Thinking outside the Euclidean box: Riemannian Geometry and Inter-temporal Decision-making*

Himanshu Mishra †1 and Arul Mishra1

1University of Utah

*forthcoming PLOS ONE
†himanshu.mishra@utah.edu
Inter-temporal decisions involves assigning values to various payoffs occurring at different temporal distances. Past research has used different approaches to study these decisions made by humans and animals. For instance, considering that people discount future payoffs at a constant rate (e.g., exponential discounting) or at variable rate (e.g., hyperbolic discounting). In this research, we question the widely assumed, but seldom questioned, notion across many of the existing approaches that the decision space, where the decision-maker perceives time and monetary payoffs, is a Euclidean space. By relaxing the rigid assumption of Euclidean space, we propose that the decision space is a more flexible Riemannian space of Constant Negative Curvature. We test our proposal by deriving a discount function, which uses the distance in the Negative Curvature space instead of Euclidean temporal distance. The distance function includes both perceived values of time as well as money, unlike past work which has considered just time. By doing so we are able to explain many of the empirical findings in inter-temporal decision-making literature. We provide converging evidence for our proposal by estimating the curvature of the decision space utilizing manifold learning algorithm and showing that the characteristics (i.e., metric properties) of the decision space resembles those of the Negative Curvature space rather than the Euclidean space. We conclude by presenting new theoretical predictions derived from our proposal and implications for how non-normative behavior is defined.

**Keywords:** Riemannian Geometry; Inter-temporal Decisions; Discounted Utility; Gaussian Curvature; Manifold learning; Non-Euclidean Geometries; Exponential Discounting; Hyperbolic Discounting.
1 Introduction

Inter-temporal decisions involve deciding between outcomes occurring in the present (e.g. receiving $200 now) that offer a trade-off with those occurring in the future (e.g. receiving $300 after 6 months). Inter-temporal decisions are common among humans and animals and affect many decisions such as whether to have a health check-up right now or delay and face the consequences, save versus spend now, consume indulgent food now and face the consequences later, hoard food now or wait till later. Across both human and animal studies, it has generally been found that organisms i.e. the decision-makers, are likely to overweight immediate, but smaller rewards over delayed and larger rewards. These trade-off decisions can significantly influence impulse control, addictive behavior, retirement savings, investment, procrastination, attitude towards climate change, etc. (Rae 1834; Jevons 1879; Dasgupta 2006; Akerlof 1991; Green and Myerson 1996; Laibson 2003; Frederick et al. 2002). Therefore, understanding how humans and animals make inter-temporal decisions is a much-researched area having implications in myriad domains.

Many computational models have been presented to explain inter-temporal decisions. These models use discount functions, which are decay/depreciation functions (see Rachlin 2006) where the current monetary or non-monetary reward decays as its receipt is pushed into the future. The first formal discount function came from the Discounted Utility (DU) model by Samuelson (1937). The DU model, which proposes a constant rate discount function (e.g. exponential discounting)\(^1\), was presented as the normative model suggesting how decision-makers should behave when presented with inter-temporal decisions (Strotz 1955). As an alternate to the DU model, variable rate discount models have also been proposed to explain empirical evidence in inter-temporal decision-making. Variable rate discount models, (e.g., hyperbolic discounting models) (Chung and Herrnstein 1967; Ainslie 1975), suggest that people’s discount rate decreases over time i.e. it is higher in proximate time intervals but

\(^1\)Constant rate discounting works as follows: The utility of $x$ decays at a constant rate per unit of time. $x$ today would be worth $x\delta$ after one year, $x\delta^2$ after 2 years and so on. Here $\delta$ is $< 1$ and remains constant over the time horizon.
lower in distant intervals. One simple way to understand differences between constant rate and variable rate discount models is to compare them to a bucket of water, with a hole that leaks water as one pushes it along a straight path. If we replace the water with money\(^2\) and the straight line with the temporal line then according to constant rate models, for each unit of temporal distance covered, the leakage (depreciation) is at a constant rate. On the other hand for variable rate models, the leakage is a function of temporal distance covered\(^3\). As with any line of research, questions have been raised about both the constant rate and the variable rate models, including questions about their ability to explain empirical evidence, their theoretical underpinnings and their predictive/explanatory power (Luhmann, 2013). Moreover recent research has suggested that recognizing the way we perceive time may not be the same as objective time, could help us better understand inter-temporal decisions. Specifically, some distinctions between constant and variable rate models fade if we formulate time perception to be non-linear (e.g., logarithmic) (Takahashi 2005; Attema et al. 2010).

Before we describe some commonalities between various discount models, we first need to explain a term we will be using frequently in the remainder of the manuscript i.e., “decision space”. One of the objectives of research involving any type of decision-making is to understand how the organism or the decision-maker perceives a decision-object (e.g. given information). Such a perception forms her subjective experience of decision objects and the decision hinges on this perception. We refer to the space where information is perceived and the decision is made as the decision space or decision surface of that decision-maker. If we take the context of inter-temporal decision-making, then the subjective perception of time and reward/money occurs in the inter-temporal decision space of the decision-maker.

\(^2\)In this manuscript for ease of description we use money as a reward or source of utility. One can easily replace money with any other source of utility.

\(^3\)Another way to see the difference between constant rate and variable rate models is to consider their corresponding differential equations. If we assume that at time \(t = 0\) the initial monetary value is \(y\) and \(y(t)\) represents the residual value at time \(t\) then constant rate models can be conceptualized as a solution to linear autonomous first order differential equation of the form: \(\frac{dy}{dt} = -\psi\) where \(\psi\) is the constant discount rate. On the other hand, variable rate models can be conceptualized as a solution to the linear nonautonomous first order differential equation of the form: \(\frac{dy}{dt} = -\Omega(t)\) where the discount rate is a function of time. At a broader level this captures the main difference between these two types of models.
Despite their much-discussed differences, there are some important commonalities among the various discount models. First, most existing models assume the inter-temporal decision space to be a metric space. For instance, a decision of comparing the receipt of $100 now to receiving $250 at a point in time eighteen months in the future has been studied using a discount function of temporal distance estimate (e.g., $\delta(t)$ where $t$ is the distance between these 2 points in time). The assumption of a metric space implies that by using a metric we can estimate the length of a time interval used in decision-making\textsuperscript{4}. The important aspect of considering the inter-temporal space to be a metric space is that we need to know what is the geometry of the surface on which the distance is being measured because any accurate distance estimate is entirely dependent on the surface on which the distance is being measured and is defined by the unique metric of that surface. Therefore, by thinking of temporal comparisons in terms of distance (which almost all the previous research does), a discussion of the geometry of the surface on which temporal distance is being measured is unavoidable.

This brings in the second point on which all existing models of inter-temporal decision-making converge. Existing models assume the metric of the inter-temporal decision space to be a Euclidean metric since the temporal distance between $t_1$ and $t_2$ (where $t_2 > t_1$) is calculated as: $d(t_1 \rightarrow t_2) = t_2 - t_1 = \sqrt{(t_2 - t_1)^2}$, which is similar to measuring the Euclidean distance between points $(t_1, 0)$ and $(t_2, 0)$. The decision space is assumed to be a metric space with an underlying Euclidean geometry.

Therefore, an important similarity across all existing models of inter-temporal decision-making is the assumption of a Euclidean decision space. However, despite such a fundamental assumption of a Euclidean space, to the best of our knowledge, no research has ever questioned this assumption and neither has anyone empirically tested for whether the assumption is true or not. Let us explain next why this might be so.

\textsuperscript{4}Specifically, if $t_1 < t_2 < t_3$ are three points on a temporal line then inter-temporal decision space being metric space means that (i) the distance between $t_1$ and $t_1$, $d(t_1 \rightarrow t_1) = 0$, (ii) $d(t_1 \rightarrow t_2) > 0$, (iii) $d(t_1 \rightarrow t_2) = d(t_2 \rightarrow t_1)$ and (iv) $d(t_1 \rightarrow t_2) + d(t_2 \rightarrow t_3) \geq d(t_1 \rightarrow t_3)$
Such an unquestioning assumption of a Euclidean space is not surprising. Due to its intuitive appeal Euclidean geometry has always had a pervasive influence on scientific inquiry. The best example of this influence can be seen in Kant’s (1902; 1908 : originally published in 1781) arguments suggesting that Euclidean space is a-priori intuition in human judgments. It matches with how the world appears visually to us and is incredibly accurate in many everyday predictions. Given the intuitive appeal and face-validity of Euclidean space, it is not surprising to see that in the fields that explore human preferences, whenever metric space needed to be defined, either we explicitly assumed the metric space to be Euclidean (Jehle and Reny 2006 pg 417, Mas-Colell et al. 1995 pg 674) or implicitly used it to model behavior without further testing for its veracity empirically. While Kant’s argument could be accepted before the introduction of non-Euclidean geometries, it seems less defensible now to assume without empirical investigation that the geometric space underlying inter-temporal decision can only be Euclidean. In keeping with Riemann (1854) and Helmholtz (1876), who were among the first to consider the notion that the geometry of space around us is an empirical matter that should be decided by measurements, we need to ask why should we assume that time and rewards are perceived by the decision-maker in a Euclidean space? What is the basis for such an assumption? Does empirical evidence support such an assumption? Might there be a different geometry involved when time and rewards are perceived by the decision maker?

First, we need to ask whether our conclusions, predictions, and explanations would change if we considered a Euclidean geometry when the geometry was actually non-Euclidean? The straightforward answer is that if the space is curved and we assume it to be flat, we will introduce distortions in measuring the shortest distance (called the geodesic distance). It is analogous to assuming that the earth is flat and trying to measure distance between two points. However, once we know that the surface of the earth is spherical then we can use the correct metric to estimate distance. Similarly, if we considered the wrong geometry for the inter-temporal decision space, we would be introducing distortions in our measurements and
subsequently in how we assess the discounting process. In such a situation infinite distance functions can be proposed and some of them would work some of the times. This is the current situation in inter-temporal distance estimates. Therefore, determining the correct geometry of the decision space is the essential first step needed to apply the appropriate metric to calculate distances on that space and as a consequence our distance estimates would be distortion-free. Second, it would help us build a theoretical model that would better explain the empirical findings and help make further theoretical predictions.

2 Overview

How does a decision-maker evaluate the utility from different bundles of time and corresponding monetary pay-offs for that time? He does so by perceiving those bundles at different distances in his decision space. To compare different bundles he brings each bundle to the present (origin) and calculates the discounted utility from each bundle. The discounting is a function of the perceived distance. Therefore, in this research our attempt is to find out the correct geometry of the decision space to ensure that our distance estimates are accurate.

Previewing briefly, in this research, we present a new model of inter-temporal decision-making by relaxing the current restrictive assumption that the decision space is a Euclidean space. We develop our theory on the proposal that the decision space resembles the metric properties of a more flexible Riemannian space of constant negative Gaussian curvature. This implies that the metric used by the decision maker is not Euclidean but the more general Riemannian metric. We further suggest that both temporal distance and reward influence the magnitude of discounting. We discuss this proposal in detail in section (4). In order to maintain consistency, we use the term Riemannian space throughout the manuscript. Another term that can be used is Riemannian manifold. In questioning the Euclidean nature of the decision space we follow previous research in other domains that have considered non-Euclidean spaces. Apart from the most well known use of non-Euclidean space in the general
theory of relativity (Einstein 1916), researchers in many domains, such as Embedding of networks (Shavitt and Tankel 2004), binocular vision and perception (Luneburg 1947; Indow 1979; Fernandez and Farell 2009) to name a few have questioned the Euclidean assumption and searched for solutions using non-Euclidean spaces.

Two arguments, interestingly opposing, can be raised against our research. One that inter-temporal decision-making is a relatively smaller problem that doesn’t require the mathematical complexity of curved spaces. We argue that this may be true if we narrowly define the domain of inter-temporal decisions. However, if we consider the concept of a decision space, which plays a critical role in almost every decision we make, this approach has implications for many types of decisions. Conversely, the second argument could question the need for replacing the current Euclidean assumption, which makes it easy to understand and apply existing discounting models, with an elaborate process of questioning the geometry of the decision space. Again, we argue that it is not appropriate for us as researchers to keep assuming something because it is easy to understand and not question it or test for it empirically.

The remaining part of the manuscript is structured in the following manner. First, in order to build our proposition that a Riemannian space of constant negative curvature underlies the decision space, we provide a brief overview of Riemannian space and Gaussian curvature and explain Riemannian spaces of constant negative curvature (we will refer to it as Negative Curvature spaces or negatively curved spaces)\(^5\). For a detailed explanation of their mathematical roots please refer to appendix B (page 37). Second, we present our proposed geometric theory of inter-temporal decision making. We explain how information is perceived in the decision space. Third, we use two approaches, analytical and empirical, to test our proposed theory. We test it analytically by examining whether we can explain

\(^5\)Our discussion of these topics is nowhere close to being exhaustive. These topics are active areas of research across many disciplines. For more in-depth understanding, please refer to the following sources: Do Carmo (1976, 1992); Ratcliffe (2006); Cannon et al. (1997); Anderson (2005); Needham (1998); Thurston and Levy (1997); Berger (2003). For a lucid overview see Penrose (2005); Greenburg (1993). For a discussion of non-Euclidean geometry using real projective geometry see Coxeter (1998).
the existing findings in inter-temporal decision making when we utilize a discount function that uses distance in the negatively curved space instead of the Euclidean temporal distance as used in past work. Empirically, we estimate the curvature of the decision space, utilizing inter-temporal decisions made by participants, to test whether it is Euclidean or Negative Curvature space. Finally, we conclude with further theoretical predictions that can be derived from the proposed geometric theory.

3 Geometries and Distances

Riemann (1854) proposed that spaces do not have any inherent geometry instead they are akin to a continuum where points are specified by their coordinates. Whenever a specific metric is used to measure the distance between two points, it means that an assumption has been made about the geometry of the surface since a metric is unique to a geometry. The problem is that different metrics can be used to estimate the distance and hence, different geometries can be imposed. However, which is the correct metric can be verified only when one knows what is the geometry of the surface on which the points are located. Let’s refer back to the example of a ruler (with the familiar Euclidean metric $ds = \sqrt{dx_1^2 + dx_2^2 + .. dx_n^2}$) versus a cooked spaghetti. If asked to measure the distance between any two points, we can use the ruler or the spaghetti as our measure. If we used the ruler we will get "a" distance estimate, irrespective of whether the surface is flat or curved. The estimate with the ruler would be correct for the flat surface. However, if the surface were curved then the distance estimate using the ruler would be distorted because the ruler is not faithfully representing each point in its distance estimate. The fact that we are getting "an" estimate does not mean that we have the "right" estimate. But how can we know that we have the right estimate? It is possible only if we knew the geometry of the surface on which the points are located. We can attest to the veracity of our estimate only if we empirically test for the geometry of the surface. Hence, Helmholtz (1876) had proposed that assumed geometry should always
be empirically verified.

In order to identify the right geometry we should have the ability to distinguish between infinite geometries that are possible between points. Such ability would allow us to decide, for instance, whether the right metric is a ruler or a spaghetti. The Gaussian curvature proves to be a very useful measure for such differentiation since it differentiates curved surfaces from flat surfaces. The Gaussian curvature ($K$) informs us how curved a specific surface is with respect to a flat surface, at a given point. The magnitude of ($K$) tells us how much the surface is bending. If the Gaussian curvature of a surface is the same at every point then we have a constant curvature geometry. An example is a globe on which, the Gaussian curvature is the same at each point. On the other hand, if we consider a crumpled paper, each point on it has a different Gaussian curvature. Pertinent to our research, we will be considering constant curvature geometries. Moreover, the sign, positive or negative, of the Gaussian curvature informs us of the type of geometry in a particular point $p$’s neighborhood. Specifically, at every point on a surface if $K = 0$ then the neighborhood of $p$ would resemble a flat surface and its intrinsic geometry would be Euclidean. If $K > 0$ then it would resemble a surface like a piece of sphere and its intrinsic geometry would be Elliptical. If $K < 0$ then it would resemble a saddle and its intrinsic geometry would be Gauss-Bolyai-Lobachevskian/Hyperbolic. Since $K$ can have any value between $-\infty$ and $\infty$, we can think about constant curvature geometries not as three separate geometries, but instead as a continuum of infinite geometries where the Euclidean geometry is a special case. In sum, the Gaussian curvature of the space helps distinguish among various geometries.

The geometry we assume for a surface will dictate our choice of metric since each geometry is defined by its unique metric. Therefore, a ruler ideal for measuring distances on surfaces with $K=0$ would not be right to use on a surface with $K$ greater than zero. We now discuss the model we use to calculate geodesic (the shortest path between two points) on the Negative Curvature space. Since the Negative Curvature space is abstract and hence, difficult to imagine, models are constructed that make it easier to measure distances between
points in such abstract spaces. These models tend to be Euclidean representations that faithfully embody the key features of the non-Euclidean geometry. Researchers depend on various models of Negative Curvature space such as Poincarè half space, Poincarè disk, and Beltrami-Klein model to understand the properties of the space. In our proposal we use the hyperboloid model.

To understand the problem of assuming the wrong geometry, let’s compare the Euclidean geodesic with the geodesic on a negatively curved space between two points \( p \) and \( q \). We use the geodesic (shortest distance) estimates for Euclidean distance and the distance on a hyperboloid. If \( p=(p_1,p_2,p_3,...,p_n) \) and \( q=(q_1,q_2,q_3,...,q_n) \) are two points in \( E^n \) (n-dimensional Euclidean space) then their geodesic distance in the Euclidean space is

\[
d_e(p, q) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + \cdots + (q_n - p_n)^2} = \sqrt{\sum_{i=1}^{n} (q_i - p_i)^2} \tag{1}
\]

However, if these points are on a hyperboloid \( H^n \subset E^{n,1} \) where \( K = -1 \) then

\[
d_h(p, q) = \text{Cosh}^{-1}(-<p,q>)
\]

\[
= \text{Cosh}^{-1}\left(\sqrt{1 + \sum_{i=1}^{n} p_i^2} \sqrt{1 + \sum_{i=1}^{n} q_i^2} - \sum_{i=1}^{n} q_i p_i\right) \tag{2}
\]

For any other value of \( K < 0 \)

\[
d_h(p, q) = \frac{\text{Cosh}^{-1}(K <p,q>)}{\sqrt{-K}} \tag{3}
\]

As these equations highlight, assuming the wrong geometry introduces distortions in estimates of distance. Please see appendix C.1 (page 39) for the detailed derivation of
equation (2) and (3).

4 The Proposed Model

In this section, we first provide details of the proposed model and then derive various mathematical relationships. We propose that the given information about time and money is perceived in (or mapped on to) the Negative Curvature decision space. Specifically, the following points capture the essence of our proposal.

1. We suggest that the decision space where the perception of information and the subsequent decision takes place resembles the metric properties of Negative Curvature space. The proposal that the decision space is a Negative Curvature space in no way implies that the decision space appears like a hyperboloid embedded in a higher dimension pseudo-Euclidean space to a decision maker. Instead, what it means is that the metric properties of the inter-temporal decision space, as perceived by the decision-maker, resemble the metric properties of the Negative Curvature space. It is important to note that our proposal does not require any new set of assumptions pertaining to the inter-temporal decision space being considered a metric space. As we discussed earlier, existing research already assumes the inter-temporal decision space to be a metric space and uses the Euclidean metric to assess inter-temporal distance. Since the Euclidean space is a special case of Riemannian spaces of constant curvature, we do not need to introduce any new assumptions. Our proposal actually relaxes a very restrictive assumption of zero Gaussian curvature, which is required for a space to be a Euclidean space.

One could question why we are proposing the decision space to have $K < 0$? Why not $K > 0$? $K > 0$ spaces are elliptical spaces. A good example of $K > 0$ space is a sphere/ball. Let’s imagine that we want to measure the distance between two points A (which is fixed) and B (which is moving) on a ball. As B moves away, we first see that the distance between A and B increases. However, later as B gets farther, its distance from A reduces; eventually B coincides with A and thus, the distance becomes 0 between A and B. This happens because a ball (like elliptical spaces) is a closed surface. If we consider the decision space to have $K > 0$, we will face some logical hurdles. For example, this would imply that as an event moves away from the present time, it’s perceived temporal distance first increases and then decreases. In other words, this would require us to think of time as circular where after a finite delay, the future coincides with the present. It is a notion that is hard to justify logically as well as experientially.
2. We further propose that the Gaussian curvature of the decision space is malleable. Although the decision space would remain one with a negative curvature, the value of the Gaussian curvature could change depending on various factors. We argue consistent with the view espoused by Riemann [see Weyl (1922) pg 98] that “the space in itself is nothing more than a three dimensional manifold devoid of all form; it acquires a definite form only through the advent of the material content filling it and determining its metric relations.” We propose that dispositional factors of the decision-maker, as well as contextual factors in which the decision is being made, have the ability to change the Gaussian curvature of the decision space and thus, in turn the metric relationship among objects perceived in it.

3. In our proposed model we consider the influence of both time and money on discounting. Therefore, instead of using time as the sole influencer on the discounting process, we use the distance in the time-money decision space. Specifically, when a decision maker is considering two bundles of time and money, \((m_1, t_1)\) and \((m_2, t_1)\), she is doing so from the vantage point of her subjective origin. The best way she can compare the two bundles is by bringing the bundles, by discounting their value appropriately, to the origin. We propose that the discounting factor she uses is the distance \(d_h\) of a specific bundle from the point on which it is situated in her inter-temporal decision space to the origin. The difference between our method and previous ones is that previous methods use just time in the discounting factor e.g. \(m_1\delta^{t_1}\) or \(\frac{m_1}{1+\gamma t_1}\). In our model the discounted utility of the \((m, t)\) bundle is not due to the Euclidean distance between the origin and just \(t\) but due to the distance between the origin and \((\phi(m), \eta(t))\) in the Riemannian space, where \(\phi(m)\) and \(\eta(t)\) are the perceived values of money and time respectively.

As we demonstrate later, our model explains many of the existing findings in inter-temporal literature such as those findings, which have been labeled as anomalies as well as
those that deal with time perception (e.g. the thesis that time is perceived logarithmically). Hence, our proposal is more inclusive and general. Moreover, by presenting a positive theory of inter-temporal decision-making, we not only explain the existing findings in the literature, we also present predictions (section 7) based on the theory that can be empirically tested. Although one could argue that we are relaxing geometric assumptions of the inter-temporal framework, we are not adding any additional variables to the framework. Hence, our model is as parsimonious as the others proposed previously.

In the next section we derive various mathematical relationships to understand how information is perceived in the inter-temporal decision space.

4.1 Distance Perception in the Decision Space

We use the distance estimate on a hyperboloid to formulate the distance estimate on the inter-temporal decision space. Let’s consider a decision-maker who is contemplating an inter-temporal choice problem from his vantage point at the origin (i.e. the present). The decision-maker is asked to make a choice between receiving $m_1$ after $t_1$ delay, versus receiving $m_2$ after $t_2$ delay (where $m_2 > m_1$ and $t_2 > t_1$). To decide, he has to compute a discounted value of $m_1$ and $m_2$. Let $\phi(m)$ denote the value(utility) one perceives in any given amount of money (here $\phi(\cdot)$ is a monotonically increasing function). According to existing literature, if we assume $\phi(m) = m$, we would predict that the decision-maker would choose $m_2$ after $t_2$ delay if $m_1\delta^{t_1} < m_2\delta^{t_2}$ (exponential discounting) or $\frac{m_1}{1+\gamma t_1} < \frac{m_2}{1+\gamma t_2}$ (hyperbolic discounting).

In our proposal, we suggest that the decision maker would choose $m_1$ if $\phi(m_1)f(d_{h1}) < \phi(m_2)f(d_{h2})$. This implies that discounting is not a function of $t$ (temporal distance) instead it is a function of $d_h$; the perceived distance between the origin and any money-time bundle in the Negative Curvature decision space. $\eta(t)$ is a monotonically increasing function of $t$ and denotes the magnitude one perceives of any given time interval. Our proposed model

\footnote{$\delta < 1$ and $\gamma$ is the discounting parameter. We acknowledge that there are various discount functions. The use of exponential and hyperbolic is just to give an example.}
can be described as a two-stage model just like any standard discounting model. In the first stage both \( \phi(m) \) and \( \eta(t) \) are used to assess distance from the origin. While in the second stage that distance is used to discount the values of \( \phi(m) \). The following equation (3), for a particular time/money bundle \((m \text{ at time } t)\), \( d_h \) represents the distance of that bundle from the origin (status quo) in the inter-temporal decision space according to

\[
d_h = \frac{Cosh^{-1}(\sqrt{-K}\sqrt{\phi^2(m) + \eta^2(t) - \frac{1}{K}})}{\sqrt{-K}}, \quad K < 0
\]  

(4)

Please see appendix C.2 (page 40) for a detailed derivation of equation (4).

In order to make our exposition simpler we adopted the following steps. First, we assumed the Gaussian curvature \( K \) to be -1 (however, we can certainly generalize this for any value of \( K \in (0, -\infty) \)). Second, in order to present the equations in the manuscript concisely we specify \( \Gamma \) (gamma) to be as follows \( \Gamma = \sqrt{1 + \phi^2(m) + \eta^2(t)} \) (for a quick reference to symbols used in this manuscript, please refer to appendix A (page 36). Substituting \( K = -1 \) in equation (4), we can write the simplified version of it as

\[
d_h = Cosh^{-1}(\Gamma)
\]  

(5)

Please see appendix C.3 (page 41) for a verification of equation (5) using a different approach. Next we will describe some properties of \( d_h \) that will be used at different points in the manuscript. In the subsequent subsection, we derive the discount function.

Inferring from equation(5), the following are some relevant properties of \( d_h \)

\[
\frac{\partial d_h}{\partial t} = \frac{\eta'(t)}{\Gamma \sqrt{\Gamma^2 - 1}}
\]  

(6)

\[
\frac{\partial d_h}{\partial m} = \frac{\phi'(m)}{\Gamma \sqrt{\Gamma^2 - 1}}
\]  

(7)

Since \( \phi(m) \) and \( \eta(t) \) are monotonic functions of \( m \) and \( t \) respectively:
This means that for a constant \( m \), as \( t \) approaches infinity, the rate of change of \( d_h \) approaches zero. In other words, \( d_h \) remains largely unchanged for increasing values of \( t \).

\[
\lim_{t \to \infty} \frac{\partial d_h}{\partial t} = 0 \tag{8}
\]

Similar to the interpretation of (8), (9) shows that for any constant value of \( t \), \( d_h \) remains largely unchanged for higher values of \( m \). While we can increase \( m \), its perceived magnitude remains unchanged in the Negative Curvature decision space.

\[
\lim_{m \to \infty} \frac{\partial d_h}{\partial m} = 0 \tag{9}
\]

(10) simply shows that \( d_h \) is a concave function with respect to \( t \).

We suggest that one way to account for the loss of water is to consider that the surface

\[
\frac{\partial^2 d_h}{\partial t^2} < 0 \tag{10}
\]

4.2 Leaking bucket and Discount Function

Let’s revisit the Leaking bucket analogy introduced earlier. If a bucket leaks water only when we walk, then inter-temporal decision-making is analogous to predicting how much water will be left in a leaking bucket after walking a certain distance. That is, what would be the discounted values of money after a delay. The approach in existing research has been to assume that the surface on which we walk is a Euclidean surface with time being the only dimension and all the functions that have been developed to explain the loss of water are associated to the variable/constant size of the hole in the bucket.

We suggest that one way to account for the loss of water is to consider that the surface
on which we walk has a constant negative Gaussian curvature, instead of the more restrictive assumption of zero Gaussian curvature (i.e., a flat Euclidean surface). We further propose that the size of the hole is constant (i.e., the same amount of water leaks for every unit of distance travelled). Hence, we can replace the water in the bucket with money and let $y_0$ be the initial amount of money in the bucket and $\psi$ be the fraction of money lost per unit of distance travelled on the Negative Curvature surface (note, that we are conceptualizing time to be a continuous variable, however, one would arrive at similar conclusions by treating time as a discrete variable. The only modification in that instance would be the use of first order difference equations rather than differential equations). Formally, this can be written in the form of a differential equation as

$$\frac{d\phi(y)}{dd_h} = -\psi\phi(y) \quad \text{(11)}$$

$$\Rightarrow \phi(y_{d_h}) = \phi(y_0)e^{-\psi d_h} \quad \text{(12)}$$

Please see appendix C.4 (page 42) for a complete solution of equation (11).

One way to understand $\phi(y_{d_h})$ and $\phi(y_0)$ in equation (12) is to think of the former as the discounted value of the latter. If we substitute $d_h$ with time $t$ in equation (12) and assume $\phi(y) = y$ (i.e., linear function), we get the exponential discounting function used in the DU models (Green and Myerson 1996). Now if we assume $\psi = 1$ (unit loss/constant unit discounting) and substitute the value of $d_h$ from equation (5), we get

$$\phi(y_{d_h}) = \phi(y_0)e^{-\cosh^{-1}(|\Gamma|)} \quad \text{(13)}$$

More generally, if the Gaussian curvature of the surface is $K < 0$ then

$$\phi(y_{d_h}) = \phi(y_0)e^{-\frac{-\cosh^{-1}(\sqrt{-K}\sqrt{\phi_0^2 + \eta^2(t) - \frac{1}{K}})}{\sqrt{-K}}} \quad \text{(14)}$$
To maintain consistency we will use equation (13) quite frequently in the remaining portion of the manuscript. An easy way to understand what distance means in this formulation is to keep either \( m \) or \( t \) constant. When \( m \) is kept constant then distance indicates the perceived magnitude of time. Similarly, if \( t \) is kept constant then distance shows the perceived magnitude of money.

Next, we provide support for the notion that the decision space is Negative Curvature space through two different methods. First, if our proposal is true then using (13), we should be able to explain the existing inter-temporal findings documented in literature (Loewenstein and Prelec 1992; Read 2001). Second, empirically we should find that the metric properties of the decision space resemble the metric properties of the Negative Curvature space more than those of the Euclidean space.

5 Explaining Inter-temporal Findings

Some of the findings that we discuss are inconsistent with only the DU (constant rate discount) model and some with both the DU and the hyperbolic discounting (variable rate discount) model.

5.1 Common Difference Effect

If \((x, t)\) denotes receiving \(x\) at time \(t\) then the stationarity property of DU model requires that when \((x_1, t_1) \succ (x_2, t_2)\) where \(x_2 > x_1\) and \(t_2 > t_1\) then \((x_1, t_1 + n) \succ (x_2, t_2 + n)\) for \(n > 0\) also known as constant willingness to wait (Attema et al., 2010). However, empirical evidence suggests that \((x_1, t_1 + n) \prec (x_2, t_2 + n)\) (please see Stewart et al. (2014) for a description of this effect). In other words, violation of stationarity property leads to a preference reversal such that a smaller, sooner reward is preferred over a larger, later reward in the near-future but a larger, later reward is preferred over a smaller, sooner reward in the distant-future (see Ainslie 1975; Green et al. 1994). It has been argued that the DU model
is unable to explain the preference reversal because it assumes a time independent, constant rate of discounting. In order to explain preference reversal, the model has to have a variable discount rate which changes (decreases) with time. Would decisions made in the Negative Curvature decision space be able to explain such preference reversals?

The discount function is $e^{-\cosh^{-1}(\Gamma)}$ (from 13). Therefore, the discount rate can be described as

$$
\psi(m, t') = \frac{-\partial e^{-\cosh^{-1}(\Gamma)}}{\partial t} e^{-\cosh^{-1}(\Gamma)}
$$

$$
= \frac{\eta'(t)}{\Gamma \sqrt{\Gamma^2 - 1}}
$$

From (6) and (8), $\lim_{t \to \infty} \psi(m, t') = 0$ i.e., as $t \to \infty$, $\frac{\eta'(t)}{\Gamma \sqrt{\Gamma^2 - 1}} = 0$. This means that the rate at which the value of reward depreciates in Negative Curvature space is variable and decreases with increasing time. Hence, preference reversals in inter-temporal decisions can be explained by our Negative Curvature model.

5.2 Temporal Sub-additivity

Temporal sub additivity (or sub additive time discounting) is steeper discounting when a delay is divided into parts compared to when it is undivided (Read 2001). For instance, assume the decision-maker is indifferent between $100$ now and $150$ in 1 year and is also indifferent between $150$ in 1 year and $200$ in 2 years. However, he appears to prefer $200$ in 2 years over $100$ now. Therefore, we observe that when the same duration of 2 years is presented as a whole it leads to less discounting as compared to when it is divided into parts. It is inconsistent with both the DU and hyperbolic discounting models. We now show that if we consider the decision space to be negatively curved we can explain temporal sub additivity.
From (10), we know that $d_h$ is a concave function. Therefore, if we keep $m$ constant,

$$d_h(t_1 + t_2) \leq d_h(t_1) + d_h(t_2)$$

(16)

In other words, dividing a fixed duration of time into smaller intervals makes it appear more than the same undivided duration.

How would this perception change discounting? Let’s take a duration $t$ and divide it into two subdivisions $t_1$ and $t_2$ where $t_1 + t_2 = t$ while keeping the monetary value constant. Now consider two situations A and B. In A, an initial amount $y_0$ is first discounted over $t_1$. If the discounted values at the end of $t_1$ is $y_1$ then from (12), $y_1 = y_0 e^{-\psi d_{h_1}}$. If we further discount $y_1$ over $t_2$ then the discounted value at the end of $t_2$ would be $y_2 = (y_0 e^{-\psi d_{h_1}}) e^{-\psi d_{h_2}} \Rightarrow \frac{y_2}{y_0} = e^{-\psi (d_{h_1} + d_{h_2})}$. On the other hand, lets consider the other situation B where $y_0$ has been discounted for the entire duration $t$. If $y_t$ is the discounted value at the end of $t$, then $y_t = y_0 e^{-\psi d_{ht}} \Rightarrow \frac{y_t}{y_0} = e^{-\psi d_{ht}}$. From equation, (16) $d_{ht} < d_{h_1} + d_{h_2} \Rightarrow \frac{y_t}{y_0} > \frac{y_2}{y_0}$. In other words, $y_0$ will discounted less over the entire time period $t$ than when $t$ is divided in intervals $t_1$ and $t_2$. Hence, our Negative Curvature model can explain temporal sub-additivity.

5.3 Logarithmic time perception

Following the Weber-Fechner and Steven’s law (Stevens 1957), it has been proposed that the perception of time follows logarithmic pattern and the perceived time intervals of the same objective duration gradually shrink as one considers the distant future (Takahashi 2005). Such a shrinkage is a novel notion that is incompatible with both the DU and hyperbolic models. More recently, (Attema et al., 2010, §4) also suggested that variable rate discounting may be caused by a non-linear (logarithmic) perception of time by the decision-maker.

However, we can explain such a shrinking in a negatively curved decision space.

Since $Cosh^{-1}(\omega) = \ln (\omega + \sqrt{\omega^2 - 1})$ where $\omega > 1$, using equation (5)
Equation (17) shows that perceived distances in the Negative Curvature decision space follow a logarithmic pattern. Unlike past research which has just shown that time follows a logarithmic pattern, our model predicts such a pattern for both time and money. From (8) and (9), we can see that the perceived magnitude of an extra unit of time/money in the decision space decreases as the objective value of time/money increases. Therefore, the predictions of Weber-Fechner law or previous logarithmic functions are subsumed in our model.

5.4 Absolute Magnitude Effect

This effect suggests that larger monetary amounts are discounted less steeply than smaller amounts (Green and Myerson 1996; Loewenstein and Prelec 1992). Consider a decision-maker who is indifferent between receiving $100 now versus $200 after 6 months and is also indifferent between receiving $3000 now versus $4500 in 6 months. This example highlights that for small amounts the discount rate is higher than for large amounts.

We can explain why the absolute magnitude effect occurs if we consider the decision space to be negatively curved. From (15), we see that \( \lim_{m \to \infty} \psi(m, t') = 0 \) which means the discount rate declines as the amount increases. In other words, in the negatively curved decision space for the same duration of time, the discount rate is lower for high monetary outcomes than for low monetary outcomes.

5.5 Preference for Improving Sequences

It has been empirically shown that people prefer improving sequences (where the worst outcome occurs temporally first and the best outcome last) to diminishing sequences (where the best outcome occurs temporally first and the worst outcome last, see Loewenstein and
Again, using a negatively curved decision space we can explain why an improving sequence is preferred. If \( m_3 > m_2 > m_1 \) where \( m_1, m_2, \) and \( m_3 \) are different values of money, we know from equation (15) that \( \psi(m, t') \) declines with increasing values of \( m \). That is, as suggested in explaining the Absolute Magnitude effect, the discount rate decreases with increasing amount of money such that \( m_3 \) would depreciate the least and \( m_1 \) would depreciate the most. Therefore, if the decision-maker has to temporally order consumption of \( m_1, m_2, m_3 \), she would prefer to delay \( m_3 \) since it will depreciate the least and consume \( m_1 \) the soonest.

Until now, analytically we provided support for our proposal that the inter-temporal decision space resembles the metric properties of Riemannian space of constant negative curvature. We also were able to explain the findings in inter-temporal decision-making using our Negative Curvature model. We next present the second set of evidence where we attempt to learn the Gaussian curvature of the decision space to determine whether it is Euclidean or not.

6 Learning the Curvature of the Decision Space

Any attempt to empirically learn the curvature of the decision space poses an intriguing question: if we cannot see the shape of the decision space, how can we infer its curvature. For example, we know that a sphere is not a Euclidean surface because we can observe its shape and find out that it is not a flat surface. However, we don’t have such a vantage point for the inter-temporal decision space, so how can we infer its curvature? The answer to this question lies in Gauss’s “Theorema Egregium” which proves that the Gaussian curvature of a surface, while defined with respect to the higher dimension space that the surface is embedded in, is an intrinsic property of the surface. Our understanding of the earth’s geometry illustrates this very elegantly. Although it has only been a few decades, since we were able to rise above the earth into space and actually observe that the earth is spherical, scientists were
able to infer quite accurately from measurements taken on the surface of the earth that its shape was spherical and not flat. More generally, “Theorema Egregium" implies that for any surface a two-dimensional bug living on it, who is unable to holistically view the surface from afar, can still measure the curvature of the surface. Thus, for surfaces with a constant Gaussian curvature, measurements on that surface itself can reveal its nature i.e. whether the surface is elliptical, hyperbolic, or Euclidean. Utilizing this property, we do not need to rise above the decision space or observe it from a distance in order to infer its shape. By taking measurements that inform us about distances on the decision space, we can infer its curvature.

We used Riemannian space learning method to estimate the Gaussian curvature of the inter-temporal decision space. In using this method we face two challenges. First, since the distance estimation process is happening in the mind of the decision maker we cannot visibly see his distance estimates. Second, we cannot provide the decision maker with an objective yardstick for measuring distances i.e. a direct metric assessment of distances. Hence, we use latent distance estimates that are inferred through the inter-temporal choices/tradeoffs the decision-maker makes (we discuss this in detail in §6.2).

The Riemannian manifold learning method offers a distinct advantage since we do not need to specify the function that maps the time and money information to the decision space. Similar to any MDS we don’t need to know how objective points are mapped into subjective points. All we need is a measure of distance among each combination of points. Hence, by freeing us from mapping constraints, the procedure provides confirmatory evidence as to the nature of the decision surface. The only input that the method needs in order to determine the nature of the surface is the distance a decision maker perceives between various combinations of money and time. First, we discuss the algorithms to assess the nature of the decision surface (Negative Curvature or Euclidean) then discuss the procedure for collecting the data and how this data was used to infer perceived distances in the decision space.
6.1 Algorithms

We used two different algorithms to test whether the decision surface was Euclidean or Negative Curvature. The input for each algorithm was the $n(n-1)/2$ values of the inter-point distance $d_{ij}$ (e.g., if $n = 5$, we had ten values of $d_{ij}$). These algorithms fit values of $d_{ij}$ to assess if the decision space is Negative Curvature or Euclidean.

For the Negative Curvature algorithm, we utilized Weierstrass coordinates to parametrically represent points on the surface of a hyperboloid (Faber 1983). Here $(r, \theta)$ are the polar coordinates in $E^3$ and $K$ is the Gaussian curvature.

\[
\begin{pmatrix}
x_n \\
y_n \\
z_n
\end{pmatrix} = \begin{pmatrix}
\frac{1}{\sqrt{K}} Sinh[r_n\sqrt{K}] * Cos[\theta_n] \\
\frac{1}{\sqrt{K}} Sinh[r_n\sqrt{K}] * Sin[\theta_n] \\
\frac{1}{\sqrt{K}} Cosh[r_n\sqrt{K}]
\end{pmatrix}
\]

The Negative Curvature algorithm calculated the following:

1. $n \times 3$ values of the Weierstrass coordinates $[x_n, y_n, z_n]$ from $n \times 2$ values of the polar coordinates $(\hat{\theta}_n, \hat{r}_n)$ and the Gaussian Curvature $\hat{K}$.

2. $n(n-1)/2$ values of $\hat{d}_h(i, j)$ between all pairs of points $i$ and $j$ using Weierstrass coordinates from the previous step and equation (3).

3. $\epsilon^2 = \sum_i \sum_j (d_{ij} - \hat{d}_h(i, j))^2$

These steps were repeated until $\epsilon^2$ was minimized. In each repetition, the values of $(\hat{\theta}_n, \hat{r}_n)$ and the Gaussian Curvature $\hat{K}$ were modified using simulated annealing (Kirkpatrick et al. (1983)). Once it converged, the Negative Curvature algorithm provided values of the polar coordinates $(\hat{\theta}_n, \hat{r}_n)$ and $\hat{K}$.

The Euclidean algorithm worked in a similar manner except that the polar coordinates were modified by simulated annealing to approximate the radial distance. The algorithms used here are similar to Lindman and Caelli (1978) but with three differences. First, for the
Negative Curvature algorithm, we used Weierstrass coordinates to parametrically represent points on the surface, second, we directly measured the Negative Curvature distance (instead of estimating it indirectly from a pseudo Euclidean distance) and third, we used simulated annealing instead of Powell’s method. Please refer to appendix D (page 43) for a test of these algorithms with simulated data.

6.2 Procedure

To test our proposal, we ran two separate studies. In study 1, forty participants were recruited from Amazon mechanical turk and completed the study for monetary compensation. In study 2, forty-four undergraduate participants took part in the study for partial course credit. In both studies participants were instructed that they would be making inter-temporal decisions.

Across both studies, a matching procedure was used to elicit participants’ responses to various inter-temporal choices. For instance, participants were asked to fill an amount such that they would be indifferent between the following 2 payoffs: $100 in 1 month versus $\_\_\_\_ in 18 months. In the given instance if $y_j$ is the amount filled by the participants, $y_i$ is $100 and $d_{ij}$ is the perceived distance between $y_i$ and $y_j$ in the decision space, then $d_{ij}$ can be calculated in the following manner:

We know from equation 13 that

\[ y_i = y_j e^{-d_{ij}} \]

\[ d_{ij} = \ln\left(\frac{y_j}{y_i}\right) \] (19)

In both studies we used 5 temporal points to elicit different inter-temporal choices (In study 1 these points were: now, 1 year from now, 2 years from now, 5 years from now, and 7 years from now. In study 2 they were: now, 1 month from now, 9 months from now, 18 months from now, and 36 months from now). We also used two different values of $y_i$: $100
and $200. For one set of the matching task we kept $y_i$ as $100$ and elicited values of $y_j$ for the five different points in time. We repeated this procedure for another value of $y_i$ as $200$. The matching procedure questions with $y_i = 100$ or $200$ were randomized (participants provided a total of 20 values of $y_j$, 10 each for $y_i = 100$ and $y_i = 200$). Finally, we calculated the inter-point distance $d_{ij}$ by averaging $\frac{y_j}{y_i}$ across $y_i = 100$ and $200$ and using it as the input in (19). The rationale for using two values of $y_i$ was to reduce the impact of outlier responses by using the average value of $\frac{y_j}{y_i}$.

The data was then used to determine the nature of the decision surface by examining how inter-temporal distances were perceived in the decision space.

### 6.3 Results

Given the assumptions in past work, our default hypothesis was that the decision space is Euclidean, hence, the base model for our comparison was the Euclidean model. To test whether the Negative Curvature or the Euclidean algorithm provided a better solution, for each participant, we compared their root mean squared errors divided by the standard deviation of the distances. We used the following criteria to exclude participants’ responses from the analysis.

1. In order to test whether participants were paying full attention to the inter-temporal choices, in both studies, we had included a test question, which asked participants to not respond to that question. If participants answered that specific question it would indicate that they were not reading the instructions completely. The program automatically terminated the data collection for such participants.

2. We excluded responses from participants that displayed more than 1 instance of preferring a smaller, later reward over a larger, sooner reward. An example would be someone who is indifferent between receiving $100$ today and $200$ after 1 year. He is also indifferent between receiving $100$ today and $150$ after 2 years.
6.3.1 Study 1

Of the 40 participants, responses from 5 participants were not used based on the two criteria specified above. Responses from the remaining 35 participants were analyzed by subjecting each participant’s response to both algorithms (detailed results for each participant are summarized in appendix (E.1) on page 44). The Negative Curvature algorithm provided a better fit (lower RMSE/standard deviation ratio) for 25 participants, the Euclidean and the Negative Curvature algorithms provided near identical fit for 6 participants, the Euclidean algorithm provided a better fit for 1, the Negative Curvature fit was less than 1% better than the Euclidean fit for 2 participants, and for 1 participant (participant #29) while the Negative Curvature algorithm provided a better fit, the estimated Gaussian curvature was near zero (indicating a Euclidean solution). Taking a more conservative approach, we classified the cases in which the fit from the Negative Curvature algorithm was less than 1% better than the Euclidean algorithm as a tie and considered all ties as evidence against the negatively curved decision space. This means out of 35 participants, the Negative Curvature algorithm provided a better fit in 25 cases ($\chi^2(1, N = 35) = 6.42, p < .01$)\(^8\)

6.3.2 Study 2

Of the 44 participants, responses from 6 were not used based on the two criteria discussed above. For the remaining 38 participants, the pattern of result was similar to study 1. The Negative Curvature algorithm provided a better fit (lower RMSE/standard deviation ratio) for 34 participants while the Euclidean algorithm provided a better fit for 3 participants; no difference (or a tie) in fit emerged for the remaining 1 participant. The results are

---

\(^8\)One could argue that the better fit of the Negative Curvature algorithm is because it has one extra parameter ($K$) than the Euclidean algorithm (where $K$ is always zero). The following is one way to address this concern. As shown in appendix (D), when interpoint distances are sampled from a Euclidean surface, the Negative Curvature algorithm accurately recovers the Gaussian curvature as zero. In other words, this algorithm can identify the Gaussian curvature from both the Euclidean as well as the negative curvature data. If we utilize only the output of this algorithm, we come up with the same conclusion as we arrived at by comparing the Negative Curvature algorithm with the Euclidean algorithm: for 10 participants the Gaussian curvature was near zero while for 25 it was less than zero.
summarized in appendix (E.2) on page 45. Again taking a more conservative approach, we classified the cases in which the fit from the Negative Curvature algorithm was less than 1% better than the Euclidean algorithm as a tie. By doing this, we obtained 31 cases where the Negative Curvature algorithm provided a fit better than 1%, 4 cases where the hyperbolic and the Euclidean algorithm provided largely the same fit (less than 1% improvement) and 3 cases where the Euclidean algorithm provided a better fit. A $\chi^2$ test shows that the Negative Curvature algorithm was a better fit for the data, $\chi^2(1, N = 38) = 15.158, p < .0001$.

In sum, the results from the two studies provide empirical support to our proposal that the inter-temporal decision space, where the information about time and money is perceived, metrically resembles a Negative Curvature space rather than a Euclidean space.

6.3.3 Robustness Check

One criticism of our findings could be that we are assuming a linear transformation of money (i.e., $\phi(y) = y$) in equation (19). We utilized a linear transformation because in much of past research be it animal behavior or human time discounting, researchers have utilized a linear transformation of reward to utility. Therefore, if we consider precedence in terms of empirical testing, the most common functional form was linear. Moreover, it has been found that utilities of small monetary amounts are generally linear (Wakker and Deneffe 1996) and since we were using small amounts such as $100 and $200 we felt that a linear transformation would suffice.

One option to test robustness of our findings is to calculate each participants’ unique $\phi(m)$ and then calculate the Gaussian curvature $K$. Abdellaoui et al. (2010) discuss a procedure for estimating utility functions in inter-temporal contexts. However, that procedure requires a very different data collection method where time weights are elicited in the first phase and then those weights are used in the temporal discounting task. While this procedure may help us in calculating an individual’s utility function, it may not provide a way to estimate distance $d_{ij}$ that is needed for our algorithm.
The second option is to utilize an empirically established functional form of \( \phi(m) \), which happens to be the functional form based on Steven’s law (Stevens 1957) and prospect theory (Hastie and Dawes 2001, page 294). So we considered \( \phi(m) = m^\alpha \). In the gain domain, past work has shown that \( \alpha = .88 \) and if \( \alpha = 1 \) then \( \phi(m) = m \) similar to what we assumed in our existing analysis. This resulted in using the equation \( d_{ij} = \ln[\frac{y_j^\alpha}{y_i^\alpha}] \). Next, to test the robustness of our conclusion that the Negative Curvature space depicts the inter-temporal decision space better than a Euclidean space, we varied alpha from .2 to 1.4 in increments of .2 and used those values to calculate \( d_{ij} \). \( d_{ij} \)'s were then used as inputs in both algorithms. The two graphs in appendix (E.3, page 47) depict the results we obtained across both studies. For all the values of \( \alpha \), the data still showed that the Negative Curvature algorithm provided a better fit across studies 1 and 2 compared to the Euclidean algorithm. As expected, since the scale changes (because changing functional forms changes the scaling), we find that the values of \( K \) change. However, and importantly, the conclusions that a Negative Curvature algorithm provides a better fit to the data remain unchanged.

7 Predictions

Analytically and empirically, we have provided evidence to support our proposal that the inter-temporal decision space resembles the metric properties of negative curvature space i.e. it is not Euclidean. As we had mentioned earlier, given that our geometric model of inter-temporal decision making has theoretical underpinning, not only does it have the ability to explain existing instances of inter-temporal decision making, it also has the ability to make future predictions that can be verified. In this section we put forth some predictions that can be derived from our geometric model.
7.1 Loss/Gain and Gaussian Curvature

Past work has shown that delaying (accelerating) a gain (loss) is discounted more steeply than accelerating (delaying) a gain (loss). The prevailing explanation is rooted in prospect theory, which suggests that since losses loom larger than gains, delaying (accelerating) a gain (loss) is an aversive outcome and thus, felt more acutely than accelerating (delaying) a gain (loss).

Using our geometric model, if $\psi(m, t')$ is the discount rate then the loss/gain discounting asymmetry can be expressed as $\psi(m, t')_{\text{loss}} > \psi(m, t')_{\text{gain}}$. If in equations (14) and (15), we substitute $K_L$ and $K_G$ as the Gaussian curvatures of the decision space in the loss and gain domain respectively, then we can depict the loss/gain discounting asymmetry in the Negative Curvature decision space as

$$
-\frac{\partial e}{\partial t} \frac{\cosh^{-1}\left(\sqrt{-K_L}\sqrt{\phi^2(m) + \eta^2(t) - \frac{1}{K_L}}\right)}{\sqrt{-K_L}} > -\frac{\partial e}{\partial t} \frac{\cosh^{-1}\left(\sqrt{-K_G}\sqrt{\phi^2(m) + \eta^2(t) - \frac{1}{K_G}}\right)}{\sqrt{-K_G}}
$$

This inequality would hold only if $K_G < K_L$ where both $K_G$ and $K_L$ are less than zero. Please see appendix (E.4, page 48) for the complete solution.

It means for delayed-gains to be discounted more than accelerated-gains, the Gaussian curvature of the decision space has to be less when a person considers a delayed-gain than when he considers an accelerated-gain (or an accelerated-loss than a delayed-loss).

Predicting that the Gaussian curvatures for gains and losses would be different is an intriguing conjecture because it is analogous to the description of a malleable geometry suggested by Riemann that we had discussed earlier. It would mean that the decision space is not a rigid container with a fixed geometry; instead its geometry (based on different curvatures) evolves with changes in the decision context. It leads to the proposal that inter-temporal decisions in the loss and gain domain diverge because loss and gain differentially
alter the curvature of the decision space.

We can extend this to make some interesting predictions in other domains where dispositional or visceral factors are known to influence the discount rate. For instance, Mischel and Moore (1973) showed that children are able to display higher self-control with symbolic vs real reward (e.g., more likely to be tempted by a real cake than an image of it). Similarly, Loewenstein and Prelec (1992, page 595) discuss how proximity to sensory contact with a choice object increases the discount rate. Our proposal would predict that proximity to sensory rewards influences the discount rate by altering the curvature (and hence, the geometry) of the decision space.

7.2 Monetary subadditivity

In our geometric model, we can see that equation (5) is sub-additive not just for time but also for money. In other words, for a constant time duration

\[ d_h(m_1 + m_2) \leq d_h(m_1) + d_h(m_2) \] (21)

Equation (21) would predict that the same amount of money over a fixed duration will be discounted more when divided in parts than as a whole. In other words, if $x_1 + x_2 = X$ then $X$ will be discounted less than the aggregate of $x_1$ and $x_2$ over a fixed time duration $t$. For example, our model would predict that the discounted value of $500 in 1 year will be more than the sum of discounted values of $200 and $300 in 1 year.

7.3 Subadditivity at distant time periods

As we discussed earlier, inter-temporal subadditivity is inconsistent with both the DU and hyperbolic discounting models, but it can be explained if we conceptualize the decision space to be negatively curved. However, very little is known about inter-temporal subadditivity at distant time periods. In other words, for the distant-future, would we observe a similar
subadditive influence by dividing the time duration into smaller intervals as we observe for proximate time or would temporal subadditivity attenuate at distant times?

We know from equation (16) that if we divide a time duration $t$ into $\frac{t}{n}$ and $\frac{(n-1)t}{n}$ then

$$d_h\left(\frac{t}{n} + \frac{(n-1)t}{n}\right) \leq d_h\left(\frac{t}{n}\right) + d_h\left(\frac{(n-1)t}{n}\right), \forall n > 0$$

If we add interval $g$ to $t$

$$d_h\left(\frac{(t+g)}{n} + \frac{(n-1)(t+g)}{n}\right) \leq d_h\left(\frac{(t+g)}{n}\right) + d_h\left(\frac{(n-1)(t+g)}{n}\right), \forall n > 0, g \geq 0$$

$$\Rightarrow \frac{\partial d_h\left(\frac{(t+g)}{n} + \frac{(n-1)(t+g)}{n}\right)}{\partial t} \leq \frac{\partial d_h\left(\frac{(t+g)}{n}\right)}{\partial t} + \frac{\partial d_h\left(\frac{(n-1)(t+g)}{n}\right)}{\partial t}$$

Now if $g \to 0$ then

$$\lim_{g \to 0} \left[ \frac{\partial d_h\left(\frac{(t+g)}{n} + \frac{(n-1)(t+g)}{n}\right)}{\partial t} \right] \leq \lim_{g \to 0} \left[ \frac{\partial d_h\left(\frac{(t+g)}{n}\right)}{\partial t} + \frac{\partial d_h\left(\frac{(n-1)(t+g)}{n}\right)}{\partial t} \right]$$

(22)

Consistent with existing research, equation (22) shows that in the near-future discounting will be more if the time duration is divided into intervals compared to when the same duration is kept undivided.

However, if $g \to \infty$ then

$$\lim_{g \to \infty} \left[ \frac{\partial d_h\left(\frac{(t+g)}{n} + \frac{(n-1)(t+g)}{n}\right)}{\partial t} \right] = \lim_{g \to \infty} \left[ \frac{\partial d_h\left(\frac{(t+g)}{n}\right)}{\partial t} + \frac{\partial d_h\left(\frac{(n-1)(t+g)}{n}\right)}{\partial t} \right]$$

(23)

Therefore, equation (23) predicts that at distant time periods, the subadditive influence on discounting would be attenuated.
8 Conclusions

In this research we question the geometry underlying the inter-temporal decision space and build a theoretical model based on an appropriate geometry that better explains the existing empirical findings (anomalous or not) and helps make further theoretical predictions. Past research has assumed that the decision space (where time and money are perceived/experienced by the decision-maker) is an Euclidean space. However, we propose that the decision space is a Riemannian space with constant Negative Curvature.

We support our proposition through two approaches. First, we provide evidence, analytically, by deriving a new discount function, which uses the distance in the Negative Curvature space instead of Euclidean temporal distance. By doing so we are able to explain the empirical anomalies that have been shown in the inter-temporal literature. In other words, empirical anomalies (such as common difference effect, absolute magnitude effect, temporal subadditivity, logarithmic time perception, preference for improving sequence) that at times are inconsistent with the DU model and/or the hyperbolic discounting model can be explained by considering the geometry of the decision space to be non-Euclidean. Second, when we measure the Gaussian curvature of the decision space through surface learning algorithms we find that the metric properties of the decision space resemble those of the Negative Curvature space rather than the Euclidean space.

By building a geometric model of the inter-temporal decision space, we question a widely accepted notion that the surface underlying the decision space is Euclidean. By relaxing this rigid assumption, we propose that a more flexible approach should be adopted such that we take into account dispositional factors of the decision-maker, as well as contextual factors in which the decision is being made, to influence the geometry of the decision space and thereby the metric relationship among the decision objects. Finally, by considering both time and money in the distance function we are suggesting that both, together, influence inter-temporal decisions. Such an integration of money helps us get richer insights rather than when time alone is considered to be the sole influencer.
If we look at the discount function in equation (12), we find that it is analogous to the DU discount function with one difference. Instead of assuming the inter-temporal space (where discounting happens) to be a Euclidean space, we are assuming it be a more flexible negatively curved space. One intriguing outcome of relaxing the rigid Euclidean assumption is that the DU model, much criticized for its inability to explain anomalies, can now explain many anomalies. It also raises the question about how we define non-normative behavior and anomalies. By assuming the wrong geometry we may conclude that the decision maker is behaving in a non-normative manner. However, our conclusion is flawed as it is based on the wrong assumption. Let’s consider the analogy of a bug moving on a transparent globe to illustrate this point. Assume that the bug is traveling along the great circle, which is the shortest (geodesic) path on a globe. However, assume further that we cannot observe the bug’s actual movements. All we can see is the shadow of the bug’s movement, including the start and end points, on the floor caused by a light bulb kept at the top of the globe. We are unaware of the shape of the object on which the bug is actually moving, whether it is a globe, a cylinder, a saddle or a flat surface. Since, the only thing we observe is the shadow on the floor, we could erroneously assume that the bug is moving on a flat surface and try to predict the shortest distance it should move to go from the start to the end point. By this error in our assumption, we would find the bug’s path to be quite irrational (at times following a straight line when moving along the prime meridian on the globe and at times a curved path as we see its shadow moving along the equator) since it would not be moving by the shortest distance predicted by a flat surface. Unfortunately, it will be our conclusion that is wrong since we are presuming the movement on the wrong surface; the bug is quite consistent in its movement as it follows the shortest path on the globe, which is the great circle.

In a similar manner, if we replace the globe with the decision space and the bug’s path with how humans estimate inter-temporal distances, we see that it is our assumption of the geometry of the space to be Euclidean which is at fault, rather than the decision-maker’s
inter-temporal choices. If we erroneously use the Euclidean distance between two inter-temporal points to estimate how much an outcome needs to be discounted, we would reach the incorrect conclusion that the decision-maker is non-normative. However, in reality the decision makers are correctly estimating the distance along the shortest path but their decision space is negatively curved. Therefore, if we identify the right geometry underlying their decision space, we would see that they are actually responding normatively. In sum, a more flexible approach is to recognize that the decision objects don’t fit into some pre-specified Euclidean space (with its established geometric metric) but that the decision space is defined by the factors present during the decision making process and can be more malleable than the rigid Euclidean space.
A Symbols

\[ E^n \quad n \text{-dimensional Euclidean space} \]
\[ H^n \quad n \text{-dimensional Negative Curvature space} \]
\[ E^{n,1} \quad n + 1 \text{-dimensional pseudo-Euclidean space which embeds a hyperboloid} \]
\[ ds \quad \text{Infinitesimal distance} \]
\[ d_e(p, q) \quad \text{Euclidean distance between points } p \text{ and } q \]
\[ d_h(p, q) \quad \text{Distance between points } p \text{ and } q \text{ in a negatively curved space} \]
\[ K \quad \text{Gaussian curvature} \]
\[ < x, y > \quad \text{Lorentzian inner product} \]
\[ m \quad \text{Amount (payoff)} \]
\[ t \quad \text{Time} \]
\[ \phi(m) \quad \text{Monotonically increasing function of } m \]
\[ \eta(t) \quad \text{Monotonically increasing function of } t \]
\[ \Gamma \quad \sqrt{1 + \phi^2(m) + \eta^2(t)} \]
\[ L[f(x)] \quad \text{Laplace transformation of } f(x) \]
\[ y(d_h) \quad \text{The discounted value of the initial value } y(0) \]
\[ \psi(m, t') \quad \text{Discount rate (keeping } m \text{ constant)} \]
\[ d_{ij} \quad \text{Interpoint distance} \]
\[ (r, \theta) \quad \text{Polar coordinates} \]
\[ \Omega_G \quad -\frac{1}{K_G} + \phi^2(m) + \eta^2(t) \]
\[ \Omega_L \quad -\frac{1}{K_L} + \phi^2(m) + \eta^2(t) \]
B Riemannian Space

Riemann (1854) proposed that spaces do not have any inherent geometry instead they are akin to a continuum where points are specified by their coordinates. The metric relationships among these points determine what kind of space it is. In other words, depending upon the metric relationships among points infinite geometric structures can be imposed on them. Formally, Riemannian space (or manifold) can be defined in a coordinate system $x_1, x_2, x_3...x_n$ by its metric tensor $g_{ij}$. A Metric Tensor uniquely defines the space (stated differently, any metric tensor relates only to a particular space). This metric tensor determines how distance is calculated between two points in a given space. It is symmetric and positive definite. We can write the metric tensor either in terms of local derivatives as
\[ ds^2 = g_{11}dx_1^2 + g_{12}dx_1dx_2 + ... \]
or as
\[ ds = \sqrt{g_{ij}dx_i dx_j} \] (see Stoker 1969 for a detailed description).

Note that in a Euclidean space $g_{ij} = \delta_{ij}$. Where $\delta_{ij}$ is the Kronecker delta (it is 0 when $i \neq j$ and 1 when $i = j$). This reduces the metric tensor to the form of a familiar Euclidean metric: $ds = \sqrt{dx_1^2 + dx_2^2 + ... dx_n^2}$ making the Euclidean space a special form of the more general Riemannian space. In small regions (locally) the Riemannian space is Euclidean. This is the reason why despite the earth’s curved surface, in small areas, geometric measurements on its surface are very similar to Euclidean measurements. The Riemannian space is smooth, which means it has no creases, no punctures, no separate pieces stuck together, and no tears [Green (1999) pg. 263, see Lee (2012) for a formal definition]. A critical parameter for a Riemannian space is its Gaussian curvature (or curvature tensor for $n$-dimensional space).

We next explain what is the Gaussian curvature.

B.1 Gaussian Curvature

Essentially, the Gaussian curvature ($K$) informs us how curved a specific surface is with respect to a flat surface, at a given point. It is a parameter associated with each point of the surface which measures the intrinsic curvature of the surface, which is invariant under bending (by Gauss’s Theorema Egregium) or the coordinate system used. Formally, assume $\Pi$ to be a plane containing the normal vector $n$ that intersects the surface at point $p$. Let $k$ be the curvature of the curve at $p$ that is the intersection of the surface with $\Pi$. The sign of $k$ is determined by whether the center of the curvature is in the direction of $n$ or $-n$. In this situation if $\Pi$ rotates about $n$, we would obtain various values of $k$. The Gaussian curvature is calculated as $K = k_{\min} * k_{\max}$ (please see Needham 1998). The Riemannian curvature tensor is a generalization of the Gaussian curvature applied to higher dimensions. 9

For any point $p$ on a surface, the Gaussian curvature tells us two things. First, the magnitude of the Gaussian curvature tells us how much the surface is bending at $p$. Second, its sign (positive or negative) tells us what the surface is like in $p$’s immediate neighborhood (Needham, 1998, pg 274). If $K$ is constant at every point of the surface, then the inherent geometries of different constant curvatures can be defined through the value of $K$ in the following manner: At every point of the surface if $K = 0$ then the neighborhood of $p$ would resemble any Euclidean i.e. flat surface. If $K > 0$ then it would resemble a surface like a piece

\[ K = -\frac{\nabla^2 \log(\varsigma)}{\varsigma^2} \] where $\nabla^2(\cdot)$ is Laplacian, $\nabla^2(\varsigma) = \frac{\partial^2 \varsigma}{\partial x_1^2} + \frac{\partial^2 \varsigma}{\partial x_2^2} + ... \frac{\partial^2 \varsigma}{\partial x_n^2}$.

9The Gaussian curvature relates to a metric induced by an element of arc length. If $ds = \varsigma(x)dx$ then

\[ K = -\frac{\nabla^2 \log(\varsigma)}{\varsigma^2} \]
of sphere and its intrinsic geometry would be Elliptical. If \( K < 0 \) then it would resemble a saddle and its intrinsic geometry would be Gauss-Bolyai-Lobachevskian/Hyperbolic. Since \( K \in (-\infty, +\infty) \), we can think about constant curvature geometries not as three separate geometries, but instead as a continuum of infinite geometries. The Gaussian curvature of the space distinguishes the various geometries.

Since we propose that the decision space of inter temporal decision-making resembles the metric properties of a Riemannian space of constant negative curvature (Negative Curvature space), we next specifically discuss features of this space and how to calculate distance between any two points in such spaces.

### B.2 Constant Negative Curvature Space

Unlike the Euclidean space, which is flat, easy to represent, and easy for us to imagine, the Negative Curvature space is more abstract and hence, difficult to imagine. Therefore, models are constructed that make it easier to study their properties in further detail. These models make it possible to measure distances between points in such difficult to imagine spaces. They tend to be Euclidean representations that faithfully embody the key features of the non-Euclidean geometry. While it is feasible to represent spaces of constant positive curvature (e.g., a sphere or an Ellipsoid) in \( E^n \) (n-dimensional Euclidean space), it is not possible to isometrically represent spaces with constant negative curvature in a Euclidean space of the same dimension \( E^n \) (as shown in Hilbert Theorem, see Do Carmo 1976). However, Negative Curvature spaces can be embedded in Euclidean or pseudo-Euclidean Spaces of higher dimensions, creating models of Negative Curvature space. Therefore, we depend on various models of Negative Curvature space to understand the properties of such curved spaces. Some such models are Poincaré half space, Poincaré disk, and Beltrami-Klein model. These three models represent Negative Curvature spaces in a 2-dimensional unit disk on a Euclidean surface. While all models of Negative Curvature spaces are isometric, in this manuscript we use the hyperboloid model\(^{10}\). The reasons are manifold: the hyperboloid model is an analytic distortion-free model of Negative Curvature space, it can easily be extended to higher dimensions, it can be embedded in a higher dimension pseudo Euclidean space, and all other models of Negative Curvature space can be derived from a hyperboloid model through various stereographic projections. Using a hyperboloid model compared to disk models of hyperbolic space is akin to studying the geometry of earth on a globe versus on a two-dimensional map (Reynolds 1993).

A hyperboloid model of the \( n \)-Dimensional negatively curved space with \( K = -1 \) is described on the surface of a sphere with an imaginary radius \( r = i = \sqrt{-1} \). Although, the hyperboloid model is embedded in a pseudo-Euclidean space, it does not inherit the Euclidean inner product. Instead the inner product on a hyperboloid is the Lorentzian inner product described as \( < x, y >= x_1y_1 + x_2y_2 + \ldots + x_ny_n - x_{n+1}y_{n+1} \). The \( n \)-dimensional hyperbolic space \( H^n \) with \( K = -1 \) is realized as \( (H^n = x \in E^{n,1} : < x, x >= -1, x_{n+1} > 0) \) i.e., of the upper sheet of a unit hyperboloid in Minkowski \( n \)-space, which has a metric tensor \( ds = \sqrt{dx_1^2 + dx_2^2 + \ldots dx_n^2 - dx_{n+1}^2} \) (see Cannon et al. 1997; Reynolds 1993).

\(^{10}\)Due to this analytic model constant, negative curvature spaces with \( |K| = -1 \) are called hyperbolic spaces (Cannon et al. 1997).
Since the hyperboloid model does not inherit the Euclidean inner product, the geodesics (paths of shortest distance between two points, e.g. in a Euclidean space the geodesics are straight lines) on a hyperboloid are calculated differently than on a Euclidean Space. The geodesics in this model are the intersection of the 2-dimensional Euclidean plane through the origin of $E^{n,1}$ with the hyperboloid. Equation (3) is the metric to measure distance in such a space.

The following is a summary of the main takeaways from our review of Riemannian space.

1. The metric used to measure the distance between points informs us about what kind of space it is. For instance, if the metric relationship (distance) among a set of points can be represented by the Euclidean metric, we can conclude that the underlying geometry is Euclidean.

2. The Gaussian Curvature informs us how curved a space is with respect to a flat Euclidean space. It also helps us distinguish among various geometries.

3. The metric used to measure distance varies depending on what geometry we assume to underlie a given set of points (as evident in equations 1, 2, and 3). Therefore, assuming the wrong geometry introduces distortions in estimates of distance.

C Mathematical Details

C.1 Derivation of the distance equation for the hyperboloid model (page 11)

In order to find the geodesic (shortest path) distance between two points on a hyperboloid model we begin with the instance of geodesic distance calculation using a sphere. A sphere is a good starting point since it is also a two-dimensional surface that is embedded in three-dimensional Euclidean space (please see Diacu (2011) for a detailed description of estimating distance on a hyperboloid).

If $R$ is the radius of the sphere and $[x, y, z]$ represent three of its coordinates then the equation of a sphere is given by $x^2 + y^2 + z^2 = R^2$.

For a sphere the geodesic distance between 2 points $p : (x_1, y_1, z_1)$ and $q : (x_2, y_2, z_2)$ follows a great circle and is given by

$$d_s(p, q) = R \theta$$

If $p \odot q$ denotes the Euclidean inner product of $p$ and $q$ (i.e., $p \odot q = x_1x_2 + y_1y_2 + z_1z_2$), then

$$\theta = \cos^{-1}\left(\frac{p \odot q}{|p||q|}\right)$$

Since $|p| = |q| = R$
\[ d_h(p, q) = R \ast Cos^{-1}(\frac{p \circ q}{R^2}) \]

For the hyperboloid model with an imaginary radius, \( R \) is replaced with \( ir \) where \( i = \sqrt{-1} \) (Peebles (1993); Roukema (2001); Wilson et al. (2010)). The equation of a hyperboloid is given by

\[ x^2 + y^2 - z^2 = -r^2 \]  

This assumes two sheets (upper and lower) of the hyperboloid; one sheet above the z-axis and one sheet below it. Since both sheets are identical, only the sheet above is considered in the derivation where \( z > 0 \). If we assume \( r = 1 \) then we have a unit hyperboloid with \( x^2 + y^2 - z^2 = -1 \). Unlike a sphere, a hyperboloid is not embedded in a Euclidean space, instead it is embedded in a space where one axis (here it is the z-axis) is imaginary. Hence, this space is called the pseudo-Euclidean (pE) space.

An important difference in having a (pE) space rather than a Euclidean space is in the inner product, which the surface inherits from the ambient space. If the ambient space is Euclidean then the inner product is the familiar \( x \circ y = x_1x_2 + y_1y_2 + z_1z_2 \) (as we saw with a sphere). However in the case of a hyperboloid the inner product is defined as \( <p.q> = x_1x_2 + y_1y_2 - z_1z_2 \) or geometrically it is given by \( <p.q> = -r^2Cosh(\theta) \), which is the Lorentzian inner product (Ratcliffe 2006). Therefore, the hyperboloid model of constant negative curvature space is also known as the Lorentz model or the Weierstrass model.

As on a sphere, the distance between two points in this space is \( d_h(p, q) = r \ast \theta \) where \( d_h(p, q) \) is the geodesic distance between points p and q on the hyperboloid. After appropriate replacements

\[ d_h(p, q) = r \ast Cosh^{-1}(\frac{<p.q>}{-r^2}) \]  

Since \( -r^2 = \frac{1}{K} \), we get equation (3)

\[ d_h(p, q) = \frac{Cosh^{-1}(K <p.q>)}{\sqrt{-K}} \]

Replacing \( K = -1 \) in the above equation we get equation (2)

\[ d_h(p, q) = Cosh^{-1}(- <p.q>) \]

### C.2 Deriving the distance equation for the inter-temporal decision space (page 15)

Utilizing equation (3) as the basis, we now apply it to an inter-temporal context. Let’s consider a decision-maker who is asked to make a choice between receiving $m_1 after \( t_1 \) delay, versus receiving $m_2 after \( t_2 \) delay (where \( m_2 > m_1 \) and \( t_2 > t_1 \)). We suggest that discounting is not a function of \( t \) (temporal distance) instead it is a function of \( d_h \); the perceived distance between the origin and any money-time bundle in the Negative Curvature decision space. Discounting in a negatively curved space will occur in the following manner. Let’s assume
that $\eta(t)$ and $\phi(m)$ are monotonically increasing functions of $t$ and $m$ respectively. Here $\eta(t)$ denotes the magnitude one perceives of any given time interval and $\phi(m)$ denotes the value one perceives in any given amount of money. If the decision maker is perceiving the time/money bundle of $\phi(m)$ and $\eta(t)$ at point $P$ from the origin, $d_h$ represents the distance of that bundle from the origin (status quo) in the negatively curved decision space. We can calculate $d_h$ in the following manner:

Following equation (24), if the radius of the hyperboloid is $ir$ then the origin $O$ would be at $(0, 0, r)$ and point $P$ would be at $(\phi(m), \eta(t), \sqrt{\phi^2(m) + \eta^2(t) + r^2})$. The Lorentzian inner product would be the following

$$< O.P >= -r\sqrt{\phi^2(m) + \eta^2(t) + r^2}$$

substituting this in (25) we get

$$d_h = r \cdot Cosh^{-1}\left(\frac{\sqrt{\phi^2(m) + \eta^2(t) + r^2}}{r}\right)$$

since $r^2 = \frac{1}{K}$ where $K$ is the Gaussian curvature and is $< 0$

$$d_h = \frac{Cosh^{-1}(\sqrt{-K} \sqrt{\phi^2(m) + \eta^2(t) + \frac{1}{K}})}{\sqrt{-K}}, \quad K < 0$$

This is equation (4). To simplify the equation for the case of $K = -1$ (i.e., unit hyperboloid) yields the following equation

$$d_h = Cosh^{-1}(\sqrt{\phi^2(m) + \eta^2(t) + 1})$$

Further, to present the equations in the manuscript concisely we specify $\Gamma$ (gamma) to be as follows $\Gamma = \sqrt{1 + \phi^2(m) + \eta^2(t)}$. Therefore, we can write the simplified version as

$$d_h = Cosh^{-1}(\Gamma)$$

which is actually equation (5).

C.3 Verification of equation 5 (page 15)

Let’s assume $\phi(m) = 0$ and $\eta(t) = \sqrt{p^2 - 1}$ where $p > 1$. By making appropriate replacements, Equation (5) would then become $d_h = Cosh^{-1}(p)$. Now for $|K| = 1$, if the upper sheet of the hyperboloid is represented by $Q(x) = x_1^2 + x_2^2 - x_3^2 = -1$, we will directly use the metric tensor $ds = \sqrt{dx_1^2 + dx_2^2 - dx_3^2}$ to measure the distance between two points. Let’s assume that the the piecewise path is represented by $f : [a, b] \rightarrow U$ then

$$d_h = \int_a^b \sqrt{Q(f'(t))} dt = \int_a^b \sqrt{\left(\frac{dx_1}{dt}\right)^2 + \left(\frac{dx_2}{dt}\right)^2 - \left(\frac{dx_3}{dt}\right)^2} dt$$
where $f'(t)$ is tangent to $U$ at $f(t)$. Let the origin of the inter-temporal decision space be represented as $[0, 0, 1]$ in a hyperboloid. Assume we want to measure the hyperbolic distance between the origin and a point $A[\sqrt{p^2 - 1}, 0, p]$. The line segment joining the origin and $A$ can be parametrized by a path $\xi : [1, r] \rightarrow U$, $\xi(t) = [\sqrt{t^2 - 1}, 0, t]$. Then

$$d_h = \int_1^p \sqrt{\frac{t^2}{t^2 - 1} - 1} dt = \int_1^p \frac{1}{\sqrt{t^2 - 1}} dt = \ln[p + \sqrt{p^2 - 1}] = \text{Cosh}^{-1}[p]$$

which is similar to the value we obtained from equation (5), hence, verifying it.

C.4 Solution of Equation 11 (page 17)

If $\psi$ is the fraction of money lost per unit of distance travelled on the Negative Curvature surface, the discounted value of money can be calculated by using the following differential equation

$$\frac{dy}{dd_h} = -\psi y$$

We can solve this equation by means of a straightforward Laplace transformation.

If $\mathcal{L}[f(x)] = \int_0^\infty e^{-sx} f(x) dx$ represents the Laplace transformation of $f(x)$ for some $s$ and $y(0)$ represents undiscounted amount of money at $d_h = 0$. Then equation (11) can be transformed as

$$(s + \psi)\mathcal{L}[y] - y(0) = 0$$

$$\mathcal{L}[y] = \frac{y(0)}{s + \psi}$$

From Laplace transformation table, $\mathcal{L}[e^{-ax}] = \frac{1}{s+a}$. Therefore,

$$\mathcal{L}[y] = y(0)\mathcal{L}[e^{-\psi d_h}]$$

Therefore, the discounted value of money $y(0)$ after traveling $d_h$ distance is

$$y(d_h) = y(0)e^{-\psi d_h}$$

If we use the functional form $\phi(y)$

$$\phi(y_{d_h}) = \phi(y_0)e^{-\psi d_h}$$

---

11 Please see Anderson (2005, pages 196-208) for more details.
D Testing the Algorithms

We tested the proposed algorithms by generating inter-point distances between 6 distinct points. We generated 2 sets of distances. One set of distances were the Negative Curvature distances on a surface with constant Gaussian curvature ($K$) of $-5$. The other set had Euclidean distances.

Negative Curvature Inter-point Distances:

\[
\begin{bmatrix}
0 & 0.7999 & 1.4 & 0.2061 & 0.3282 & 0.4065 \\
0.7999 & 0 & 1.762 & 1.006 & 0.8965 & 1.201 \\
1.4 & 1.762 & 0 & 1.493 & 1.09 & 1.576 \\
0.2061 & 1.006 & 1.493 & 0 & 0.4031 & 0.2117 \\
0.3282 & 0.8965 & 1.09 & 0.4031 & 0 & 0.5011 \\
0.4065 & 1.201 & 1.576 & 0.2117 & 0.5011 & 0
\end{bmatrix}
\]

Euclidean Inter-point Distances:

\[
\begin{bmatrix}
0 & 0.6762 & 1.4 & 0.1745 & 0.3244 & 0.341 \\
0.6762 & 0 & 1.691 & 0.8166 & 0.8732 & 0.976 \\
1.4 & 1.691 & 0 & 1.488 & 1.087 & 1.556 \\
0.1745 & 0.8166 & 1.488 & 0 & 0.4026 & 0.1684 \\
0.3244 & 0.8732 & 1.087 & 0.4026 & 0 & 0.5009 \\
0.341 & 0.976 & 1.556 & 0.1684 & 0.5009 & 0
\end{bmatrix}
\]

Our aim was two-fold: first, we wanted to test the efficacy of the proposed algorithms. Second, we wanted to understand what is the optimal number of points that would accurately recover the Gaussian curvature from data. To achieve these objectives, we sampled different subsets of distances from the above matrices and obtained best Negative Curvature and Euclidean solutions. First, we sampled 4 out of 6 points i.e., the interpoint distances from the first 4 rows and columns of the matrices. Then we sampled 5 out of 6 points and finally, all 6 points. The following table summarizes how each algorithm performed.

| Distances          | Points Sampled | Algorithm              | Fit [RMSE/SD]          | Estimated Gaussian Curvature $|K|$ |
|--------------------|----------------|------------------------|------------------------|----------------------------|
| Negative Curvature | 4              | Negative Curvature     | $2.37 \times 10^{-7}$  | 4.288                      |
|                    |                | Euclidean              | .01                    | -                          |
|                    | 5              | Negative Curvature     | $1.96 \times 10^{-8}$  | 5                          |
|                    |                | Euclidean              | .022                   | -                          |
|                    | 6              | Negative Curvature     | $2.36 \times 10^{-8}$  | 5                          |
|                    |                | Euclidean              | .034                   | -                          |
These results show that:

1. For Negative Curvature distances, the Negative Curvature algorithm performed better (i.e., offered a better fit) than the Euclidean algorithm.

2. By sampling distances between 4 points, the Negative Curvature algorithm was not able to accurately recover the Gaussian curvature. However, it recovered the Gaussian curvature well with 5 or 6 points.

3. For Euclidean distance, both the Negative Curvature and Euclidean algorithms performed well.

4. As we discussed earlier, for Euclidean spaces, the Gaussian curvature is zero. We can see from the output of the algorithm that for Euclidean distances, regardless of the points sampled, the Negative Curvature algorithm always recovered near zero values of the Gaussian curvature.

## E Results

### E.1 Study-1

| Participant # | RMSE/ Standard Dev (Negative Curvature) | RMSE/ Standard Dev (Euclidean) | Gaussian Curvature $|K|$ | % improvement over Euclidean fit |
|---------------|----------------------------------------|--------------------------------|-----------------------------|--------------------------------|
| 1             | 0.107                                  | 0.107                          | $8.175 \times 10^{-9} \sim 0$ | 0                             |
| 2             | 0.036                                  | 0.576                          | 36.210                      | 93.820                        |
| 3             | 0.154                                  | 0.362                          | 17.714                      | 57.504                        |
| 4             | 0.003                                  | 0.003                          | 0.080                       | 0.936                         |
| 5             | 0.158                                  | 0.363                          | 15.714                      | 56.566                        |
| 6             | 0.074                                  | 0.418                          | 156.683                     | 82.319                        |
| 7             | 0.096                                  | 0.115                          | 14.388                      | 16.387                        |
### E.2 Study-2

| Participant # |  |  |  |  |
|---------------|---------|-----------------|----------------|----------------|
| 8 | 0.083 | 0.246 | 53.843 | 66.386 |
| 9 | 0.131 | 0.305 | 12.924 | 57.090 |
| 10 | 0.260 | 0.253 | 63.506 | -2.461 |
| 11 | 0.115 | 0.276 | 84.877 | 58.230 |
| 12 | 0.171 | 0.286 | 31.042 | 40.321 |
| 13 | 0.104 | 0.119 | 35.359 | 12.287 |
| 14 | 0.060 | 0.353 | 106.052 | 82.979 |
| 15 | 0.101 | 0.128 | 122.127 | 21.008 |
| 16 | 0.068 | 0.426 | 1.691 | 84.115 |
| 17 | 0.180 | 0.278 | 660.030 | 35.323 |
| 18 | 0.066 | 0.066 | 6.615 x 10^{-10} \sim 0 | 0.0001 |
| 19 | 0.155 | 0.428 | 63.875 | 63.784 |
| 20 | 0.067 | 0.181 | 53.128 | 62.997 |
| 21 | 0.052 | 1.009 | 42.175 | 94.854 |
| 22 | 0.356 | 0.382 | 1.435 x 10^{-7} \sim 0 | 6.682 |
| 23 | 0.194 | 0.194 | 9.934 x 10^{-9} \sim 0 | 0 |
| 24 | 0.175 | 0.631 | 75.538 | 72.188 |
| 25 | 0.050 | 0.210 | 6.357 | 76.068 |
| 26 | 0.109 | 0.109 | 0.177 | 0.015 |
| 27 | 0.088 | 0.382 | 20.174 | 77.027 |
| 28 | 0.160 | 0.201 | 40.125 | 20.187 |
| 29 | 0.060 | 0.079 | 1.452 x 10^{-8} \sim 0 | 23.836 |
| 30 | 0.210 | 0.267 | 30.506 | 21.349 |
| 31 | 0.124 | 0.124 | 1.926 x 10^{-8} \sim 0 | 0 |
| 32 | 0.155 | 0.664 | 66.848 | 76.688 |
| 33 | 0.114 | 0.114 | 2.525 x 10^{-9} \sim 0 | -0.001 |
| 34 | 0.195 | 0.467 | 53.813 | 58.260 |
| 35 | 0.137 | 0.137 | 3.184 x 10^{-9} \sim 0 | 0 |
|   |     |     |     |     |
|---|-----|-----|-----|-----|
| 8 | 0.030 | 0.300 | 0.665 | 90.152 |
| 9 | 0.059 | 0.220 | 152.212 | 73.227 |
| 10 | 0.530 | 0.533 | 360.720 | 0.600 |
| 11 | 0.145 | 0.151 | 596.090 | 3.958 |
| 12 | 0.279 | 0.297 | 2.408 | 5.949 |
| 13 | 0.550 | 0.551 | 66.231 | 0.184 |
| 14 | 0.352 | 0.238 | 308.575 | -47.898 |
| 15 | 0.063 | 0.163 | 7.934 | 61.571 |
| 16 | 0.211 | 0.206 | 2.482 | -2.168 |
| 17 | 0.047 | 0.058 | 30.248 | 19.113 |
| 18 | 0.067 | 0.398 | 26.643 | 83.294 |
| 19 | 0.070 | 0.194 | 44.622 | 63.763 |
| 20 | 0.193 | 0.221 | 1.186 | 12.544 |
| 21 | 0.345 | 0.447 | 0.961 | 22.765 |
| 22 | 0.256 | 0.500 | 45.399 | 48.822 |
| 23 | 0.020 | 0.021 | 0.404 | 4.035 |
| 24 | 0.280 | 0.394 | 1.180 | 28.915 |
| 25 | 0.162 | 0.193 | 0.598 | 16.443 |
| 26 | 0.071 | 0.222 | 1.472 | 68.094 |
| 27 | 0.182 | 0.186 | 1.124 | 2.259 |
| 28 | 0.381 | 0.381 | 2.913 | 0.000 |
| 29 | 0.156 | 0.414 | 28.160 | 62.392 |
| 30 | 0.184 | 0.184 | 3.549 | 0.037 |
| 31 | 0.129 | 0.153 | 3.547 | 15.597 |
| 32 | 0.184 | 0.203 | 2.408 | 9.116 |
| 33 | 0.012 | 0.858 | 15.503 | 98.652 |
| 34 | 0.011 | 0.045 | 1.615 | 76.217 |
| 35 | 0.151 | 0.159 | 19.018 | 5.045 |
| 36 | 0.515 | 0.536 | 876.980 | 3.868 |
| 37 | 0.027 | 0.171 | 8.420 | 84.495 |
| 38 | 0.123 | 0.146 | 0.563 | 15.592 |
E.3 Robustness Check

Figure 1: Study 1 (Euclidean vs. Negative Curvature Algorithm fit)

Figure 2: Study 2 (Euclidean vs. Negative Curvature Algorithm fit)
E.4 Prediction in the loss and gain domain

If $L$ and $G$ represent discount functions in loss and gain domains respectively, i.e.,

$$L = e^{-\cosh^{-1}\left(\sqrt{-K_L}\sqrt{\phi^2(m)+\eta^2(t)}-\frac{1}{K_L}\right)}$$

$$G = e^{-\cosh^{-1}\left(\sqrt{-K_G}\sqrt{\phi^2(m)+\eta^2(t)}-\frac{1}{K_G}\right)}$$

Then inequality (20) can be written as

$$\frac{-\partial L}{\partial t} > \frac{-\partial G}{\partial t}$$

(26)

$$\frac{-\partial L}{L} = \frac{\eta'(t)}{\sqrt{\Omega_L} - \sqrt{-1 + \sqrt{-K_L}\sqrt{\Omega_L}}\sqrt{1 + \sqrt{-K_L}\sqrt{\Omega_L}}}$$

where $\Omega_L = -\frac{1}{K_L} + \phi^2(m) + \eta^2(t)$

$$\frac{-\partial G}{G} = \frac{\eta'(t)}{\sqrt{\Omega_G} - \sqrt{-1 + \sqrt{-K_G}\sqrt{\Omega_G}}\sqrt{1 + \sqrt{-K_G}\sqrt{\Omega_G}}}$$

$$\Omega_G = -\frac{1}{K_G} + \phi^2(m) + \eta^2(t)$$

By substituting these values, inequality (20) becomes

$$\frac{1}{\sqrt{\Omega_G}\sqrt{-K_G}\Omega_G - 1} > \frac{1}{\sqrt{\Omega_G}\sqrt{-K_G}\Omega_G - 1}$$

(27)

This inequality would hold only if $K_G < K_L$ where both $K_G$ and $K_L < 0$. 
References

Abdellaoui, M., Attema, A. E., and Bleichrodt, H. (2010). Intertemporal tradeoffs for gains and losses: An experimental measurement of discounted utility. *The Economic Journal*, 120(545):845–866.

Ainslie, G. (1975). Specious reward: a behavioral theory of impulsiveness and impulse control. *Psychological Bulletin*, 82(4):463–496.

Akerlof, G. A. (1991). Procrastination and obedience. *The American Economic Review*, 81(2):1–19.

Anderson, J. W. (2005). *Hyperbolic Geometry*. Springer.

Attema, A. E., Bleichrodt, H., Rohde, K. I., and Wakker, P. P. (2010). Time-tradeoff sequences for analyzing discounting and time inconsistency. *Management Science*, 56(11):2015–2030.

Berger, M. (2003). *A Panoramic View of Riemannian Geometry*. Springer.

Cannon, J. W., Floyd, W. J., Kenyon, R., and Parry, W. R. (1997). Hyperbolic geometry. *Flavors of Geometry*, 31:59–115.

Chapman, G. B. (2000). Preferences for improving and declining sequences of health outcomes. *Journal of Behavioral Decision Making*, 13(2):203–218.

Chung, S.-H. and Herrnstein, R. J. (1967). Choice and delay of reinforcement1. *Journal of the Experimental Analysis of Behavior*, 10(1):67–74.

Coxeter, H. S. M. (1998). *Non-Euclidean Geometry*. Cambridge University Press.

Dasgupta, P. (2006). *Comments on the Stern Review’s Economics of Climate Change*. University of Cambridge.

Diacu, F. (2011). Relative equilibria in the 3-dimensional curved n-body problem. *arXiv preprint arXiv:1108.1229*.

Do Carmo, M. P. (1976). *Differential Geometry of Curves and Surfaces*, volume 2. Prentice-Hall Englewood Cliffs.

Do Carmo, M. P. (1992). *Riemannian Geometry*. Springer.

Einstein, A. (1916). Die grundlage der allgemeinen relativitätstheorie. *Annalen der Physik*, 354(7):769–822.

Faber, R. L. (1983). *Foundations of Euclidean and non-Euclidean Geometry*. Marcel Dekker New York.

Fernandez, J. M. and Farell, B. (2009). Is perceptual space inherently non-euclidean? *Journal of mathematical psychology*, 53(2):86–91.
Frederick, S., Loewenstein, G., and O’donoghue, T. (2002). Time discounting and time preference: A critical review. *Journal of economic literature*, pages 351–401.

Green, B. (1999). *The Elegant Universe: Superstrings, hidden dimensions, and the quest for the ultimate theory*. Nueva York: Norton.

Green, L., Fristoe, N., and Myerson, J. (1994). Temporal discounting and preference reversals in choice between delayed outcomes. *Psychonomic Bulletin & Review*, 1(3):383–389.

Green, L. and Myerson, J. (1996). Exponential versus hyperbolic discounting of delayed outcomes: Risk and waiting time. *American Zoologist*, 36(4):496–505.

Greenburg, M. J. (1993). *Euclidean and non-Euclidean Geometries: Development and History*. New York: WH Freemand and Compnay.

Hastie, R. and Dawes, R. (2001). *Rational choice in an uncertain world: The psychology of judgment and decision making*. Thousand Oaks, CA: Sage.

Helmholtz, H. V. (1876). The origin and meaning of geometrical axioms. *Mind*, (3):301–321.

Indow, T. (1979). Alleys in visual space. *Journal of Mathematical Psychology*, 19(3):221–258.

Jehle, G. A. and Reny, P. J. (2006). *Advanced Microeconomic Theory, 2/e*. Addison Wesley.

Jevons, W. S. (1879). *The Theory of Political Economy*. Macmillan and Co.

Kant, I. (1902). *Prolegomena: To Any Future Metaphysics that Can Qualify as Science*. Number 53. Open Court Publishing.

Kant, I. (1908). *Critique of Pure Reason, 1781*.

Kirkpatrick, S., Jr., D. G., and Vecchi, M. P. (1983). Optimization by simulated annealing. *Science*, 220(4598):671–680.

Laibson, D. (2003). Intertemporal decision making. *Encyclopedia of Cognitive Science*, 1:915–919.

Lee, J. M. (2012). *Introduction to smooth manifolds*, volume 218. Springer.

Lindman, H. and Caelli, T. (1978). Constant curvature riemannian scaling. *Journal of Mathematical Psychology*, 17(2):89–109.

Loewenstein, G. and Prelec, D. (1992). Anomalies in intertemporal choice: Evidence and an interpretation. *The Quarterly Journal of Economics*, 107(2):573–597.

Loewenstein, G. and Sicherman, N. (1991). Do workers prefer increasing wage profiles? *Journal of Labor Economics*, pages 67–84.

Luhmann, C. C. (2013). Discounting of delayed rewards is not hyperbolic. *Journal of experimental psychology: learning, memory, and cognition*, 39(4):1274.
Luneburg, R. K. (1947). *Mathematical Analysis of Binocular Vision*. Princeton University Press.

Mas-Colell, A., Whinston, M. D., and Green, J. R. (1995). *Microeconomic Theory*. Oxford University Press New York.

Mischel, W. and Moore, B. (1973). Effects of attention to symbolically presented rewards on self-control. *Journal of Personality and Social Psychology*, 28(2):172.

Needham, T. (1998). *Visual Complex Analysis*. Oxford University Press.

Peebles, P. J. E. (1993). *Principles of Physical Cosmology*. Princeton University Press.

Penrose, R. (2005). *The Road to Reality: A Complete Guide to the Laws of the Universe*. New York: AA Knopf.

Rachlin, H. (2006). Notes on discounting. *Journal of the Experimental Analysis of Behavior*, 85(3):425–435.

Rae, J. (1834). *The Sociological Theory of Capital*. The Macmillan Company.

Ratcliffe, J. G. (2006). *Foundations of Hyperbolic Manifolds*, volume 149. Springer.

Read, D. (2001). Is time-discounting hyperbolic or subadditive? *Journal of risk and uncertainty*, 23(1):5–32.

Reynolds, W. F. (1993). Hyperbolic geometry on a hyperboloid. *The American Mathematical Monthly*, 100(5):442–455.

Riemann, B. (1854). *Ueber die Hypothesen, welche der Geometrie zu Grunde liegen*. J. Springer (Published in 1921).

Roukema, B. F. (2001). On the comoving distance as an arc-length in four dimensions. *Monthly Notices of the Royal Astronomical Society*, 325(1):138–142.

Samuelson, P. A. (1937). A note on measurement of utility. *The Review of Economic Studies*, 4(2):155–161.

Shavitt, Y. and Tankel, T. (2004). On the curvature of the internet and its usage for overlay construction and distance estimation. In *INFOCOM 2004. Twenty-third Annual Joint Conference of the IEEE Computer and Communications Societies*, volume 1. IEEE.

Stevens, S. S. (1957). On the psychophysical law. *Psychological review*, 64(3):153–181.

Stewart, N., Reimers, S., and Harris, A. J. (2014). On the origin of utility, weighting, and discounting functions: How they get their shapes and how to change their shapes. *Management Science*, (Forthcoming).

Stoker, J. J. (1969). *Differential Geometry*. Wiley-Interscience, New York.
Strotz, R. H. (1955). Myopia and inconsistency in dynamic utility maximization. *The Review of Economic Studies*, 23(3):165–180.

Takahashi, T. (2005). Loss of self-control in intertemporal choice may be attributable to logarithmic time-perception. *Medical Hypotheses*, 65(4):691–693.

Thurston, W. P. and Levy, S. (1997). *Three-dimensional Geometry and Topology*, volume 1. Oxford University Press.

Wakker, P. and Deneffe, D. (1996). Eliciting von neumann-morgenstern utilities when probabilities are distorted or unknown. *Management Science*, 42(8):1131–1150.

Weyl, H. (1922). *Space, Time, Matter*. Dutton.

Wilson, R. C., Hancock, E. R., Pekalska, E., and Duin, R. P. (2010). Spherical embeddings for non-euclidean dissimilarities. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 1903–1910. IEEE.