

Mixtures of Mean-Preserving Contractions*

Joseph Whitmeyer[†] Mark Whitmeyer[‡]

September 20, 2020

Abstract

Given any purely atomic probability distribution with support on n points, P , any mean-preserving contraction (mpc) of P , Q , with support on $m > n$ points is a mixture of mpcs of P , each with support on at most n points. We illustrate several applications of this result to Bayesian persuasion and information design.

Keywords: Mean-Preserving Contraction, Information Design, Bayesian Persuasion, Fusion of a Probability Distribution

JEL Classifications: C72; D82; D83

1 Introduction

The mean-preserving contraction (mpc) of a probability distribution, in which a probability distribution is altered by collapsing portions of its measure to their respective barycenters, is an important concept in economics. In recent years there has been a surge of papers in which

*We thank Ying Chen and two anonymous referees for their useful suggestions. We are also grateful to Gill Grindstaff and Jan Schlupp for their help and advice. Mark Whitmeyer's work was generously funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy-GZ 2047/1, Projekt-ID 390685813.

[†]Department of Sociology, University of North Carolina at Charlotte

[‡]Institute for Microeconomics and Hausdorff Center for Mathematics, University of Bonn

Email: mark.whitmeyer@gmail.com

information structures are endogenous, and many works involve optimization problems in which one or many agents choose from the set of mpcs of a given distribution.

The purpose of this paper is to establish a result that is useful both in clarifying the economic intuition behind the solutions to, as well as simplifying the solving of, such problems. Our main finding, Theorem 1.4, establishes that given a purely atomic probability measure P on \mathbb{R} with support on n points, any mpc of P with support on $n + 1$ points, Q , is the convex combination of two mpcs of P , Q' and Q'' , each of which have support on at most n points. An important corollary of this result, Corollary 1.6, is that any mpc, Q , of P is a mixture of mpcs with support on at most n points.

These discoveries imply several compelling results related to Bayesian persuasion and information design,¹ a literature that studies how a principal (or principals) can design information structures in various environments in order to achieve some objective. One particularly tractable class of problems are those in which the state is a random variable and the sender and receiver's utilities are linear in the state (see e.g. Gentzkow and Kamenica (2016); Kolotilin et al. (2017); Kolotilin (2018); and Dworzak and Martini (2019)). Corollary 1.6 allows for an elementary proof for the upper bound of the number of messages needed by the persuader in an optimal mechanism in such problems (Proposition 2.1).

Corollary 1.6 also has useful ramifications for competitive Bayesian persuasion problems, in which multiple principals compete by designing information structures. Our result implies that when constructing equilibria, instead of checking deviations to *any* mpc of P , one need check only deviations to mpcs with support on n points. In Subsection 2.2 we present a simple competitive persuasion problem in which this result is used, and Whitmeyer (2018) and Jain and Whitmeyer (2019) both appeal to Corollary 1.6 in similar settings to this example. The existence of a symmetric mixed strategy equilibrium with support on n (or fewer)-point mpcs (Proposition 2.2) also follows from this corollary.

This paper contributes to the literature in economics and mathematics on statistical ex-

¹The number of papers in this area has grown rapidly in the decade following the writing of Rayo and Segal (2010) and Kamenica and Gentzkow (2011), the two papers that first sparked the widespread interest in the topic. Bergemann and Morris (2019) provides a recent and relatively self-contained overview of the topic and Kamenica (2019) provides a comprehensive survey of the literature.

periments and majorization. The first foray into the area is the seminal work of Hardy et al. (1929) (see also Hardy et al. (1959)), who establish a fundamental result on majorization, which was shortly followed and generalized by Blackwell (1953) and Strassen (1965).² The equivalent notion of a mean-preserving spread (mps)—the opposite of an mpc in which portions of a probability measure are “spread” out—was subsequently introduced by Rothschild and Stiglitz (1970).

More recently, Elton and Hill (1992) introduce the concept of a “fusion” of a probability distribution. A fusion, Q , of a distribution P is any distribution that can be obtained by collapsing parts of the mass of P to their respective barycenters. In the setting we analyze in this paper, where P is a measure on \mathbb{R} (with finite first moment), the notions of fusion and mean-preserving contraction are equivalent. A key contribution of Elton and Hill (1992) is the construction of the set of fusions: just as a measurable function can be approximated by simple functions, which are themselves finite linear combinations of indicator functions, a fusion can be approximated by simple fusions, which are themselves finite compositions of elementary fusions.³

In this spirit, the contribution of our paper can be seen as it providing an alternate characterization for the set of fusions of a (finitely supported) probability measure P (on \mathbb{R}), one that we believe should be particularly compelling (and useful) to economists. Rather than taking elementary fusions as the building blocks, as Elton and Hill (1992) do, our basic objects are the fusions of P with support on n points. Consequently, Corollary 1.6 implies that any fusion can be approximated by finite *mixtures* of fusions with n -point support.

As we discuss later on in the paper, the set of all mpcs of a discrete probability measure is a compact and convex set. By the Krein-Milman theorem, such a set is merely the closed and convex hull of its extreme points. This is valuable in well-behaved linear optimization problems, in which the objective function obtains a maximum at an extreme point of the constraint set. This theorem (Krein-Milman) and its extension, Choquet’s theorem, provide

²Another related paper is Elton and Hill (1998), who provide a constructive proof of Hardy et al. (1929)’s main result and show how it can be generalized to infinite-dimensional spaces.

³In their parlance, an elementary fusion is just the result of collapsing part of a single (Borel) subset of the probability measure to the barycenter of the subset.

part of the impetus behind the remarkable recent paper by Kleiner et al. (2020), who characterize the extreme points of the set of monotonic functions that majorize (or are majorized by) a given monotonic function such as, e.g., a cumulative distribution function (cdf). They thus characterize the set of mpcs and mpss for an arbitrary distribution.

We believe that our paper can be seen as complementary to Kleiner et al. (2020). Indeed their main result in which they characterize the extreme points of the set of mpcs of a random variable with cdf F —Theorem 2—requires that F be continuous. In turn, Corollary 1.6 in this paper provides a necessary condition for a distribution to be an extreme point of the set of mpcs of a distribution with n -point support—it must have support on n points or fewer.

1.1 The Main Result

Throughout, bold font identifies vectors. Let $X = \mathbb{R}$ and denote by \mathcal{B} the Borel subsets of X . Moreover, let \mathcal{P} denote the set of Borel probability measures on (X, \mathcal{B}) . Let $P \in \mathcal{P}$ be a purely atomic probability measure with support on n points i.e. $\text{supp}P = \mathbf{a} := \{a_1, a_2, \dots, a_n\}$, with respective masses p_1, p_2, \dots, p_n , where $p_i > 0$ for all $i = 1, \dots, n$ and $\sum_i^n p_i = 1$. Denote $\mathbf{p} := (p_1, p_2, \dots, p_n)$. Without loss of generality, $a_1 < a_2 < \dots < a_n$. Let $Q \in \mathcal{P}$ be a purely atomic probability measure with support on m points: $\text{supp}Q = \mathbf{b} := \{b_1, b_2, \dots, b_m\}$, with respective masses q_1, q_2, \dots, q_m , where $q_i > 0$ for all $i = 1, \dots, m$ and $\sum_i^m q_i = 1$. Denote $\mathbf{q} := (q_1, q_2, \dots, q_m)$. Finally, denote $\mathbf{pa} := (p_1 a_1, p_2 a_2, \dots, p_n a_n)$ and analogously for \mathbf{qb} .

We adopt the following definition from Elton and Hill (1992) and Elton and Hill (1998) for simple fusions:⁴

Definition 1.1. Q is a **Simple Mean-Preserving Contraction (SMPC)** of P if there exists a non-negative row-stochastic (Markov) $n \times m$ matrix F that satisfies

$$\mathbf{p}F = \mathbf{q}$$

and

$$(\mathbf{pa})F = \mathbf{qb}$$

Denote by $m(P)$ the set of all smpcs of P .

⁴Recall that the notions of a fusion and an mpc of a probability measure P are equivalent.

Consequently, an smpc with $n + 1$ atoms, \mathbf{b} , of an arbitrary n -atom probability measure P with atoms, \mathbf{a} , can be defined by the $n \times (n + 1)$ Markov matrix F . Each row of F , F_i , sums to 1 and denotes the partition of the weight P_i at atom a_i across the atoms \mathbf{b} .

To better grasp the concept (and notation) consider the following example:

Example 1.2. Let P be a discrete probability distribution with support on three points:

$$P = \begin{Bmatrix} 0 & \frac{1}{2} & 1 \\ \frac{3}{10} & \frac{3}{10} & \frac{2}{5} \end{Bmatrix}$$

where the top row is the vector of support, and the bottom row is the vector of associated probability weights. Then, Q is a smpc of P , with an associated Markov matrix, F :

$$Q = \begin{Bmatrix} \frac{1}{6} & \frac{1}{2} & \frac{3}{4} & \frac{5}{6} \\ \frac{3}{10} & \frac{1}{5} & \frac{1}{5} & \frac{3}{10} \end{Bmatrix} \quad \& \quad F = \begin{pmatrix} \frac{2}{3} & \frac{1}{3} & 0 & 0 \\ \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{pmatrix}$$

After Elton and Hill (1992), we construct the set of mean-preserving contractions of P , $\mathcal{M}(P)$, by taking the weak- $*$ closure of $m(P)$:

Definition 1.3. $R \in \mathcal{D}$ is a **Mean-Preserving Contraction (MPC)** of P if there exists a sequence of smpcs $\{Q_m\}_{m=1}^{\infty} \subset m(P)$ that satisfies $Q_m \rightarrow_w R$.⁵ Denote by $\mathcal{M}(P)$ the set of all mpcs of P .

Next, we say that that $Q = \alpha Q' + (1 - \alpha)Q''$ for smpcs Q , Q' , and Q'' if and only if

$$F = \alpha F' + (1 - \alpha)F'' \tag{1}$$

for Markov matrices F , F' , and F'' , where 0-columns are inserted into the Markov matrices corresponding to the zero-probability atoms in Q' and Q'' but otherwise the ratios within columns in F' and F'' are the same as in F . Using this, we may now state the main theorem:

Theorem 1.4. *Let $Q \in m(P)$ be any smpc of P with support on $n + 1$ points, $n \geq 2$. Then Q is the convex combination of two purely atomic probability measures Q' and Q'' ; $Q', Q'' \in m(P)$, each with support on at most n points. Q' and Q'' are unique.*

⁵ \rightarrow_w denotes the standard idea of weak convergence.

The essence of the proof is as follows: consider the Markov matrix F corresponding to a smpc Q where $m = n + 1$. Because F is an $n \times (n + 1)$ matrix, there is some subset of column vectors that is linearly dependent, and because all atoms are distinct, this subset has a minimal size of three. Hence, any of these column vectors can be expressed as a linear sum of the others. This means we can write a new matrix with any of these column vectors redistributed to the other vectors, such that the horizontal sum across vectors remains the same (1 for the whole matrix), the chosen vector is reduced to the 0-vector, and each other column vector is either the 0-vector or is a multiple of its original self, i.e., preserves its internal ratios. We refer to this procedure as “zeroing” a vector of the matrix.

Thus, in the proof it suffices to show that we can always choose two different column vectors from the linearly dependent set such that zeroing either of them gives entries in the remaining vectors from 0 to 1. We obtain the uniqueness of Q' and Q'' by showing that there is no third such vector that can be zeroed independently of the other two, that is, without also zeroing one of the other two vectors.

Proof. Let $\mathbf{0}$ and $\mathbf{1}$ denote the column vectors of all zeroes and ones, respectively. Without loss of generality, suppose the linearly dependent set consists of the entire $n \times n + 1$ matrix F , that is, $n + 1$ column vectors \mathbf{f}_j , where $\mathbf{0} \leq \mathbf{f}_j \leq \mathbf{1}$ and $\sum_{j=1}^{n+1} f_{ij} = 1$ for all i or $\sum_{j=1}^{n+1} \mathbf{f}_j = \mathbf{1}$, the component-wise sum of the vectors \mathbf{f}_j . Then there are scalars \hat{c}_j , not all 0, such that $\sum_{j=1}^{n+1} \hat{c}_j \mathbf{f}_j = \mathbf{0}$. Because the vectors are non-negative, some of the \hat{c}_j will be positive and some negative. Re-index the vectors so that

$$\sum_{j=1}^k -|\hat{c}_j| \mathbf{f}_j + \sum_{j=k+1}^{n+1} \hat{c}_j \mathbf{f}_j = \mathbf{0}$$

for some $k \leq n$. Set $c_j = |\hat{c}_j|$ for all j . Then,

$$\sum_{j=1}^k c_j \mathbf{f}_j = \sum_{j=k+1}^{n+1} c_j \mathbf{f}_j$$

where $c_j \geq 0$ for all j . Define

$$c_{j^*} := \max_{1 \leq j \leq k} \{c_j\}, \quad \& \quad c_{j^{**}} := \max_{k+1 \leq j \leq n+1} \{c_j\}$$

Then, the corresponding \mathbf{f}_{j^*} and \mathbf{f}_{j^*} can be zeroed. To see that take one, say \mathbf{f}_{j^*} :

$$\mathbf{f}_{j^*} = - \sum_{\substack{j \leq k \\ j \neq j^*}} \frac{c_j}{c_{j^*}} \mathbf{f}_j + \sum_{j \geq k+1} \frac{c_j}{c_{j^*}} \mathbf{f}_j$$

In the new Markov matrix for the new mpc, set

$$\mathbf{f}'_j = \begin{cases} \left(1 - \frac{c_j}{c_{j^*}}\right) \mathbf{f}_j, & j \leq k, j \neq j^* \\ \left(1 + \frac{c_j}{c_{j^*}}\right) \mathbf{f}_j, & j \geq k+1 \\ \mathbf{0}, & j = j^* \end{cases}$$

We have produced matrix F' by zeroing vector \mathbf{f}_{j^*} . By construction, for $j \leq k, j \neq j^*$:

$$0 \leq \frac{c_j}{c_{j^*}} \leq 1$$

and so

$$0 \leq \left(1 - \frac{c_j}{c_{j^*}}\right) f_{ij} \leq 1, \quad \text{for all } i, j$$

Also,

$$\mathbf{0} \leq \sum_{\substack{j \leq k \\ j \neq j^*}} \left(1 - \frac{c_j}{c_{j^*}}\right) \mathbf{f}_j < \sum_{\substack{j \leq k \\ j \neq j^*}} \mathbf{f}_j < \mathbf{1}$$

Then

$$\mathbf{1} = \sum_{j=1}^{n+1} \mathbf{f}_j = \sum_{j=1}^{n+1} \mathbf{f}'_j = \sum_{\substack{j \leq k \\ j \neq j^*}} \left(1 - \frac{c_j}{c_{j^*}}\right) \mathbf{f}_j + \sum_{j \geq k+1} \left(1 + \frac{c_j}{c_{j^*}}\right) \mathbf{f}_j$$

This means

$$\sum_{j \geq k+1} \left(1 + \frac{c_j}{c_{j^*}}\right) \mathbf{f}_j \leq \mathbf{1}$$

i.e. for all i ,

$$\sum_{j \geq k+1} \left(1 + \frac{c_j}{c_{j^*}}\right) f_{ij} \leq 1$$

and since

$$\left(1 + \frac{c_j}{c_{j^*}}\right) \geq 1$$

for all j , we have

$$0 \leq \left(1 + \frac{c_j}{c_{j^*}}\right) f_{ij} \leq 1$$

As a result, for the two zeroed column vectors we create the two distinct $n \times n + 1$ Markov matrixes, F' and F'' , containing different zero vectors, required for Equation 1. From the linear dependence the coefficient α in Equation 1 can be calculated. Q' and Q'' are given, of course, by the partitions obtained by removing the zero column vectors from F' and F'' .

To see that F' and F'' , and therefore Q' and Q'' , are unique, suppose for the sake of contradiction that independently of \mathbf{f}_j and \mathbf{f}_{j^*} there is a third \mathbf{f}_{j^\dagger} that can be zeroed, where $j^* \neq j^\dagger \neq j^{**}$. Without loss of generality suppose $1 \leq j^\dagger \leq k$. Then by definition $c_{j^\dagger} \leq c_{j^*}$. Without loss of generality let $c_{j^\dagger} < c_{j^*}$ (in the case of an equality zeroing \mathbf{f}_j would also zero \mathbf{f}_{j^\dagger} , giving F' and F'' again). Thus,

$$\mathbf{f}_{j^\dagger} = - \sum_{\substack{j \leq k \\ j \neq j^\dagger}} \frac{c_j}{c_{j^\dagger}} \mathbf{f}_j + \sum_{j \geq k+1} \frac{c_j}{c_{j^\dagger}} \mathbf{f}_j$$

Then, because

$$\frac{c_{j^*}}{c_{j^\dagger}} > 1$$

the other zeroing produces the following component,

$$f''_{ij^*} = \left(1 - \frac{c_{j^*}}{c_{j^\dagger}}\right) f_{ij^*} < 0$$

which is not permitted. ■

Next, let us revisit the distribution P and smpc Q from Example 1.2 and the two smpcs, Q' and Q'' , whose convex combination yields Q :

Example 1.5. Recall P , Q , and F from Example 1.2. Then,

$$F' = \begin{pmatrix} \frac{5}{6} & \frac{1}{6} & 0 & 0 \\ \frac{5}{12} & 0 & 0 & \frac{7}{12} \\ 0 & \frac{1}{8} & 0 & \frac{7}{8} \end{pmatrix}, \quad \& \quad F'' = \begin{pmatrix} \frac{4}{9} & \frac{5}{9} & 0 & 0 \\ \frac{2}{9} & 0 & \frac{7}{9} & 0 \\ 0 & \frac{5}{12} & \frac{7}{12} & 0 \end{pmatrix}$$

Thus,

$$Q' = \begin{Bmatrix} \frac{1}{6} & \frac{1}{2} & \frac{5}{6} \\ \frac{3}{8} & \frac{1}{10} & \frac{21}{40} \end{Bmatrix}, \quad \& \quad Q'' = \begin{Bmatrix} \frac{1}{6} & \frac{1}{2} & \frac{3}{64} \\ \frac{1}{5} & \frac{1}{3} & \frac{7}{15} \end{Bmatrix}$$

and $\alpha = 4/7$.

Corollary 1.6. *Any mpc $Q \in \mathcal{M}(P)$ is a mixture of smpcs with support on at most n points.*

Proof. From Theorem 1.4, given a purely atomic probability measure P with support on n points, any smpc $Q \in \mathcal{M}(P)$ with support on $n + 1$ points is a convex combination of two smpcs $Q', Q'' \in \mathcal{M}(P)$ with support on at most n points. If $m = n + 1$ then the result is simply Theorem 1.4. If not, simply iterate backward until the desired mixture of smpcs of P with support on n points is obtained.

Taking weak limits, this result holds for mpcs of P that are not purely atomic. ■

2 Applications

Here we illustrate the usefulness of Corollary 1.6 in Bayesian persuasion problems. The first subsection illustrates how this result yields easily an upper bound for the number of messages required for the optimal mechanism in a (linear) persuasion problem. The second subsection contains a pair of results related to competitive persuasion. There, we observe that Corollary 1.6 simplifies the task of solving for an equilibrium and aids us in the construction of mixed strategy equilibria in such games.

2.1 A Bound for Linear Persuasion Problems

One particularly tractable class of persuasion problems are those in which the state is a real-valued random variable and the sender's and receiver's optimal actions are linear in the state, and hence depend only on the expected state (the posterior mean). It is a standard result (see, e.g., Dworzak and Martini (2019)) that this allows us to simplify the persuader's problem to one of choosing any distribution of posterior means that is an mpc of the prior (or equivalently, any cdf over values that second-order stochastically dominates the prior).

In such an environment, given prior P , the sender's persuasion problem is thus

$$\max_Q E_Q[u(x)], \quad \text{subject to } Q \in \mathcal{M}(P)$$

where $u(x)$ is the *ex-post* utility of the sender from inducing posterior mean x .

Kamenica and Gentzkow (2011) use the Fenchel-Bunt extension (see e.g. Theorem 1.3.7 of Hiriart-Urruty and Lemaréchal (2001)) of Carathéodory's theorem to show that an optimal

signal in an n -state persuasion problem requires at most n signal realizations. Our results above allow us to derive an analog of this result for linear persuasion problems with ease.

By Elton and Hill (1992) the set $\mathcal{M}(P)$ is convex and closed. By the Riesz Representation theorem, the space of all probability measures with the same (compact) support is compact in the weak-* topology. Since $\mathcal{M}(P)$ is a closed subset of this space it must also be weak-* compact and therefore Hausdorff.

Bauer's Maximum Principle (see e.g. Corollary 7.70 of Aliprantis and Border (2006)) implies that a linear functional such as the expectation operator on a (non-empty) compact and convex set A of a locally-convex Hausdorff space attains its maximum at an extreme point of A . By Corollary 1.6 a necessary condition for Q to be an extreme point of $\mathcal{M}(P)$ is that it have support on at most n points. Thus, we have proved the following result.

Proposition 2.1. *In linear persuasion problems in which the prior, P , has support on n points, the optimal signal requires at most n messages to be used.*

2.2 Simplifying Competitive Persuasion

Now consider a simple competitive persuasion problem as studied by Au and Kawai (2020). There are k competing sellers who want a single risk-neutral buyer to purchase their good. The sellers sell products with quality given by an i.i.d. random variable X_i , $i = 1, \dots, k$, distributed according to the purely atomic measure P .

Each seller, without knowing the quality of her (or the other sellers') goods, simultaneously chooses a Blackwell experiment conditioned on her type. The buyer observes the realization of the k experiments and selects the seller whose product's posterior expected value is highest (and randomizes fairly when indifferent). A seller gets a payoff normalized to one if the buyer selects her and zero otherwise. Because it is only the expected value of the product that matters to the buyer (and because the sellers' payoffs are state-independent), this is a linear persuasion problem and so again each seller's problem is to choose a distribution of posterior means that is an mpc of the prior. That is, each seller's set of pure strategies is $\mathcal{M}(P)$.

As Au and Kawai (2020) show, the unique symmetric equilibrium in pure strategies re-

quires an uncountable number of signal realizations—each seller’s equilibrium distribution, Q , is continuous except possibly at the upper bound of its support. However, such an equilibrium may be infeasible practically, either because the firms are unable to construct a signal of the necessary complexity or due to cognitive constraints suffered by the citizens. Corollary 1.6 allows us to sidestep this issue:

Proposition 2.2. *For the k player competitive persuasion game in which the prior distribution over qualities has support on n points, there exists a symmetric mixed strategy equilibrium supported on mpcs with support on at most n points.*

Proof. Theorem 4 in Au and Kawai (2020) establishes that a symmetric pure strategy equilibrium exists. Let Q^* denote an equilibrium pure strategy distribution over (expected) qualities. By Corollary 1.6, Q^* can be obtained as a mixture of mpcs, each with support on at most n points. Since this mixture results in the same distribution over qualities and the same payoff to each player as the pure strategy, Q^* , there must also be an equilibrium of the game in which each player mixes. ■

As the next example illustrates, Corollary 1.6 can also be used to easily establish whether candidate pairs of strategies are equilibria.

Example 2.3. Consider the two player competitive persuasion problem with

$$P = \begin{Bmatrix} 0 & \frac{1}{2} & \frac{3}{4} \\ \frac{1}{6} & \frac{1}{2} & \frac{1}{3} \end{Bmatrix}$$

We claim that there is a Nash Equilibrium in which both sellers choose the cdf

$$F(x) = \begin{cases} \frac{2}{3}x, & 0 \leq x \leq \frac{1}{2} \\ \frac{8}{3}x - 1, & \frac{1}{2} \leq x \leq \frac{3}{4} \end{cases}$$

From Corollary 1.6 it suffices to check that there is no profitable deviation to an mpc with support on three points. To that end, suppose that seller 2 chooses such a distribution, distribution Q :

$$Q = \begin{Bmatrix} a & b & c \\ p & q & r \end{Bmatrix}$$

where $Q \in \mathcal{M}(P)$, $pa + qb + rc = 1/2$, and $p + q + r = 1$. Without loss of generality $a \leq 1/2$ and $c \geq 1/2$. First suppose that $b \leq 1/2$. Then, seller 2's payoff from deviating is

$$u_2 = pF(a) + qF(b) + rF(c) = \frac{1}{3} + r(2c - 1)$$

Because $c \leq 1/2 + 1/(12r)$,

$$u_2 \leq \frac{1}{3} + r \left(1 + \frac{1}{6r} - 1 \right) = \frac{1}{2}$$

In a similar manner, for $b \geq 1/2$ we have

$$u_2 = pF(a) + qF(b) + rF(c) = \frac{1}{3} + p(1 - 2a)$$

where we use the fact that $p + q + r = 1$ and that $pa + qb + rc = 1/2$. Then, since $a \geq (1 - 1/(6p))/2$

$$u_2 \leq \frac{1}{3} + p \left(1 + \frac{1}{6p} - 1 \right) = \frac{1}{2}$$

Thus, there is no profitable deviation to any mpc with three point support and so no profitable deviation to any mpc.

References

- Charalambos D. Aliprantis and Kim Border. *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Springer, 2006.
- Pak Hung Au and Keiichi Kawai. Competitive information disclosure by multiple senders. *Games and Economic Behavior*, 119:56 – 78, 2020.
- Dirk Bergemann and Stephen Morris. Information design: A unified perspective. *Journal of Economic Literature*, 57(1):44–95, March 2019.
- David Blackwell. Equivalent comparisons of experiments. *The Annals of Mathematical Statistics*, 24(2):265–272, 1953.
- Piotr Dworzak and Giorgio Martini. The simple economics of optimal persuasion. *Journal of Political Economy*, 127(5):1993–2048, 2019.

- J. Elton and T. P. Hill. Fusions of a probability distribution. *The Annals of Probability*, 20(1): 421–454, 1992.
- J Elton and T.P Hill. On the basic representation theorem for convex domination of measures. *Journal of Mathematical Analysis and Applications*, 228(2):449 – 466, 1998.
- Matthew Gentzkow and Emir Kamenica. A rothschild-stiglitz approach to bayesian persuasion. *American Economic Review*, 106(5):597–601, 5 2016.
- G.H. Hardy, J.E. Littlewood, and Polya G. Some simple inequalities satisfied by convex functions. *Messenger of Mathematics*, 58:145–152, 1929.
- G.H. Hardy, J.E. Littlewood, and Polya G. *Inequalities*. Cambridge University Press, 1959.
- Jean-Baptiste Hiriart-Urruty and Claude Lemaréchal. *Fundamentals of Convex Analysis*. Springer, 2001.
- Vasudha Jain and Mark Whitmeyer. Competing to Persuade a Rationally Inattentive Agent. *ArXiv e-prints*, 04 2019.
- Emir Kamenica. Bayesian persuasion and information design. *Annual Review of Economics*, 11(1):249–272, 2019.
- Emir Kamenica and Matthew Gentzkow. Bayesian persuasion. *The American Economic Review*, 101(6):2590–2615, 2011.
- Andy Kleiner, Benny Moldovanu, and Philipp Strack. Extreme points and majorization: Economic applications. *Mimeo*, 2020.
- Anton Kolotilin. Optimal information disclosure: A linear programming approach. *Theoretical Economics*, 13(2):607–635, 2018.
- Anton Kolotilin, Tymofiy Mylovanov, Andriy Zapechelnyuk, and Ming Li. Persuasion of a privately informed receiver. *Econometrica*, 85(6):1949–1964, 2017.
- Luis Rayo and Ilya Segal. Optimal information disclosure. *Journal of Political Economy*, 118 (5):949–987, 2010.

Michael Rothschild and Joseph E Stiglitz. Increasing risk: I. a definition. *Journal of Economic Theory*, 2(3):225 – 243, 1970.

Volker Strassen. The existence of probability measures with given marginals. *Annals of Mathematical Statistics*, 36(2):423–439, 04 1965.

Mark Whitmeyer. Dynamic Competitive Persuasion. *ArXiv e-prints*, 11 2018.