7-17-2019

The Price of Independence in an Echo Chamber with Dependence Ambiguity

Lorán Chollete
Sacred Heart University, cholletel@sacredheart.edu

Michael Klass
University of California - Berkeley, klass@stat.berkeley.edu

Victor de la Peña
Columbia University, vp@stat.columbia.edu

The Price of Independence in an Echo Chamber with Dependence Ambiguity

Lorán Chollete, Michael Klass and Victor de la Peña *

July 17, 2019

Abstract

How much should we pay to remove the interdependence of biased information sources? This question is relevant in both statistics and political economy. When there are many information sources or variables, their dependence may be unknown, which creates multivariate ambiguity. One approach to answer our leading question involves use of decoupling inequalities from probability theory. We present a new inequality, designed to cope with this question, which holds for any type of dependence across information sources. We apply our method to a simple formalization of a political echo chamber. For a given set of marginal information, this bound is the sup over all possible joint distributions connecting the marginals. Our method highlights a price to pay for facing summed dependent (multivariate) data, similar to the probability premium required for univariate data. We show that a conservative decisionmaker will pay approximately 50% more than if the data were independent, in order to freely neglect the correlations.

Keywords: Multivariate Ambiguity; Decoupling Inequality; Echo Chamber; Independence; Probability Premium

*Chollete is at the Welch College of Business at Sacred Heart University, email cholletel@sacredheart.edu. de la Peña is at Columbia University, email vp@stat.columbia.edu. Klass is at UC Berkeley, email klass@stat.berkeley.edu.

1 Introduction

How much would a pollster pay to extract truly independent data from citizens that are highly correlated (e.g. from the same political party)? Statistically, this question may be posed as: what is the price of independence? In this paper we will see that the price of independence is a constant proportion, which provides a useful benchmark in some economic contexts.

In particular, in Section 3 we show that the price of independence PD may be expressed as the sup over all possible joint distributions, once you know the marginals. How much would you pay for the information on the dependence? Our main Theorem in Section 3 shows that, given a concave loss function, it is roughly 50% more than you would pay if the data were truly independent.[1]

The term dependence is used in several ways in the statistical and economics literature. In this paper, we will generally refer to dependence with Kendall's tau, denoted τ, which measures the difference between concordance and discordance.[2] We utilize τ because it measures dependence between random variables in a way that is robust to rescaling, and unlike the Pearson correlation, does not confuse the marginals with the dependence structure. Moreover, experimental evidence suggests that although individuals tend to estimate average dependence fairly well, they incorrectly estimate high or low dependence (Clemen et al. (2000)). We therefore require a measure such as τ that can apply overall and in the tails of the distribution. Our main result does not depend on the particular measure of dependence that we utilize.

Dependence in a Statistical Echo Chamber. In section 3 below, we will define a statistical echo chamber $C^k(\mathbf{d})$ as a collection of m-dimensional random variables $\mathbf{d}$, which is observed k different times. Each observation on $\mathbf{d}$ has a different level of dependence, and we order the observations in terms of their dependence $\tau(k)$, such that each observation features greater dependence than the previous observation, i.e. $\tau(k) \geq \tau(k-1)$, for $k = 1, \ldots, K$. This problem has inherent ambiguity because in general, the researcher will not know what

[1] The term 'price of independence' signifies the amount that a researcher would pay to work with data that are truly independent, see discussion in Section 3 below. Our main Theorem of Section 2 shows that PD equals $\frac{e}{e-1}$, which is approximately 1.58.

[2] τ is defined for two variables in the following manner: Consider a random vector (X_1, X_2) and an independent vector $(\tilde{X}_1, \tilde{X}_2)$ from the same distribution. Then τ is the probability of concordance minus the probability of discordance for these pairs:

$$\tau(X_1, X_2) = P((X_1 - \tilde{X}_1)(X_2 - \tilde{X}_2) > 0) - P((X_1 - \tilde{X}_1)(X_2 - \tilde{X}_2) < 0).$$

For a multivariate version of τ, see McNeil et al. (2015).

k is; hence there is ambiguity about the increase in dependence. When a decisionmaker faces ambiguity, she may worry about the worst-case scenario (Gilboa and Schmeidler (1989)), and make decisions based on that (by selecting univariate probability distributions consistent with the worst-case). We apply the theorem in this setting, the ‘worst-case scenario’ is now a dependence bound on a multivariate (unknown) distribution.

This setup resembles real world situations. One example is where the same set of assets becomes increasingly dependent, as the economy moves from a boom market to a bust market. Another example is in politics, where voters within a given political party increasingly coalesce in their opinions on a particular issue, as the voting date comes closer. We apply our result in a proposition, which allows us to evaluate conservatively, the price of information loss due to being in a political echo chamber.

Related Literature. Our paper builds on the statistical literature on decoupling, much of which is summarized in de la Peña and Giné (1999). It is also complementary to the following research in economics. Meyer and Strulovici (2012) study how to order interdependence in economic settings. They analyze five orderings of interdependence. The authors assess when the orderings can be meaningfully ranked, depending on the dimension of the variables involved. Levy and Razin (2015) examine the effect on policy outcomes when voters neglect correlations. This cognitive bias can improve political outcomes. They characterize conditions on the distribution of preferences under which this induces higher vote shares for the optimal policies and better information aggregation. Ellis and Piccione (2017) illustrate that improper perception of correlations has a role in individual choice. In their setup, agents choose assets, and may misperceive the joint returns. Maximizing subjective expected utility, using these beliefs, might lead to under-valuing or over-valuing particular asset streams. Levy and Razin (2018) suggest a framework for how sophisticated decision makers combine multiple sources of information to form predictions, under multivariate ambiguity. They show that there is a tradeoff betwen correlation neglect and an information-based correlation bound, which affects choice behavior in nontrivial ways.

In this paper we generally focus on bivariate dependence or concordance. However, given that our main result is in terms of arbitrary dependence, our analysis can also apply to multivariate dependence of any form. For example, tail dependence may be relevant in political economy, where a researcher is interested in analyzing the dependence among the extreme left (or extreme right) partisans in a particular election cycle. Moreover, we use both dependence and payoff function, which may be interpreted as a first step towards incorporating not just the level of dependence, but how a decisionmaker feels about dependence.

The rest of the paper is organized as follows: Section 2 provides an overview of decoupling, including our own new result. Section 3 develops an application to political economy. Section 4 concludes.

2 Background on Decoupling Results

In this section, we first illustrate the basic idea behind decoupling, since this approach is not commonly used in economics. In order to build intuition and to provide a sense of the scope of decoupling, we will provide 3 example areas of decoupling that may be relevant in diverse branches of economics. We then develop a new result for application to political economy. Heuristically speaking, the term 'decoupling' means taking dependent variables and transforming them so that they become (conditionally) independent. In this way, we can then examine dependent data, using standard techniques that are designed for independent data. For background on decoupling, see the monograph by de la Peña and Giné (1999).

Intuition for Decoupling. The main idea behind decoupling is to create additional copies of the dependent data, then rearrange these copies to extract conditional dependence. To begin with, consider two sets of variables, d_i and y_i, for $i = 1, ..., n$, with the same marginal distributions for each i. $d_1, ..., d_n$ are **dependent** random variables; and $y_1, ..., y_n$ **independent** random variables. We can now show our first result.

Example 1 of Decoupling: Linearity of Expectation. Note that for each i, y_i has the same marginal distribution as x_i and,

$$E\sum_{i=1}^{n} d_i = E\sum_{i=1}^{n} y_i. \tag{1}$$

Remark: The above expression is a basic decoupling result. The left side of (1) is the expectation of a sum of dependent random variables, while the right side is the sum of independent random variables. This expression is useful in statistics and economics, when we require average effects on a particular set of dependent random variables, but all we know is the corresponding result for independent counterparts of those variables. Also, the marginal distributions are often observed or estimable, while the dependence is trickier to estimate, see McNeil et al. (2015). So it is tractable to work with results for independent variables.

How, in practice, can we apply decoupling? Put another way, given the dependent d_i's, how can we obtain the independent y_i's? One way of constructing the y_i's is by producing n independent copies of $d_1, d_2, ..., d_n$, in the following manner:[3]

$$\mathbf{d_1^{(1)}}, d_2^{(1)}, d_3^{(1)}, ..., d_n^{(1)}$$

$$d_1^{(2)}, \mathbf{d_2^{(2)}}, d_3^{(2)}, ..., d_n^{(2)}$$

$$d_1^{(3)}, d_2^{(3)}, \mathbf{d_3^{(3)}}, ..., d_n^{(3)}$$

.

.

$$d_1^{(n)}, d_2^{(n)}, d_3^{(n)}, ..., \mathbf{d_n^{(n)}}$$

Letting $y_i = d_i^{(i)}$ it is easy to see that the following relations hold:

$$E\sum_{i=1}^{n} d_i = \sum_{i=1}^{n} Ed_i =$$

$$\sum_{i=1}^{n} Ed_i^{(i)} =$$

$$\sum_{i=1}^{n} Ey_i = E\sum_{i=1}^{n} y_i.$$

Example 2 of Decoupling: Sampling with Replacement. Applied researchers often deal with not only dependent data but also functions of dependent data. Hence a natural step to move beyond the linear model in (1) is to consider functions of dependent data. In this regard, an important early result is that of Hoeffding (1963), who provided a complete decoupling inequality, comparing sampling without replacement to sampling with replacement. The argument is as follows. First, let $\{S\}$ be a set with N elements . Let $d_1, ..., d_n$ be a sample without replacement and $y_1, ..., y_n$ a sample with replacement from the same population. Then, for every continuous and convex function Φ, Hoeffding shows that the following inequality holds for all $n \leq N$:

$$E\Phi(\sum_{i=1}^{n} d_i) \leq E\Phi(\sum_{i=1}^{n} y_i). \quad (2)$$

[3]For further details on this approach, see de la Peña (1990).

Note that y_i has the same marginal distribution as d_i, and that $\{d_i\}$ is a sequence of dependent variables while $\{y_i\}$ is a set of dependent random variables.[4]

Example 3 of Decoupling: Unions and Maxima. Some branches of economics rely on order statistics in order to develop appropriate pricing, for example, in auction implementation.[5] We therefore develop decoupling for the more general case of finite unions of dependent variables, and then specialize it to a particular order statistic, namely the maximum, and functions of the maximum.

Theorem 2.1. (Decoupling Finite Unions of Dependent Variables). Let $\{D_i\}$ be a sequence of arbitrarily dependent events and $\{Y_i\}$ a sequence of independent events with $P(D_i) = P(Y_i)$ for all $i \geq 1$. Then,

$$P\left(\bigcup_{i=1}^{n} D_i\right) \leq \frac{e}{e-1} P\left(\bigcup_{i=1}^{n} Y_i\right), \tag{3}$$

and the result is sharp.

Corollary 2.1 (Decoupling the Maximum)**.** Let $\{d_i\}$ be a sequence of arbitrarily dependent variables and $\{y_i\}$ a sequence of independent variables with Y_i having the same distribution as $\{d_i\}$ for all $i \geq 1$. Then

$$P\left(\max_{1\leq i\leq n} d_i > t\right) \leq \frac{e}{e-1} P\left(\max_{1\leq i\leq n} y_i > t\right). \tag{4}$$

Proposition 2.2 (Decoupling Functions of the Maximum)**.** Let $f(x)$, $x \geq 0$, be a continuous increasing function. Let $\{d_i\}$ be a sequence of arbitrarily dependent variables and $\{y_i\}$ a sequence of independent variables with y_i having the same distribution as $\{d_i\}$ for all $i \geq 1$. Then

$$Ef\left(\max_{1\leq i\leq n} d_i\right) \leq \frac{e}{e-1} Ef\left(\max_{1\leq i\leq n} y_i\right). \tag{5}$$

With the above examples in hand, we now turn to our main development of decoupling, namely, decoupling concave functions of summed dependent variables. This development has straightforward implications for microeconomics and political economy, as we shall see in Section 3.

Decoupling with a Concave Function of Dependent Variables. In economic analysis, it often falls on the researcher to move beyond treating the raw dependent data, and instead work with an objective function attached to these data. These functions often have

[4] Shao (2000) extended the result in (2) to the case when the d_i's are negatively associated random variables and the y_i's is a sequence of independent variables with d_i having the same distribution as y_i for every i.

[5] We are grateful to Ronny Razin for pointing out the auction context.

shape restrictions, and may include utility functions in economic theory, or loss functions in econometrics. For example, in the echo chamber context of Section 3 below, voter polls may feature dependent data, and the pollster is interested in the mean, variance or some utility function attached to these data. Sometimes the function is concave (in consumer theory), at other times it is convex (in producer theory). We present our result for concave functions.[6] Motivated by these considerations, our main new result is below:

Theorem 2.2. Let $d_1,...,d_n$ be a sequence of arbitrarily dependent random variables. Let $y_1,...,y_n$ be a sequence of independent random variables such that for all i, y_i has the same distribution as d_i. Then, for all concave increasing functions $\Phi(\cdot)$ with $\Phi(0) = 0$ we have

$$E\Phi(\sum_{i=1}^{n} d_i) \leq \frac{e}{e-1} E\Phi(\sum_{i=1}^{n} y_i). \tag{6}$$

Remarks. Note that inequality (6) holds for all n, and hence can be used in situations when the number of variables is arbitrarily large. Moreover, the inequality features a universal constant: it does not matter the type of dependence. Our result has natural applications in statistics and economics, because concave utility functions are often used to capture risk aversion (Mas-Colell et al. (1995), Chapter 6), and dependence is linked to lack of diversification (Markowitz (1952)). Heuristically, equation (6) can be thought of as capturing the following type of situation. When faced with summed, dependent data, a risk averse researcher's payoff must exceed what it would be if faced with independent data. However, this payoff is restricted to be at most 1.5 times higher. Hence there is a price to pay for facing summed dependent (multivariate) data, similar to the probability premium required for variable (univariate) data.[7]

Illustration in Statistics and Bank Portfolios. We now present a simple illustration of the above Theorem. We provide the main application in the next section. Here, we focus on the chi-square distribution, since it is widely used in statistical inference, for example in hypothesis testing.[8] Let the d_i's be dependent chi-square random variables with mean 1. Under the Theorem's maintained assumptions, then, the y_i's are independent chi-square random variables with mean 1. Hence $\sum_{i=1}^{n} y_i$ is distributed as chi-square with mean n. Now

[6] The result for convex functions is similar to (6) and omitted for brevity. It is available from the authors upon request. A version of (6) was first introduced by de la Peña (1990).

[7] For the premium required by a risk averse individual facing univariate data, see Mas-Colell et al. (1995), Chapter 6.

[8] The chi-square distribution arises naturally when analyzing averages that are squared. For details on the chi-square's role in the sampling distribution of estimators, see DeGroot and Schervish (2012), Chapter 8. For the chi-square in hypothesis testing, see Johnston and DiNardo (1997), Chapter 5. For its role in goodness of fit tests for categorical data, see Agresti (1996).

applying the theorem, we have that for concave increasing functions Φ, the following bound holds:

$$\begin{aligned} E\Phi(d_1 + \ldots + d_n) &\leq \frac{e}{e-1} E\Phi(y_1 + \ldots + y_n) \\ &\leq \frac{e}{e-1} E\Phi(\chi^2(n)). \end{aligned} \tag{7}$$

The above bound reduces analysis of multiple data, with unknown dependence, to a single chi-square with known structure. This construction is useful in a purely statistical setting when assessing dependent data.

Perhaps more interestingly for the current paper, the bound in (7) can be applied in economic settings. Consider a situation where a regulator requires related banks to sum up their (dependent) losses, and the regulator or bank is risk averse and therefore has a concave loss function. To be concrete, let $n = 2$, so that there are two banks, say, Citibank and Deutsche Bank. And let the payoff function Φ be quadratic utility. In this situation, the bound in (7) says that, under the maintained assumptions, a risk averse regulator can get the worst case scenario by the right hand side of (7). This is helpful because the joint distribution and dependence of the left hand side (the raw losses of Citibank and Deutsche Bank) are unknown. Nevertheless, the regulator can place a meaningful, worst-case scenario bound, no matter the dependence. This worst case scenario corresponds to just a scalar $\frac{e}{e-1}$ times the payoff she would require in a known case, if Citibank and Deutsche Bank were truly diversified.

Note that although the right-hand side proportion $\frac{e}{e-1}$ is constant, sample size n still matters. The reason is that sample size shows up in the right side, namely, the degrees of freedom of the chi-square distribution.[9] Thus, the present example has a benchmark chi-square with mean $n = 2$. But if the portfolio comprised, say, 100 banks, then the benchmark would be a chi-square with mean $n = 100$. Nevertheless, the import is that, relative to the independence or fully-diversified case (right hand side) the price of independence is larger by a constant factor of $\frac{e}{e-1}$.

[9]More generally, although the proportion on the right hand side is constant, the level in this type of decoupling bound will reflect sample size n. The reason is that the independent benchmark will be related to the product of n independent marginals.

3 Application to An Echo Chamber

We now define a statistical echo chamber, and demonstrate our main application of decoupling in political economy. Intuitively, an echo chamber comprises a set of variables that become more dependent each time they are observed. Standard methods of inference and estimation require independence of the relevant variables. Hence, researchers who face echo chambers may wish to glean a sense of the loss in subjective payoff from assuming the data are independent, when in fact there is dependence. The problem is compounded by the fact that the researcher will not know what k is, i.e. how much dependence has grown. Hence there is *ambiguity* about the amount of increase in dependence. For example, we might consider a utility function for a risk-averse policymaker (e.g. Prime Minister David Cameron on the eve of the 2016 Brexit vote) who faces voters with correlation neglect. We will apply the above new theorem to establish a proposition.

As mentioned in the literature review of Section 1, our approach is complementary to related research. Some of that research treats interdependence via stochastic ordering (see Meyer and Strulovici (2015)). Other research assesses divergence between dependent data and a benchmark using entropy-related multivariate divergence measures (Miller and Liu (2002)).

Price of Independence. We use the term price of independence to denote the proportional increase in resources that a researcher with a given payoff or loss function f would pay, at most, in order to deal with independent data, instead of dependent data.

For a given payoff function f, we define the *price of independence PI* as an upper bound. This bound is based on the ratio of expected values of this function of dependent variables d_i to the same function of independent variables y_i. Specifically,

$$PI \leq \frac{Ef(d_1,...d_i,...d_n)}{Ef(y_1,...,y_i,...,y_n)}. \tag{8}$$

Our definition of the price of independence PI encompasses general functions of dependent variables, such as order statistics and sums. In particular, it may be of interest in some branches of finance to consider $f(d_1,...d_i,...d_n) = max_i(d_i)$ and $f(d_1,...d_i,...d_n) = \sum_i d_i$, in which case the corresponding prices of independence are $\frac{Emax_i(d_i)}{Emax_i(y_i)}$ and $\frac{Ef(\sum_i d_i)}{Ef(\sum_i y_i)}$, respectively. The fact that the price of independence for maxima can be computed in this manner for the maximum is surprising, given that $Emax_i(d_i)$ can be a large quantity. In our application below, we will focus on the sum. PI is unit free, and therefore can be expressed in currency such as dollars or pounds. Moreover, PI does not depend on sample size. We can also extend the above definition of PI to be the sup over marginals, or sup over concave functions. Specifically, we can define a normalized price of independence $\hat{PI}$, as

$$\hat{PI} = \sup_{\{g_i\}_{i=1}^n} PI, \tag{9}$$

where $\{g_i\}_{i=1}^n$ denotes the set of common marginal distributions for d_i and y_i. This extension may be relevant for conservative decisionmakers facing multivariate ambiguity.

If evaluating the same set of m dependent variables $\{d_i\}_{i=1}^m$ that are observed at repeated periods indexed by k, for $k = 1,..,K$, we can also define the total price of independence. Specifically, we define the price of independence PI^{total} for the entire set of K observations, as simply the sum of the prices of independence over all periods:

$$PI^{\text{total}} = \sum_{k=1}^{K} PI(k). \tag{10}$$

Intuition for Echo Chamber. Before defining an echo chamber, we provide a few words of intuition. The echo chamber definition in (11) below is a basic attempt to capture rudimentary features of the following scenario: a researcher or policymaker observes data on citizens or assets at periodic intervals, say, monthly, where the variables become increasingly correlated. For example, she may observe assets in a period of systemic risk. Alternatively, she may observe monthly polling data in the buildup to an important vote, where the environment features biased perceptions such as correlation neglect, and increasing partisanship (Ortoleva and Snowberg (2015); Levy and Razin (2015)). In the latter, voting, case, our echo chamber formalizes a salient feature of partisan politics, where members of a political party (say, Labour in the UK) read the same newspapers and receive similar news feeds. Suppose the voters were polled at the beginning of each month, shortly after the government's plan to hold a vote on Brexit was announced. At the first poll, say, in January 2016, members of the Labour party would have some dependence in their views. In February, after one month of discussion among themselves, members of the political party would plausibly display stronger views, having had time to harden their resolve. A similar increase in dependence would occur in March, and so on. This scenario embodies the type of increasing dependence that we study. In such situations of unknown dependence, a policymaker will face difficulty in evaluating the polling data, and in ascertaining how much the echo chamber effect costs her.

Definition: Statistical Echo Chamber. Consider an m-dimensional vector of dependent, nonnegative variables $\mathbf{d} \equiv d_1,...,d_m$, and a vector of independent variables $\mathbf{y} \equiv y_1,...,y_m$, with the same marginal distributions as $\mathbf{d}$, i.e. $f_i(d_i) = g_i(y_i)$, for all $i = 1,...,m$. To fix ideas, $\mathbf{d}$

may refer to the returns on m different stocks in a portfolio, or m different voters that are polled periodically. As in the Introduction, let $\tau(k)$ denote average dependence during month k, across the marginals $\{d_i\}_{i=1}^m$ of $\mathbf{d}$. Let $\mathbf{d}^k$ denote the kth observation of the vector $\mathbf{d}$. We define an echo chamber $C^k(\mathbf{d})$ as a set of $K > 1$ observations of $\mathbf{d}$, where at each observation $k = 1,...,K$ on the vector $\mathbf{d}$, dependence $\tau(k)$ exceeds that of the previous observation $k-1$. That is, $\tau(k) \geq \tau(k-1)$, for all $k > 1$. In particular, a k-replica *echo chamber* $C^k(\mathbf{d})$ of dependent variables $\mathbf{d}$ is defined as

$$C^k(\mathbf{d}) = \{\mathbf{d}^k | \tau(k) \geq \tau(k-1), \quad k = 1,...,K\}. \tag{11}$$

The echo chamber definition in (11) features a weak inequality for the increase in dependence. We do not require that dependence among voters strictly increases despite their gathering more information. We merely indicate that dependence does not decrease.

Challenges in an Echo Chamber. Suppose that the data $\{d_i\}_{i=1}^m$ are observed monthly. The echo chamber researcher or policymaker is interested in the quantity $\sum_{i=1}^m d_i$, the sum of numbers in $\{d_i\}_{i=1}^m$, each month. The sum is a key quantity in many economic contexts. For example, in voting, for each poll month k in the year 2016, the quantity $\sum_{i=1}^m d_i$ measures the total number of individuals that say they would vote 'Yes' for Brexit in the UK (or vote for Hillary Clinton or Donald Trump in the USA).[10] Perhaps more naturally, rather than providing a 'Yes-No' vote, the voters could be report an intensity of preference for each candidate or platform, by responding to the following questionaire: 'on a scale of 1 to 5 (with 5 being the highest level of interest), report how interested you are in voting 'Yes' for this proposal or candidate'. Thus $\sum_{i=1}^m d_i$ represents the amount of 'expressed interest' in a particular vote, where the sample includes groups of voters with large or increasing (unknown) dependence.[11] Building on our Introduction and above discussion, below is a problem a policymaker might wish to address, if operating in an echo chamber environment.

Echo Chamber Problem: How much is it worth to a conservative researcher or policymaker, to analyze independent data, instead of echo chamber data?

In order to address the above challenge, we require a behavioral result, which characterizes choices of individuals faced with ambiguity. This Lemma summarizes a widely used approach (see Gilboa and Schmeidler (1989)) to ameliorate ambiguity, and places it in the context of ambiguous dependence. Essentially, a decisionmaker facing ambiguous depen-

[10]Alternatively, the term $\sum_{i=1}^m d_i$ can represent equally weighted returns in a portfolio of m assets.

[11]Note that the above formulation is in terms of average dependence measured by Kendall's tau. Since the result in (6) applies to all types of dependence, we could also include downside risk or tail dependence, i.e. dependence at the extremes of the distribution. Such dependence corresponds to far-right or far-left partisan dependence in a voting context from politics.

dence computes for each k, the price of independence. Being ambiguity-averse, she takes the worst evaluation as the one to be counted.

Lemma 3.1 (Choice of conservative individual facing ambiguous dependence). *When faced with ambiguity on dependence, a max-min or conservative individual will choose the worst-case dependence and optimize according to that scenario. Similarly, if faced with unknown time to a boundary crossing event with dependent data, a conservative individual will choose the expected upper limit of the boundary crossing time.*

The following Proposition describes how a researcher or policymaker may price the unknown dependence in an echo chamber. A major challenge is that the policymaker will not know the form of dependence (e.g. symmetric vs asymmetric), nor how much time dependence has been increasing (the value of k relative to K). The Proposition shows that despite these considerations, a policymaker can establish a meaningful premium for dependence.

Proposition 3.1 (Decoupling and Price of Independence in an Echo Chamber). *Suppose that a risk-averse policymaker with continuous payoff function u observes an echo chamber C^k as defined in (11) above. The policymaker does not know the value of k, and displays a conservative attitude towards ambiguous dependence. Then the policymaker's total price of independence is proportional to K, the maximal echo chamber period.*

Proof: Observe that a risk averter's objective function u is concave (see Definition 6.C.1, Mas-Colell et al. (1995)), and has $u(0) = 0$. Therefore we can apply Theorem 2.2, with $\Phi = u$.
· To compute the price of independence at each month k, expression (8) says $PI(k) \leq \frac{Eu(\sum_{i=1}^k d_i)}{Eu(\sum_{i=1}^k y_i)} = \frac{e}{e-1}$, each k, where the latter equality follows from Theorem 1, expression (6).
· By Lemma 1, if the policymaker does not know how long the echo chamber has been developing (i.e. does not know what the current month k is), she will apply her worst-case scenario, namely, that dependence has plateaud to the maximal τ, where $\tau(k) = \tau(K)$. Hence she assumes $k = K$, when calculating the total price of independence.
To compute the total price of independence for the conservative policymaker, we use the expression for PI^{total} from (10). This involves multiplying the monthly $PI(k)$ from above by the maximal index: $PI^{\text{total}} = \sum_{k=1}^K PI(k) = K \cdot \frac{e}{e-1}$. ■

Discussion of Proposition 3.1. The price of independence is similar to the probability premium required by risk averse individuals facing risky bets. Intuitively, if dependence is unknown and harmful for inference, a conservative researcher will choose the worst case scenario and hence the maximal price of independence. If the policymaker has information that a dependence plateau is far away, i.e. it is very early and correlation neglect has not set in, then she may not require as high a premium. Conversely, if the echo chamber is known to have been developing for some time, the bound is more likely to be sharp.

4 Concluding Remarks

Echo chambers are potentially problematic to analyze since they display unknown, potentially increasing dependence with each observation. Such phenomena occur in politics (see Levy and Razin (2015); Ortoleva and Snowberg (2015)), where partisan voters increasingly coalesce in their opinions, as time passes. We offer an approach to deal with evaluating the ambiguous dependence in echo chambers, based on the theory of decoupling in probability. Our approach is complementary to other approaches, such as those involving stochastic orders (Meyer and Strulovici (2015)) and multivariate divergence measures based on entropy considerations (Miller and Liu (2002)).

We develop a new inequality, which we then use to establish a maximal price to pay to compensate a researcher or decisionmaker for being in a political echo chamber. We show that a conservative decisionmaker would pay roughly 50% more than if the data were in fact independent. Hence there is a price to pay for facing summed dependent (multivariate) data, similar to the probability premium required for variable (univariate) data. For a given set of marginal information, this bound can be reformulated as the sup over all possible joint distributions connecting the marginals. This latter interpretation may tie our result to theories of decisionmaking under (multivariate) ambiguity.

Our main result in (6) treats arbitrary dependence, hence our analysis can also apply to multivariate dependence of any form. For example, tail dependence may be relevant in political economy, where a researcher is interested in analyzing the dependence among the extreme left (or extreme right) partisans in a particular election cycle. The decoupling approach may have applications to other areas of decision theory and political economy.

References

Agresti, A., 1996. An Introduction to Categorical Data Analysis. John Wiley and Sons, New Jersey.

Clemen, R., Fisher, G., Winkler, R., 2000. Assessing dependence: Some experimental results. Management Science 46 (8), 1100–1115.

de la Peña, V., Giné, E., 1999. Decoupling: From Dependence to Independence. Springer, New York.

de la Peña, V. H., 1990. Bounds on the expectation of functions of martingales and sums of positive rv's in terms of norms of sums of independent random variables. Proceedings of the American Mathematical Society 108 (1), 233–239.

DeGroot, M., Schervish, M., 2012. Probability and Statistics, 4th Edition. Pearson Education, Boston, MA.

Ellis, A., Piccione, M., 2017. Correlation misperception in choice. American Economic Review 107 (4), 1264–1292.

Gilboa, I., Schmeidler, D., 1989. Maxmin expected utility with non-unique prior. Journal of Mathematical Economics 18, 141–153.

Hoeffding, W., 1963. Probability inequalities for sums of bounded random variables. Journal of the American Statistical Association 58, 13–30.

Johnston, J., DiNardo, J., 1997. Econometric Methods, 4th Edition. McGraw-Hill, New York.

Levy, G., Razin, R., 2015. Correlation neglect, voting behavior and information aggregation. American Economic Review 105 (4), 1634–1635.

Levy, G., Razin, R., 2018. Combining forecasts in the presence of ambiguity over correlation structures. Tech. rep., London School of Economics.

Makarychev, K., Sviridenko, M., 2018. Solving optimization problems with diseconomies of scale via decoupling. Journal of the ACM 65 (6: 42), 1–27.

Markowitz, H., 1952. Portfolio selection. Journal of Finance 7 (1), 77–91.

Mas-Colell, A., Whinston, M., Green, J., 1995. Microeconomic Theory. Oxford Press.

McNeil, A., Embrechts, P., , Frey, R., 2015. Quantitative Risk Management: Concepts, Techniques and Tools, Revised Edition. Princeton Series in Finance. Princeton University Press, New Jersey.

Meyer, M., Strulovici, B., 2012. Increasing interdependence of multivariate distributions. Journal of Economic Theory 147, 1460–1489.

Meyer, M., Strulovici, B., 2015. Beyond correlation: measuring interdependence through complementarities. Tech. rep., University of Oxford.

Miller, D., Liu, W., 2002. On the recovery of joint distributions from limited information. Journal of Econometrics 107, 259–274.

Ortoleva, P., Snowberg, E., 2015. Overconfidence in political behavior. American Economic Review 105 (2), 504–535.

Shao, Q., 2000. A comparison theorem on moment inequalities between negatively-associated and independent random variables. Journal of Theoretical Probability 13 (2), 343–356.

A Proofs and Additional Results

Proof of Theorem 2.1: Let $P(D_i) = P(Y_i) = p_i$. $P\left(\bigcup_{i=1}^n D_i\right) \leq \min\left\{\sum_{i=1}^n p_i,\ 1\right\} = s$. Without loss of generality, we can assume that $\sum_{i=1}^n p_i \leq 1$. Therefore,

$$\begin{aligned} P\left(\bigcup_{i=1}^n Y_i\right) &= 1 - P\left(\bigcap_{i=1}^n Y_i^c\right) = 1 - \prod_{i=1}^n P(Y_i^c) = 1 - \prod_{i=1}^n (1-p_i) = \\ &= 1 - \prod_{i=1}^n (1-p_i) \leq 1 - \prod_{i=1}^n e^{\log(1-p_i)} \leq 1 - e^{-\sum_{i=1}^n p_i}, \end{aligned}$$

by a Taylor expansion. After verifying that $\frac{s}{1-e^{-s}}$ is increasing. To see that $f(s) \equiv \frac{s}{1-e^{-s}}$ is increasing for the stated range $s \in (0,1)$, we differentiate it to obtain

$$f'(s) = \frac{(1)\cdot(1-e^{-s}) - s\cdot(-1)(-e^{-s})}{(1-e^{-s})^2} = \frac{1-e^{-s}-se^{-s}}{(1-e^{-s})^2}.$$

The function $f(s)$ is increasing if $f'(s) \geq 0$. Examining the expression for $f'(s)$, its denominator is always positive. Therefore $f(s)$ is increasing once the numerator is positive. The numerator is positive if $1 - e^{-s}(1+s) \geq 0$, or equivalently, since $s \geq \ln(1+s)$. for $0 < s < 1$, we get

$$\frac{P\left(\bigcup_{i=1}^n D_i\right)}{P\left(\bigcup_{i=1}^n E_i\right)} \leq \frac{s}{1-e^{-s}} \leq \frac{e}{e-1}.$$

The sharpness of the inequality (as n tends to infinity) can be easily verified by letting $P(Y_i) = \frac{1}{n}$ for $i = 1, ..., n$.

Proof of Theorem: 2.2 We will show the proof for expression (6), concave increasing functions with $\Phi(0) = 0$.

Let d_i's be a sequence of arbitrarily dependent non-negative random variables. Let y_i independent of the sequence of d_i's and that the y_i's are independent with y_i having the same distribution as d_i for all i. From de la Peña (1990), and Makarychev and Sviridenko (2018) it follows that for all concave functions $\Phi(\cdot)$,

$E\Phi(\pi\sum_{i=1}^n d_i) \leq E\Phi(\sum_{i=1}^n y_i)$.

The left hand side can be bounded from below by using conditional expectations.

$$E\Phi(\pi\sum_{i=1}^{n} d_i) = \sum_{k=0}^{\infty} \frac{E\Phi(k\sum_{i=1}^{n} d_i)e^{-1}}{k!} \geq$$

$$\sum_{k=1}^{\infty} \frac{E\Phi(\sum_{i=1}^{n} d_i)e^{-1}}{k!} = E\Phi(\sum_{i=1}^{n} d_i)e^{-1}\sum_{k=1}^{\infty}\frac{1}{k!} =$$

$$E\Phi(\sum_{i=1}^{n} d_i)e^{-1}(e-1).$$