

# DISTRIBUTIONALLY ROBUST PORTFOLIO MAXIMISATION AND MARGINAL UTILITY PRICING IN ONE PERIOD FINANCIAL MARKETS

JAN OBLÓJ\* AND JOHANNES WIESEL\*+

**ABSTRACT.** We consider the optimal investment and marginal utility pricing problem of a risk averse agent and quantify their exposure to a small amount of model uncertainty. Specifically, we compute explicitly the first-order sensitivity of their value function, optimal investment policy and Davis' option prices to model uncertainty. To achieve this, we capture model uncertainty by replacing the baseline model  $\mathbb{P}$  with an adverse choice from a small Wasserstein ball around  $\mathbb{P}$  in the space of probability measures. Our sensitivities are thus fully non-parametric. We show that the results entangle the baseline model specification and the agent's risk attitudes. The sensitivities can behave in a non-monotone way as a function of the baseline model's Sharpe's ratio, the relative weighting of assets in an agent's portfolio can change and marginal prices can increase when an agent faces model uncertainty.

*This paper is dedicated to the memory of Mark H.A. Davis. I remain forever grateful to Mark who introduced me to the world of mathematical finance when I joined his group at Imperial College London in 2006. He was a mentor and a friend. His intellectual curiosity combined with a generosity of spirit and good humour were hugely enriching. His capacity to ask key questions and find elegant and insightful answers set my gold standard. I hope he would have enjoyed this study.*

*Jan Oblój*

## 1. INTRODUCTION

Modern theory of finance enables powerful and versatile quantitative analysis of optimal investment problems. It would be impossible to give justice here to the body of relevant literature. However, since the seminal works of Markowitz [1959], Merton [1969], at its heart are mathematical techniques which compute an agent's investment decisions from two inputs: the agent's risk attitudes and a probabilistic description of the financial market. The former is considered subjective and an important stream of research looks at ways to elucidate an agents' preferences and whether these can be assumed to be *rational*, i.e., to satisfy certain axiomatic properties. The latter input, a probability measure  $\mathbb{P}$ , captures the agent's model

---

*Date:* July 24, 2021.

*Key words and phrases.* Distributionally robust optimisation, Davis marginal utility price, optimal investment, Wasserstein distance, robust finance, model uncertainty, sensitivity analysis.

\*The authors would like to thank Daniel Bartl and Samuel Drapeau for helpful discussions. The authors declare that there are no conflicts of interests. JW acknowledges support from the German Academic Scholarship Foundation.

+Corresponding author: Johannes Wiesel, Department of Statistics, Columbia University, 1255 Amsterdam Avenue, New York, NY 10027.

for the financial market and results from numerous considerations and tradeoffs, such as calibration to empirical data, reproduction of certain stylised features and analytical tractability, among others.

With a fixed baseline model  $\mathbb{P}$ , a rational decision maker, following Merton [1969] and in line with the classical decision theory, see e.g., Von Neumann and Morgenstern [1953], Savage [1951], finds her optimal investment policy  $\pi^*$  by maximising her expected utility of wealth

$$\max_{\pi \in \mathcal{A}} \mathbb{E}_{\mathbb{P}}[u(X^{\pi})],$$

and computes the prices of derivative instruments through marginal utility pricing, see Davis [1997]. However, in practice, the agent is bound to have a degree of uncertainty about their choice of  $\mathbb{P}$ . This is often referred to as *Knightian uncertainty* after Knight [1921]. Following Anderson et al. [2003], we are interested here in misspecifications of  $\mathbb{P}$  which are small, in a statistical sense, but unrestricted otherwise (e.g., these do not need to be absolutely continuous with respect to  $\mathbb{P}$ ). Our main contribution is to understand and quantify the sensitivity of an agent's decisions to such model uncertainty. We compute explicitly the first order change to the key quantities – the agent's value function, the agent's optimal portfolio allocation and their marginal utility prices – in response to small levels of model uncertainty.

From an axiomatic point of view, see Gilboa and Schmeidler [1989], Maccheroni et al. [2006], Schied [2007], Föllmer et al. [2009], an agent who is averse both to risk and model uncertainty, considers a max-min criterion

$$\max_{\pi \in \mathcal{A}} \min_{\tilde{\mathbb{P}} \in \mathcal{P}} \mathbb{E}_{\tilde{\mathbb{P}}}[u(X^{\pi})],$$

see section 3 for discussion and generalisations. While appealing from a decision theoretic point of view, this criterion only introduces an additional subjective input: the family  $\mathcal{P}$  of plausible models for reality. Furthermore, it also often renders the problem intractable with little hope for analytic formulae. In this paper we offer a tried and tested mathematical solution to this problem: we compute the first-order approximation to the mapping giving outputs in function of  $\mathcal{P}$ , i.e., we compute their derivative with respect to  $\mathcal{P}$  at the point  $\mathcal{P} = \{\mathbb{P}\}$ . Specifically, as  $\mathcal{P}$  varies in an infinite dimensional space, we consider  $\mathcal{P}$  given as a ball around  $\{\mathbb{P}\}$  with radius  $\delta$  and compute explicitly the first order expansion of all of the outputs in  $\delta$ . This, we believe, captures the essence of model misspecification as described above, see Anderson et al. [2003].

We use the Wasserstein metric to capture balls around  $\{\mathbb{P}\}$ . We argue that this metric is the natural choice both because of its theoretical properties as well as from a data-driven perspective, see section 4. Our analysis is fully non-parametric and reveals that the sensitivity to model uncertainty is a complex issue, which entangles the baseline model and the agent's risk attitudes. We provide explicit formulae to compute these sensitivities, thus enabling an agent to quantify their exposure to model uncertainty. We believe that these sensitivities offer a more robust and structural tool than any existing, parametric or ad-hoc, methods.

Naturally, the agent's value function is decreasing in  $\delta$ : the introduction of model uncertainty combined with a max-min approach means the agent will never welcome the new source of uncertainty. However, the agent's optimal investment policy will react in a much more complex way. Already in a one dimensional setting, we show that the reduction in the agent's position is not monotone in the baseline model's

Sharpe ratio. In higher dimensions, as the uncertainty affects the individual assets as well as their co-dependence, also the relative weighting of different assets changes. Finally, the agent's marginal utility price of an option, also known as the Davis price after Davis [1997], can both decrease or increase as model uncertainty appears. It always decreases if an agent was not trading in the baseline model. But if the agent trades, then the option can effectively provide insurance for certain adverse model changes and may become more valuable than before.

The rest of the paper is organised as follows. We first briefly introduce the baseline setting and its classical analysis. Then, in section 3, we give a high-level overview of the literature and topics related to model uncertainty. After that, we proceed to introducing our approach and discuss our main findings. These are then illustrated with worked-out examples in section 5. A careful presentation of all the assumptions and theorems then ensues. We look at the optimal investment problem and the marginal utility pricing of options in sections 6 and 7 respectively. Appendix A presents proofs of the classical results recalled in section 2.2, while the supplement Oblój and Wiesel [2021] contains detailed explicit computations for all of the examples of section 5.

## 2. THE BASELINE OPTIMAL INVESTMENT AND OPTION PRICING PROBLEMS

We consider a one-period financial market with  $d$  risky assets, for some  $d \geq 1$ . We assume there is also a deterministic riskless asset and, without loss of generality, express all prices in its discounted units. The prices of risky assets today and at time one are given by the vectors  $S_0$  and  $S_1$  and we let  $X = S_1 - S_0$ . We fix a closed convex state space  $\mathcal{S} \subseteq \mathbb{R}^d$  and we only consider  $X \in \mathcal{S}$ . This can be used, e.g., to ensure that in all considered models the prices remain positive. Throughout the paper, a *model* is synonymous to the distribution of  $X$ : a probability measure supported on  $\mathcal{S}$ . We fix one such  $\mathbb{P}$  and refer to it as the *baseline model*. To simplify the notations, we work on the canonical probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  with  $\Omega = \mathcal{S}$ ,  $X(\omega) = \omega$ ,  $\mathcal{F}_0 = \{\emptyset, \Omega\}$ ,  $\mathcal{F}_1 = \mathcal{F} = \sigma(X)$ .

An agent is able to trade in the market using strategies  $\pi \in \mathcal{A} \subseteq \mathbb{R}^d$ , where  $\mathcal{A}$  is closed and convex. In particular, if the first  $k$  assets are traded with no restrictions and the remaining assets are not traded, then  $\mathcal{A} = \mathbb{R}^k \times \{0\}^{d-k}$ . The scalar product of  $X$  and  $\pi$  is denoted  $\langle X, \pi \rangle = \sum_{i=1}^d \pi_i X_i$ .

**2.1. Optimal investment.** We consider an agent endowed with a utility function  $u$ . Throughout the paper we make the following:

**Assumption 2.1** (Standing Assumption). *The state space  $\mathcal{S}$  and the action space  $\mathcal{A}$  are closed convex subsets of  $\mathbb{R}^d$ . The baseline model  $\mathbb{P}$  is supported on  $\mathcal{S}$  and satisfies the following no-arbitrage and non-degeneracy condition:*

$$\forall \pi \in \mathbb{R}^d \setminus \{0\} \quad \mathbb{P}(\langle X, \pi \rangle > 0) > 0.$$

$\mathcal{D}$  is an open convex subset of  $\mathbb{R}$  and  $u : \mathcal{D} \rightarrow \mathbb{R}$  is strictly concave, strictly increasing, continuously differentiable and bounded from above. The state, action and domain sets are compatible in that there exists  $\varepsilon_0 > 0$  such that

$$\{\langle x, \pi \rangle + \eta : x \in \mathcal{S}, \pi \in \mathcal{A}, |\eta| < \varepsilon_0\} \subseteq \mathcal{D}.$$

We fix a function  $g : \mathcal{S} \rightarrow \mathbb{R}$  and assume that either  $\mathcal{D} = \mathbb{R}$  or  $g$  is bounded.

Consider the optimal investment problem for an agent who is maximising their expected utility of wealth, i.e.,

$$(1) \quad V = \sup_{\pi \in \mathcal{A}} \mathbb{E}_{\mathbb{P}} [u(\langle X, \pi \rangle)].$$

Let us remark that in order to simplify our notation we have not explicitly accounted for interest rates or an initial investment capital in the above formulation. These can be easily incorporated by considering discounted stock prices and composing  $u$  with an affine function. Under our assumptions, the optimal portfolio  $\pi^*$  for  $V$  exists, is unique and, for most choices of  $\mathcal{A}$  considered in the literature, belongs to the (relative) interior  $\mathcal{A}^o$  of  $\mathcal{A}$ . It is characterised through the first order condition

$$\mathbb{E}_{\mathbb{P}} [X \cdot u'(\langle X, \pi^* \rangle)] = 0.$$

We can express this differently by saying that  $X$  is a  $\mathbb{Q}_u$ -martingale, where

$$(2) \quad \frac{d\mathbb{Q}_u}{d\mathbb{P}} = \frac{u'(\langle X, \pi^* \rangle)}{\mathbb{E}_{\mathbb{P}} [u'(\langle X, \pi^* \rangle)]}.$$

This idea was exploited by Rogers [1994] to obtain a proof of the Dalang-Morton-Willinger theorem in discrete time using a dynamic programming approach. It is also crucial for marginal utility pricing which we discuss next.

**2.2. Marginal utility price.** Davis [1997] proposed an innovative way of option pricing based on marginal utility. Consider an option with payoff  $\tilde{g}(S_0, S_1)$ . Note that since  $S_0$  is known and  $X = S_1 - S_0$ , we can put  $g(x) = \tilde{g}(S_0, x + S_0)$  and with no loss of generality simply consider an option with payoff  $g(X)$ . Our standing assumptions on  $g$  were included in Assumption 2.1 and further growth assumptions will be added later. We now consider the maximisation problem

$$(3) \quad V(\varepsilon, p_d) := \sup_{\pi \in \mathcal{A}} \mathbb{E}_{\mathbb{P}} \left[ u \left( -\varepsilon + \langle X, \pi \rangle + \frac{\varepsilon}{p_d} g(X) \right) \right],$$

where  $p_d \geq 0, \varepsilon \in \mathbb{R}$ , and note that  $V(0, p_d) = V$ . We note that  $\varepsilon \mapsto V(\varepsilon, p_d)$  is well-defined for  $\varepsilon > 0$  small enough since

$$\left\{ -\varepsilon + \langle x, \pi \rangle + \frac{\varepsilon}{p_d} g(x) : x \in \mathcal{S}, \pi \in \mathcal{A} \right\} \subseteq \mathcal{D}$$

for  $|\varepsilon|$  small enough by Assumption 2.1. We denote strategies attaining the supremum in the expression above by  $\pi^*(\varepsilon)$ . For notational simplicity and in order to emphasize that  $\pi^*(0)$  does not depend on  $p_d$  (nor on  $g$ ), we still write  $\pi^*$  for the  $V$ -optimiser. The marginal utility price is the one which makes the agent indifferent between buying  $\varepsilon$  options  $g$  and the investment problem without the option  $g$ , where  $\varepsilon$  is small. This is made precise in the following definition.

**Definition 2.2** (Marginal utility price of Davis [1997]). *Suppose that for each  $p_d > 0$ , the function  $\varepsilon \mapsto V(\varepsilon, p_d)$  is differentiable at  $\varepsilon = 0$  and  $\hat{p}_d$  is a solution to*

$$\partial_{\varepsilon} V(0, p_d) := \partial_{\varepsilon} V(\varepsilon, p_d)|_{\varepsilon=0} = 0$$

*Then  $\hat{p}_d$  is called a marginal utility price of the option  $g$ .*

Under the assumption on  $u$  above and with suitable uniform integrability,  $\pi^*(\varepsilon)$  is unique and  $\pi^*(\varepsilon) \rightarrow \pi^*(0)$ . The key result of Davis [1997] is that  $\hat{p}_d$  is also unique and satisfies

$$(4) \quad \hat{p}_d = \mathbb{E}_{\mathbb{Q}_u} [g(X)],$$

where the precise statements and proofs of these results are given in Appendix A. The Davis price can thus be interpreted as the arbitrage-free model price under the martingale measure  $\mathbb{Q}_u$ . It is a linear pricing rule, a consequence of considering *marginal* pricing. An indifference pricing problem for a given quantity of options would naturally lead to a non-linear pricing rule. We refer to Musiela and Zariphopoulou [2008], Henderson and Hobson [2004] and the references therein for a detailed discussion.

While Davis' original article focuses on classical continuous time models (see also Karatzas and Kou [1996], Henderson and Hobson [2004], Hugonnier et al. [2005]), Davis pricing in discrete time models has been investigated, e.g., in Schäl [2000b,a, 2002], Kallsen [2002], Rásonyi and Stettner [2005]. More broadly, (4) gives a way to select a martingale measure used for pricing and can be seen as part of a broader literature considering such measures for incomplete markets. In the case of exponential utility,  $u(x) = -\exp(-\gamma x)$ ,  $\mathbb{Q}_u$  is also the *minimal relative entropy* martingale measure of Frittelli [2000], Rouge and El Karoui [2000]. Indeed, for any  $\mathbb{Q} \in \mathcal{P}(\mathcal{S})$ , using the explicit density in (2), we have

$$H(\mathbb{Q} \mid \mathbb{P}) := \mathbb{E}_{\mathbb{Q}}[\ln(d\mathbb{Q}/d\mathbb{P})] = H(\mathbb{Q} \mid \mathbb{Q}_u) - \mathbb{E}_{\mathbb{Q}}[\gamma \langle X, \pi^* \rangle] - \ln \mathbb{E}_{\mathbb{P}}[u'(\langle X, \pi^* \rangle) / \gamma].$$

Restricting to martingale measures  $\mathbb{Q}$ , the middle term on the right disappears and  $H(\mathbb{Q} \mid \mathbb{Q}_u) \geq 0$  so that

$$\inf_{\mathbb{Q}} H(\mathbb{Q} \mid \mathbb{P}) \geq -\ln \mathbb{E}_{\mathbb{P}}[u'(\langle X, \pi^* \rangle) / \gamma],$$

with equality when  $\mathbb{Q} = \mathbb{Q}_u$ , as required. In the literature on pricing in incomplete markets, other martingale measures have also been discussed, such as the *minimal martingale measure* of Föllmer and Schweizer [1991] or the *variance optimal measure* of Schweizer [1996].

Despite the abundant literature on this topic, to the best of our knowledge, even in the simple one-period framework we discuss here, the impact of model uncertainty on the Davis price has not been investigated. This is the main motivation behind our work which we now discuss in more detail.

### 3. MODEL MISSPECIFICATION

Model uncertainty and robustness to model perturbations or misspecification are of paramount importance in any modelling context. In particular, in quantitative finance and economics, model uncertainty has been an active topic of research, not least in the wake of the 2008 financial crisis. While we can not hope to do justice here to all the relevant works, we will highlight some key developments which put our contributions in their historical perspective.

To guide our discussion, it is helpful to establish, following Hansen and Marinacci [2016], a taxonomy of levels of model uncertainty. Going back to Knight [1921], but also Keynes [1921], Arrow [1951], we speak of *risk* as the probabilistic (uncertain) nature of the future states of the world captured *within a model* and we speak of *uncertainty* where this nature is not captured by the proposed model. This could be because we picked a wrong model from a given ensemble: if the question is about which model from a given class to pick we speak of *model ambiguity*. Finally, if we are uncertain about which class of models to use or how to model future outcomes at all, we speak of *model misspecification*. The distinction between ambiguity and misspecification may appear artificial. The former is a special case of the latter

and, by taking a large infinite dimensional class of models, we can recast the latter as the former, see also [Cont et al., 2010, Remark 4.1]. Nevertheless, the distinction often serves as a useful taxonomy to guide a general discussion.

Without being too prescriptive, we think of model ambiguity when the class of models is restricted to a specific, often parametric, family. The main advantage of this approach is its tractability. Parameters are typically constants, or processes in a dynamic setting, taking values in some finite dimensional compact set. In a seminal contribution Merton [1969] considered the optimal investment problem of an agent trading in risky asset modelled using a geometric Brownian motion. Specifically, he considered an agent endowed with a power utility and wanting to maximise the expected utility of their consumption, a criterion justified by the axiomatic decision-theoretic works of Von Neumann and Morgenstern [1953], Savage [1951]. Merton [1969] derived the optimal portfolio value and trading strategy explicitly. In consequence, even if Merton worked with a fixed baseline model, his results included *implicitly* a sensitivity analysis with respect to model parameters. Since then, parameter uncertainty has been considered explicitly in a great number of papers, see for example Rogers [2001], Chen and Epstein [2002], Maenhout [2004], Kerkhof et al. [2010], Hernández-Hernández and Schied [2006], Biagini and Pinar [2017], Balter and Pelsser [2020] and the references therein. The classical expected utility maximisation problem, as considered by Merton [1969], is typically replaced by a maxmin formulation. This paradigm is also widely adopted in the robust control literature, see Sirbu et al. [2014], Bayraktar et al. [2016], and can be seen as a two-player game setting: the agent picks their best strategy to play against the nature who decides on adverse choice of model  $\mathbb{P}$  from a set  $\mathcal{P}$ . An axiomatic decision-theoretic justification of this criterion was provided by Gilboa and Schmeidler [1989], building on earlier contributions, including Anscombe et al. [1963], Ellsberg [1961] and Schmeidler [1989].

We note that in a dynamic multi-period setting, the agent can learn and reduce their model uncertainty. This led Bielecki et al. [2019] to develop an adaptive robust control approach and was exploited in a forward utility setting by Källblad et al. [2018]. It is also the cornerstone to the Bayesian model averaging and updating approach, see Hoeting et al. [1999] or Karatzas and Zhao [2001] for a continuous-time approach, where the drift is filtered out. In this approach one fixes a family of models  $\mathcal{M}$  whose corresponding parameters are denoted by  $\theta \in \Theta(M)$ , for  $M \in \mathcal{M}$ . The Bayesian observer has two levels of prior beliefs, namely a prior distribution on the models denoted by  $\mu_{\text{prior}}$  and a prior distribution of the model parameters given the model, denoted by  $\mathbb{P}(\theta | M)$ . The posterior probability distribution  $\mu_{\text{post}}$  on the space  $\mathcal{M}$  is then computed according to Bayes' rule

$$d\mu_{\text{post}}(M | y) = \frac{\int_{\Theta(M)} \mathbb{P}(y | \theta, M) d\mathbb{P}(\theta | M) \cdot d\mu_{\text{prior}}(M)}{\int_{\mathcal{M}} \int_{\Theta(N)} \mathbb{P}(y | \theta, N) d\mathbb{P}(\theta | N) d\mu_{\text{prior}}(N)}.$$

This gives rise to Bayesian updating of optimal investment problems. While appealing because of its theoretical simplicity, this modelling approach is often too sophisticated from a practical point of view, both to specify the priors and to compute the posteriors, see Cont [2006].

We classed the works discussed so far as dealing with model ambiguity. We think of model misspecification when possible perturbations of the baseline probability measure are not restricted to a parametric family, although they could be constrained,

e.g., via calibration to given market quoted prices of assets. This typically results in an infinite dimensional set of models to consider. To quote [Anderson et al., 2003, Section 10], “we envision this approximating model to be analytically tractable, yet to be regarded by the decision maker as not providing a correct model of the evolution of the state vector. The misspecifications we have in mind are small in a statistical sense but can otherwise be quite diverse.” We thus believe that the parametric ambiguity is rarely going to capture the true nature of model uncertainty facing a modeller. When we speak of *model uncertainty* we mean it in the broadest meaning of *model misspecification* and this is our prime interest here.

We note that Gilboa and Schmeidler [1989] did not provide specific insights into how the set of models  $\mathcal{P}$  in the maxmin criterion should be selected. In most of the works mentioned so far, its elements formed a parametric family of probability measures. We start our discussion of the literature dealing with model misspecification using the generic set of models which are absolutely continuous with respect to the baseline measure  $\mathbb{P}$ :  $\mathcal{P} = \{\tilde{\mathbb{P}} : \tilde{\mathbb{P}} \ll \mathbb{P}\}$ . It is natural to consider choosing attainable wealth  $\langle X, \pi \rangle$  to maximise a Lagrangian criterion

$$(5) \quad \inf_{\tilde{\mathbb{P}} \in \mathcal{P}} \left[ u(\langle X, \pi \rangle) + \lambda L \left( \frac{d\tilde{\mathbb{P}}}{d\mathbb{P}} \right) \right]$$

for some penalty function  $L : \mathbb{R} \rightarrow \mathbb{R}$ . In decision theory, this criterion is known as multiplier preferences, see Maccheroni et al. [2006], and includes the maxmin criterion (for a penalty function taking only the values of zero and infinity), see also Hansen and Sargent [2001], Uppal and Wang [2003], Lam [2016]. In continuous time models, such an approach can also be used to cover drift uncertainty for SDEs. Building on the earlier results of Kramkov and Schachermayer [1999] and the works on risk measures, Schied and Wu [2005] developed this robust optimal investment problem and subsequent works focused on more particular setups, see Hernández-Hernández and Schied [2006]. More recently, motivated by the above, Cohen [2017] considers the problem

$$\inf_{\tilde{\mathbb{P}} \in \mathcal{P}} \left( \mathbb{E}_{\tilde{\mathbb{P}}} [u(\langle X, \pi \rangle, \mathbf{x})] - \alpha_{\mathcal{P}|\mathbf{x}}(\tilde{\mathbb{P}}) \right)$$

for some observed data  $\mathbf{x}$ , a set of measures  $\mathcal{P}$  and a likelihood function  $\alpha_{\mathcal{P}|\mathbf{x}}$ , depending both on the choice of models  $\mathcal{P}$  and the data  $\mathbf{x}$ . This offers a novel approach to finding a robust estimate for  $\mathbb{E}_{\mathbb{P}}[u(\langle X, \pi \rangle)]$ , including both the likelihood of a particular model  $\tilde{\mathbb{P}}$  and the information coming from the data  $\mathbf{x}$ . Finally, note that in the particularly well studied case of  $L(x) = \log(x)$ , i.e., the relative entropy penalty, an essentially equivalent criterion is given by

$$\inf_{\tilde{\mathbb{P}} \in B_{\delta}^{\text{KL}}(\mathbb{P})} \mathbb{E}_{\tilde{\mathbb{P}}} [u(\langle X, \pi \rangle)], \quad \text{where } B_{\delta}^{\text{KL}}(\mathbb{P}) = \{\tilde{\mathbb{P}} \ll \mathbb{P} : H(\tilde{\mathbb{P}} | \mathbb{P}) \leq \delta\}$$

known as *constraint preferences*, see Hansen and Sargent [2001], Lam [2016] and section 4.1 below. While often tractable, approaches of this kind are restrictive since any measure in a ball in  $\mathcal{P}$  around  $\mathbb{P}$  has to be absolutely continuous with respect to  $\mathbb{P}$ .

Recognising the important limitations of the above approaches, a rich stream of literature in mathematical finance focuses on the so-called non-dominated setting for model uncertainty, i.e., considering  $\mathcal{P}$  whose elements are not all absolutely continuous with respect to one baseline model. This started with the so-called

uncertain volatility models of Avellaneda et al. [1995], Lyons [1995]. The main focus has since been on pricing and hedging questions from specific option types, e.g., Hobson [1998], Cox and Oblój [2011], Galichon et al. [2014], to more holistic duality results, e.g., Beiglböck et al. [2013], Bouchard and Nutz [2015], Hou and Oblój [2018]. It is impossible to do this research justice here and we refer the reader to the discussion in Burzoni et al. [2019]. Equally, the maxmin approach to the expected utility maximisation in a non-dominated setting has been considered in a number of works, see Denis and Kervarec [2013], Rásonyi and Stettner [2005], Nutz [2014], Neufeld and Nutz [2018], Carassus et al. [2019], usually focusing on existence and uniqueness of optimal investment strategies and dual representations.

Finally, we close this short discussion by noting that another stream of literature in which the *de facto used* model is a small perturbation of the baseline model is found among papers on expected utility maximisation under transaction costs, see Cvitanic and Karatzas [1996], Kallsen et al. [2010], Czichowsky et al. [2016]. Using a convex duality approach these authors derive a solution to the optimal investment problem under proportional transaction cost in continuous time by constructing a so-called shadow price process, i.e., a semi-martingale process taking values within the bid-ask spread region, whose solution to the optimal investment problem in the frictionless market exists and coincides with the solution of the original problem under transaction cost. In this sense the shadow price can be interpreted as a process living in a small neighbourhood of the original price process and the problem with transaction costs is included in a suitable maxmin formulation of the one without transaction costs.

#### 4. DISTRIBUTIONALLY ROBUST APPROACH - SUMMARY OF THE MAIN RESULTS

We propose to capture model misspecification by considering a ball, in the space of probability measures, around the baseline model  $\mathbb{P}$ . As noted above, this has most often been considered in the economics literature using the Kullback–Leibler divergence, i.e., the relative entropy, see Lam [2016] for general sensitivity results, Calafiore [2007] for applications in portfolio optimisation and Hansen and Sargent [2001] and the references therein for a broader context. KL-divergence has good analytic properties and often leads to closed-form solutions. However, it only allows one to consider measures which are absolutely continuous with respect to the baseline model  $\mathbb{P}$ . In particular, if the latter is the empirical measure supported on  $N$  observations, any meaningful analysis requires ad hoc structural assumptions and modifications, creating an additional layer of uncertainty.

In this paper we propose to use the Wasserstein distance. Coming from the optimal transport theory, this distance is known to lift the natural distance on the state space  $\mathcal{S}$  to the space of measures on  $\mathcal{S}$ . It is widely used in many application fields, from image processing, see for example Swoboda and Schnorr [2013], Tartavel et al. [2016] to statistical analysis of the Wasserstein barycenters, see for example Bigot and Klein [2018]. We refer to Peyré et al. [2019] and the references therein for a comprehensive overview of the field. While harder to handle analytically, it is more versatile and does not require any additional structural assumptions. It is well motivated from statistical theory and asymptotic consistency results, see Fournier and Guillin [2015], Esfahani and Kuhn [2018].



We first introduce some additional notation. We endow  $\mathbb{R}^d$  with the Euclidean norm  $|\cdot|$  and write  $\Gamma^\circ$  for the interior of a set  $\Gamma$ . We let  $\mathcal{P}(\mathcal{S})$  denote the set of all (Borel) probability measures on  $\mathcal{S}$  and, for  $p \in [1, \infty)$ ,

$$\mathcal{P}_p(\mathcal{S}) := \{\mathbb{P} \in \mathcal{P}(\mathcal{S}) : \mathbb{E}_{\mathbb{P}}[|X|^p] < \infty\}.$$

For  $\mathbb{P}, \tilde{\mathbb{P}} \in \mathcal{P}_p(\mathcal{S})$ , we define the  $p$ -Wasserstein distance via

$$W_p(\mathbb{P}, \tilde{\mathbb{P}}) = \inf \left\{ \mathbb{E}_{\gamma}[|X - Y|^p] : \gamma \in \text{Cpl}(\mathbb{P}, \tilde{\mathbb{P}}) \right\}^{1/p},$$

where  $\text{Cpl}(\mathbb{P}, \tilde{\mathbb{P}})$  is the set of all probability measures  $\gamma$  on  $\mathcal{S} \times \mathcal{S}$  with first marginal  $\gamma_1 := \gamma(\cdot \times \mathcal{S}) = \mathbb{P}$  and second marginal  $\gamma_2 := \gamma(\mathcal{S} \times \cdot) = \tilde{\mathbb{P}}$  and  $(X, Y)$  is the canonical process on  $\mathcal{S} \times \mathcal{S}$ . We also define the  $W_\infty$  distance for  $\mathbb{P}, \tilde{\mathbb{P}} \in \mathcal{P}(\mathcal{S})$  via

$$\begin{aligned} W_\infty(\mathbb{P}, \tilde{\mathbb{P}}) &:= \inf_{\gamma \in \text{Cpl}(\mathbb{P}, \tilde{\mathbb{P}})} \gamma\text{-ess-sup } |X - Y| \\ (6) \quad &= \inf \left\{ \varepsilon > 0 \mid \mathbb{P}(B) \leq \tilde{\mathbb{P}}(B^\varepsilon), \tilde{\mathbb{P}}(B) \leq \mathbb{P}(B^\varepsilon) \forall B \in \mathcal{B}(\mathbb{R}_+^d) \right\}. \end{aligned}$$

The Wasserstein ball of size  $\delta \geq 0$  around  $\mathbb{P}$  is denoted

$$B_\delta(\mathbb{P}) = \left\{ \tilde{\mathbb{P}} \in \mathcal{P}(\mathcal{S}) : W_p(\mathbb{P}, \tilde{\mathbb{P}}) \leq \delta \right\}.$$

Let us fix some  $p \in (1, \infty]$  and let  $q := p/(p-1)$  so that  $1/p + 1/q = 1$ , with the usual convention that  $q = 1$  if  $p = \infty$ .

**Remark 4.1.** *The choice of  $p$  is part of the agent's preferences, like the utility function  $u$ , and has important consequences. As the Wasserstein distances  $W_p$  are increasing in  $p$ , the choice of  $p$  influences the size of the Wasserstein ball  $B_\delta(\mathbb{P})$  considered. A more ambiguity-averse agent might choose a lower Wasserstein order than a less ambiguity-averse one. They would then obtain a ball  $B_\delta(\mathbb{P})$  which is larger and thus allows for greater perturbations to the baseline model. On the technical side, we will see in section 6 that the choice of  $p$  is linked to the growth rate of the utility function  $u$ .*

Throughout the rest of the paper we make the following assumption:

**Assumption 4.2.**  $\mathbb{P} \in \mathcal{P}_p(\mathcal{S})$  and the boundary of  $\mathcal{S} \subset \mathbb{R}^d$  has  $\mathbb{P}$ -zero measure.

We consider an agent with variational preferences with respect to  $B_\delta(\mathbb{P})$ , i.e., the agent considers the maxmin criterion

$$\sup_{\pi \in \mathcal{A}} \inf_{\tilde{\mathbb{P}} \in B_\delta(\mathbb{P})} \mathbb{E}_{\tilde{\mathbb{P}}}[u(\langle X, \pi \rangle)]$$

of Gilboa and Schmeidler [1989] and picks their best strategy against the nature who picks the worst model from within the ball  $B_\delta(\mathbb{P})$ . We are interested in understanding how this changes the classical problems considered in Section 2. Specifically, we want to quantify how the value and the actions of the agent change for small  $\delta$ . In this section, we only give an overview of the main results and highlight some of their consequences. The rigorous statements require assumptions which, even if not surprising or restrictive, are often cumbersome to state. Consequently, we defer these to later sections.

**4.1. Wasserstein variational preferences.** We start with a simple remark on variational preferences of Gilboa and Schmeidler [1989]. Consider an agent using a strategy  $\pi$  leading to wealth  $\langle X, \pi \rangle$ . The agent's utility function  $u$  and their choice of  $p$  and  $\delta$  imply a preference relation on the space of models  $\mathcal{P}_p(\mathcal{S})$ , i.e., distributions of  $X$ , via

$$(7) \quad \mathbb{P} \succeq_W \tilde{\mathbb{P}} \Leftrightarrow \min_{\tilde{\mathbb{P}} \in B_\delta(\mathbb{P})} \mathbb{E}_{\tilde{\mathbb{P}}} [u(\langle X, \pi \rangle)] \geq \min_{\tilde{\mathbb{P}} \in B_\delta(\tilde{\mathbb{P}})} \mathbb{E}_{\tilde{\mathbb{P}}} [u(\langle X, \pi \rangle)].$$

Assuming that  $u$  is differentiable and applying [Bartl et al., 2020, Theorem 2] we obtain that, up to  $o(\delta)$ ,  $\mathbb{P} \succeq_W \tilde{\mathbb{P}}$  if and only if

$$\mathbb{E}_{\mathbb{P}} [u(\langle X, \pi \rangle)] - \delta |\pi| (\mathbb{E}_{\mathbb{P}} [|u'(\langle X, \pi \rangle)|^q])^{1/q} \geq \mathbb{E}_{\tilde{\mathbb{P}}} [u(\langle X, \pi \rangle)] - \delta |\pi| (\mathbb{E}_{\tilde{\mathbb{P}}} [|u'(\langle X, \pi \rangle)|^q])^{1/q}.$$

In particular these preferences first take the expectation of  $u$  under  $\mathbb{P}$  and  $\tilde{\mathbb{P}}$  into account. If these align, then the norm of the first derivative of  $u$ , or the subjective pricing kernel (stochastic discount factor) is decisive for small  $\delta > 0$ . We believe that the fact that marginal quantities appear and control the assessment of small perturbations of the model seems both natural and desirable.

To the best of our knowledge, (7) is the first instance of using the Wasserstein balls to define preferences. Instead, in the economics literature, balls with respect to relative entropy have been suggested. Hansen and Sargent [2001] call these *constraint preferences* and observed that these are equivalent to using multiplier preferences with a penalty proportional to the relative entropy, see also Maccheroni et al. [2006]. Accordingly, if we consider

$$(8) \quad \mathbb{P} \succeq_{KL} \tilde{\mathbb{P}} \Leftrightarrow \min_{\tilde{\mathbb{P}}: H(\tilde{\mathbb{P}}|\mathbb{P}) \leq \delta^2} \mathbb{E}_{\tilde{\mathbb{P}}} [u(\langle X, \pi \rangle)] \geq \min_{\tilde{\mathbb{P}}: H(\tilde{\mathbb{P}}|\tilde{\mathbb{P}}) \leq \delta^2} \mathbb{E}_{\tilde{\mathbb{P}}} [u(\langle X, \pi \rangle)]$$

then using the results of Lam [2016], assuming finiteness of exponential moments, it follows that, up to  $o(\delta)$ ,  $\mathbb{P} \succeq_{KL} \tilde{\mathbb{P}}$  if and only if

$$\mathbb{E}_{\mathbb{P}} [u(\langle X, \pi \rangle)] - \delta \sqrt{2 \text{Var}_{\mathbb{P}}(u(\langle X, \pi \rangle))} \geq \mathbb{E}_{\tilde{\mathbb{P}}} [u(\langle X, \pi \rangle)] - \delta \sqrt{2 \text{Var}_{\tilde{\mathbb{P}}}(u(\langle X, \pi \rangle))},$$

where  $\text{Var}_{\mathbb{P}}(u(\xi))$  is the variance of  $u(\xi)$  under  $\mathbb{P}$ . In particular, in contrast to  $\succeq_W$ , marginal quantities do not appear but instead the second moment of the utility  $u(\xi)$  is decisive.

**4.2. Distributionally robust expected utility maximisation.** Consider a distributionally robust expected utility maximisation problem of the form

$$(9) \quad V(\delta) := \sup_{\pi \in \mathcal{A}} \inf_{\tilde{\mathbb{P}} \in B_\delta(\mathbb{P})} \mathbb{E}_{\tilde{\mathbb{P}}} [u(\langle X, \pi \rangle)].$$

This problem can be written in a number of equivalent ways, using different representations of the  $B_\delta(\mathbb{P})$ -ball. In particular, when  $\mathbb{P}$  is non-atomic, denseness of Monge couplings among all couplings, see Pratelli [2007], implies that

$$V(\delta) = \sup_{\pi \in \mathcal{A}} \inf_{f: \mathbb{E}_{\mathbb{P}} [|f(X)|^p] \leq \delta^p} \mathbb{E}_{\mathbb{P}} [u(\langle X + f(X), \pi \rangle)].$$

For  $\delta \geq 0$  we denote by  $\pi_\delta^* \in \mathcal{A}_\delta^*$  the optimizers for  $V(\delta)$ . In particular,  $V(0) = V$  and  $\pi_0^* = \pi^*$  as defined in Section 2.1. In Theorem 6.2 we show that

$$V'(0) = \lim_{\delta \rightarrow 0} \frac{V(\delta) - V(0)}{\delta} = -|\pi^*| (\mathbb{E}_{\mathbb{P}} [|u'(\langle X, \pi^* \rangle)|^q])^{1/q},$$

where  $q = p/(p-1)$ . In particular, the value function always decreases when model misspecification occurs. This is intuitively obvious: the agent considers a

max-min problem and hence is model-misspecification-averse. The loss in value is proportional to the number of shares held  $|\pi^*|$  and to the norm of the agent's pricing kernel. The first accounts for the agent's total exposure to the market and the second for the relative distance between the physical measure and the agent's subjective pricing measure. We note also that  $V'(0)$  is decreasing in  $q$ , i.e., is increasing in  $p$ . Again, this is intuitive: larger  $p$  corresponds to a stronger metric and hence smaller balls  $B_\delta(\mathbb{P})$ , i.e., less uncertainty about the baseline model. However,  $V'(0)$  does not appear to have any other simple monotonicity properties in relation to the baseline model. In particular, changing  $\mathbb{P}$  to increase the Sharpe ratio or the value  $V$  may lead both to increasing or decreasing  $V'(0)$ . Arguing purely formally, we expect that  $\pi^*$  increases with the Sharpe ratio and thus the monotonicity of  $V'(0)$  depends on the tradeoff between  $|\pi^*|$  and  $|u'(\langle X, \pi^* \rangle)|$ , which is usually decreasing in  $\pi^*$ . This tradeoff can lead to different behaviour for different utility functions and baseline models, see section 5 for examples.

Next, considering the optimal trading strategy, in Theorem 6.4 we show that

$$(10) \quad \pi^*(\delta) = \pi^* + (\pi^*)'(0)\delta + o(\delta),$$

where the gradient  $(\pi^*)'(0)$  is given by

$$(\pi^*)'(0) = (\nabla_\pi^2 V(0))^{-1} \cdot \frac{\pi^*}{|\pi^*|} \cdot \kappa_u, \quad \text{with}$$

$$\kappa_u = \|u'(\langle X, \pi^* \rangle)\|_{L^q(\mathbb{P})}^{1-q} \cdot \left( \mathbb{E}_\mathbb{P} \left[ \frac{\langle X, \pi^* \rangle u''(\langle X, \pi^* \rangle) + u'(\langle X, \pi^* \rangle)}{|u'(\langle X, \pi^* \rangle)|^{1-q}} \right] \right).$$

The first two terms above decide about the relative adjustments to the components in  $\pi^*$ , while  $\kappa_u$  is a constant multiplier. The first term is the inverse Hessian matrix, analogue to the inverse Fisher information matrix in a statistical problem, see Bartl et al. [2020]. It is multiplied by the relative weights in the portfolio, the second term.

**4.3. Distributionally robust marginal utility price.** We introduce now a robust version of the marginal utility price of Davis [1997]. We recall that  $g$ , see Assumption 2.1, denotes the payoff of the option we want to price. Define

$$(11) \quad V(\delta, \varepsilon, p_d) = \sup_{\pi \in \mathcal{A}} \inf_{\tilde{\mathbb{P}} \in B_\delta(\mathbb{P})} \mathbb{E}_{\tilde{\mathbb{P}}} \left[ u \left( -\varepsilon + \langle X, \pi \rangle + \frac{\varepsilon}{p_d} g(X) \right) \right].$$

**Definition 4.3.** Suppose that for each  $p_d > 0$  the function  $\varepsilon \mapsto V(\delta, \varepsilon, p_d)$  is differentiable. A number  $\hat{p}_d(\delta)$ , which satisfies

$$\nabla_\varepsilon V(\delta, 0, \hat{p}_d(\delta)) = 0.$$

is called a robust marginal utility price for the uncertainty level  $\delta$ .

Note that for  $\delta = 0$  this notion agrees with the marginal utility price of Davis [1997],  $\hat{p}_d(0) = \hat{p}_d$  from Definition 2.2, and can thus be considered as its natural distributionally robust counterpart. Furthermore, in Theorem 7.8 below we show that it is still computed via an expectation under a subjective martingale measure, only now this choice of measure also depends on the level of uncertainty  $\delta$ :

$$\hat{p}_d(\delta) = \frac{\mathbb{E}_{\mathbb{P}^*} [u'(\langle X, \pi_\delta^* \rangle) g(X)]}{\mathbb{E}_{\mathbb{P}^*} [u'(\langle X, \pi_\delta^* \rangle)]},$$

where  $\mathbb{P}^*$  is a minimising measure in  $B_{\mathbb{P}}(\delta)$  in (9), the nature's optimal response to the agent's best strategy  $\pi_{\delta}^*$ . Interestingly, the marginal price  $\hat{p}_d(\delta)$  is not monotone in  $\delta$ , in particular the price can both increase or decrease as model uncertainty is introduced. The behaviour of  $\hat{p}_d(\delta)$  is specific to the agent and the option's payoff. This again is intuitive: there is an interplay between an agent's trading intent and their valuation of the option, while uncertainty affects both. The only case when one would expect the marginal price to always decrease when uncertainty is introduced, is when the agent does not see trading as profitable to start with, i.e., when  $\pi^* = 0$ . This is confirmed by our results on the first order sensitivity in  $\delta$  given in Theorem 7.9. We find that if  $\pi^* = 0$ , then

$$\hat{p}'_d(0) = -(\mathbb{E}_{\mathbb{P}}[|\nabla g(x)|^q])^{1/q}.$$

If  $\pi^* \neq 0$  this sensitivity is more involved and, in particular, can be both positive or negative. Remarkably, we can still compute it in a closed form:

$$\hat{p}'_d(0) = \mathbb{E}_{\mathbb{Q}_u}[R_u(\langle X, \pi^* \rangle) (\langle T(X), \pi^* \rangle - \langle X, (\pi^*)'(0) \rangle) \cdot (g(X) - \hat{p}_d) - \langle \nabla g(X), T(X) \rangle],$$

where  $\mathbb{Q}_u$  was given in (2),  $R_u(x) = -\frac{u''(x)}{u'(x)}$  is the agent's absolute risk aversion coefficient and

$$T(x) := \frac{\pi^*}{|\pi^*|} |u'(\langle x, \pi^* \rangle)|^{q-1} (\mathbb{E}_{\mathbb{P}}[|u'(\langle X, \pi^* \rangle)|^q])^{1/q-1}.$$

## 5. EXAMPLES

In this section we consider a number of simple examples to illustrate the notions and results presented above. We selected the examples so that most of the computations can be derived in two ways: either through a direct brute force computation, see Oblój and Wiesel [2021], or by using results presented in section 4 and stated in more detail in sections 6-7 below.

Specifically, throughout this section we take  $d = 1$  and mainly focus on  $\mathbb{P}$  which is either binomial or Gaussian. The former is the only complete model in this one-period setting so that, in particular, the martingale measure  $\mathbb{Q}_u$  is unique and independent of the utility function  $u$ . Introducing model uncertainty however, we lose market completeness, and sensitivity of Davis' price is subjective. In all of the figures, we plot sensitivities as functions of the Sharpe ratio  $\mu/\sigma$  of the baseline model  $\mathbb{P}$ , where  $\mu = \mathbb{E}_{\mathbb{P}}[X]$  and  $\sigma^2 = \text{Var}_{\mathbb{P}}(X)$ .

**5.1. Binomial model.** Fix  $a \in (0, 1/2)$ . Consider the baseline model  $\mathbb{P} = a\delta_{-1} + (1-a)\delta_1$  and a log investor with initial capital equal to one, so that  $u(x) = \log(1+x)$ . For concreteness we take  $\mathcal{D} = (-1, \infty)$ ,  $\mathcal{A} = (-1+a, 1-a)$ ,  $\mathcal{S} = [-1-a, 1+a]$  and note that for  $\varepsilon_0 = a^2/2$  we have

$$\{\langle x, \pi \rangle : x \in \mathcal{S}, \pi \in \mathcal{A}\}^{\varepsilon_0} \subset (-1+a^2, 1-a^2)^{\varepsilon_0} = [-1+a^2/2, 1-a^2/2] \subseteq \mathcal{D}$$

and that any continuous function  $g$  is bounded on  $\mathcal{S}$ . The unique optimiser  $\pi^*$  is then given by  $\pi^* = 1 - 2a$  and  $V = a \log(2a) + (1-a) \log(2-2a)$ .

We now consider the robust optimal investment problem  $V(\delta)$ . For the case  $p = \infty$  we explicitly calculate

$$\pi_{\delta}^* = a \log\left(\frac{2a}{1-\delta}\right) + (1-a) \log\left(\frac{-2+2a}{-1-\delta}\right),$$

which in turn yields

$$V'(0) = -\pi^*.$$

The same result follows directly from Theorem 6.2, which also covers the case of finite  $p$ . We can compare this to Lam [2016], where uncertainty is quantified by balls in KL-divergence. Indeed, an application of [Lam, 2016, Theorem 3.1] yields

$$\begin{aligned} \tilde{V}'(0) &= -\sqrt{2\text{Var}(\log(1 + \pi^* X))} \\ &= -\left(2\left(a \log(2a)^2 + (1-a) \log(2(1-a))^2\right.\right. \\ &\quad \left.\left.- \left(a \log(2a) + (1-a) \log(2(1-a))\right)^2\right)^{1/2}, \end{aligned}$$

for

$$\tilde{V}(\delta) := \sup_{\pi \in \mathcal{A}} \inf_{H(\tilde{\mathbb{P}}|\mathbb{P}) \leq \delta^2} \mathbb{E}_{\tilde{\mathbb{P}}} [u(\langle X, \pi \rangle)],$$

see Figure 1 for a comparison.

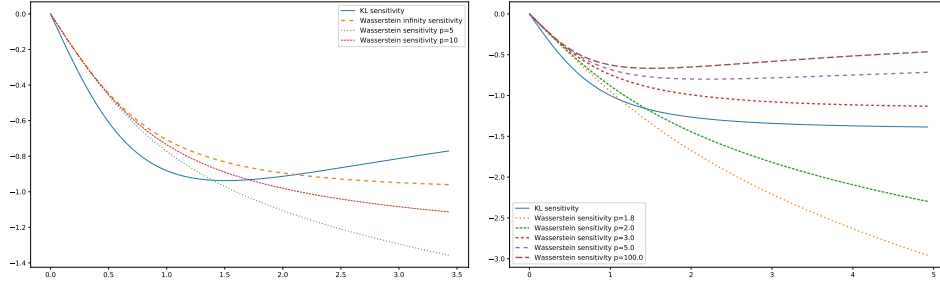


FIGURE 1. Wasserstein-sensitivity  $V'(0)$  as a function of  $\mu/\sigma$  for different  $p$  and KL-sensitivity  $\tilde{V}'(0)$  for  $u(x) = \log(1+x)$  (left) and  $u(x) = -\exp(-x)$  (right).

This plot raises the question if  $V'(0)$  is generally a decreasing function of the Sharpe ratio  $\mu/\sigma$  for all  $p$  in the binomial model. This turns out not to be the case. Indeed if we choose the exponential utility function  $u(x) = -\exp(-\gamma x)$  for  $\gamma > 0$ , then

$$\pi^* = \frac{\log\left(\frac{a}{1-a}\right)}{-2\gamma}$$

and in particular

$$V'(0) = -\left(a \left(\frac{a}{1-a}\right)^{-q/2} + (1-a) \left(\frac{a}{1-a}\right)^{q/2}\right)^{1/q} \cdot \pi^*.$$

Clearly,  $V'(0)|_{a=\frac{1}{2}} = 0$ , as expected since  $\pi^*|_{a=\frac{1}{2}} = 0$ . Consider now the asymptotics as  $a \rightarrow 0$ . Clearly  $\pi^* \rightarrow \infty$ . Considering the leading behaviour for the first term (in parentheses), we see that for  $q \geq 2$  it diverges so that  $V'(0) \rightarrow -\infty$ , but for  $1 < q < 2$  it converges to zero and dominates so that  $V'(0) \rightarrow 0$ . In particular, we see that the monotonicity of  $V'(0)$  depends on  $p$  and, for  $p > 2$ ,  $V'(0)$  is in fact

increasing for large Sharpe ratio  $\mu/\sigma$ , see Figure 1. Finally, for comparison, we note that [Lam, 2016, Theorem 3.1] yields

$$\tilde{V}'(0) = -\sqrt{2(1 - 4a(1 - a))}.$$

We continue our investigation with the sensitivity of the optimiser  $\pi^*(\delta)$ : a direct calculation yields  $(\pi^*)'(0) = -1$ . Alternatively we can use Theorem 6.4 to obtain

$$(\pi^*)'(0) = -a(1 - a) (a^{-q+1} + (1 - a)^{-q+1})^{1/q-1} \cdot (a^{-q} + (1 - a)^{-q}),$$

so that the results coincide for  $q = 1$ . See Figure 2 for a plot of  $(\pi^*)'(0)$  for different values of  $p$ .

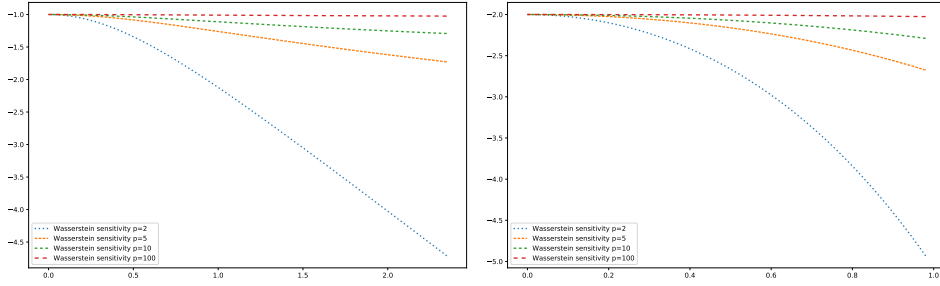


FIGURE 2. Wasserstein-sensitivity  $(\pi^*)'(0)$  of  $\pi_\delta^*$  (left) and Wasserstein-sensitivity  $\hat{p}_d'(0)$  of the Davis price  $\hat{p}_d$  (right) for different  $p$ . Both are plotted as functions of  $\mu/\sigma$ .

Let us now compare the Davis price of the baseline model with its robust counterpart. According to (4) it is given by  $\hat{p}_d = \mathbb{E}_{\mathbb{Q}_u}[g(X)]$ . We consider first  $g(x) = x^3$  as a concrete example. An easy computation shows  $\hat{p}_d = 0$ , while we can explicitly calculate the robust Davis price  $\hat{p}_d(\delta)$  as

$$\hat{p}_d(\delta) = -2\delta + 2\delta^3.$$

In particular,  $\hat{p}_d'(0) = -2$  as can also be derived from Theorem 7.9. For a general  $p \in (1, \infty]$  we obtain

$$\begin{aligned} (\hat{p}_d)'(0) &= -\frac{1}{4a} \left( (2a)^{1-q}(1-2a) \cdot 2^{q-1} (a^{-q+1} + (1-a)^{-q+1})^{1/q-1} - 1 \right) \\ &\quad + \frac{1}{4(1-a)} \left( (2(1-a))^{1-q}(1-2a) \cdot 2^{q-1} (a^{-q+1} + (1-a)^{-q+1})^{1/q-1} + 1 \right) \\ &\quad - \frac{3}{2} (a^{1-q} + (1-a)^{1-q}) \cdot (a^{1-q} + (1-a)^{1-q})^{1/q-1}, \end{aligned}$$

see Figure 2. Next consider  $g(x) = x^+$  and  $p = \infty$ . Note that  $g$  is not differentiable, but as  $\mathbb{P}$  is only supported in  $\{-1, 1\}$  we can take a suitable  $C^\infty$ -approximation instead. We compute  $\hat{p}_d = 0.5$  and

$$\hat{p}_d(\delta) = \frac{1 - \delta^2}{2}.$$

In particular  $\hat{p}_d'(0) = 0$ , which is also readily seen using Theorem 7.9.

As we remarked before, it is however not always true that  $\hat{p}_d(\delta) \leq \hat{p}_d$ . Take, e.g.,  $g(x) = |x + x_0|$  for some  $x_0 \in (0, 1)$ . For  $p = \infty$ , as  $\hat{p}_d(\delta)$  is the expectation under the unique martingale measure concentrated on  $-1 - \delta$  and  $1 - \delta$ , we conclude that

$$\hat{p}_d(\delta) = 1 - \delta^2 + \delta x_0 > 1 = \hat{p}_d$$

for small enough  $\delta > 0$ .

**5.2. Normal model.** We now set  $\mathbb{P} = \mathcal{N}(\mu, \sigma^2)$  and consider the utility function  $u(x) = -e^{-\gamma x}$  for  $\gamma > 0$  and  $\mathcal{D} = \mathbb{R}$ . We obtain

$$\pi^* = \frac{\mu}{\gamma \sigma^2} \quad \text{and} \quad V = -\exp\left(-\frac{\mu^2}{2\sigma^2}\right).$$

For  $p = \infty$  we find

$$V(\delta) = -\exp\left(-\frac{(\mu - \delta)^2}{2\sigma^2}\right).$$

A direct calculation thus gives

$$V'(0) = \exp\left(-\frac{\mu^2}{2\sigma^2}\right) \frac{\mu}{\sigma^2},$$

which can be recovered by Theorem 6.2. Figure 3 compares the sensitivity  $V'(0)$  as a function of  $\mu/\sigma$  for different  $\mu$ . We also remark that in this case we can not compare with model uncertainty in the sense of relative entropy balls as this problem is degenerate. In fact

$$\mathbb{E}_{\mathbb{P}} [\exp(\theta(u(\langle X, \pi^* \rangle)))] = \mathbb{E}_{\mathbb{P}} \left[ \exp\left(\theta \exp\left(\frac{\mu}{\sigma^2} X\right)\right) \right] = \infty$$

for  $\theta, \mu \neq 0$ , so that [Lam, 2016, Assumption 3.1] is not satisfied. Next we find that

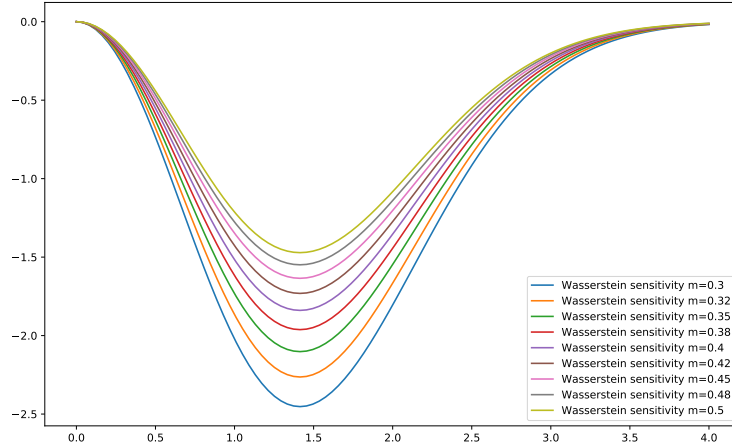


FIGURE 3. Wasserstein- $\infty$  sensitivity  $V'(0)$  as a function of  $\mu/\sigma$  for different  $\mu$ .

the distributionally robust optimiser is

$$\pi_\delta^* = \frac{\mu - \delta}{\gamma \sigma^2}, \quad \text{and thus } (\pi^*)'(0) = -\frac{1}{\gamma \sigma^2},$$

which can again be recovered by Theorem 6.4.

We can also compute the Davis price explicitly. E.g., for the case  $g(x) = x^2$  we obtain  $\hat{p}_d = \sigma^2$ . Interestingly enough this result remains unchanged for any positive  $\delta > 0$ , i.e.,  $\hat{p}_d(\delta) = \sigma^2$ . More generally one can show that

$$\hat{p}_d(\delta) = \mathbb{E}_{\tilde{\mathbb{P}}}[g(X)]$$

for any  $\delta > 0$ , where  $\tilde{\mathbb{P}} = \mathcal{N}(0, \sigma^2)$ . This implies in particular  $\hat{p}'_d(0) = 0$  for any payoff  $g$ , a result which can again be recovered from Theorem 7.9, albeit through a tedious calculation, see [Oblój and Wiesel, 2021, Section 1].

**5.3. Discussion of normal model in the case  $p \in (1, \infty)$ .** Note that in Section 5.2 the exponential function  $u(x) = -e^{-\gamma x}$  does not satisfy Assumption 6.1.(i) for any  $p \in (1, \infty)$ , so Theorem 6.2 is not applicable. Indeed it is not hard to see that  $V'(0) = -\infty$  in this case: this follows by noting that for any  $\pi \neq 0$

$$W_p(\delta_{-\text{sign}(\pi) \cdot n}, \mathbb{P}) \leq n + (\mathbb{E}_{\mathbb{P}}[|X|^p])^{1/p} =: f(n),$$

where  $f(n) = O(n)$ . In particular for all  $n \in \mathbb{N}$  large enough, the probability measures

$$\tilde{\mathbb{P}}^n = \left(\frac{\delta}{f(n)}\right)^p \delta_{-\text{sign}(\pi) \cdot n} + \left(1 - \left(\frac{\delta}{f(n)}\right)^p\right) \mathbb{P}$$

satisfy

$$W_p(\tilde{\mathbb{P}}^n, \mathbb{P}) \leq \frac{\delta}{f(n)} W_p(\delta_{-\text{sign}(\pi) \cdot n}, \mathbb{P}) \leq \delta$$

so that  $\tilde{\mathbb{P}}^n \in B_\delta(\mathbb{P})$ . Thus

$$\begin{aligned} \inf_{\tilde{\mathbb{P}} \in B_\delta(\mathbb{P})} \mathbb{E}_{\tilde{\mathbb{P}}}[u(\langle X, \pi \rangle)] &\leq \mathbb{E}_{\tilde{\mathbb{P}}^n}[u(\langle X, \pi \rangle)] \\ &= -\left(\frac{\delta}{f(n)}\right)^p \exp(\gamma|\pi|n) - \left(1 - \left(\frac{\delta}{f(n)}\right)^p\right) \mathbb{E}_{\mathbb{P}}[\exp(-\gamma\pi X)], \end{aligned}$$

which goes to  $-\infty$  with  $n \rightarrow \infty$ , using  $f(n) = O(n)$ . In particular

$$\begin{aligned} V(\delta) &= \sup_{\pi \in \mathcal{A}} \inf_{\tilde{\mathbb{P}} \in B_\delta(\mathbb{P})} \mathbb{E}_{\tilde{\mathbb{P}}}[u(\langle X, \pi \rangle)] = \sup_{\pi \in \mathcal{A}} \inf_{\tilde{\mathbb{P}} \in B_\delta(\mathbb{P})} \mathbb{E}_{\tilde{\mathbb{P}}}[-\exp(-\gamma\langle X, \pi \rangle)] \\ &= \mathbb{E}_{\tilde{\mathbb{P}}}[-\exp(-\gamma\langle X, 0 \rangle)] = -1 \end{aligned}$$

and so

$$V'(0) = \lim_{\delta \rightarrow 0} \frac{V(\delta) - V(0)}{\delta} = \lim_{\delta \rightarrow 0} \frac{-1 + \exp\left(-\frac{\mu^2}{2\sigma^2}\right)}{\delta} = -\infty$$

for  $\mu \neq 0$ . As we have seen above, Wasserstein- $p$ -balls do not allow for enough control over the tails of the distribution when considering a utility function decreasing exponentially. There are two ways to remedy this:

- (i) Consider an approximating sequence of utility functions  $(u_\kappa)_{\kappa \in (0,1)}$  satisfying Assumption 6.1, such that  $u_\kappa \rightarrow u$  for  $\kappa \rightarrow 0$ .
- (ii) Use a different Wasserstein distance adapted to the utility function under investigation.

We will briefly comment on both approaches. For (i) we can formally consider

$$u_\kappa(x) := \mathbb{1}_{\{x < -1/\kappa\}} \left[ -e^{\gamma/\kappa} + \gamma e^{\gamma/\kappa} (x + 1/\kappa) \right] + \mathbb{1}_{\{x \geq -1/\kappa\}} u(x)$$



for  $\kappa > 0$  and note that  $u_\kappa(x) \geq u(x)$  for all  $x \in \mathcal{S}$ , so that

$$u(x) = \inf_{\kappa > 0} u_\kappa(x) = \lim_{\kappa \rightarrow 0} u_\kappa(x).$$

In particular  $(u_\kappa)'(x) \leq \gamma e^{\gamma/\kappa}$  and  $(u_{\tilde{\kappa}})'(x) \geq (u_\kappa)'(x)$  for all  $\tilde{\kappa} \leq \kappa$  and all  $x \in \mathcal{S}$ . Thus defining

$$V_\kappa(\delta) = \sup_{\pi \in \mathcal{A}} \inf_{\tilde{\mathbb{P}} \in B_\delta(\mathbb{P})} \mathbb{E}_{\tilde{\mathbb{P}}} [u_\kappa(\langle X, \pi \rangle)]$$

we can apply the monotone convergence theorem to obtain

$$\begin{aligned} \lim_{\kappa \rightarrow 0} V'_\kappa(0) &= \lim_{\kappa \rightarrow 0} -(\mathbb{E}_{\mathbb{P}} [|(u_\kappa)'(\langle X, \pi^* \rangle)|^q])^{1/q} |\pi^*| \\ &= -\gamma (\mathbb{E}_{\mathbb{P}} [\exp(-\gamma \pi^* X)^q])^{1/q} |\pi^*| \\ &= -\gamma \left( \exp \left( -\mu q \gamma \pi^* + \frac{\sigma^2 (q \gamma \pi^*)^2}{2} \right) \right)^{1/q} |\pi^*| \\ &= -\gamma \exp \left( -\mu \gamma \pi^* + \frac{\sigma^2 q (\gamma \pi^*)^2}{2} \right) \left( \frac{\mu}{\gamma \sigma^2} \right) \\ &= -\exp \left( -\frac{\mu^2}{\sigma^2} + \frac{q \mu^2}{2 \sigma^2} \right) \left( \frac{\mu}{\sigma^2} \right). \end{aligned}$$

This formula aligns with section 5.2 for the case  $q = 1$ . Similarly, we can also calculate

$$(\pi^*)'(0) = -\frac{\gamma^{1/q+q-3}}{\sigma^2} \exp \left( \frac{q-1}{2} \right) \left( -\frac{\mu^2(1-q)}{\sigma^2} + 1 \right),$$

which again matches the result for  $q = 1$  in section 5.2, see Figure 4.

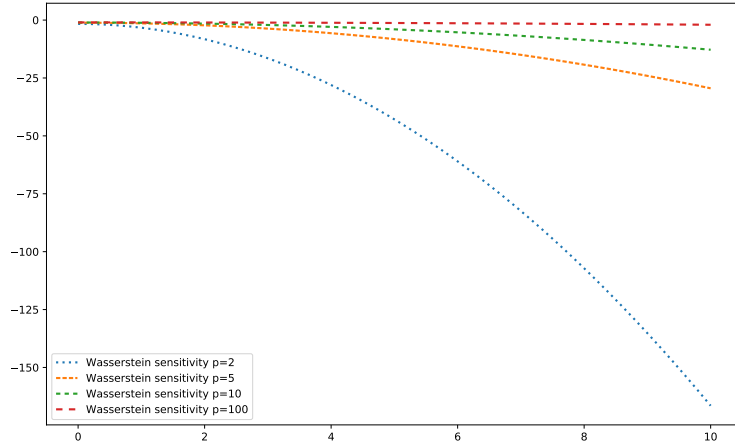


FIGURE 4. Wasserstein-sensitivity  $(\pi^*)'(0)$  of  $\pi^*_\delta$  as a function of  $\mu/\sigma$  for  $\sigma = \gamma = 1$ .

Turning to (ii), assuming  $u(0) = 0$  we could work on generalised Orlicz hearts

instead and define the Wasserstein-Luxemburg metric

$$\begin{aligned} W_u(\mathbb{P}, \tilde{\mathbb{P}}) &:= \inf \left\{ \int u(|x - y|) \pi(dx, dy) : \pi \in \text{Cpl}(\mathbb{P}, \tilde{\mathbb{P}}) \right\} \\ &= \inf \left\{ b > 0 : \int \frac{u(|x - y|)}{b} \pi(dx, dy) \leq 1, \text{ where } \pi \in \text{Cpl}(\mathbb{P}, \tilde{\mathbb{P}}) \right\}. \end{aligned}$$

As  $\|x\| := u(|x|)$  is a norm,  $W_u$  is still metric. A more general version of this definition was investigated in Sturm [2011]. We further remark that the above definition can be connected to results of Frittelli and Gianin [2002], Cheridito and Li [2009]. From [Bartl et al., 2020, Theorem 2] we know that

$$V'(0) = -\|u'(\langle X, \pi^* \rangle)\|_{u, \mathbb{P}}^* |\pi^*|,$$

where  $\|f\|_{u, \mathbb{P}}^*$  denotes a dual norm of  $f$ , which can be construed as a conjugate Orlicz norm of  $f$  with respect to the probability measure  $\mathbb{P}$ . Alas, we were not able to compute either of these norms explicitly, even for the simple benchmark cases described above.

**5.4. Lognormal model.** Let us lastly consider the shifted lognormal distribution  $\mathbb{P} = (1 + x)_{\#} \exp(\mathcal{N}(\mu, \sigma^2))$  with  $\mu < -\sigma^2/2$  as well as  $\mathcal{A} = [0, 1]$ ,  $\mathcal{D} = (-1, \infty)$  and  $u(x) = \log(x + 1)$ . Here  $(1 + x)_{\#} \exp(\mathcal{N}(\mu, \sigma^2))$  denotes the push forward of the distribution  $\exp(\mathcal{N}(\mu, \sigma^2))$  through the function  $x \mapsto (1 + x)$ . In this case  $\pi^* = 0$  and for the butterfly payoff  $g(x) = (x + K)^+ - 2x^+ + (x - K)^+$  with strike  $K > 0$  we compute the distributionally robust Davis price directly as

$$\hat{p}_d(\delta) = \text{BS}(1, 1, -K + \delta, \mu, \sigma) - 2\text{BS}(1, 1, \delta, \mu, \sigma) + \text{BS}(1, 1, K + \delta, \mu, \sigma),$$

where  $\delta \geq 0$  and  $\text{BS}(T - t, S_0, K, \mu, \sigma)$  is the Black-Scholes call price. In consequence,  $\hat{p}'_d(0)$  corresponds to the partial derivative of  $\text{BS}(1, 1, -K, \mu, \sigma) - 2\text{BS}(1, 1, 0, \mu, \sigma) + \text{BS}(1, 1, K, \mu, \sigma)$  w.r.t.  $K$ .

## 6. DISTRIBUTIONALLY ROBUST EXPECTED UTILITY MAXIMISATION

We return now to the discussion of our main results and present rigorous statements of the results discussed in section 4.2 above.

**6.1. The value function and its sensitivity analysis.** In order to quantify the first-order sensitivity of (9), we start our analysis by calculating the sensitivity of the value function  $V(\delta)$  using the general results obtained in Bartl et al. [2020]. Remark that  $\mathcal{A}_0^* = \{\pi^*\}$  follows simply from Assumption 2.1. We make the following further assumption, where we recall that  $\mathcal{A}^o$  denotes the interior of  $\mathcal{A}$ :

**Assumption 6.1.** *The following hold:*

- (i) (a) *If  $p < \infty$ , then for every  $r > 0$  there exists  $c > 0$  such that*

$$u'(\langle x, \pi \rangle) \leq c(1 + |x|^{p-1})$$

*for all  $x \in \mathcal{S}$  and  $\pi \in \mathcal{A}$  with  $|\pi| \leq r$ .*

- (b) *If  $p = \infty$ , then there exists  $\tilde{\delta} > 0$  such that for each  $r > 0$*

$$\mathbb{E}_{\mathbb{P}} \left[ \sup_{\pi \in \mathcal{A}, |\pi| \leq r} u'(\langle X, \pi \rangle - r\tilde{\delta}) \right] < \infty.$$

- (ii) *The optimiser  $\pi^* \in \mathcal{A}_0^*$  satisfies  $\pi^* \in \mathcal{A}^o$ .*

In particular, we see that agent's ambiguity aversion, as captured by the choice of  $p$ , and their risk aversion, as captured by  $u$ , have to be compatible, see also Remark 4.1. We obtain the following result:

**Theorem 6.2.** *Suppose the utility function  $u$  and the baseline model  $\mathbb{P}$  satisfy Assumption 6.1. Then*

$$V'(0) = \lim_{\delta \rightarrow 0} \frac{V(\delta) - V(0)}{\delta} = -(\mathbb{E}_{\mathbb{P}} [|u'(\langle X, \pi^* \rangle)|^q])^{1/q} |\pi^*|.$$

*Proof of Theorem 6.2.* By Assumptions 2.1 and 6.1, we have that the optimizer  $\pi^*$  is unique,  $\mathcal{A}_0^* = \{\pi^*\}$ , and belongs to the interior of  $\mathcal{A}$ ,  $\pi^* \in \mathcal{A}^o$ .

*Step 1:* In the case  $p \in (1, \infty)$  we can simply apply [Bartl et al., 2020, Theorem 2] for the function  $f(x, \pi) = u(\langle x, \pi \rangle)$  to obtain

$$V'(0) = -(\mathbb{E}_{\mathbb{P}} [|\nabla_x f(X, \pi^*)|^q])^{1/q} = -(\mathbb{E}_{\mathbb{P}} [|u'(\langle X, \pi^* \rangle)|^q])^{1/q} |\pi^*|.$$

*Step 2:* If  $p = \infty$ , then following [Bartl et al., 2020, proof of Theorem 2] line by line and replacing [Bartl et al., 2020, equation (7)] by

$$\begin{aligned} \sup_{\tilde{\mathbb{P}} \in B_\delta(\mathbb{P})} \mathbb{E}_{\tilde{\mathbb{P}}} [u'(\langle X, \pi^* \rangle)] &\leq \mathbb{E}_{\mathbb{P}} \left[ \sup_{|a| \leq \delta} u'(\langle X + a, \pi^* \rangle) \right] \\ &= \mathbb{E}_{\mathbb{P}} [u'(\langle X, \pi^* \rangle - |\pi^*| \delta)] \leq \mathbb{E}_{\mathbb{P}} [u'(\langle X, \pi^* \rangle - r \delta)] < \infty \end{aligned}$$

for  $|\pi^*| \leq r$  we realise that

$$\lim_{\delta \rightarrow 0} \mathbb{E}_{\pi^*} [u'(\langle X + t(Y - X), \pi^* \rangle)] = \mathbb{E}_{\mathbb{P}} [u'(\langle X, \pi^* \rangle)]$$

still holds by the dominated convergence theorem. This concludes the proof of the “ $\leq$ ” inequality. For the “ $\geq$ ” inequality we can again follow the steps of [Bartl et al., 2020, proof of Theorem 2]. In particular we note that  $T(x) = T = \text{sign}(\pi^*)$  so that  $\|T\|_\infty \leq 1$  and again by the dominated convergence theorem we conclude that

$$\begin{aligned} \lim_{\delta \rightarrow 0} \mathbb{E}_{\mathbb{P}} [u'(\langle X - t\delta T(X), \pi_\delta^* \rangle) \langle \pi_\delta^*, T(X) \rangle] &= \mathbb{E}_{\mathbb{P}} [u'(\langle X, \pi^* \rangle) \cdot \langle \pi^*, T \rangle] \\ &= \mathbb{E}_{\mathbb{P}} [u'(\langle X, \pi^* \rangle) \cdot |\pi^*|], \end{aligned}$$

which concludes the proof.  $\square$

**6.2. The optimising strategy and its sensitivity analysis.** Next, we calculate the sensitivity of optimisers  $\pi^*(\delta) \in \mathcal{A}_\delta^*$ . In order to carry this out, we impose the following additional assumptions:

**Assumption 6.3.** *The following hold:*

- (i) *The function  $u : \mathbb{R} \rightarrow \mathbb{R}$  is twice continuously differentiable and*
- (a) *if  $p < \infty$ ,*

$$|x| |u''(\langle x, \pi \rangle)| + |u'(\langle x, \pi \rangle)| \leq c(1 + |x|^{p-1-\varepsilon})$$

*for some  $\varepsilon > 0$ ,  $c > 0$  and all  $\pi \in \mathcal{A}$  close to  $\pi^*$  and all  $x \in \mathcal{S}$ .*

- (b) *if  $p = \infty$ , then there exists  $\tilde{\delta} > 0$  such that*

$$\mathbb{E}_{\mathbb{P}} \left[ \sup_{|z| \leq \tilde{\delta}, |\pi| \leq r} |x + z| |u''(\langle x + z, \pi \rangle)| + |u'(\langle x + z, \pi \rangle)| \right] < \infty.$$

*for all  $\pi \in \mathcal{A}$  close to  $\pi^*$  and all  $x \in \mathcal{S}$ .*

(ii) *The matrix*

$$\nabla_{\pi}^2 V(0, \pi^*) = \mathbb{E}_{\mathbb{P}} [XX^{\top} u''(\langle X, \pi^* \rangle)]$$

*is negative definite.*

With these additional assumptions at hand, we can state the main result of this section as follows:

**Theorem 6.4.** *Suppose the utility function  $u$  and the baseline model  $\mathbb{P}$  satisfy Assumptions 6.1 and 6.3, and that  $\pi^* \neq 0$ . Take  $\pi^*(\delta) \in \mathcal{A}_{\delta}^*$ . Then*

$$\pi^*(\delta) = \pi^* + (\pi^*)'(0)\delta + o(\delta),$$

where

$$(\pi^*)'(0) = \|u'(\langle X, \pi^* \rangle)\|_{L^q(\mathbb{P})}^{1-q} \cdot (\nabla_{\pi}^2 V(0))^{-1} \cdot \frac{\pi^*}{|\pi^*|} \cdot \left( \mathbb{E}_{\mathbb{P}} \left[ \frac{\langle X, \pi^* \rangle u''(\langle X, \pi^* \rangle) + u'(\langle X, \pi^* \rangle)}{|u'(\langle X, \pi^* \rangle)|^{1-q}} \right] \right).$$

*Proof of Theorem 6.4. Step 1:* Let us first consider the case  $p \in (1, \infty)$ . We note that uniqueness of  $\pi^*$  and [Bartl et al., 2020, Lemma 19] implies that  $\pi^*(\delta) \rightarrow \pi^*$  as  $\delta \rightarrow 0$ . Furthermore Assumptions 6.1 and 6.3 are sufficient to apply [Bartl et al., 2020, Theorem 4]. We thus conclude

$$\pi^*(\delta) = \pi^* + (\pi^*)'(0)\delta + o(\delta)$$

where

$$(\pi^*)'(0) = (\mathbb{E}_{\mathbb{P}} [|\nabla_x f(X, \pi^*)|^q])^{\frac{1}{q}-1} (\nabla_{\pi}^2 V(0, \pi^*))^{-1} \cdot \mathbb{E}_{\mathbb{P}} \left[ \frac{\nabla_x \nabla_{\pi} f(X, \pi^*) \nabla_x f(X, \pi^*)}{|\nabla_x f(X, \pi^*)|^{2-q}} \right]$$

for  $f(x, \pi) = u(\langle x, \pi \rangle)$ . Writing  $u' := u'(\langle x, \pi^* \rangle)$  and  $u'' := u''(\langle x, \pi^* \rangle)$  in order to simplify notation, an explicit computation yields

$$\begin{aligned} \nabla_x f(x, \pi^*) &= \pi^* u', \\ \nabla_{\pi} \nabla_x f(x, \pi^*) &= \pi^* x^{\top} u'' + I u', \\ \nabla_{\pi} \nabla_x f(x, \pi^*) \nabla_x f(x, \pi^*) &= (\langle x, \pi^* \rangle u' u'' + (u')^2) \cdot \pi^*, \\ \nabla_{\pi}^2 f(x, \pi^*) &= x x^{\top} u'', \\ \nabla_{\pi}^2 V(0) &= \mathbb{E}_{\mathbb{P}} [X X^{\top} u''] \end{aligned}$$

where  $I \in \mathbb{R}^{d \times d}$  denotes the identity matrix and by Assumption 6.3 the matrix

$$\mathbb{E}_{\mathbb{P}} [X X^{\top} u'']$$

is negative definite. In particular we have

$$\begin{aligned} (\pi^*)'(0) &= (\mathbb{E}_{\mathbb{P}} [|\nabla_x f(X, \pi^*)|^q])^{\frac{1}{q}-1} (\nabla_{\pi}^2 V(0, \pi^*))^{-1} \cdot \mathbb{E}_{\mathbb{P}} \left[ \frac{\nabla_x \nabla_{\pi} f(X, \pi^*) \nabla_x f(X, \pi^*)}{|\nabla_x f(X, \pi^*)|^{2-q}} \right] \\ &= (\mathbb{E}_{\mathbb{P}} [|u'|^q])^{\frac{1}{q}-1} |\pi^*|^{1-q} (\nabla_{\pi}^2 V(0, \pi^*))^{-1} \mathbb{E}_{\mathbb{P}} \left[ \frac{\langle X, \pi^* \rangle u' u'' + (u')^2}{|\pi^* u'|^{2-q}} \right] \cdot \pi^* \\ &= \|u'\|_{L^q(\mathbb{P})}^{1-q} \left( \mathbb{E}_{\mathbb{P}} \left[ \frac{\langle X, \pi^* \rangle u'' + u'}{|u'|^{1-q}} \right] \right) (\nabla_{\pi}^2 V(0))^{-1} \cdot \frac{\pi^*}{|\pi^*|}, \end{aligned}$$

which concludes the proof for the case  $p \in (1, \infty)$ .

*Step 2:* The case  $p = \infty$  follows again by going through [Bartl et al., 2020, Proof of Theorem 4] line by line, but some arguments can be cut short in this case: indeed, let us first note that there is only one measure in the ball  $B_\delta(\mathbb{P})$  attaining the value

$$\inf_{\tilde{\mathbb{P}} \in B_\delta(\mathbb{P})} \mathbb{E}_{\tilde{\mathbb{P}}} [u(\langle X, \pi \rangle)],$$

namely  $\mathbb{P}_\delta^* = (x - \delta \text{sign}(\pi^*))_{\#} \mathbb{P}$ . On the other hand, by optimality of  $\pi_\delta^*$  we have

$$\mathbb{E}_{\mathbb{P}_\delta^*} [Xu'(\langle X, \pi_\delta^* \rangle)] = 0,$$

and under the uniform integrability condition stated in Assumption 6.3(i)(b) one can directly compute

$$\begin{aligned} - \lim_{\delta \rightarrow 0} \frac{\mathbb{E}_{\mathbb{P}} [Xu'(\langle X, \pi_\delta^* \rangle)]}{\delta} &= \lim_{\delta \rightarrow 0} \frac{\mathbb{E}_{\mathbb{P}_\delta^*} [Xu'(\langle X, \pi_\delta^* \rangle)] - \mathbb{E}_{\mathbb{P}} [Xu'(\langle X, \pi_\delta^* \rangle)]}{\delta} \\ &= \lim_{\delta \rightarrow 0} \left( \frac{\mathbb{E}_{\mathbb{P}} [(X - \delta \text{sign}(\pi^*))u'(\langle X - \delta \text{sign}(\pi^*), \pi_\delta^* \rangle)]}{\delta} \right. \\ &\quad \left. - \frac{\mathbb{E}_{\mathbb{P}} [Xu'(\langle X, \pi_\delta^* \rangle)]}{\delta} \right) \\ &= -\mathbb{E}_{\mathbb{P}} [\langle X, \pi^* \rangle u''(\langle X, \pi^* \rangle) + u'(\langle X, \pi^* \rangle)] \cdot \text{sign}(\pi^*). \end{aligned}$$

Lastly, as the matrix  $\nabla_\pi^2 V(0)$  is invertible, we can again follow the same arguments as in [Bartl et al., 2020, Proof of Theorem 4] to obtain

$$\begin{aligned} (\pi^*)'(0) &= (\nabla_\pi^2 V(0))^{-1} \lim_{\delta \rightarrow 0} \frac{\mathbb{E}_{\mathbb{P}} [Xu'(\langle X, \pi_\delta^* \rangle)]}{\delta} \\ &= (\nabla_\pi^2 V(0))^{-1} \mathbb{E}_{\mathbb{P}} [\langle X, \pi^* \rangle u''(\langle X, \pi^* \rangle) + u'(\langle X, \pi^* \rangle)] \cdot \text{sign}(\pi^*). \end{aligned}$$

This shows the claim for the case  $p = \infty$ .  $\square$

## 7. DISTRIBUTIONALLY ROBUST MARGINAL UTILITY PRICING

We move now to the discussion and rigorous statements of results announced in section 4.3 above.

**7.1. Existence and regularity.** Before we state the main theorem of this section, we set up some useful notation and state some immediate consequences of our setup. For this, we will first keep  $p_d > 0$  fixed. We recall  $V(\delta, \varepsilon, p_d)$  defined in (11).

**Definition 7.1.** *Let us define*

$$(12) \quad v(\varepsilon, x, \pi) = u \left( -\varepsilon + \langle x, \pi \rangle + \frac{\varepsilon}{p_d} g(x) \right)$$

and

$$V(\delta, \varepsilon, \pi) := \inf_{\tilde{\mathbb{P}} \in B_\delta(\mathbb{P})} \mathbb{E}_{\tilde{\mathbb{P}}} [v(\varepsilon, X, \pi)].$$

Given a strategy  $\pi \in \mathcal{A}$ , write

$$B_\delta^{*, \varepsilon}(\mathbb{P}, \pi) := \{\mathbb{P}^* \in B_\delta(\mathbb{P}) \text{ such that } \mathbb{E}_{\mathbb{P}^*} [v(\varepsilon, X, \pi)] = V(\delta, \varepsilon, \pi)\}$$

for the set of minimising measures. Lastly, denote the set of optimising vectors by

$$\mathcal{A}_\delta^{*, \varepsilon} := \{\pi \in \mathcal{A} : V(\delta, \varepsilon, \pi) = V(\delta, \varepsilon, p_d)\}.$$

We denote a generic element of  $\mathcal{A}_\delta^{*, \varepsilon}$  by  $\pi^*(\delta, \varepsilon)$  and recall that  $\pi^* = \pi^*(0, 0)$  as well as  $\pi_\delta^* = \pi^*(\delta, 0)$ .

In particular we note that the sets  $\mathcal{A}_\delta^{*,0}$  and  $B_\delta^{*,0}(\mathbb{P}, \pi)$  are independent of  $p_d$  and the payoff  $g$ . We now make the following assumption:

**Assumption 7.2.** *There exists  $k \in (0, \infty)$  such that the function  $v : [-k, k] \times \mathcal{S} \times \mathcal{A} \rightarrow \mathbb{R}$  in (12) is continuously differentiable. Furthermore,*

(i) *If  $p \in (1, \infty)$  then for each  $r > 0$  there exist  $c, \gamma > 0$  such that*

- (a)  $|v(\varepsilon, x, \pi)| \leq c(1 + |x|^{p-\gamma}),$
- (b)  $|\nabla_\varepsilon v(\varepsilon, x, \pi)| \leq c(1 + |x|^{p-\gamma})$

*for all  $x \in \mathcal{S}, \pi \in \mathcal{A}$  such that  $|\pi| \leq r$  and  $\varepsilon \in [-k, k]$ .*

(ii) *If  $p = \infty$ , then there exists some  $\tilde{\delta} > 0$  such that for each  $r > 0$*

$$\mathbb{E}_{\mathbb{P}} \left[ \sup_{\pi \in \mathcal{A}, |\pi| \leq r} \sup_{|z| \leq \tilde{\delta}} \sup_{\varepsilon \in [-k, k]} |\nabla_\varepsilon v(\varepsilon, X + z, \pi)| \right] < \infty.$$

**Remark 7.3.** *Recall that  $u : \mathbb{R} \rightarrow \mathbb{R}$  is assumed to be continuously differentiable. Thus*

$$\nabla_\varepsilon v(\varepsilon, x, \pi) = u' \left( -\varepsilon + \langle x, \pi \rangle + \frac{\varepsilon}{p_d} g(x) \right) \left( -1 + \frac{g(x)}{p_d} \right).$$

*The assumption that  $v$  is continuously differentiable thus implies in particular that  $g : \mathcal{S} \rightarrow \mathbb{R}$  is continuous.*

**Remark 7.4.** *Note that Assumption 7.2 is the natural robust counterpart of Assumption A.1 in appendix A and is thus stronger than Assumption A.1. We also remark that when setting  $\varepsilon = 0$  in item (i) of Assumption 7.2 we obtain in particular*

$$\begin{aligned} |u(\langle x, \pi \rangle)| &\leq c(1 + |x|^{p-\gamma}), \\ \left| u'(\langle x, \pi \rangle) \left( -1 + \frac{g(x)}{p_d} \right) \right| &\leq c(1 + |x|^{p-\gamma}), \end{aligned}$$

*which is usually stronger than Assumption 6.1.*

Before we consider a characterisation of the distributionally robust Davis price in the spirit of (4) we first establish existence and uniqueness of optimisers  $\pi^*(\delta, 0) \in \mathcal{A}_\delta^{*,0}$  and  $\mathbb{P}^* \in B_\delta^{*,0}(\mathbb{P}, \pi(\delta, 0))$ .

**Lemma 7.5.** *Let Assumption 7.2 be satisfied. Then the following hold:*

(i) *The set  $\mathcal{A}_\delta^{*,\varepsilon} \neq \emptyset$  for all  $\delta \geq 0$  and there exists a compact set  $K \subseteq \mathbb{R}^d$  such that the set of optimisers is contained in  $K$ , i.e.*

$$\bigcup_{\varepsilon \in [-k, k]} \bigcup_{\delta \geq 0} \mathcal{A}_\delta^{*,\varepsilon} \subseteq K.$$

(ii) *For every sequence  $(\varepsilon_n)_{n \in \mathbb{N}}$  such that  $\lim_{n \rightarrow \infty} \varepsilon_n = 0$ , fixed  $\delta \geq 0$  and  $(\pi_n)_{n \in \mathbb{N}}$  such that  $\pi_n \in \mathcal{A}_\delta^{*,\varepsilon_n}$  for all  $n \in \mathbb{N}$  there exists a subsequence which converges to some  $\pi^* \in \mathcal{A}_\delta^{*,0}$ .*

(iii) *For every sequence  $(\delta_n)_{n \in \mathbb{N}}$  such that  $\lim_{n \rightarrow \infty} \delta_n = 0$  and  $(\pi_n)_{n \in \mathbb{N}}$  such that  $\pi_n \in \mathcal{A}_{\delta_n}^{*,0}$  for all  $n \in \mathbb{N}$  there exists a subsequence which converges to  $\pi^* \in \mathcal{A}_0^{*,0}$ .*

*Proof.* The first claim follows exactly as in the proof of [Bartl et al., 2020, Lemma 19] exchanging the “inf” and the “sup” and noting that

$$\inf_{\tilde{\mathbb{P}} \in B_\delta(\mathbb{P})} \mathbb{E}_{\tilde{\mathbb{P}}} [v(\varepsilon, X, \pi)] \leq \mathbb{E}_{\mathbb{P}} [v(\varepsilon, X, \pi)]$$

for all  $\delta \geq 0, \pi \in \mathcal{A}$  and  $\varepsilon \in [-k, k]$ , see also the proof of Lemma A.2, noting that  $V(\delta, \varepsilon) > -\infty$  by Assumption 7.2.(i). For (ii) we argue by contradiction: let us assume that (possibly after taking a subsequence)  $\pi_n$  converges to a limit  $\tilde{\pi} \notin \mathcal{A}_\delta^{*,0}$ . Since  $\tilde{\pi}$  is not an optimiser, and using the reverse Fatou lemma, we have

$$\begin{aligned} V(\delta, 0, p_d) &> \inf_{\tilde{\mathbb{P}} \in B_\delta(\mathbb{P})} \mathbb{E}_{\tilde{\mathbb{P}}} [v(0, X, \tilde{\pi})] \geq \limsup_{n \rightarrow \infty} \inf_{\tilde{\mathbb{P}} \in B_\delta(\mathbb{P})} \mathbb{E}_{\tilde{\mathbb{P}}} [v(\varepsilon_n, X, \pi_n)] \\ &= \limsup_{n \rightarrow \infty} V(\delta, \varepsilon_n, p_d). \end{aligned}$$

On the other hand, plugging  $\pi^* \in \mathcal{A}_\delta^{*,0}$  into  $V(\delta, \varepsilon_n, p_d)$  implies

$$\begin{aligned} \liminf_{n \rightarrow \infty} V(\delta, \varepsilon_n, p_d) &\geq \liminf_{n \rightarrow \infty} \inf_{\tilde{\mathbb{P}} \in B_\delta(\mathbb{P})} \mathbb{E}_{\tilde{\mathbb{P}}} [v(\varepsilon_n, X, \pi^*)] \\ &\geq \liminf_{n \rightarrow \infty} \left( \inf_{\tilde{\mathbb{P}} \in B_\delta(\mathbb{P})} \mathbb{E}_{\tilde{\mathbb{P}}} [v(0, X, \pi^*)] \right. \\ &\quad \left. - \varepsilon_n \sup_{\tilde{\mathbb{P}} \in B_\delta(\mathbb{P})} \mathbb{E}_{\tilde{\mathbb{P}}} \left[ \int_0^1 \nabla_\varepsilon v(t\varepsilon_n, X, \pi^*) dt \right] \right) \\ &= V(\delta, 0, p_d) \end{aligned}$$

as  $|\nabla_\varepsilon v(\varepsilon, x, \pi^*)| \leq c(1 + |x|^p)$  for some  $c > 0$ , all  $\varepsilon$  small enough and all  $x \in \mathcal{S}$ . This gives the desired contradiction concludes the proof of the second part. The proof of (iii) follows again as in the proof of [Bartl et al., 2020, Lemma 19] exchanging the “inf” and the “sup”. This concludes the proof.  $\square$

While existence of optimisers is thus guaranteed, their uniqueness is more delicate. Let us start with the specific case  $\pi^* = 0$ . Interestingly, it turns out that under certain assumptions, this case fully characterises the martingale property of  $\mathbb{P}$  (recall that  $X = S_1 - S_0$ ):

**Lemma 7.6.** *Assume that  $0 \in \mathcal{A}^\circ$ . Then  $\sup_{\pi \in \mathcal{A}} \mathbb{E}_{\mathbb{P}} [u(\langle X, \pi \rangle)]$  is attained for  $\pi^* = 0$  if and only if  $\mathbb{E}_{\mathbb{P}} [X] = 0$ . In that case, for all  $\delta \geq 0$ ,*

$$\sup_{\pi \in \mathcal{A}} \inf_{\tilde{\mathbb{P}} \in B_\delta(\mathbb{P})} \mathbb{E}_{\tilde{\mathbb{P}}} [u(\langle X, \pi \rangle)]$$

*is attained for  $\pi^*(\delta, 0) = 0$ . Furthermore, the optimiser is unique,  $\mathcal{A}_\delta^{*,0} = \{0\}$ , and  $B_\delta^{*,0}(\mathbb{P}, \pi^*(\delta, 0)) = B_\delta(\mathbb{P})$ .*

*Proof.* Note that the first-order condition for  $\delta = 0$  at  $\pi^* = 0$  implies

$$0 = \mathbb{E}_{\mathbb{P}} [X u'(\langle X, \pi^* \rangle)] = u'(0) \mathbb{E}_{\mathbb{P}} [X],$$

and thus  $\mathbb{E}_{\mathbb{P}} [X] = 0$ , as  $u$  is strictly increasing. On the other hand, if  $\mathbb{E}_{\mathbb{P}} [X] = 0$ , Jensen's inequality implies

$$\sup_{\pi \in \mathcal{A}} \mathbb{E}_{\mathbb{P}} [u(\langle X, \pi \rangle)] \leq \sup_{\pi \in \mathcal{A}} u(\mathbb{E}_{\mathbb{P}} [\langle X, \pi \rangle]) = u(0),$$

and the supremum is attained for  $\pi^* = 0$ . Furthermore, if  $\pi^* = 0$  then

$$\begin{aligned} (13) \quad u(0) &= \inf_{\tilde{\mathbb{P}} \in B_\delta(\mathbb{P})} \mathbb{E}_{\tilde{\mathbb{P}}} [u(\langle X, 0 \rangle)] \leq \sup_{\pi \in \mathcal{A}} \inf_{\tilde{\mathbb{P}} \in B_\delta(\mathbb{P})} \mathbb{E}_{\tilde{\mathbb{P}}} [u(\langle X, \pi \rangle)] \\ &\leq \sup_{\pi \in \mathcal{A}} \mathbb{E}_{\mathbb{P}} [u(\langle X, \pi \rangle)] = u(0), \end{aligned}$$

so in fact equality holds in (13) and the supremum is again attained for  $\pi^*(\delta, 0) = 0$ . Lastly, we note that for any  $\pi \in \mathcal{A}, \pi \neq 0$  we have

$$\mathbb{P}(\{X \in \mathcal{S}: \langle X, \pi \rangle > 0\}) > 0$$

as stated in Assumption 2.1. Thus it is always possible to find  $\hat{\mathbb{P}} \in B_\delta(\mathbb{P})$  such that  $\mathbb{E}_{\hat{\mathbb{P}}}[\langle X, \pi \rangle] < 0$ , and so

$$\begin{aligned} \inf_{\hat{\mathbb{P}} \in B_\delta(\mathbb{P})} \mathbb{E}_{\hat{\mathbb{P}}}[u(\langle X, \pi \rangle)] &\leq \mathbb{E}_{\hat{\mathbb{P}}}[u(\langle X, \pi \rangle)] \\ &\leq u(\mathbb{E}_{\hat{\mathbb{P}}}[\langle X, \pi \rangle]) < u(0), \end{aligned}$$

where we used Jensen's inequality again. In conclusion  $\pi_\delta^* = 0$  is the unique optimiser. This concludes the proof.  $\square$

More generally, the lemma below shows that  $\pi_\delta^* \in \mathcal{A}_\delta^{*,0}$  is always unique. In the case  $0 \notin \mathcal{A}^*$  we also obtain uniqueness of  $\mathbb{P}^* \in B_\delta^{*,0}(\mathbb{P}, \pi_\delta^*)$ .

**Lemma 7.7.** *Let Assumption 7.2 hold. For  $\delta > 0$  small enough the optimal strategy is unique,  $\mathcal{A}_\delta^{*,0} = \{\pi_\delta^*\}$ . Furthermore  $\pi_\delta^* \rightarrow \pi^*$  for  $\delta \rightarrow 0$ , where  $\pi^*$  is again unique. If in addition  $0 \notin \mathcal{A}^*$ , then  $B_\delta^{*,0}(\mathbb{P}, \pi^*) = \{\mathbb{P}^*\}$  is a singleton.*

*Proof.* Take any  $\pi, \tilde{\pi} \in \mathcal{A}$  such that  $\pi \neq \tilde{\pi}$ . Assumption 7.2 and [Bartl et al., 2020, Lemma 20] implies that the set  $B_\delta^{*,0}(\mathbb{P}, \lambda\pi + (1-\lambda)\tilde{\pi}) \neq \emptyset$  for all  $\lambda \in [0, 1]$ . Thus for all  $\lambda \in (0, 1)$  we have for any  $\hat{\mathbb{P}} \in B_\delta^{*,0}(\mathbb{P}, \lambda\pi + (1-\lambda)\tilde{\pi})$  that

$$\begin{aligned} V(\delta, 0, \lambda\pi + (1-\lambda)\tilde{\pi}) &= \mathbb{E}_{\hat{\mathbb{P}}}[v(0, x, \lambda\pi + (1-\lambda)\tilde{\pi})] \\ &> \lambda \mathbb{E}_{\hat{\mathbb{P}}}[v(0, X, \pi)] + (1-\lambda) \mathbb{E}_{\hat{\mathbb{P}}}[v(0, X, \tilde{\pi})] \\ &\geq \lambda V(\delta, 0, \pi) + (1-\lambda) V(\delta, 0, \tilde{\pi}) \end{aligned}$$

for all  $\delta \geq 0$  small enough using Assumption 2.1 and the fact  $u$  is strictly concave. In conclusion  $\pi \mapsto V(\delta, 0, \pi)$  is strictly concave on  $\mathcal{A}$  and thus the optimiser  $\pi^* \in \mathcal{A}_\delta^{*,0}$  is unique. Next, Lemma 7.5 implies that  $\pi_\delta^* \rightarrow \pi^*$  for  $\delta \rightarrow 0$ , where  $\pi^*$  is again unique by strict concavity of  $u$ .

On the other hand, for  $\pi_\delta^* \neq 0$  and  $1 < p < \infty$ , uniqueness of  $\mathbb{P}^* \in B_\delta^{*,0}(\mathbb{P}, \pi_\delta^*)$  follows from strict convexity of  $L^p$ -spaces: indeed if there exists  $\mathbb{P}' \in B_\delta^{*,0}(\mathbb{P}, \pi_\delta^*)$ , then

$$\frac{1}{2}(\mathbb{P}^* + \mathbb{P}') \in (B_\delta(\mathbb{P}))^o$$

and

$$\frac{1}{2}(\mathbb{E}_{\mathbb{P}^*}[v(0, X, \pi_\delta^*)] + \mathbb{E}_{\mathbb{P}'}[v(0, X, \pi_\delta^*)]) = V(\delta, 0, \pi_\delta^*) = V(\delta).$$

However, as  $0 \notin \mathcal{A}^*$ , we conclude that  $V(\delta) < V(0)$ . This leads to a contradiction. Lastly, for  $p = \infty$  and  $\pi_\delta^* \neq 0$ , it can be directly seen that  $\mathbb{P}^* = (x - \text{sign}(\pi_\delta^*))_{\#} \mathbb{P}$  is unique as well.  $\square$

We now compute the derivative  $\nabla_\varepsilon V(\delta, 0, p_d)$  for fixed  $\delta \geq 0$  and  $p_d > 0$ .

**Theorem 7.8.** *Fix  $\delta \geq 0, p_d > 0$  and let Assumption 7.2 hold. Then*

$$\begin{aligned} \nabla_\varepsilon V(\delta, 0, p_d) &= \sup_{\pi \in \mathcal{A}_\delta^{*,0}} \inf_{\mathbb{P}^* \in B_\delta^{*,0}(\mathbb{P}, \pi)} \mathbb{E}_{\mathbb{P}^*}[\nabla_\varepsilon v(0, X, \pi)] \\ &= \sup_{\pi \in \mathcal{A}_\delta^{*,0}} \inf_{\mathbb{P}^* \in B_\delta^{*,0}(\mathbb{P}, \pi)} \mathbb{E}_{\mathbb{P}^*} \left[ u'(\langle X, \pi \rangle) \left( -1 + \frac{g(X)}{p_d} \right) \right]. \end{aligned}$$



In particular the robust marginal utility price  $\hat{p}_d(\delta)$  is a solution to

$$(14) \quad \inf_{\mathbb{P}^* \in B_{\delta}^{*,0}(\mathbb{P}, \pi_{\delta}^*)} \mathbb{E}_{\mathbb{P}^*} \left[ u'(\langle X, \pi_{\delta}^* \rangle) \left( -1 + \frac{g(X)}{p_d} \right) \right] = 0$$

and is thus given by

$$\hat{p}_d(\delta) = \frac{\mathbb{E}_{\mathbb{P}^*} [u'(\langle X, \pi_{\delta}^* \rangle) g(X)]}{\mathbb{E}_{\mathbb{P}^*} [u'(\langle X, \pi_{\delta}^* \rangle)]}$$

for any  $\mathbb{P}^* \in B_{\delta}^{*,0}(\mathbb{P}, \pi^*(\delta, 0))$ .

*Proof of Theorem 7.8. Step 1:* We first consider the case  $p \in (1, \infty)$ . We start by proving the inequality

$$(15) \quad \lim_{\varepsilon \downarrow 0} \frac{V(\delta, \varepsilon, p_d) - V(\delta, 0, p_d)}{\varepsilon} \leq \sup_{\pi \in \mathcal{A}_{\delta}^{*,0}} \inf_{\mathbb{P}^* \in B_{\delta}^{*,0}(\mathbb{P}, \pi)} \mathbb{E}_{\mathbb{P}^*} [\nabla_{\varepsilon} v(0, X, \pi)].$$

Indeed, for small  $\varepsilon > 0$  let us take optimisers  $\pi^*(\delta, \varepsilon) \in \mathcal{A}_{\delta}^{*,\varepsilon}$ . Then, we can apply Lemma 7.5.(ii), so that after taking a subsequence (without relabelling) there exists some  $\pi \in \mathcal{A}_{\delta}^{*,0}$  such that  $\lim_{\varepsilon \rightarrow 0} \pi^*(\delta, \varepsilon) = \pi$ . Take  $\mathbb{P}^* \in B_{\delta}^{*,0}(\mathbb{P}, \pi)$ . As  $\mathbb{P}^* \in B_{\delta}(\mathbb{P})$  we then have

$$\lim_{\varepsilon \downarrow 0} \frac{V(\delta, \varepsilon, p_d) - V(\delta, 0, p_d)}{\varepsilon} \leq \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \mathbb{E}_{\mathbb{P}^*} [v(\varepsilon, X, \pi^*(\delta, \varepsilon)) - v(0, X, \pi)].$$

By assumption we have  $|\nabla_{\varepsilon} v(\tilde{\varepsilon}, x, \pi)| \leq c(1 + |x|^{p-\gamma})$ . Then arguing as in (18), the envelope theorem for arbitrary choice sets of [Milgrom and Segal, 2002, Theorem 3] implies that

$$\lim_{\varepsilon \rightarrow 0} \frac{V(\delta, \varepsilon, p_d) - V(\delta, 0, p_d)}{\varepsilon} \leq \mathbb{E}_{\mathbb{P}^*} [\nabla_{\varepsilon} v(0, X, \pi)].$$

As  $\mathbb{P}^* \in B_{\delta}^{*,0}(\mathbb{P}, \pi)$  was arbitrary, we have shown (15).

We proceed to show that

$$(16) \quad \lim_{\varepsilon \rightarrow 0} \frac{V(\delta, \varepsilon, p_d) - V(\delta, 0, p_d)}{\varepsilon} \geq \sup_{\pi \in \mathcal{A}_{\delta}^{*,0}} \inf_{\mathbb{P}^* \in B_{\delta}^{*,0}(\mathbb{P}, \pi)} \mathbb{E}_{\mathbb{P}^*} [\nabla_{\varepsilon} v(0, X, \pi)].$$

Take  $\pi \in \mathcal{A}_{\delta}^{*,0}$ . For every sufficiently small  $\varepsilon > 0$ , let  $\mathbb{P}^{*,\varepsilon} \in B_{\delta}^{*,\varepsilon}(\mathbb{P}, \pi)$ , so that  $V(\delta, \varepsilon, \pi^*) = \mathbb{E}_{\mathbb{P}^{*,\varepsilon}} [v(\varepsilon, X, \pi^*)]$ . The existence of such  $\mathbb{P}^{*,\varepsilon}$  is guaranteed by the assumption  $|v(\varepsilon, x, \pi)| \leq c(1 + |x|^{p-\gamma})$  for all  $x \in \mathcal{S}$  and [Bartl et al., 2020, Lemma 20], which also guarantees that (possibly after passing to a subsequence) there is  $\mathbb{P}^* \in B_{\delta}(\mathbb{P})$  such that  $\mathbb{P}^{*,\varepsilon} \rightarrow \mathbb{P}^*$  in  $W_{p-\gamma}$ . We claim that  $\mathbb{P}^* \in B_{\delta}^{*,0}(\mathbb{P}, \pi)$ . Indeed, as  $|\nabla_{\varepsilon} v(\tilde{\varepsilon}, x, \pi)| \leq c(1 + |x|^{p-\gamma})$  and  $|v(0, x, \pi)| \leq c(1 + |x|^{p-\gamma})$  one has

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0} V(\delta, \varepsilon, \pi) &= \lim_{\varepsilon \rightarrow 0} \mathbb{E}_{\mathbb{P}^{*,\varepsilon}} [v(\varepsilon, X, \pi)] \\ &= \lim_{\varepsilon \rightarrow 0} \left( \mathbb{E}_{\mathbb{P}^{*,\varepsilon}} [v(0, X, \pi)] + \varepsilon \mathbb{E}_{\mathbb{P}^{*,\varepsilon}} \left[ \int_0^1 \nabla_{\varepsilon} v(t\varepsilon, X, \pi) dt \right] \right) \\ &= \mathbb{E}_{\mathbb{P}^*} [v(0, X, \pi)] \geq V(\delta, 0, \pi). \end{aligned}$$

On the other hand, for any choice  $\tilde{\mathbb{P}} \in B_{\delta}^{*,0}(\mathbb{P}, \pi)$  one has by Assumption 7.2 and dominated convergence

$$\lim_{\varepsilon \rightarrow 0} V(\delta, \varepsilon, \pi) \leq \lim_{\varepsilon \rightarrow 0} \mathbb{E}_{\tilde{\mathbb{P}}} [v(\varepsilon, X, \pi)] = \mathbb{E}_{\tilde{\mathbb{P}}} [v(0, X, \pi)] = V(\delta, 0, \pi).$$

This implies  $V(\delta, 0, \pi) = \mathbb{E}_{\mathbb{P}^*} [v(0, X, \pi)]$  and in particular  $\mathbb{P}^* \in B_\delta^{\star, 0}(\mathbb{P}, \pi)$ . At this point we expand  $v(\varepsilon, x, \pi) = v(0, x, \pi) + \varepsilon \int_0^1 \nabla_\varepsilon v(t\varepsilon, x, \pi) dt$  so that

$$\begin{aligned} V(\delta, \varepsilon, \pi^*(\delta, \varepsilon)) - V(\delta, 0, \pi) &\geq V(\delta, \varepsilon, \pi) - V(\delta, 0, \pi) \\ &= \mathbb{E}_{\mathbb{P}^{\star, \varepsilon}} \left[ \left( v(0, X, \pi) + \varepsilon \int_0^1 \nabla_\varepsilon v(t\varepsilon, X, \pi) dt \right) \right] - \mathbb{E}_{\mathbb{P}^*} [v(0, X, \pi)] \\ &\geq \mathbb{E}_{\mathbb{P}^{\star, \varepsilon}} \left[ \left( v(0, X, \pi) + \varepsilon \int_0^1 \nabla_\varepsilon v(t\varepsilon, X, \pi) dt \right) \right] - \mathbb{E}_{\mathbb{P}^{\star, \varepsilon}} [v(0, X, \pi)] \\ &= \varepsilon \mathbb{E}_{\mathbb{P}^{\star, \varepsilon}} \left[ \int_0^1 \nabla_\varepsilon v(t\varepsilon, X, \pi) dt \right], \end{aligned}$$

where we used  $\mathbb{P}^* \in B_\delta^{\star, 0}(\mathbb{P}, \pi)$  for the first equality. Fix  $t \in [0, 1]$  and recall that  $|\nabla_\varepsilon v(\tilde{\varepsilon}, x, \pi)| \leq c(1 + |x|^{p-\gamma})$  for all  $x \in \mathcal{S}$ ,  $\pi \in K$ ,  $\tilde{\varepsilon} > 0$  small enough and  $(\tilde{\varepsilon}, x) \mapsto \nabla_\varepsilon v(\tilde{\varepsilon}, x, \pi)$  is continuous. As  $t\varepsilon \rightarrow 0$  and  $\mathbb{P}^{\star, \varepsilon}$  converge to  $\mathbb{P}^*$  in  $W_{p-\gamma}$  for  $\varepsilon \rightarrow 0$ , we conclude that the measures  $\delta_{t\varepsilon} \otimes \mathbb{P}^{\star, \varepsilon}$  converge (on the product space) in  $W_{p-\gamma}$  to the measure  $\delta_0 \otimes \mathbb{P}^*$  (here  $\otimes$  denotes the product measure). As

$$\mathbb{E}_{\mathbb{P}^{\star, \varepsilon}} [\nabla_\varepsilon(t\varepsilon, X, \pi)] = \mathbb{E}_{\delta_{t\varepsilon} \otimes \mathbb{P}^{\star, \varepsilon}} [\nabla_\varepsilon(\tilde{\varepsilon}, X, \pi)]$$

and similarly

$$\mathbb{E}_{\mathbb{P}^*} [\nabla_\varepsilon(0, X, \pi)] = \mathbb{E}_{\delta_0 \otimes \mathbb{P}^*} [\nabla_\varepsilon(\tilde{\varepsilon}, X, \pi)],$$

we conclude that

$$\mathbb{E}_{\mathbb{P}^{\star, \varepsilon}} [\nabla_\varepsilon(t\varepsilon, X, \pi)] \rightarrow \mathbb{E}_{\mathbb{P}^*} [\nabla_\varepsilon(0, X, \pi)].$$

Using the above together with Fubini's theorem and the dominated convergence theorem, we finally conclude that

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \mathbb{E}_{\mathbb{P}^{\star, \varepsilon}} \left[ \int_0^1 \nabla_\varepsilon v(t\varepsilon, X, \pi) dt \right] &= \lim_{\varepsilon \rightarrow 0} \int_0^1 \mathbb{E}_{\mathbb{P}^{\star, \varepsilon}} [\nabla_\varepsilon v(t\varepsilon, X, \pi)] dt \\ &= \int_0^1 \mathbb{E}_{\mathbb{P}^*} [\nabla_\varepsilon v(0, X, \pi)] dt \\ &= \mathbb{E}_{\mathbb{P}^*} [\nabla_\varepsilon v(0, X, \pi)], \end{aligned}$$

which ultimately shows (16) for  $p \in (1, \infty)$ .

*Step 2:* Let us now consider the case  $p = \infty$ . Inequality (15) follows as before, noting that the condition

$$(17) \quad \mathbb{E}_{\mathbb{P}} \left[ \sup_{\pi \in \mathcal{A}, |\pi| \leq r} \sup_{|a| \leq s} \sup_{\varepsilon \in [-k, k]} |\nabla_\varepsilon v(\varepsilon, X + a, \pi)| \right] < \infty$$

implies by the dominated convergence theorem that

$$\nabla_\varepsilon \mathbb{E}_{\mathbb{P}^*} [v(\varepsilon, X, \pi)] = \mathbb{E}_{\mathbb{P}^*} [\nabla_\varepsilon v(\varepsilon, X, \pi)]$$

for all  $\varepsilon \in [-k, k]$ . Now we can apply again the envelope theorem for arbitrary choice sets of [Milgrom and Segal, 2002, Theorem 3] to conclude that

$$\lim_{\varepsilon \rightarrow 0} \frac{V(\delta, \varepsilon, p_d) - V(\delta, 0, p_d)}{\varepsilon} \leq \mathbb{E}_{\mathbb{P}^*} [\nabla_\varepsilon v(0, X, \pi)].$$

We now show inequality (16). Note that in this case the existence of  $\mathbb{P}^{\star, \varepsilon} \in B_\delta^{\star, \varepsilon}(\mathbb{P}, \pi)$ , so that  $V(\delta, \varepsilon, \pi^*) = \mathbb{E}_{\mathbb{P}^{\star, \varepsilon}} [v(\varepsilon, X, \pi^*)]$  is guaranteed by Prokhorov's theorem, the Portmanteau lemma, (17) and the dominated convergence theorem.

In the same way it can be shown that (possibly after passing to a subsequence) there is  $\mathbb{P}^* \in B_\delta(\mathbb{P})$  such that

$$\lim_{\varepsilon \rightarrow 0} \mathbb{E}_{\mathbb{P}^*, \varepsilon} [v(\varepsilon, X, \pi)] = \mathbb{E}_{\mathbb{P}^*} [v(0, X, \pi)].$$

In particular the same arguments as in *Step 1* above imply  $V(0, \pi) = \mathbb{E}_{\mathbb{P}^*} [v(0, X, \pi)]$ . Lastly we conclude using Fubini's theorem, (17) and the dominated convergence theorem as above that

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0} \frac{V(\delta, \varepsilon, \pi^*(\delta, \varepsilon)) - V(\delta, 0, \pi)}{\varepsilon} &\geq \lim_{\varepsilon \rightarrow 0} \mathbb{E}_{\mathbb{P}^*, \varepsilon} \left[ \int_0^1 \nabla_\varepsilon v(t\varepsilon, X, \pi) dt \right] \\ &= \mathbb{E}_{\mathbb{P}^*} [\nabla_\varepsilon v(0, X, \pi)]. \end{aligned}$$

This concludes the proof.  $\square$

In conclusion, Theorem 7.8 states that for any pair of optimisers  $(\pi_\delta^*, \mathbb{P}^*) \in \mathcal{A}_\delta^{*,0} \times B_\delta^{*,0}(\mathbb{P}, \pi)$  – and we recall that apart from the extreme case  $\pi^* = 0$  this pair is in fact unique – the robust marginal utility price  $\hat{p}_d(\delta)$  can be written as

$$\hat{p}_d(\delta) = \frac{\mathbb{E}_{\mathbb{P}^*} [u'(\langle X, \pi_\delta^* \rangle) g(X)]}{\mathbb{E}_{\mathbb{P}^*} [u(\langle X, \pi_\delta^* \rangle)]}.$$

We remark that the same arguments as in the classical case given in Section 2.2 imply that  $\hat{p}_d(\delta)$  is the expectation under a martingale measure equivalent to  $\mathbb{P}^*$ , in particular the optimiser  $\mathbb{P}^*$  does not allow for arbitrage and is close in Wasserstein sense to  $\mathbb{P}$ .

**7.2. Sensitivity of distributionally robust marginal utility price.** We already know from results in Bartl et al. [2020], that for small  $\delta > 0$ , the maximising measure  $\mathbb{P}^* \in B_\delta(\mathbb{P})$  is built from  $\mathbb{P}$  as a push forward in the direction  $-\pi^* u'(\langle X, \pi^* \rangle)$ , which is  $\pi^*$  times the Radon-Nikodym derivative  $d\mathbb{Q}_u/d\mathbb{P}$  of the agent's subjective martingale measure  $\mathbb{Q}_u$  defined in (2). For small  $\delta > 0$  we thus have that

$$\begin{aligned} \hat{p}_d(\delta) &\approx \frac{1}{\mathbb{E}_{\mathbb{P}} [u'(\langle X - \delta \pi^* u'(\langle X, \pi^* \rangle), \pi^* + \delta(\pi^*)'(0) \rangle)]} \\ &\quad \cdot \mathbb{E}_{\mathbb{P}} \left[ u'(\langle X - \delta \pi^* u'(\langle X, \pi^* \rangle), \pi^* + \delta(\pi^*)'(0) \rangle) \right. \\ &\quad \left. \cdot g(X - \delta \pi^* u'(\langle X, \pi^* \rangle)) \right] \\ &= \mathbb{E}_{\mathbb{Q}_u^\delta} [g(X - \delta \pi^* u'(\langle X, \pi^* \rangle))], \end{aligned}$$

where

$$\frac{d\mathbb{Q}_u^\delta}{d\mathbb{P}}(x) = \frac{u'(\langle x - \delta \pi^* u'(\langle x, \pi^* \rangle), \pi^* + \delta \pi'(0) \rangle)}{\mathbb{E}_{\mathbb{P}} [u'(\langle X - \delta \pi^* u'(\langle X, \pi^* \rangle), \pi^* + \delta(\pi^*)'(0) \rangle)]},$$

where we recall from (10) that

$$(\pi^*)'(0) = \|u'(\langle X, \pi^* \rangle)\|_{L^q(\mathbb{P})}^{1-q} \cdot (\nabla_\pi^2 V(0))^{-1} \cdot \frac{\pi^*}{|\pi^*|} \cdot \left( \mathbb{E}_{\mathbb{P}} \left[ \frac{\langle X, \pi^* \rangle u''(\langle X, \pi^* \rangle) + u'(\langle X, \pi^* \rangle)}{|u'(\langle X, \pi^* \rangle)|^{1-q}} \right] \right)$$

was computed in Section 6.2.

We now conduct a first-order asymptotic analysis of the robust marginal utility price  $\hat{p}_d(\delta)$ . This will enable us to quantify a first-order premium paid for model-uncertainty. We derive the following result:

**Theorem 7.9.** *Assume  $g$  is continuously differentiable.*

(i) *Suppose  $\pi^* = 0$  and the following are satisfied:*

(a) *if  $p \in (1, \infty)$ , then there exists  $c > 0$  such that*

$$|\nabla g(x)| \leq c(1 + |x|^{p-1})$$

*for all  $x \in \mathcal{S}$ ,*

(b) *if  $p = \infty$ , then there exists  $\tilde{\delta} > 0$  such that*

$$\mathbb{E}_{\mathbb{P}} \left[ \sup_{|z| \leq \tilde{\delta}} |\nabla g(x + z)| \right] < \infty.$$

*Then the Davis price  $\hat{p}_d(\delta)$  satisfies*

$$\hat{p}'_d(0) = -(\mathbb{E}_{\mathbb{P}} [|\nabla g(x)|^q])^{1/q}.$$

(ii) *Suppose  $\pi^* \neq 0$  and Assumptions 6.1, 6.3, 7.2 hold. Define*

$$T(x) := \frac{\pi^*}{|\pi^*|} \frac{1}{|u'(\langle x, \pi^* \rangle)|^{1-q}} (\mathbb{E}_{\mathbb{P}} [|u'(\langle X, \pi^* \rangle)|^q])^{1/q-1}.$$

*Then*

$$\hat{p}'_d(0) = \mathbb{E}_{\mathbb{Q}_u} [R_u(\langle X, \pi^* \rangle) (\langle T(X), \pi^* \rangle - \langle X, (\pi^*)'(0) \rangle) \cdot (g(X) - \hat{p}_d) - \langle \nabla g(X), T(X) \rangle],$$

*where  $\mathbb{Q}_u$  was given in (2) and  $R_u(x) = -\frac{u''(x)}{u'(x)}$ .*

*Proof of Theorem 6.2.* For the first claim, note that Lemma 7.6 implies  $\pi_\delta^* = 0 \in \mathcal{A}_\delta^{*,0}$  is the unique optimiser for all  $\delta \geq 0$  and thus by equation (14) in Theorem 7.8 the robust marginal utility price  $\hat{p}_d(\delta)$  is given by

$$\hat{p}_d(\delta) = \inf_{\tilde{\mathbb{P}} \in B_\delta(\mathbb{P})} \mathbb{E}_{\tilde{\mathbb{P}}} [g(X)].$$

Thus the claim follows from [Bartl et al., 2020, Theorem 2].

Assume now that  $\pi^* \neq 0$  and fix the sequence  $(\pi_\delta^*)_{\delta \geq 0}$  in  $\mathcal{A}_\delta^{*,0}$ . By Lemma 7.5.(iii) we have  $\lim_{\delta \rightarrow 0} \pi_\delta^* = \pi^*$ . Furthermore we know from Theorem 7.8 that

$$\hat{p}_d(\delta) = \frac{\mathbb{E}_{\mathbb{P}_\delta^*} [u'(\langle X, \pi_\delta^* \rangle) g(X)]}{\mathbb{E}_{\mathbb{P}_\delta^*} [u'(\langle X, \pi_\delta^* \rangle)]}$$

for  $\mathbb{P}_\delta^* \in B_\delta^{*,0}(\mathbb{P})$  and all  $\delta \geq 0$ . Note that as

$$\begin{aligned} & \frac{\pi^*}{|\pi^*|} \frac{1}{|u'(\langle x, \pi^* \rangle)|^{1-q}} (\mathbb{E}_{\mathbb{P}} [|u'(\langle X, \pi^* \rangle)|^q])^{1/q-1} \\ &= \frac{\pi^* u'(\langle x, \pi^* \rangle)}{(|\pi^*| u'(\langle x, \pi^* \rangle))^{2-q}} |\pi^*|^{1-q} (\mathbb{E}_{\mathbb{P}} [|u'(\langle X, \pi^* \rangle)|^q])^{1/q-1} \end{aligned}$$

we have

$$T(x) = \frac{\nabla_x f(x, \pi^*)}{|\nabla_x f(x, \pi^*)|^{2-q}} (\mathbb{E}_{\mathbb{P}} [|\nabla_x f(X, \pi^*)|^q])^{1/q-1}$$

for  $f(x, \pi) = u(\langle x, \pi \rangle)$ . By [Bartl et al., 2020, proof of Theorem 2] we may write  $\mathbb{P}_\delta^* = (x - \delta T(x) + o(\delta))_{\#} \mathbb{P}$ . Applying the quotient rule to

$$\begin{aligned} \hat{p}_d(\delta) &= \frac{\mathbb{E}_{\mathbb{P}_\delta^*} [u'(\langle X, \pi_\delta^* \rangle) g(X)]}{\mathbb{E}_{\mathbb{P}_\delta^*} [u'(\langle X, \pi_\delta^* \rangle)]} \\ &= \frac{\mathbb{E}_{\mathbb{P}} [u'(\langle X - \delta T(X) + o(\delta), \pi^* + (\pi^*)'(0)\delta + o(\delta) \rangle) g(X - \delta T(X) + o(\delta))]}{\mathbb{E}_{\mathbb{P}} [u'(\langle X - \delta T(X) + o(\delta), \pi^* + (\pi^*)'(0)\delta + o(\delta) \rangle)]} \end{aligned}$$

yields

$$\begin{aligned} \hat{p}'_d(0) = & \frac{1}{(\mathbb{E}_{\mathbb{P}}[u'(\langle X, \pi^* \rangle)])^2} \cdot \left( \mathbb{E}_{\mathbb{P}}[u'(\langle X, \pi^* \rangle)g(X)] \cdot \mathbb{E}_{\mathbb{P}}[u''(\langle X, \pi^* \rangle)(\langle T(X), \pi^* \rangle - \langle X, (\pi^*)'(0) \rangle)] \right. \\ & - \mathbb{E}_{\mathbb{P}}[u'(\langle X, \pi^* \rangle)] \cdot \mathbb{E}_{\mathbb{P}} \left[ u''(\langle X, \pi^* \rangle)g(X) \left[ \langle T(X), \pi^* \rangle - \langle X, (\pi^*)'(0) \rangle \right] \right. \\ & \left. \left. + u'(\langle X, \pi^* \rangle) \langle \nabla g(X), T(X) \rangle \right] \right), \end{aligned}$$

which equals

$$\begin{aligned} & \frac{1}{\mathbb{E}_{\mathbb{P}}[u'(\langle X, \pi^* \rangle)]} \left( \mathbb{E}_{\mathbb{P}}[u''(\langle X, \pi^* \rangle)(\langle T(X), \pi^* \rangle - \langle X, (\pi^*)'(0) \rangle)] \cdot (\mathbb{E}_{\mathbb{Q}_u}[g(X)] - g(X)) \right) \\ & - \mathbb{E}_{\mathbb{Q}_u}[\langle \nabla g(X), T(X) \rangle] \end{aligned}$$

and thus shows the claim.  $\square$

## 8. DATA AVAILABILITY STATEMENT

Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

## APPENDIX A. PROOFS FOR SECTION 2.2

For completeness, we present here rigorous statements and proofs of the classical results on marginal utility pricing in absence of model uncertainty which we discussed in section 2.2. We start with uniform integrability assumptions which yield the necessary control over  $V(\varepsilon, p_d)$ .

**Assumption A.1.** *For each fixed  $p_d > 0$  the following holds:*

- (1) *For all  $\pi \in \mathcal{A}$  there exists  $k \in (0, \infty)$  such that the function*

$$x \mapsto u \left( -\varepsilon + \langle x - X_0, \pi \rangle + \frac{\varepsilon}{p_d} g(x) \right)$$

*is dominated by a  $\mathbb{P}$ -integrable function uniformly for all  $\varepsilon \in [-k, k]$ .*

- (2) *There exists a compact set  $K$  such that  $\pi^* \in K^\circ$  and a constant  $k \in (0, \infty)$  such that  $\mathbb{P}$ -a.s.*

$$u' \left( -\varepsilon + \langle x - X_0, \pi \rangle + \frac{\varepsilon}{p_d} g(x) \right) \leq f(x)$$

*for all  $x \in \mathbb{R}^d$ ,  $\pi \in K$ ,  $\varepsilon \in [-k, k]$ , where  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  is chosen in such a way that*

$$x \mapsto f(x) \left( -1 + \frac{g(x)}{p_d} \right)$$

*is a  $\mathbb{P}$ -integrable function.*

Assumptions 2.1 and A.1 imply the following:

**Lemma A.2.** *Let Assumptions 2.1 and A.1.(1) hold. For fixed  $\varepsilon \in \mathbb{R}$ ,  $p_d > 0$ , the supremum in (3) is attained at a unique point, which we call  $\pi^*(\varepsilon)$ . Furthermore  $\pi^*(\varepsilon) \rightarrow \pi^*(0)$  for  $\varepsilon \rightarrow 0$ .*

*Proof of Lemma A.2.* We first show existence of a maximiser. We follow ideas outlined in [Bartl et al., 2020, Lemma 19] (Note that we are only working in the case  $\delta = 0$  here): consider a maximising sequence  $(\pi_n)_{n \in \mathbb{N}}$  for  $V(\varepsilon, p_d)$ , i.e.

$$V(\varepsilon, p_d) = \lim_{n \rightarrow \infty} \mathbb{E}_{\mathbb{P}} \left[ u \left( -\varepsilon + \langle X, \pi_n \rangle + \frac{\varepsilon}{p_d} g(X) \right) \right].$$

If  $(\pi_n)_{n \in \mathbb{N}}$  is bounded, then after passing to a subsequence there is a limit, and the reverse Fatou lemma (recall that  $u$  is bounded above) shows that this limit is a maximiser. It remains to argue why  $(\pi_n)_{n \in \mathbb{N}}$  is bounded. Heading for a contradiction, assume that  $|\pi_n| \rightarrow \infty$  as  $n \rightarrow \infty$ . After passing to a (not relabelled) subsequence, there exists  $\tilde{\pi} \in \mathbb{R}^d$  with  $|\tilde{\pi}| = 1$  such that  $\pi_n / |\pi_n| \rightarrow \tilde{\pi}$  as  $n \rightarrow \infty$ . As stated in Assumption 2.1 we have

$$\mathbb{P}(\{X \in \mathcal{S} : \langle X, \tilde{\pi} \rangle < 0\}) > 0.$$

As  $u$  is bounded from above this shows that

$$\lim_{n \rightarrow \infty} \mathbb{E}_{\mathbb{P}} [u(\langle X, \pi_n \rangle + g(X))] = -\infty,$$

a contradiction. Uniqueness of optimisers follows again by Assumption 2.1 and strict concavity of  $u$ .

Now we prove the last claim. Heading for a contradiction, we assume that there exists a sequence  $(\varepsilon_n)_{n \in \mathbb{N}}$  converging to zero, such that  $\pi^*(\varepsilon_n)$  does not converge to  $\pi^*$ . The exact same reasoning as above implies that  $(\pi^*(\varepsilon_n))_{n \in \mathbb{N}}$  is bounded, so that possibly after passing to a not relabelled subsequence there exists a limit  $\tilde{\pi} \neq \pi^*$ . The reverse Fatou lemma again implies that

$$\begin{aligned} V(0, p_d) &> \mathbb{E}_{\mathbb{P}} [u(\langle X, \tilde{\pi} \rangle)] \\ &\geq \limsup_{n \rightarrow \infty} \mathbb{E}_{\mathbb{P}} \left[ u \left( -\varepsilon_n + \langle X, \pi^*(\varepsilon_n) \rangle + \frac{\varepsilon_n}{p_d} g(X) \right) \right] = \limsup_{n \rightarrow \infty} V(\varepsilon_n, p_d). \end{aligned}$$

On the other hand, plugging in  $\pi^*$  yields

$$\begin{aligned} \liminf_{n \rightarrow \infty} V(\varepsilon_n, p_d) &\geq \liminf_{n \rightarrow \infty} \mathbb{E}_{\mathbb{P}} \left[ u \left( -\varepsilon_n + \langle X, \pi^* \rangle + \frac{\varepsilon_n}{p_d} g(X) \right) \right] \\ &= \mathbb{E}_{\mathbb{P}} [u(\langle X, \pi^* \rangle)] \\ &= V(0, p_d), \end{aligned}$$

where we have used the dominated convergence theorem together with Assumption A.1.(1). This yields a contradiction and concludes the proof.  $\square$

With the above notational conventions, Mark Davis characterised the marginal utility price as follows:

**Theorem A.3** (cf. [Davis, 1997, Theorem 3]). *Let Assumptions 2.1 and A.1 hold and assume*

$$\mathbb{E}_{\mathbb{P}} [u'(\langle X - X_0, \pi^* \rangle)] > 0.$$

*Then  $\hat{p}_d$  is unique and given by (4).*

*Proof of Theorem A.3.* Recall that  $x \mapsto u'(x)$  is continuous as stated in the introduction. Assumption A.1 then enables the use of the dominated convergence

theorem, so that

$$\begin{aligned} & \nabla_\varepsilon \left( \mathbb{E}_{\mathbb{P}} \left[ u \left( -\varepsilon + \langle X, \pi \rangle + \frac{\varepsilon}{p_d} g(X) \right) \right] \right) \\ &= \mathbb{E}_{\mathbb{P}} \left[ u' \left( -\varepsilon + \langle X, \pi \rangle + \frac{\varepsilon}{p_d} g(X) \right) \left( -1 + \frac{g(X)}{p_d} \right) \right]. \end{aligned}$$

In particular the function

$$(\varepsilon, \pi) \mapsto \nabla_\varepsilon \left( \mathbb{E}_{\mathbb{P}} \left[ u \left( -\varepsilon + \langle X, \pi \rangle + \frac{\varepsilon}{p_d} g(X) \right) \right] \right)$$

is uniformly continuous on  $[-k, k] \times K$ , as it is continuous and the set  $[-k, k] \times K$  is compact. Furthermore we find that

$$\begin{aligned} (18) \quad & \left| \nabla_\varepsilon \left( \mathbb{E}_{\mathbb{P}} \left[ u \left( -\varepsilon + \langle X, \pi \rangle + \frac{\varepsilon}{p_d} g(X) \right) \right] \right) \right| \\ & \leq \mathbb{E}_{\mathbb{P}} \left[ \left| u' \left( -\varepsilon + \langle X, \pi \rangle + \frac{\varepsilon}{p_d} g(X) \right) \right| \left| -1 + \frac{g(X)}{p_d} \right| \right] \\ & \leq \mathbb{E}_{\mathbb{P}} \left[ |f(X)| \left| -1 + \frac{g(X)}{p_d} \right| \right]. \end{aligned}$$

Thus the envelope theorem for arbitrary choice sets [Milgrom and Segal, 2002, Theorem 2] applies and so

$$\nabla_\varepsilon V(0, p_d) = \mathbb{E}_{\mathbb{P}} \left[ u'(\langle X, \pi^* \rangle) \left( -1 + \frac{g(X)}{p_d} \right) \right].$$

As the above formula holds for any  $p_d > 0$ , we conclude that

$$\hat{p}_d = \frac{\mathbb{E}_{\mathbb{P}} [u'(\langle X, \pi^* \rangle) g(X)]}{\mathbb{E}_{\mathbb{P}} [u'(\langle X, \pi^* \rangle)]},$$

which proves the result.  $\square$

## REFERENCES

- E. W. Anderson, L. P. Hansen, and T. J. Sargent. A quartet of semigroups for model specification, robustness, prices of risk, and model detection. *Journal of the European Economic Association*, 1(1):68–123, 2003.
- F. J. Anscombe, R. J. Aumann, et al. A definition of subjective probability. *Annals of mathematical statistics*, 34(1):199–205, 1963.
- K. J. Arrow. Alternative approaches to the theory of choice in risk-taking situations. *Econometrica: Journal of the Econometric Society*, pages 404–437, 1951.
- M. Avellaneda, A. Levy, and A. Parás. Pricing and hedging derivative securities in markets with uncertain volatilities. *Applied Mathematical Finance*, 2(2):73–88, 1995.
- A. G. Balter and A. Pelsser. Pricing and hedging in incomplete markets with model uncertainty. *European Journal of Operational Research*, 282(3):911–925, 2020.
- D. Bartl, S. Drapeau, J. Obłój, and J. Wiesel. Robust uncertainty sensitivity analysis. *arXiv preprint arXiv:2006.12022*, 2020.
- E. Bayraktar, A. Cosso, and H. Pham. Robust feedback switching control: dynamic programming and viscosity solutions. *SIAM Journal on Control and Optimization*, 54(5):2594–2628, 2016.

- M. Beiglböck, P. Henry-Labordère, and F. Penkner. Model-independent bounds for option prices—a mass transport approach. *Finance and Stochastics*, 17(3):477–501, 2013.
- S. Biagini and M. Ç. Pınar. The robust merton problem of an ambiguity averse investor. *Mathematics and Financial Economics*, 11(1):1–24, 2017.
- T. R. Bielecki, T. Chen, I. Cialenco, A. Cousin, and M. Jeanblanc. Adaptive robust control under model uncertainty. *SIAM Journal on Control and Optimization*, 57(2):925–946, 2019.
- J. Bigot and T. Klein. Characterization of barycenters in the Wasserstein space by averaging optimal transport maps. *ESAIM: Probability and Statistics*, 22:35–57, 2018.
- B. Bouchard and M. Nutz. Arbitrage and duality in nondominated discrete-time models. *The Annals of Applied Probability*, 25:823–859, 2015.
- M. Burzoni, M. Frittelli, Z. Hou, M. Maggis, and J. Oblój. Pointwise arbitrage pricing theory in discrete time. *Mathematics of Operations Research*, 44(3):1034–1057, 2019.
- G. C. Calafiore. Ambiguous risk measures and optimal robust portfolios. *SIAM Journal on Optimization*, 18(3):853–877, 2007.
- L. Carassus, J. Oblój, and J. Wiesel. The robust superreplication problem: a dynamic approach. *SIAM Journal on Financial Mathematics*, 10(4):907–941, 2019.
- Z. Chen and L. Epstein. Ambiguity, risk, and asset returns in continuous time. *Econometrica*, 70(4):1403–1443, 2002.
- P. Cheridito and T. Li. Risk measures on orlicz hearts. *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics*, 19(2):189–214, 2009.
- S. N. Cohen. Data-driven nonlinear expectations for statistical uncertainty in decisions. *Electronic Journal of Statistics*, 11(1):1858–1889, 2017.
- R. Cont. Model uncertainty and its impact on the pricing of derivative instruments. *Mathematical finance*, 16(3):519–547, 2006.
- R. Cont, R. Deguest, and G. Scandolo. Robustness and sensitivity analysis of risk measurement procedures. *Quantitative finance*, 10(6):593–606, 2010.
- A. Cox and J. Oblój. Robust pricing and hedging of double no-touch options. *Finance and Stochastics*, 15(3):573–605, 2011.
- J. Cvitanić and I. Karatzas. Hedging and portfolio optimization under transaction costs: a martingale approach. *Mathematical finance*, 6(2):133–165, 1996.
- C. Czichowsky, W. Schachermayer, et al. Duality theory for portfolio optimisation under transaction costs. *Annals of Applied Probability*, 26(3):1888–1941, 2016.
- M. H. Davis. Option pricing in incomplete markets. In M. A. H. Dempster and S. R. Pliska, editors, *Mathematics of derivative securities*, pages 216–226. Cambridge University Press, 1997.
- L. Denis and M. Kervarec. Utility functions and optimal investment in non-dominated models. *SIAM Journal on control and optimization*, 51(3):1803–1822, 2013.
- D. Ellsberg. Risk, ambiguity, and the savage axioms. *The quarterly journal of economics*, pages 643–669, 1961.
- P. M. Esfahani and D. Kuhn. Data-driven distributionally robust optimization using the wasserstein metric: Performance guarantees and tractable reformulations.



- Mathematical Programming*, 171(1):115–166, 2018.
- H. Föllmer and M. Schweizer. Hedging of contingent claims under incomplete information. *Applied stochastic analysis*, 5(389–414):19–31, 1991.
- H. Föllmer, A. Schied, , and S. Weber. Robust preferences and robust portfolio choice. *Handbook of Numerical Analysis*, 15:29–87, 2009.
- N. Fournier and A. Guillin. On the rate of convergence in Wasserstein distance of the empirical measure. *Probability Theory and Related Fields*, 162(3–4):707–738, 2015.
- M. Frittelli. The minimal entropy martingale measure and the valuation problem in incomplete markets. *Mathematical Finance*, 10(1):39–52, 2000.
- M. Frittelli and E. R. Gianin. Putting order in risk measures. *Journal of Banking & Finance*, 26(7):1473–1486, 2002.
- A. Galichon, P. Henry-Labordère, and N. Touzi. A stochastic control approach to no-arbitrage bounds given marginals, with an application to lookback options. *The Annals of Applied Probability*, 24(1):312–336, 2014.
- I. Gilboa and D. Schmeidler. Maxmin expected utility with non-unique prior. *Journal of mathematical economics*, 18(2):141–153, 1989.
- L. Hansen and T. J. Sargent. Robust control and model uncertainty. *American Economic Review*, 91(2):60–66, 2001.
- L. P. Hansen and M. Marinacci. Ambiguity aversion and model misspecification: An economic perspective. *Statistical Science*, 31(4):511–515, 2016.
- V. Henderson and D. Hobson. Utility indifference pricing: An overview. In R. Carmona, editor, *Indifference pricing: theory and applications*. Princeton University Press, Princeton, NJ, USA, 2004.
- D. Hernández-Hernández and A. Schied. Robust utility maximization in a stochastic factor model. *Statistics & Risk Modeling*, 24(1):109–125, 2006.
- D. Hobson. Robust hedging of the lookback option. *Finance and Stochastics*, 2(4):329–347, 1998.
- J. A. Hoeting, D. Madigan, A. E. Raftery, and C. T. Volinsky. Bayesian model averaging: a tutorial. *Statistical science*, pages 382–401, 1999.
- Z. Hou and J. Oblój. Robust pricing–hedging dualities in continuous time. *Finance and Stochastics*, 22(3):511–567, 2018.
- J. Hugonnier, D. Kramkov, and W. Schachermayer. On utility-based pricing of contingent claims in incomplete markets. *Math. Finance*, 15(2):203–212, 2005.
- S. Källblad, J. Oblój, and T. Zariphopoulou. Dynamically consistent investment under model uncertainty: the robust forward criteria. *Finance and Stochastics*, 22(4):879–918, 2018.
- J. Kallsen. Utility-based derivative pricing in incomplete markets. In *Mathematical Finance—Bachelier Congress 2000*, pages 313–338. Springer, 2002.
- J. Kallsen, J. Muhle-Karbe, et al. On using shadow prices in portfolio optimization with transaction costs. *Annals of Applied Probability*, 20(4):1341–1358, 2010.
- I. Karatzas and S. G. Kou. On the pricing of contingent claims under constraints. *Ann. Appl. Prob.*, pages 321–369, 1996.
- I. Karatzas and X. Zhao. Bayesian adaptive portfolio optimization. *Option pricing, interest rates and risk management*, pages 632–669, 2001.
- J. Kerkhof, B. Melenberg, and H. Schumacher. Model risk and capital reserves. *Journal of Banking & Finance*, 34(1):267–279, 2010.
- J. M. Keynes. *A treatise on probability*. Macmillan and Company, limited, 1921.

- F. Knight. *Risk, uncertainty and profit*. Boston: Houghton Mifflin, 1921.
- D. Kramkov and W. Schachermayer. The asymptotic elasticity of utility functions and optimal investment in incomplete markets. *Annals of Applied Probability*, pages 904–950, 1999.
- H. Lam. Robust sensitivity analysis for stochastic systems. *Mathematics of Operations Research*, 41(4):1248–1275, 2016.
- T. J. Lyons. Uncertain volatility and the risk-free synthesis of derivatives. *Applied mathematical finance*, 2(2):117–133, 1995.
- F. Maccheroni, M. Marinacci, and A. Rustichini. Ambiguity aversion, robustness, and the variational representation of preferences. *Econometrica*, 74(6):1447–1498, 2006.
- P. J. Maenhout. Robust portfolio rules and asset pricing. *Review of Financial Studies*, 17(4):951–983, 2004.
- H. Markowitz. Portfolio selection, 1959.
- R. C. Merton. Lifetime portfolio selection under uncertainty: The continuous-time case. *The review of Economics and Statistics*, pages 247–257, 1969.
- P. Milgrom and I. Segal. Envelope theorems for arbitrary choice sets. *Econometrica*, 70(2):583–601, 2002.
- M. Musiela and T. Zariphopoulou. The single period binomial model. In R. Carmona, editor, *Indifference pricing: theory and applications*. Princeton University Press, 2008.
- A. Neufeld and M. Nutz. Robust utility maximization with lévy processes. *Mathematical Finance*, 28(1):82–105, 2018.
- M. Nutz. Utility maximization under model uncertainty in discrete time. *Mathematical Finance*, 2014.
- J. Oblój and J. Wiesel. Supplement to “Distributionally robust portfolio maximization and marginal utility pricing in one period financial markets”. 2021.
- G. Peyré, M. Cuturi, et al. Computational optimal transport: With applications to data science. *Foundations and Trends® in Machine Learning*, 11(5-6):355–607, 2019.
- A. Pratelli. On the equality between Monge’s infimum and Kantorovich’s minimum in optimal mass transportation. In *Annales de l’Institut Henri Poincaré (B) Probability and Statistics*, volume 43, pages 1–13. Elsevier, 2007.
- M. Rásonyi and L. Stettner. On utility maximization in discrete-time financial market models. *Ann. Appl. Prob.*, 15(2):1367–1395, 2005.
- L. C. Rogers. Equivalent martingale measures and no-arbitrage. *Stochastics*, 51(1-2):41–49, 1994.
- L. C. G. Rogers. The relaxed investor and parameter uncertainty. *Finance and Stochastics*, 5(2):131–154, 2001.
- R. Rouge and N. El Karoui. Pricing via utility maximization and entropy. *Mathematical Finance*, 10(2):259–276, 2000.
- L. J. Savage. The theory of statistical decision. *Journal of the American Statistical association*, 46(253):55–67, 1951.
- M. Schäl. Portfolio optimization and martingale measures. *Math. Finance*, 10(2):289–303, 2000a.
- M. Schäl. Price systems constructed by optimal dynamic portfolios. *Math. Methods Oper. Res.*, 51(3):375–397, 2000b.

- M. Schäl. Markov decision processes in finance and dynamic options. In *Handbook of Markov decision processes*, pages 461–487. Springer, 2002.
- A. Schied. Optimal investments for risk-and ambiguity-averse preferences: a duality approach. *Finance and Stochastics*, 11(1):107–129, 2007.
- A. Schied and C.-T. Wu. Duality theory for optimal investments under model uncertainty. *Statistics & Decisions*, 23(3/2005):199–217, 2005.
- D. Schmeidler. Subjective probability and expected utility without additivity. *Econometrica: Journal of the Econometric Society*, pages 571–587, 1989.
- M. Schweizer. Approximation pricing and the variance-optimal martingale measure. *The Annals of Probability*, 24(1):206–236, 1996. ISSN 0091-1798. doi: 10.1214/aop/1042644714.
- M. Sirbu et al. A note on the strong formulation of stochastic control problems with model uncertainty. *Electronic Communications in Probability*, 19, 2014.
- K.-T. Sturm. Generalized orlicz spaces and wasserstein distances for convex–concave scale functions. *Bulletin des sciences mathématiques*, 135(6-7):795–802, 2011.
- P. Swoboda and C. Schnorr. Convex variational image restoration with histogram priors. *SIAM Journal on Imaging Sciences*, 6(3):1719–1735, 2013.
- G. Tartavel, G. Peyré, and Y. Gousseau. Wasserstein loss for image synthesis and restoration. *SIAM Journal on Imaging Sciences*, 9(4):1726–1755, 2016.
- R. Uppal and T. Wang. Model misspecification and underdiversification. *The Journal of Finance*, 58(6):2465–2486, 2003.
- J. Von Neumann and O. Morgenstern. *Theory of games and economic behavior*. Princeton University Press, 1953.

JAN OBLÓJ

MATHEMATICAL INSTITUTE AND ST JOHN’S COLLEGE, UNIVERSITY OF OXFORD  
WOODSTOCK ROAD  
OXFORD, OX2 6GG

JOHANNES WIESEL

DEPARTMENT OF STATISTICS, COLUMBIA UNIVERSITY  
1255 AMSTERDAM AVENUE  
NEW YORK, NY 10027