of bids (as defined above), the optimal bid (at scaled budget B) is:

$$u^\star(v,B) = \min\{b \in U : b \geq \frac{v}{1+V'(B)}\}. \tag{14}$$

Numerical example. In Figures 1-4, we illustrate the analytical results obtained in the case of repeated second-price auctions. The agent may use any bid in $\mathbb{R}_+$. The highest opponent bid is distributed over $\{1,2,\ldots,10\}$ with probability vector proportional to $\{1,2,3,1,1,1,2,3,2,1\}$ and the self valuation has a distribution supported on $\{4,5,6\}$ with probabilities $\{0.2,0.6,0.2\}$. Plots are obtained for different values of the income a.

3.3.2. Application to the repeated GSP auction. We now apply Theorem 3.1 to repeated GSP auctions as described in Subsection 3.2.2. In what follows we just consider a continuous set of feasible bids, although the analysis can be readily extended to the case of a discrete set of bids. We first construct the corresponding function ϕ:

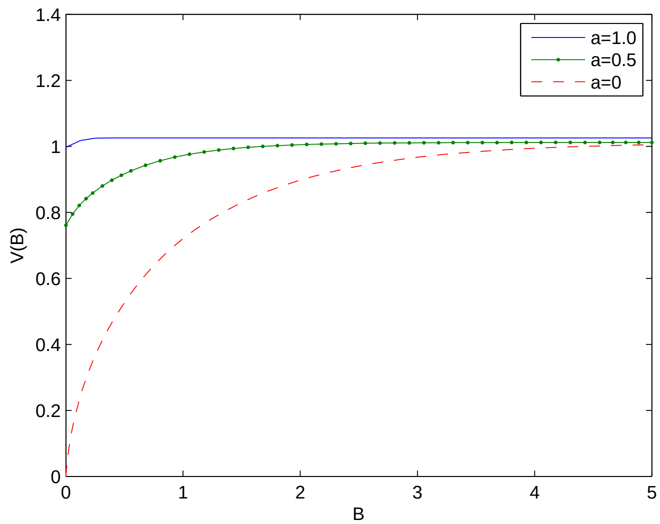

Fig. 1. Value functions $V(B)$ plotted for different income rates a.

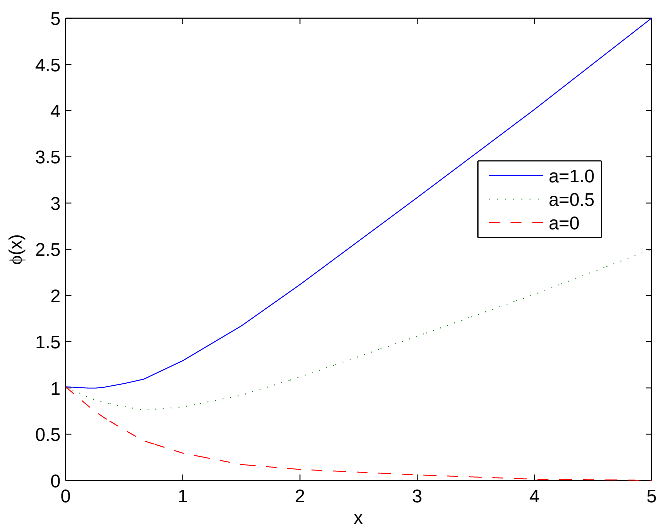

Fig. 2. Plots of $\phi(x)$ (used to define the ODE that specifies the value function) for different income rates a.

$$\phi(x) = ax + \sup_{u \in \mathbb{R}_+} \mathbb{E}\left[\sum_{l=1}^{L} \mathbb{1}_{\{b'_{l-1} < u \leq b'_l\}} \chi_l (v - b'_{l-1}(1+x))\right] \tag{15}$$

$$= ax + (1+x) \sup_{u \in \mathbb{R}_+} \mathbb{E}\left[\sum_{l=1}^{L} \mathbb{1}_{\{b'_{l-1} < u \leq b'_l\}} \chi_l \left(\frac{v}{1+x} - b'_{l-1}\right)\right]. \tag{16}$$

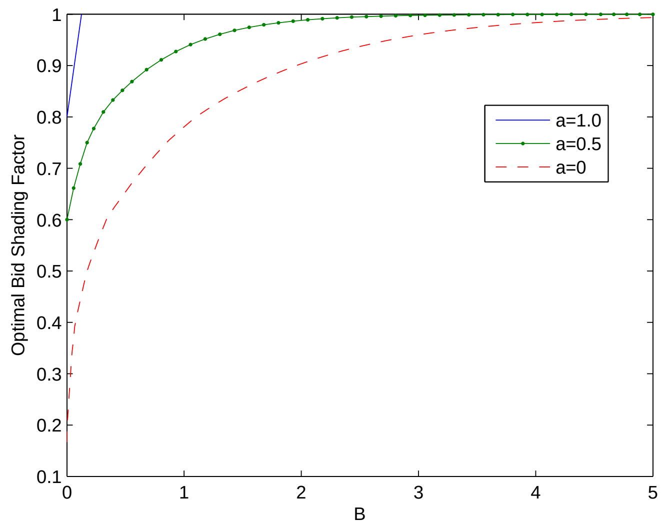

Fig. 3. The optimal bid shading factor, $\frac{1}{1+V'(B)}$ as a function of the budget B for different income rates a.

The scaled value function $V(B)$ is then computed using the inverse of function ϕ as before. Now the optimal bid depends on the (scaled) budget only, and one can easily check that:

$$u^\star(B) = \arg\max_{u\in\mathbb{R}_+} \mathbb{E}\left[\sum_{l=1}^{L} \mathbb{1}_{\{b'_{l-1}<u\le b'_l\}}\chi_l\left(\frac{v}{1+V'(B)} - b'_{l-1}\right)\right].$$

Let $u^\star_{GSP}(v)$ denote the optimal bid maximizing the expected utility in a stand-alone GSP auction with the given opponent bid distribution. Then, once the above value function has been computed, we can express the optimal bid in terms of optimal bid in a stand-alone GSP auction at current balance B:

$$u^\star(B) = u^\star_{GSP}\left(\frac{v}{1+V'(B)}\right) \tag{17}$$

Although the GSP was described assuming that v is unobservable, the above characterization also holds even if v was observable. In this case, $u^\star(B,v)$, the optimal bid at current budget B and for an observed valuation v, is given by (17). Therefore, both repeated second-price and GSP auctions under budget constraints share the property that the optimal bidding strategy corresponds to the optimal bid in a corresponding single (stand-alone) auction where the valuation is shaded by a factor that depends on the current budget.

4. MEAN FIELD EQUILIBRIA

In this section, we introduce a Mean Field game for a repeated second price auction. The analysis and result presented below can be readily extended to the case of repeated GSP auctions. The game is played by a very large population of bidders with different valuations, initial budgets, and incomes. We assume that the set of bids and valuations are finite: $U = \{u_1 = 0, \ldots, u_K\}$ and the set of valuations is $\mathcal{V} = \{v_1, \ldots, v_J\}$. Let W denote a probability measure on $\Delta^J \times \mathbb{R}^2_+$ that defines the valuation distribution, the initial (scaled) budget, and income of a bidder sampled uniformly at random; where for any integer m, Δ^m denotes the $m-1$ simplex $\Delta^m = \{x \in [0,1]^m : \sum_{j=1}^m x_j = 1\}$.

The discount factor $e^{-\beta}$ used in previous sections to compute the value function of a bidder may be interpreted as follows. After each auction, the bidder leaves the system with probability $1-e^{-\beta}$, and remains active with probability $e^{-\beta}$. Optimizing the discounted utility over an infinite time-horizon is then equivalent to optimizing the un-discounted utility over the time the bidder remains active, i.e., over her *life-time*. We use this interpretation to define our game. We consider the limiting regime where β tends to 0, and by renormalizing time, the bidder life-time is exponentially distributed with unit mean. During her life-time, the bidder is involved in a large number of auctions. At the end of her life-time, the bidder disappears and is replaced by a new bidder whose valuation distribution, initial budget, and income is drawn according to probability measure W.

The game proceeds as follows. For each auction, α active bidders are selected uniformly at random[3], and compete. The random selection of bidders is made so that each active bidder is selected (in the limiting regime) at a positive rate γ over her life-time. Each bidder then aims at maximizing the average utility over her life-time. In summary, the game is defined through the probability measure W and parameter γ.

To interpret the Mean Field game as a limit of a sequence of games with finite population of bidders, we may consider repeated second-price auctions played by N bidders, where for each auction, a given bidder is selected with probability α/N. To get the appropriate limiting games when $N \to \infty$, we require that the time-horizon of any bidder scales as $h(N)$ where $\lim_{N\to\infty} h(N)/N = \infty$. The convergence of the sequence of games with finite bidder population to the corresponding Mean Field game can be analyzed using classical Mean Field asymptotics techniques, and is out of the scope of this paper.

Assume now that in the Mean Field game, when selected a bidder faces an i.i.d. highest opponent bid distribution p. She would then apply the corresponding optimal bidding strategy as described in the previous section. Let us consider a bidder with valuation distribution f, initial budget B, income a, and life-time T. In response to the i.i.d. highest bid distribution p, this bidder applies an optimal bidding strategy as characterized in the previous section, and the resulting empirical distribution of her bid over her life-time is denoted by $\psi(p, f, B, a, \gamma T)$. Note that the life-time is reduced by a factor γ to account for the fact that each bidder is selected to compete at rate γ only. This distribution characterizes the proportions of time the bidder selects the various possible bids during her life-time (e.g., $\psi(p, f, B, a, \gamma T)_k$ is the proportion of time the bidder uses bid u_k during her life-time). Further define the long-term empirical bid distribution of a bidder characterized by $f, B, a)$ as:

$$\bar{\psi}(p, f, B, a) = \int_0^\infty \psi(p, f, B, a, \gamma T) e^{-T} dT.$$

Denote by $X(p, f, B, a)$ a random variable with distribution $\bar{\psi}(p, f, B, a)$. Finally define the mapping $\Gamma : \Delta^K \to \Delta^K$ where $\Gamma(p)$ is the distribution of the highest bid of $\alpha - 1$ bidders sampled uniformly at random (their valuation distribution, initial budget and income have distribution W) when they apply the optimal bidding strategy when facing the highest bid distribution p. In other words, $\Gamma(p)$ is the distribution of the following random variable:

$$\max\left(X_1, \ldots, X_{\alpha-1}\right)),$$

[3]This assumption is not essential, and the selection of bidders could be biased, e.g. towards bidders with higher budget or valuations.

where the random variables $(X_1, \dots, X_{\alpha-1})$ are i.i.d. copies of random variable $X(p, f, B, a)$ and where (f, B, a) has distribution W. We are now able to define the notion of a Mean Field equilibrium. In such equilibrium, each bidder faces i.i.d. highest opponent bids, and has no incentive to change her bidding strategy.

Definition 4.1. A Mean Field equilibrium is defined as a bid distribution p, such that

$$\Gamma(p) = p.$$

THEOREM 4.2. *The repeated second-price auction Mean Field games admit at least one equilibrium.*

The above theorem is proved in appendix. The theorem follows from Brouwer's theorem provided that the mapping Γ is continuous. We obtain the required continuity by showing that $p \mapsto \bar{\psi}(p, f, B, a)$ is continuous for all (f, B, a). By conditioning on the bidder life-time, the continuity of $\bar{\psi}$ results from the continuity of $p \mapsto \psi(p, f, B, a, \gamma T)$. To prove the latter continuity, we use some structural properties of the optimal bidding strategy. More precisely, it can be easily shown that this optimal bidding strategy is defined through a finite sequence of *switching* points $t_1, \dots, t_m$ where at time t_j the agent changes her bid, and actually decreases her bid. The continuity of ψ in p is obtained by showing that the switching points are continuous with respect to p. This results from observing that the function ϕ in Theorem 3.1, the value function V and its derivative V' are continuous in p. In view of the construction of the optimal bidding strategy through V', the continuity of V' in p implies that of the bid switching points, and finally that of $p \mapsto \psi(p, f, B, a, \gamma T)$.

As mentioned at the beginning of this section, the above theorem is also valid in case of repeated GSP auctions.

5. FINITE TIME HORIZON

Finally, we investigate the case where agents maximize their (un-discounted) utility over a finite time-horizon. We first characterize the asymptotically optimal strategy assuming that an agent faces an i.i.d. environment and that the time horizon t is large. We then study the corresponding Mean Field equilibria.

5.1. Model and asymptotically optimal bidding strategy

5.1.1. Model. Let us first consider a single agent wishing to maximize her un-discounted utility over a finite time horizon. The model is similar to that described in Subsection 3.1. The only difference is that the value function is un-discounted and computed over a finite time-horizon t:

$$v(b, t) = \sup_{u \in \mathcal{U}_b} \sum_{i=0}^{t-1} \mathbb{E}[g(u(i), \xi(i)].$$

In the present model, there is no income, i.e., $a = 0$. Hence, when action $u(i)$ is selected for time period i, the budget evolves according to:

$$b(i + 1) = b(i) - c(u(i), \xi(i)).$$

The initial budget is $b(0) = b$.

5.1.2. Asymptotically optimal strategy. Characterizing the value function is also difficult here, and to progress we consider scenarios with a large time-horizon. More precisely, we scale the time-horizon and initial budget by a factor $1/\beta$, and define:

$$V_\beta(B, T) = \beta v(B/\beta, \lfloor T/\beta \rfloor).$$

Now when β tends to 0, the limiting value function is characterized as follows.

THEOREM 5.1. *For all* $B, T \geq 0$, $V(B,T) = \lim_{\beta \to 0} \beta v(B/\beta, \lfloor T/\beta \rfloor)$ *is well defined, and satisfies the PDE:*

$$\frac{\partial V}{\partial B} = \phi^{-1}(\frac{\partial V}{\partial T}), \quad V(0,T) = 0, V(B,0) = 0. \tag{18}$$

The theorem is proved (see [Kushner 1990]) by observing that $V(B,T)$ is the value function of the following continuous time control problem:

$$V(B,T) = \sup_{u \in \bar{\mathcal{U}}} \left(\int_0^{\tau \wedge T} \bar{g}(u(t)) dt \right),$$

where $\bar{\mathcal{U}}$ is the set of mappings from $[0,T]$ to $\mathcal{F}_0$, $\tau = \inf\{t \geq 0 : B(t) = 0\}$, and

$$\frac{dB(t)}{dt} = -\bar{c}(u(t)), \forall t \geq 0, \quad B(0) = B.$$

In general, this problem is easier to solve than its discounted counter-part studied earlier in the paper.

We now characterize its solutions explicitly in the cases of repeated second-price and GSP auctions.

5.1.3. Repeated second-price auctions. To simplify the exposition, we restrict our attention to second-price auctions with a finite set of bids. The set of bids is $U = \{u_1 = 0, \ldots, u_K\}$ and is ordered ($u_k < u_{k+1}$). Valuations also take a finite set $\mathcal{V} = \{v_1, \ldots, v_J\}$ of values. We assume that if the agent places the same bid as the highest opponent bid, she loses the auction. Identifying the optimal bidding strategy is similar to solving a classical relaxed Knapsack (KP) problem [Lueker 1995]. Indeed, the auction taking place in time period i, with valuation $v(i)$ and opponent highest bid $b'(i)$, corresponds to an *item* in a KP problem. Accepting the item in the KP problem means that the agent wins the corresponding auction, in which case her pay-off and cost are $v(i) - b'(i)$ and $b'(i)$, respectively. Now the agent simply has to determine which auctions she should win (depending on the corresponding valuation and highest opponent bid). It is well known [Lueker 1995] that the following greedy algorithm provides the solution of the relaxed KP problem. The algorithm specifies the probability with which the agent should win an auction where the valuation is v_j and the highest opponent bid is u_k, for any couple (j,k). Denote by p_k (resp. f_j) the probability that the highest opponent bid is u_k (resp. that the valuation is v_j).

We start by ordering the auctions characterized by the couple (v_j, u_k) in decreasing payoff density defined as $(v_j - u_k)/u_k$. For any bijection $\pi : \{1, \ldots, J \times K\} \to \{1, \ldots, J\} \times \{1, \ldots, J\}$, we denote $\pi(l) = (\pi_1(l), \pi_2(l))$ for all $l \in \{1, \ldots, J \times K\}$. We construct a bijection π such that for all $l < J \times K$, $(v_{\pi_1(l)} - u_{\pi_2(l)})/u_{\pi_2(l)} \leq (v_{\pi_1(l+1)} - u_{\pi_2(l+1)})/u_{\pi_2(l+1)}$.

Greedy algorithm.
$l \leftarrow 0$, $V \leftarrow 0$, and $b \leftarrow B$.
While ($b \geq T f_{\pi_1(l)} p_{\pi_2(l)} u_{\pi_2(l)}$) and ($l < J \times K$), do
  win auctions $\pi(l)$ with probability 1;
  $b \leftarrow b - T f_{\pi_1(l)} p_{\pi_2(l)} u_{\pi_2(l)}$;
  $V \leftarrow V + T f_{\pi_1(l)} p_{\pi_2(l)} (v_{\pi_1(l)} - u_{\pi_2(l)})$.
Win auctions $\pi(l)$ with probability $B/(T f_{\pi_1(l)} p_{\pi_2(l)} u_{\pi_2(l)})$.
Loose auctions $\pi(l')$ for all $l' > l$ with probability 1.

In other words the optimal strategy is to consider winning auctions in decreasing order of their pay-off density until the budget is exhausted, or until no auction is left (if the decision maker has a sufficient budget to win all possible auctions). In turn, the above algorithm characterizes an optimal bidding strategy. Indeed, it provides a mapping from the set of valuations to a distribution of bids supported by at most two consecutive bids. For example, if for valuation v_j, the algorithm tells us that the agent should win the auction when the opponent bid is up to u_k with probability 1, then the decision maker just bids u_{k+1} with probability 1. Similarly, if for valuation v_j, the agent should win auctions with opponent bid u_{k-1} with probability 1, and u_k with probability c, then she should bid u_k with probability $(1-c)$ and u_{k+1} with probability c. Note that there exists at most one auction such that the corresponding distribution of bids has positive mass on two bids.

It is worth remarking that the optimal bidding strategy does not depend on the remaining budget, and hence does not depend on time. Indeed as mentioned above, for a given auction, an optimal strategy maps the observed valuation to a bid distribution. The agent plays the same strategy over time until her budget reaches 0, or until the end of the time-horizon is reached.

5.1.4. Repeated GSP auctions. Again we restrict our attention to finite sets of feasible bids and valuations. As it turns out, the problem of identifying the optimal bidding strategy is equivalent to solving a relaxed multi-dimensional Knapsack problem. This is due to the fact that the expected utility only depends upon the proportions of time for which the various bids are used; we define α_k to be the proportion of time the agent uses bid u_k. We also let $\bar{g}_k$ and $\bar{c}_k$ denote the expected utility and cost per time unit respectively when the agent bids u_k. The optimal bidding strategy then solves the following relaxed multi-dimensional KP problem:

$$\max \sum_{k=1}^{K} T\alpha_k \bar{g}_k,$$
$$\text{s.t.} \sum_{k=1}^{K} T\alpha_k \bar{c}_k = B, \quad \sum_{k=1}^{K} T\alpha_k = 1.$$

It is known [Kellerer et al. 2004] that the solution of the above problem has at most two non-binary components. This implies that either there exists a bid u_k such that $\alpha_k = 1$, and $\alpha_{k'} = 0$ for all $k' \neq k$, or there exist two bids $u_k, u_{k'}$ such that $\alpha_k \alpha_{k'} > 0$, and $\alpha_{k"} = 0$ for all $k" \neq k, k'$. Unfortunately, we were not able to identify additional structural properties of the problem to help us characterize the optimal bidding strategy. For example, it might well be that the bid corresponding to the highest pay-off density, defined for bid u_k as $\bar{g}_k/\bar{c}_k$, is not selected in the optimal strategy. It is also possible that the two optimal bids are not consecutive (the optimal bids are not necessarily u_k, u_{k+1} for some k). This is illustrated in the simple following example. Assume that there are three possible bids with respective expected utility and cost equal to $(10, \delta)$, $(g, 1)$, and $(12, 3)$, where $g \in (10, 12)$ and $\delta < 1$. The initial budget is equal to 2, so that the third bid has to be used to exhaust the budget. Now if $\delta = 11$, it can be checked that the optimal strategy consists in using the second and third bids, whereas when $g = 10 + \delta$, then using the first and third bids is optimal for all δ. Nevertheless, the fact that the optimal bidding strategy concentrates on two bids only makes it easy to identify.

5.2. Mean Field equilibria

We now define a Mean field game for repeated second-price auctions (we could define a similar game for repeated GSP auctions). The game is played by a very large population of bidders with different valuations, initial budgets and finite time-horizons. We assume that the set of bids and valuations are finite, see §5.1.3. Let W denote a probability measure on $\Delta^J \times \mathbb{R}_+^2$ that defines the valuation distribution, the initial (scaled) budget, and time-horizon of a bidder sampled uniformly at random. A bidder with budget B and time-horizon T reinitializes her budget periodically every T time units, i.e., at times $t + kT$ for all $k \in \mathbb{N}$ (for some fixed time t), her budget is set to B.

The game proceeds as follows. For each auction, α bidders are selected uniformly at random, and compete. We consider the following limiting regime: each bidder, say with time-horizon T, is selected to compete for a large number of auctions during any periods of length T. Equivalently, the rate at which any bidder is selected to compete is $\gamma > 0$ per time unit. In summary, the Mean Field game is defined through the probability measure W and parameter γ.

Assume now that when selected in the Mean Field game, a bidder faces a i.i.d. highest opponent bid distribution p . She would then apply the corresponding optimal bidding strategy as described in the previous subsection (remember that this strategy is stationary, i.e., it does not depend on time). Define by $\phi(p, v, f, B, \gamma T)$ the optimal bid distribution of a bidder with valuation distribution f, initial budget B, and time-horizon γT, when the realization of her valuation is v. Note that the time-horizon is γT because each bidder is assumed to be selected at rate γ. Now the empirical bid distribution of such bidder over time will be:

$$\bar{\phi}(p, f, B, \gamma T) = \sum_{j=1}^{J} f_j \phi(p, v_j, f, B, \gamma T).$$

Denote by $X(p, f, B, T)$ a random variable with distribution $\bar{\phi}(p, f, B, \gamma T)$. Finally define the mapping $\Gamma : \Delta^K \to \Delta^K$ where $\Gamma(p)$ is the distribution of the highest bid of $\alpha - 1$ bidders sampled uniformly at random (their valuation distribution, initial budget and time-horizon have distribution W) when they apply the optimal bid strategy when facing highest bid distribution p. In other words, $\Gamma(p)$ is the distribution of the following random variable:

$$\max\left(X_1, \ldots, X_{\alpha-1}\right)),$$

where the random variables $(X_1, \ldots, X_{\alpha-1})$ are i.i.d. copies of $X(p, f, B, \gamma T)$ where (f, B, T) has distribution W. We are now able to define the notion of Mean Field equilibrium. In such equilibrium, each bidder faces i.i.d. highest opponent bids, and does not have any incentive to change her bidding strategy.

Definition 5.2. A Mean Field equilibrium is defined as a bid distribution p, such that

$$\Gamma(p) = p.$$

THEOREM 5.3. *(Mean Field Equilibrium - Finite time-horizon)*
The repeated second-price auction Mean Field games admit at least one equilibrium.

Proof. Observe that by construction of the optimal bidding strategy via the Greedy algorithm, the mappings $\bar{\phi}(p, f, B, \gamma T)$ are continuous w.r.t. p. Now we deduce that Γ is continuous, and the theorem follows from Brouwer fixed point theorem. □

6. CONCLUSION AND FUTURE WORK

We have studied a general model of a repeated auction, and described the optimal bidding strategy for an agent who is constrained by her budget. Starting from a discrete time model, discounting rewards over time, and then taking the limit as the discount rate tends to 1 enables us to prove that the limiting scaled value function $V(B)$ at budget B satisfies a particular differential equation. The optimal bidding strategy in then simply related to the derivative of limiting value function. The form of this optimal strategy is particularly simple in the case of a repeated second-price or GSP auctions: bid *as if* bidding in an (unconstrained) single shot auction (SP or GSP respectively), but shading the true value by a factor $1/(1 + V'(B))$.

The limiting regime requires scaling the budgets appropriately (by $1/\beta$) as $\beta \to 0$. It is an appropriate regime when there are a large number of auctions but the gain in each auction is small. For ad-auctions in thick markets this is an appealing scaling. In other settings, the limiting behavior can be used as an approximation for when the discount factor is close to 1.

When there are not only a large number of auctions, but also a large number of potentially competing agents, we have described a corresponding Mean Field game, and proved existence of Mean Field Equilibria (MFE). The equilibria appear relatively simple to calculate. The implications are that given a distribution of valuations and/or initial budgets, the resulting MFE will give a stationary distribution of bids, and hence a straightforward calculation of shading factors that only depend upon the budget.

Of course we have assumed that the distributions of budgets and valuations are known. In the case of a single agent, she would have to infer these distributions, or simply attempt to learn the opponent's bid distribution and then react to that distribution. The challenge then is to show that asynchronous best response learning converges to a MFE. The auctioneer is in a position to calculate bid distributions, and by inferring valuations, could use the information to smooth the budgets of bidders as a way of increasing social welfare.

ACKNOWLEDGMENTS

The first author would like to thank Michael Padilla for his careful reading and feedback on the online budgeting problem.

REFERENCES

ADLAKHA, S., JOHARI, R., AND WEINTRAUB, G. 2011. Equilibria of dynamic games with many players: Existence, approximation, and market structure. Submitted for publication.

ATHEY, S. AND NEKIPELOV, D. 2010. A structural model of sponsored search advertising auctions. In *Sixth Ad Auctions Workshop*.

BERTSEKAS, D. 2007. Dynamic programming and optimal control. *Athena Scientific*.

BORGS, C., CHAYES, J., IMMORLICA, N., JAIN, K., ETESAMI, O., AND MAHDIAN, M. 2007. Dynamics of bid optimization in online advertisement auctions. In *Proceedings of the 16th international conference on World Wide Web*. WWW '07.

CARY, M., DAS, A., EDELMAN, B., GIOTIS, I., HEIMERL, K., KARLIN, A. R., MATHIEU, C., AND SCHWARZ, M. 2007. Greedy bidding strategies for keyword auctions. In *Proceedings of the 8th ACM conference on Electronic commerce*. EC '07.

CHE, Y.-K. AND GALE, I. 1998. Standard auctions with financially constrained bidders. *The Review of Economic Studies*.

EDELMAN, B., OSTROVSKY, M., AND SCHWARZ, M. 2007. Internet advertising and the generalized second price auction: Selling billions of dollars worth of keywords. *American Economic Review*, 242–259.

EVEN-DAR, E., MANSOUR, Y., MIRROKNI, V. S., MUTHUKRISHNAN, S., AND NADAV, U. 2009. Bid optimization in broad-match ad auctions. *CoRR abs/0901.3754*.

FELDMAN, J., MUTHUKRISHNAN, S., PAL, M., AND STEIN, C. 2007. Budget optimization in search-based advertising auctions. In *Proceedings of the 8th ACM conference on Electronic commerce*. EC '07. ACM, New York, NY, USA, 40–49.

GOMES, R. D. AND SWEENEY, K. S. 2009. Bayes-nash equilibria of the generalized second price auction. In *Proceedings of the 10th ACM conference on Electronic Commerce*. EC '09. New York, NY, USA.

GUMMADI, R. 2011. Coding and Scheduling in Networks for Erasures and Broadcast. Ph.D. thesis, University of Illinois at Urbana Champaign.

IYER, K., JOHARI, R., AND SUNDARARAJAN, M. 2011. Mean field equilibria of dynamic auctions with learning. In *Proceedings of the 12th ACM conference on Electronic commerce*. EC '11. ACM, New York, NY, USA, 339–340.

JANSEN, B. J. AND MULLEN, T. 2008. Sponsored search: an overview of the concept, history, and technology. *Int. J. Electronic Business 6,* 2, 114–131.

KELLERER, H., PFERSCHY, U., AND PISINGER, D. 2004. *Knapsack problems*. Springer.

KUSHNER, H. 1990. Numerical methods for stochastic control problems in continuous time. *SIAM Journal on Control and Optimization 28*.

LUEKER, G. S. 1995. Average-case analysis of off-line and on-line knapsack problems. In *Proceedings of the sixth annual ACM-SIAM symposium on Discrete algorithms*. SODA '95. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 179–188.

OSTROVSKY, M. AND SCHWARZ, M. 2011. Reserve prices in internet advertising auctions: a field experiment. In *Proceedings of the 12th ACM conference on Electronic commerce*. EC '11. ACM, New York, NY, USA, 59–60.

PIN, F. AND KEY, P. 2011. Stochastic variability in sponsored search auctions: observations and models. In *Proceedings of the 12th ACM conference on Electronic commerce*. EC '11. ACM, New York, NY, USA, 61–70.

PITCHIK, C. 2009. Budget-constrained sequential auctions with incomplete information. *Games and Economics Behavior 66*.

VARIAN, H. R. 2007. Position auctions. *International Journal of Industrial Organization*, 1163–1178.

VOROBEYCHIK, Y. AND REEVES, D. M. 2007. Equilibrium analysis of dynamic bidding in sponsored search auctions. In *Proceedings of the 3rd international Workshop on Internet and Network Economics*. WINE'07. Springer-Verlag.

ZHOU, Y., CHAKRABARTY, D., AND LUKOSE, R. M. 2008. Budget constrained bidding in keyword auctions and online knapsack problems. In *WINE*. 566–576.

ZHOU, Y. AND NARODITSKIY, V. 2008. An algorithm for stochastic multiple-choice knapsack problem and keywords bidding. In *Seventeenth International World Wide Web Conference*. 1175–1176.

A. APPENDIX

A.1. Proof of Theorem 2.1

First note that for any b, $\lambda(b', b)$ is strictly increasing in the argument b' (since $-v_\beta(b + a - b')$ is non decreasing in b' and there is an additional term that increases strictly with b'). This implies that

$$\mathbb{1}_{b \geq u > b'} = \mathbb{1}_{\lambda(b,b) \geq \lambda(u,b) > \lambda(b',b)},$$

and hence

$$u^*_\beta(b, v, a) \triangleq \arg\max_{u \in \mathbb{R}_+} \mathbb{E}_{b'} \left[\mathbb{1}_{\lambda(b,b) \geq \lambda(u,b) > \lambda(b',b)} \left(v - \lambda(b', b) \right) \right]. \tag{19}$$

Now the optimum bid in a *static* second price auction with valuation v and budget constraint b is $\min(v, b)$ and hence for any $v \in \mathbb{R}_+$:

$$\arg\max_{u \in \mathbb{R}_+} \mathbb{E}_{b'} \left[\mathbb{1}_{b \geq u > b'} \left(v - b' \right) \right] = \min(v, b). \tag{20}$$

Rewriting (20) with b, b' replaced by $\lambda(b, b), \lambda(b', b)$ respectively, gives:

$$\arg\max_{u \in \mathbb{R}_+} \mathbb{E}_{b'} \left[\mathbb{1}_{\lambda(b,b) \geq u > \lambda(b',b)} \left(v - \lambda(b', b) \right) \right] = \min(v, \lambda(b, b)).$$

By replacing the bid b by the function $\lambda(., b)$ we have:

$$\arg\max_{u \in \mathbb{R}_+} \mathbb{E}_{b'} \left[\mathbb{1}_{\lambda(b,b) \geq \lambda(u,b) > \lambda(b',b)} \left(v - \lambda(b', b) \right) \right] = x,$$

where x is the unique solution to $\lambda(x, b) = \min(v, \lambda(b, b))$. Since λ is a one to one strictly increasing map, we have $x = \min(\lambda^{-1}(v, b), b)$. From (19), this implies that:

$$u^\star_\beta(b, v, a) = \min(\lambda^{-1}(v, b), b).$$

Clearly, $u^\star_\beta(b, v, a)$ is non-decreasing in (b, v, a) if $\lambda^{-1}(v, b)$ is non-decreasing in (b, v, a).

Without loss of generality, let us assume that the maximum bid or valuation amount in a single auction is 1.

Now assume that the current balance, $b \geq 1$. Under this assumption:

$$v_\beta(b) - e^{-\beta} v_\beta(b + a) = \mathbb{E}_v \left[\max_{u \in \mathbb{R}_+} \mathbb{E}_{b'} \left[\mathbb{1}_{u > b'} \left(v - (b' + e^{-\beta}(v_\beta(b + a) - v_\beta(b + a - b'))) \right) \right] \right] \tag{21}$$

Note that $\lambda(x, b) \geq x$ and since $v \leq 1$, this implies that $\lambda^{-1}(v, b) \leq 1 \leq b$. Therefore:

$$u^\star_\beta(b, v, a) = \min(\lambda^{-1}(v, b), b) = \lambda^{-1}(v, b).$$

This implies for any fixed v, a, b:

$$\begin{aligned}
&\max_{u \in \mathbb{R}_+} \mathbb{E}_{b'} \left[\mathbb{1}_{u > b'} \left(v - (b' + e^{-\beta}(v_\beta(b + a) - v_\beta(b + a - b'))) \right) \right] \\
&\quad = \mathbb{E}_{b'} \left[\mathbb{1}_{u^\star_\beta(b,v,a) > b'} \left(v - \lambda(b', b) \right) \right] \\
&\quad = \mathbb{E}_{b'} \left[\mathbb{1}_{\lambda^{-1}(v,b) > b'} \left(v - \lambda(b', b) \right) \right] \\
&\quad = \mathbb{E}_{b'} \left[\mathbb{1}_{v > \lambda(b',b)} \left(v - \lambda(b', b) \right) \right] \\
&\quad = \mathbb{E}_{b'} \left[v - \lambda(b', b) \right]^+ \quad \text{where } [.]^+ \triangleq \max(., 0)
\end{aligned}$$

(21) now becomes:

$$v_\beta(b) = e^{-\beta} v_\beta(b + a) + \mathbb{E}\left[v - \lambda(b', b) \right]^+, \tag{22}$$

which concludes the proof.

A.2. Value iteration convergence

We illustrate the convergence rate of the value iteration procedure to compute the value function for different values of β. We consider an example where the competing opponent bid is distributed over $\{1, 2, \ldots, 10\}$ with probability vector proportional to $\{1, 2, 3, 1, 1, 1, 2, 3, 2, 1\}$ and where the bidder's valuation has a distribution supported on $\{4, 5, 6\}$ with probabilities $\{0.2, 0.6, 0.2\}$. In this example, we set $a = 0$. The convergence of value iteration is graphically depicted in the following three figures, Figures 5 to 7 for $\beta = 0.1, 0.05, 0.01$, respectively. Note that in these figures, the x and y axes are scaled by factor $1/\beta$. We clearly observe that the number of iterations required to obtain convergence significantly grows as β decreases to 0.

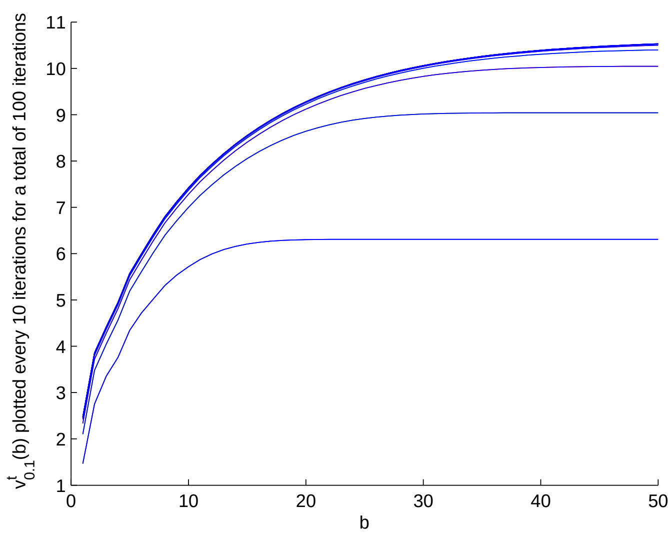

Fig. 4. $v^t_{0.1}(b)$ plotted every 10 iterations

A.3. Proof of Theorem 3.1

The proof consists of the following steps. First we decompose the MDP into two sub-problems: a first sub-problem where the exit-payoff or exit-utility is fixed (the time horizon corresponds to the time when the budget reaches 0, and at this time the agent receives an exit-utility), and a second sub-problem where the exit-utility is identified.

A.3.1. Solution of the problem with exit pay-off. For proof purposes, it is useful to introduce a variant of the initial MDP, with identical state transitions, but without the budget constraint. Instead, the variant is parametrized with an exit payoff, η, which defines the final utility obtained when the balance reaches a value less than or equal to zero for the first time. More precisely, let $\mathcal{U}'$ represent the collection of all admissible Markov policies. Therefore, $\mathcal{U}'$ represents all sequences of actions, defined by random variables $u(i) \in \mathcal{F}_0$, which can be chosen after observing the corresponding states, $b(i)$ (but without the additional restriction on actions that could lead to $b(i+1) < 0$). Let κ denote the following stopping time:

$$\kappa = \inf\{i \geq 0 : b(i) \leq 0\}$$

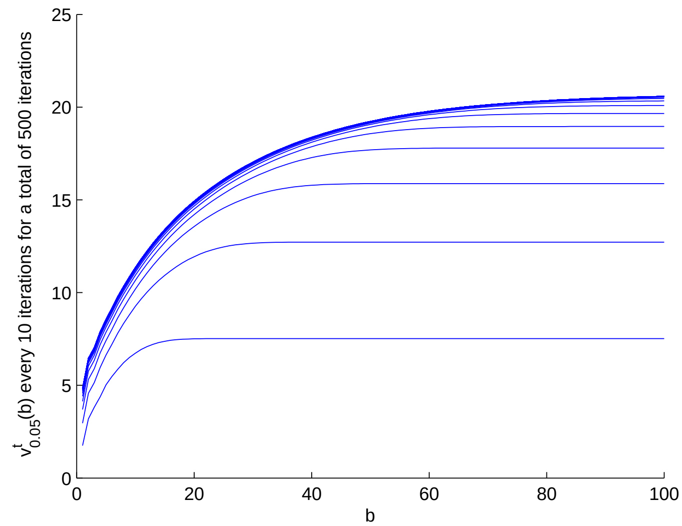

Fig. 5. $v^t_{0.05}(b)$ plotted every 10 iterations

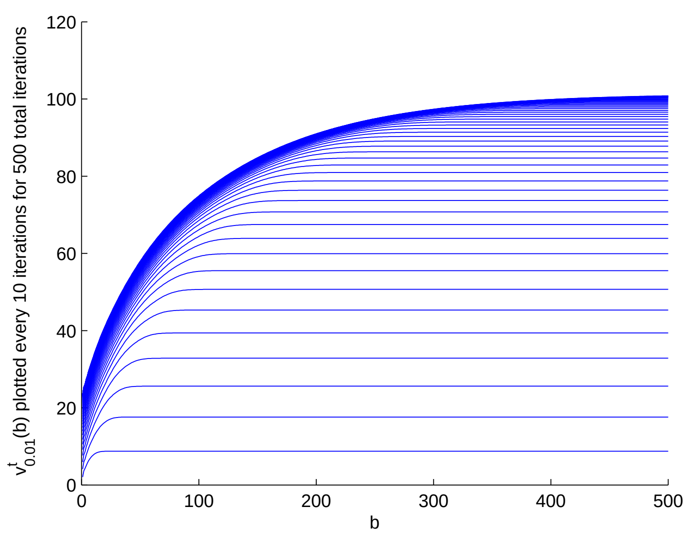

Fig. 6. $v^t_{0.01}(b)$ plotted every 10 iterations

The value function of this modified MDP is given by:

$$j_\beta(b,\eta) \triangleq \sup_{u\in\mathcal{U}'} \left(\sum_{i=0}^{\kappa} e^{-\beta i}\bar{g}(u(i)) + e^{-\kappa\beta}\eta \right).$$

Following prior convention, the scaled version of the value function is given by:

$$J_\beta(B,\eta) = \beta j_\beta(B/\beta, \eta/\beta).$$

To illustrate the convergence of the above problem to its continuous time version, consider the following change of variables/notation: $t_i \triangleq \beta i, \Delta t_i \triangleq t_{i+1} - t_i = \beta$, $B(t_i) \triangleq \beta b(i), \tau \triangleq \beta\kappa$ and $u(t_i) \triangleq u(i),\ \forall i$. Using this change of notation:

$$J_\beta(B,\eta) = \sup_{u \in \mathcal{U}'} \left(\mathbb{E}_b \sum_{t_i=0}^{\tau} e^{-t_i} \bar{g}(u(t_i))\Delta t_i + e^{-\tau}\eta \right) \tag{23}$$

where $\tau = \inf\{t_i \geq 0 : B(t_i) \leq 0\}$. The state transitions can be expressed as:

$$\Delta B(t_i) \triangleq B(t_{i+1}) - B(t_i) = (a - \bar{c}(\mathbf{u}(t_i)))\Delta t$$

From standard results on convergence of discrete to continuous time control problems [Kushner 1990], the scaled value function J_β converges to the value function of the following continuous time control problem:

$$\bar{\mathcal{U}} \triangleq \{u : u(t) \in \mathcal{F}_0,\ \forall\ t \geq 0\},$$

$$V(B,\eta) = \sup_{u \in \bar{\mathcal{U}}} \left(\int_0^\tau e^{-t}\bar{g}(u(t))\, dt + e^{-\tau}\eta \right) \tag{24}$$

where $\tau = \inf\{t \geq 0 : B(t) = 0\}$, with state evolution:

$$\frac{dB(t)}{dt} = a - \bar{c}(u(t)), \forall t \geq 0, \quad B(0) = B.$$

This is a deterministic continuous time optimal control problem, whose value function can be characterized by the corresponding Hamilton-Jacobi-Bellman (HJB) equation [Bertsekas 2007], i.e., by an ODE with a boundary condition. The boundary condition is $V(0,\eta) = \eta$, while the ODE is specified by simplifying the following dynamic programming relation: for a small Δt,

$$V(B) \geq \sup_{u \in \mathcal{F}_0} \left(\bar{g}(u)\Delta t + e^{-\Delta t} \left(V(B) + \Delta t(a - \bar{c}(u))V'(B) \right) \right)$$

This gives

$$V(B) = aV'(B) + \sup_{u \in \mathcal{F}_0} \left(\bar{g}(u) - \bar{c}(u)V'(B) \right) = \phi(V'(B)).$$

ϕ is convex because it is defined as the supremum of a family of linear functions. From the definition of $\eta_\star$ and the convexity of ϕ, it follows that ϕ is strictly decreasing on $[0, \phi^{-1}(\eta_\star)]$. From Lemma A.8, presented at the end of this proof, ϕ is continuous,[4] implying that ϕ^{-1} is well defined and continuous on $[\eta_\star, \eta_0]$ where $\eta_0 \triangleq \phi(0)$.[5] Therefore, the ODE, $\frac{dV}{dB} = \phi^{-1}(V), \quad V(0) = \eta$, has a continuously differentiable solution. And from HJB sufficient condition, it is equal to $V(B,\eta)$, the value function corresponding to the continuous time problem defined in (24).

To summarize the analysis so far, we have:

LEMMA A.1. *Let $\eta \in [\eta_\star, \eta_0]$. Then, $\lim_{\beta\to 0} J_\beta(B,\eta) = V(B,\eta)$, which satisfies the ODE: $\frac{dV}{dB} = \phi^{-1}(V)$, with initial condition, $V(0) = \eta$.*

[4]In fact, ϕ is Lipschitz though this fact is not necessary for the proof.
[5]ϕ^{-1} is also Lipschitz on $[\eta, \eta_0], \ \forall \eta \in (\eta_\star, \eta_0]$.

A.4. Characterizing the Exit Pay-off

Theorem 3.1 states that $\lim_{\beta\to 0} V_\beta(B) = V(B, \eta_\star)$. Combining Lemma A.3 and Lemma A.6, we prove that $\lim_{\beta\to 0} V_\beta(0) = \eta_\star$. This implies that: $\forall\ B \geq 0$, $\lim_{\beta\to 0} V_\beta(B) = \lim_{\beta\to 0} J_\beta(B, \eta_\star) = V(B, \eta_\star)$, proving the theorem.

LEMMA A.2. $\forall \eta, h > 0, B \geq 0$, $J_\beta(B + h, \eta) \geq V_\beta(B)$

Proof. Any optimal policy for $V_\beta(B)$ provides a lower bound for $J_\beta(B+h, \eta)$, since such policy is feasible for the control problem with exit pay-off η and initial budget $B + h$. Since the state transitions are identical in both variants, the stopping time is ∞ in the exit time variant, which implies that $V_\beta(B)$ is an achievable utility in the exit time variant starting from initial balance $B + h$ for any $h > 0$. □

LEMMA A.3.

$$\limsup_{\beta\to 0} V_\beta(0) \leq \eta_\star$$

Proof. Lemma A.1, Lemma A.2 and the fact that $V(B, \eta)$ is continuous in B for any $\eta \in [\eta_*, \eta_0]$ together imply Lemma A.3. Specifically, letting $B = 0, \eta = \eta_*$ in Lemma A.2 and letting $\beta \to 0$, we get $\forall h > 0, \quad \lim_{\beta\to 0} J_\beta(h, \eta_*) \geq \limsup_{\beta\to 0} V_\beta(0)$. From Lemma A.1, we have $\lim_{\beta\to 0} J_\beta(h, \eta_*) = V(h, \eta_*)$, implying $\forall h > 0, \limsup_{\beta\to 0} V_\beta(0) \leq V(h, \eta_*)$, which implies the lemma due to continuity of V. □

Before stating and proving Lemma A.6, we need the following preliminary analysis.

Definition A.4. [ϵ-subdifferential] For $\epsilon \geq 0$, $c \in \mathbb{R}$ is called an ϵ-subdifferential to a function f at x_0 if $f(x) + \epsilon \geq f(x_0) + c(x - x_0),\ \forall\ x$. A 'subdifferential' refers to a 0-subdifferential.

LEMMA A.5. *For any* $\epsilon > 0,\ \exists\ u^\star \in \mathcal{F}_0$ *such that* $\bar{g}(u^\star) > \eta_\star - \epsilon$ *and* $\bar{c}(u^\star) \leq a$.

Proof. Let $x^\star = \sup\{x \geq 0 : \phi(x) = \eta_\star\}$ and $x_\star = \inf\{x \geq 0 : \phi(x) = \eta_\star\}$ (both of which are finite if $a > 0$, and when $a = 0$, we have $\eta_\star = 0$, in which case the result is trivially true by choosing $u^\star = 0$). We first consider the case where $x_\star = x^\star$. For any $h > 0$, any subdifferential of ϕ at $x^\star + h$ is strictly positive because $\phi(x^\star + h) > \phi(x^\star)$. Let us choose $\theta(h) > 0$ be small enough so that any θ-subdifferential is also strictly positive for $\theta < \theta(h)$ (see Proposition A.7). Let $\theta = \min(\epsilon/2, \theta(h))$. Let $\psi(u, x) \triangleq \bar{g}(u) + (a - \bar{c}(u))x$. Choose $u^\star_+ \in \mathcal{F}_0$ such that $\psi(u^\star_+, x^\star + h) > \phi(x^\star + h) - \theta(h)$. Then, for any x, $\phi(x) \geq \psi(u^\star_+, x) = \psi(u^\star_+, x^\star + h) + (a - \bar{c}(u^\star_+))(x - x^\star - h) > \phi(x^\star + h) + (a - \bar{c}(u^\star_+))(x - x^\star - h) - \theta(h)$, implying that $a - \bar{c}(u^\star_+)$ is a $\theta(h)$-subdifferential of ϕ at $x^\star + h$. Therefore, $\bar{c}(u^\star_+) \leq a$. Furthermore, since $\eta_\star < \phi(x^\star + h) < \theta(h) + \psi(u^\star_+, x^\star + h)$, we have:

$$\eta_\star < \theta + \bar{g}(u^\star_+) + (a - \bar{c}(u^\star_+))(x^\star + h). \tag{25}$$

Analogously, we can pick $u^\star_- \in \mathcal{F}_0$ such that $\bar{c}(u^\star_-) \geq a$ and

$$\eta_\star < \theta + \bar{g}(u^\star_-) + (a - \bar{c}(u^\star_-))(x^\star - h). \tag{26}$$

Now we can choose $u^\star = \delta u^\star_- + (1 - \delta)u^\star_+$ where[6]

$$\delta = \begin{cases} 1 & \text{with probability } \alpha = \frac{a - \bar{c}(u^\star_+)}{\bar{c}(u^\star_-) - \bar{c}(u^\star_+)} \\ 0 & \text{with probability } 1 - \alpha = \frac{\bar{c}(u^\star_-) - a}{\bar{c}(u^\star_-) - \bar{c}(u^\star_+)} \end{cases}$$

[6]if $\bar{c}(u^\star_+) = \bar{c}(u^\star_-) = a$, then we can pick any $\alpha \in [0, 1]$.

so that: $\bar{c}(u^\star) = a$. Using this, inequality $(25) \times (1-\alpha) + (26) \times \alpha$ gives $\eta_\star < \bar{g}(u^\star) + \theta + 2(a+C)h$. Now choose any $h < \frac{\epsilon}{4(a+C)}$ and since $\theta < \epsilon/2$, we have: $\eta_\star < \bar{g}(u^\star) + \epsilon$.

It remains to argue for the case $x_\star < x^\star$. For given $h > 0$, we may pick $u_+^\star$ exactly as before because the subdifferential at $x^\star + h$ is again strictly positive. However, in this case, the only subdifferential for any $x_0 \in (x_\star, x^\star)$ is zero. Therefore, for any $h > 0$, $\exists\ \theta(h) > 0$ such that any $\theta-$subdifferential at $x^\star - h$ has to lie in $(-h, h)$. As before we can pick $u_-^\star \in \mathcal{F}_0$ such that Equation (26) holds and $a - \bar{c}(u_-^\star)$ is a subdifferential at $x^\star - h$, implying that $a - \bar{c}(u_-^\star) \in (-h, h)$ (where we chose $\theta = \min(\epsilon/2, \theta(h))$). If $a \leq \bar{c}(u_-^\star)$, we can go through the same construction as before to obtain $u^\star$ for which $\bar{c}(u^\star) = a$ and $\eta_\star < \bar{g}(u^\star) + \epsilon$. Now suppose $a > \bar{c}(u_-^\star)$. From Equation (26), we get $\eta_\star < \theta + \bar{g}(u_-^\star) + h(x^\star - h) > \bar{g}(u_-^\star) + \epsilon$ if we pick $h < \frac{\epsilon}{2x^\star}$. Therefore, in this case, for $u^\star = u_-^\star$, we have $\bar{c}(u^\star) < a$ and $\bar{g}(u^\star) > \eta_\star - \epsilon$, as required. □

LEMMA A.6.

$$\liminf_{\beta \to 0} V_\beta(0) \geq \eta_\star$$

Proof. It is sufficient to show that there exists a feasible policy that can achieve a scaled infinite horizon discounted utility with a lower bound that is arbitrarily close to $\eta_\star$. Let $u^\star \in \mathcal{F}_0$ with $\bar{g}(u^\star) = \eta_\star$ and $a = \bar{c}(u^\star)$. Consider a policy in which the decision maker chooses the random variable $u^\star$ in every time slot, i.e.:

$$U^\star \triangleq \{u^\star, u^\star, \ldots\}.$$

Therefore, the state transitions are defined by $b(i+1) = b(i) + w(i)$, where $w(1), w(2), \ldots$ are i.i.d. copies of their generic version, $w \triangleq a - c(u^\star, \xi)$, which is bounded by $|w| \leq 1$ without loss of generality, and has zero drift: $\mathbb{E}[w] = 0$. Therefore, $b(i)$ represents a random walk with zero drift, which has a positive probability of falling below zero. This random walk has a strictly positive probability of going below zero. Therefore, $U^\star$ is not feasible. Consider a modified walk, for which, whenever $b(i) \leq 1$, we set $u'(j) = 0,\ \ \forall j = i, i+1, \ldots i + \frac{M}{a}$, where M is a constant that will be fixed later on. For all other j, $u'(j) = u^\star$. Clearly, the modified policy is feasible for the original problem. Let x denote the utility achieved by this policy starting from $b(0) = M$. Let $\tau_M \triangleq \min\{i \geq 0 : \sum_{j=0}^{i} w(i) \leq -M + 1\}$. We have:

$$x = \mathbb{E}\left[\sum_{i=0}^{\tau} g(u^\star, \xi) + e^{-\beta(\tau + \frac{M}{a})} x\right]$$

This can be simplified to give:

$$x = \left(\frac{\eta_\star}{1 - e^{-\beta}}\right) \left(\frac{1 - \mathbb{E}[e^{-\beta(\tau_M + 1)}]}{1 - \mathbb{E}[e^{-\beta(\tau_M + M)}]}\right)$$

which implies:

$$V_\beta(0) \geq \beta e^{-\beta \frac{M}{a}} x = \eta_\star e^{-\beta \frac{M}{a}} \left(\frac{\beta}{1 - e^{-\beta}}\right) \left(\frac{1 - \mathbb{E}[e^{-\beta(\tau_M + 1)}]}{1 - \mathbb{E}[e^{-\beta(\tau_M + M)}]}\right)$$

For any fixed M, as $\beta \to 0$, the lower bound above becomes arbitrarily close to $\eta_\star \frac{\mathbb{E}[\tau_M] + 1}{\mathbb{E}[\tau_M] + M}$. For large M, the expected hitting time for the asymmetric random walk with maximum step size equal to one is at most (stochastically bounded by) the expected hitting time of a symmetric random walk with step size equal to one, which scales as $\mathbb{E}[\tau_M] = \Omega(M^2)$. Therefore, the lower bound is arbitrarily close to $\eta_\star$, proving the lemma. □

PROPOSITION A.7. *Suppose all subdifferentials of a convex function, f, at x belong to $[d_l, d_h]$. Then, for any $\delta > 0$, $\exists\, \epsilon > 0$ small enough, such that any ϵ-subdifferential of f at x belongs to $(d_l - \delta, d_h + \delta)$.*

PROOF. Draw two lines, L_1 and L_2 with slopes $d_l - \delta, d_h + \delta$ at x. Clearly, there exist $x_l < x$ and $x_h > x$ such that L_1 strictly dominates f in (x_l, x) and L_2 strictly dominates f in (x, x_h). Choose:

$$\epsilon^* = \min\left(\sup\{L_1(x) - f(x) : x \in (x_l, x)\}, \sup\{L_2(x) - f(x) : x \in (x, x_h)\}\right)$$

Then $\epsilon^* > 0$ and for any $\epsilon < \epsilon^*$, neither $d_l - \delta$ nor $d_h + \delta$ can be ϵ-subdifferentials, and therefore any ϵ-subdifferential of f has to lie in $(d_l - \delta, d_h + \delta)$. □

LEMMA A.8. *ϕ is Lipschitz.*

Proof. Let $\psi(u, x) \triangleq \bar{g}(u) + (a - \bar{c}(u))x$, so that $\phi(x) = \sup_{u \in \mathcal{F}_0} \psi(u, x)$. Let $|x - y| < \delta$. Let $u^\star(x), u^\star(y) \in \mathcal{F}_0$ such that $\psi(u^\star(x), x) \in (\phi(x) - \delta, \phi(x)]$ and $\psi(u^\star(y), y) \in (\phi(y) - \delta, \phi(y)]$ and assume $\psi(u^\star(x), x) \geq \psi(u^\star(y), y)$ without loss of generality. Then, $|\phi(x) - \phi(y)| \leq \delta + \psi(u^\star(x), x) - \psi(u^\star(y), y)$. Note that $\psi(u^\star(y), y) > \phi(y) - \delta \geq \bar{g}(u^\star(x)) + (a - \bar{c}(u^\star(x)))\, y - \delta \geq \psi(u^\star(x), x) - (1 + |a - C|)\delta$, where C is the bound on the consumption function. Therefore, $|\phi(x) - \phi(y)| < M\delta$, where $M = 2 + |a - C|$ is the Lipschitz constant. □

A.5. A remark on the rate of convergence as $\beta \to 0$

Theorem 3.1 shows that if β is small enough, the approximation error involved in using $V(B)$, to compute $\beta v_\beta(B/\beta)$ is arbitrarily small. We now provide a simple argument to estimate the rate of this convergence. We consider the specific case where the budget constraints are not effective, i.e., $B \to \infty$. Denote by $g^\star$ the maximum expected utility that the agent may receive in a single time period. When the balance is unlimited, we have the following limit for the value function:

$$\lim_{B \to \infty} v_\beta(B/\beta) = \sum_{i=0}^{\infty} e^{-\beta i} g^\star = \frac{g^\star}{1 - e^{-\beta}}.$$

This implies that the convergence when $\beta \to 0$ of $\beta v_\beta(B/\beta)$ towards $V(B)$ is as fast as that of $\frac{\beta}{1 - e^{-\beta}}$ to 1, at least when the budget B is very large.

A.6. Proof of Theorem 4.2

We show $p \mapsto \bar{\psi}(p, f, B, a)$ is continuous for all (f, B, a). To do so, we condition on the bidder left-time T, and prove the continuity of $p \mapsto \psi(p, f, B, a, \gamma T)$. The proof relies on first studying the function ϕ involved in Theorem 3.1.

$$\phi(x) = ax + \mathbb{E}\left[(v - b'(1 + x))^+\right].$$

In the case of finite sets of bids and valuations, we have:

$$\mathbb{E}\left[(v - b'(1 + x))^+\right] = \sum_{j,k} f_j p_k (v_j - u_k(1 + x)) \mathbb{1}_{\{x \leq \frac{v_j - u_k}{u_k}\}}.$$

Hence ϕ is a continuous piece-wise linear function whose slopes are linear in the components of distribution p, and that changes slopes at most $K \times J$ switching points, $x_{j,k} = \frac{v_j - u_k}{u_k}$, for any j, k. Its pseudo-inverse ϕ^{-1} has the same property: it is piece-wise linear and continuous, and its slopes and switching points depend continuously on p. We simply deduce by integrating ϕ^{-1} that the value function $V(B)$ and its derivative $V'(B)$ at scaled budget B are continuous in p. Now observe (see Section 3) that the optimal bidding strategy consists in changing bids only when the value of $V'(B)$ reaches some thresholds, and from the continuity of $V'(B)$ in p, we conclude that the times at which the optimal strategy switches bids are also continuous in p. Finally, the empirical distribution $\psi(p, f, B, a, \gamma T)$ that characterizes the proportions of time the various bids are used before time T is continuous in p.

We proved that $p \mapsto \bar{\psi}(p, f, B, a)$ is continuous for all (f, B, a). Thus Γ is continuous, and the theorem follows from Brouwer fixed point theorem.