Investment Decisions and Falling Cost of Data Analytics[*]

Jussi Keppo[†], Hong Ming Tan[‡], Chao Zhou[§]

March 24, 2018

Abstract

This paper studies investment decisions and their data analytics. We show that financially constrained or highly risk-averse investors use less data analytics. We also show that the demand of data analytics is highest for investment opportunities with high expected returns and the demand is either high or zero for opportunities with low expected returns. Furthermore, the falling cost of data analytics raises investors' leverage, which leads to higher losses during the crises.

[*]We have benefited from comments at the 2017 Supply Chain Finance & Risk Management Workshop (Washington University in St. Louis), Fifth Asian Quantitative Finance Conference (Seoul), Eleventh Annual Risk Management Conference (Singapore), 2017 INFORMS Annual Meeting, Leeds School of Business (University of Colorado Boulder), Rotman School of Management (University of Toronto), MIT, Bank of Finland, National University of Singapore, and Institut Henri Poincaré, and more specifically from Ashish Agarwal, Crystal Abidin, Vlad Babich, Diego Garcia, Goh Yan Han, Jukka Isohatala, Steven Kou, Kai Larsen, Pauli Murto, Clark Pixton, David Simchi-Levi, Lones Smith, Juuso Valimaki, Zhiwen Wang, and Chen Yang. All errors are ours.

[†]Business School, Institute of Operations Research and Analytics, and Institute of Data Science, National University of Singapore. Email: keppo@nus.edu.sg

[‡]Institute of Operations Research and Analytics, National University of Singapore. Email: thm@u.nus.edu

[§]Department of Mathematics, National University of Singapore. Email: matzc@nus.edu.sg

1 Introduction

Information reduces uncertainty and enables better decisions. Therefore, companies and individuals are willing to pay for data and its analyses (we call this as data analytics). For instance, companies invest in information systems and build analytics teams, they pay market research firms to establish the likelihood that a new product is well-received in the market, and they pay recruitment agencies to find and collect information on prospective employees. New technologies, digitalization of data, and data analytics methods have decreased the cost of information substantially during the last few decades (see e.g. Goldfarb and Tucker 2017). Figure 1 below illustrates this trend (storing data is one part of data analytics, and it has become substantially cheaper). In this paper, we study how the cost of data analytics affects firms and individuals' investment decisions, and which type of investors and investment opportunities benefit the most from the declining cost of data analytics.

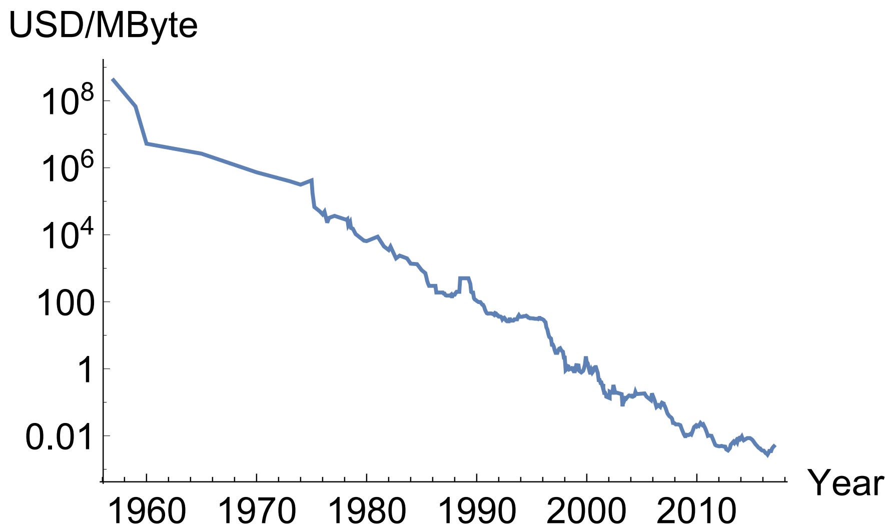


Figure 1: **Cost of data storage.** Y-axis is US$ per megabyte in a log scale. Data is from McCallum (2017).

We develop a model where a risk averse decision maker faces a joint decision of the level of data analytics and investment amount. More specifically, we model data acquisition, collection, and analytics with Bayesian learning. The decision maker has a CARA (or exponential) utility function and she has an option to buy data and analyze that before her investment decision. Information is costly due to the price of datasets and efforts in collecting and processing the data. The more the investor buys and analyzes data, the more signals she receives on the payoff of the investment opportunity. After the data analytics, the decision maker decides how much to invest and, if needed, she uses leverage in the investment.

We derive several key results. First, the more data analytics the decision maker acquires, the more she invests on average. This is because the analytics lowers the riskiness of the investment and, thus, unconditionally on the analytics outcome, it makes the investment opportunity more attractive. This indicates that data analytics is a risk management device. Second, the lower the cost of data analytics, naturally the more the decision maker acquires analytics, which decreases the riskiness of the investment opportunity and this way raises the investment amount and leverage. We show that the higher leverage, driven by the falling cost of data analytics, leads to higher losses during the crises. On the other hand, if the decision maker cannot take leverage then obviously she invests less, but also uses less data analytics. That is, financially constrained firms have less incentive to learn before their investments. Third, the lower the risk aversion of the decision maker, the more she acquires data analytics. This is because a decision maker with low risk aversion invests more in the risky investment opportunity and, therefore, she has a higher incentive to collect new information on the opportunity. Fourth, the demand of data analytics is highest with high expected return opportunities, which is surprising since the value of data analytics is highest for mediocre investment opportunities. With the high expected return opportunities the decision maker uses a high level of leverage and, therefore, she decides to thoroughly analyze the investment opportunity. Fifth, with low expected return opportunities the decision maker acquires either a lot of data analytics or not at all, so in this case there is lumpiness in the demand of data analytics.

Our model applies to, e.g., companies' real asset investments, private equity and venture capital firms' investment decisions, and mutual funds' portfolio decisions on publicly traded assets with a rare event risk that cannot be learned from the market for free. In these examples, the decision makers may acquire data analytics on the investment opportunities, and then invest according to their forecasts on the investment payoffs and the precisions of the forecasts.

We utilize the value of information model in Keppo et al. (2008), which is related to Radner and Stiglitz (1984), Moscarini and Smith (2001), Chade and Schlee (2002), Moscarini and Smith (2002), and Xie et al. (2016). In contrast to these papers, we also optimize the investment level, which is not a binary decision. In Kihlstrom (1974), a consumer faces a linear price for precision of a Gaussian signal given a Gaussian prior. For a hyperbolic utility function, he can write utility as a function of signals and avoid computing the density of posteriors. Our theory is not driven by such linearity, and we consider an investment decision with borrowing and lending. Our work is also related to Grossman and Stiglitz (1980), where traders maximize CARA utility and may buy information on the payoff of a stock, which is traded in a competitive market. The traders' private information is only partially observed from the equilibrium price as the equilibrium contains some uninformed traders. Verrecchia (1982) extends that model to include diverse information acquisition and derives an equilibrium in such a market. By utilizing Grossman and Stiglitz (1980) and Verrecchia (1982), Peress (2004) shows that as long as absolute risk aversion falls in wealth, there are rising returns to acquiring private information despite being revealed by public signals. Dugast and Foucault (2017) show that data abundance raises asset price informativeness in the short run but not necessarily in the long run, because profits from trading on more precise signals fall. Our paper is also related to studies that consider portfolio optimization when the expected returns of risky assets are unknown, but can be learned from the realized returns of the assets (e.g. Lakner 1998, Karatzas and Zhao 2001, Guéant and Pu 2017), as well as learning models, for instance, in experimental design (e.g. Chick et al. 2017), healthcare (e.g. Chick et al. 2016, Negoescu et al. 2017), and dynamic pricing (e.g. Afèche and Ata 2013, Yu et al. 2015, Cheung et al. 2017). In contrast to these studies, in the present paper we optimize the quantity of information and investment amount with borrowing and lending, and focus on the falling cost of information.

The rest of the paper is organized as follows: Section 2 introduces the decision problem and the quantity of data analytics. Section 3 derives the value of data analytics and Section 4 solves the optimal quantity of data analytics. Section 5 analyzes the optimal investment. Section 6 introduces a financial constrained DM, and finally, Section 7 concludes. Appendixes A and B extend the model to consider a nonlinear cost of data analytics and multiple investment opportunities, and show that our results are robust with respect to these extensions.

2 Model

We assume a two period model, and that the decision maker (DM) has a CARA utility function $u(\kappa) = -\exp(-\gamma\kappa)$, where $\gamma > 0$ is the risk aversion parameter. Hence, $u'(\kappa) = \gamma \exp(-\gamma\kappa) > 0$, $u''(\kappa) = -\gamma^2 \exp(-\gamma\kappa) < 0$, and the absolute values of these derivatives fall in κ. The DM has initial wealth κ and decides the quantity of data analytics (t), the level of risky investment ($w \geq 0$), and lending or borrowing amount. Quantity t of data analytics costs ct, so c is the marginal cost. Each unit of the risky investment has an initial sunk cost of \$1 and a payoff θ which is unknown to the DM. When $\kappa - w - ct < 0$, the DM uses leverage. The borrowing and lending rates are both equal to $r > 0$. Hence, borrowing \$1 gives a payoff of $\$ - e^r$ in the next period. We assume that both risky invesment and debt payoffs are discounted present values (r and θ can be calibrated to give the present value payoffs). Therefore, the DM's utility from the profit and loss in the next period is $-\exp\left(-\gamma\left[w(\theta-1) + (\kappa - w - ct)(e^r - 1)\right]\right)$.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, where Ω is a set, $\mathcal{F}$ a σ-algebra of subsets of Ω, and $\mathbb{P}$ a probability measure on $\mathcal{F}$. The investment payoff θ is constant but unknown, and this probability space captures all the uncertainties related to the payoff θ. The quantity of data analytics t corresponds to observing a noisy signal Z_t on payoff θ from 0 to t (note that time is fixed, t is quantity). The noisy signal follows $dZ_t = \theta dt + dB_t$, where B_t is a standard Wiener process.[1] $\mathcal{F}_t$ is the information level from observing the noisy signal Z_t up to quantity t. Thus, the DM can learn about the payoff θ by obtaining quantity $t \geq 0$ of data analytics at a marginal cost $c > 0$, i.e., t quantity of data analytics costs ct. Therefore, quantity t lowers the DM's initial wealth from κ to $\kappa - ct$, and the leverage becomes $\max\{w - \kappa + ct, 0\}$. Under quantity t, the expected payoff $\hat{\theta}_t = E[\theta|\mathcal{F}_t]$ and variance $s_t = E[(\theta - \hat{\theta}_t)^2|\mathcal{F}_t]$, where $\hat{\theta}_t$ and s_t are given by Kalman-Bucy filter (see Øksendal 2003). If the DM acquires analytics quantity $t \geq 0$ then the payoff can be written as follows (the proof is in Appendix C.1).

Lemma 1 (Unconditional and conditional payoffs)
Let the DM obtain analytics quantity $t \geq 0$. Unconditionally on the signal outcome, i.e., with respect to information $\mathcal{F}_0$, we have

$$\theta = \hat{\theta}_0 + \sqrt{\frac{s_0^2 t}{1 + s_0 t}}\epsilon_t + \sqrt{s_t}\epsilon,$$

and conditional on the signal outcome, i.e., with respect to information $\mathcal{F}_t$, we have

$$\theta = \hat{\theta}_t + \sqrt{s_t}\epsilon,$$

[1] If we used discrete time signals to model the quantity of information then the quantity would naturally be an integer (number of signals) and the optimization of the quantity would be more complicated.

where $\hat{\theta}_0$ and s_0 are the DM's prior mean and variance, posterior variance $s_t = \frac{s_0}{1+ts_0}$, ϵ_t is $\mathcal{F}_t$-measurable standard normal random variable, ϵ is a standard normal variable with respect to informations $\mathcal{F}_0$ and $\mathcal{F}_t$, ϵ_t and ϵ are independent with respect to information $\mathcal{F}_0$.

By Lemma 1, initially the payoff θ follows a normal distribution with mean $\hat{\theta}_0$ and variance s_0. With analytics quantity $t > 0$, variance $s_t < s_0$, i.e., analytics decreases the uncertainty of the payoff. This means that the data analytics is a risk management device. After the DM learns the uncertainty term ϵ_t through observing the signal Z_t (from 0 to t), the DM has belief $\hat{\theta}_t$ on the payoff, but the uncertainty term ϵ still remains. The effect of ϵ falls as the DM learns more ($t \to \infty$) since s_t vanishes at infinity. That is, the DM learns the true payoff with no uncertainty at infinite quantity of analytics.

In the first period, information is first bought and analyzed, and then the investment amount and its funding (own money and leverage) are decided. In the second period, the investment payoff is realized and the loans (if any) are paid back. With respect to the first period's information, the payoff is a normally distributed random variable. Thus, decisions (in the first period) are quantity of data analytics (variable t), which gives information on the investment payoff, and the investment level (variable w), which with the quantity t gives the leverage. We illustrate the structure of the model in Figure 2.

Now we can write the DM's optimization problem as follows

$$\sup_{t\geq 0} E\left[\sup_{w\geq 0} E\left[-\exp\left(-\gamma\left[w(\theta-1)+(\kappa-ct-w)(e^r-1)\right]\right)|\mathcal{F}_t\right]\right], \tag{1}$$

where $\gamma > 0$, $c > 0$, and $r > 0$. Thus, the DM maximizes her expected utility from profit and loss by acquiring analytics and making the investment decision conditional on the results of the analytics.

3 Value of data analytics

Data analytics has value if it raises the DM's expected utility. So, the value of data analytics is the difference between the expected utilities with a positive quantity of data analytics and with zero quantity of data analytics, and if this difference is positive then the data analytics brings value to the DM. Note that presumably the value of data analytics is positive since it lowers the uncertainty

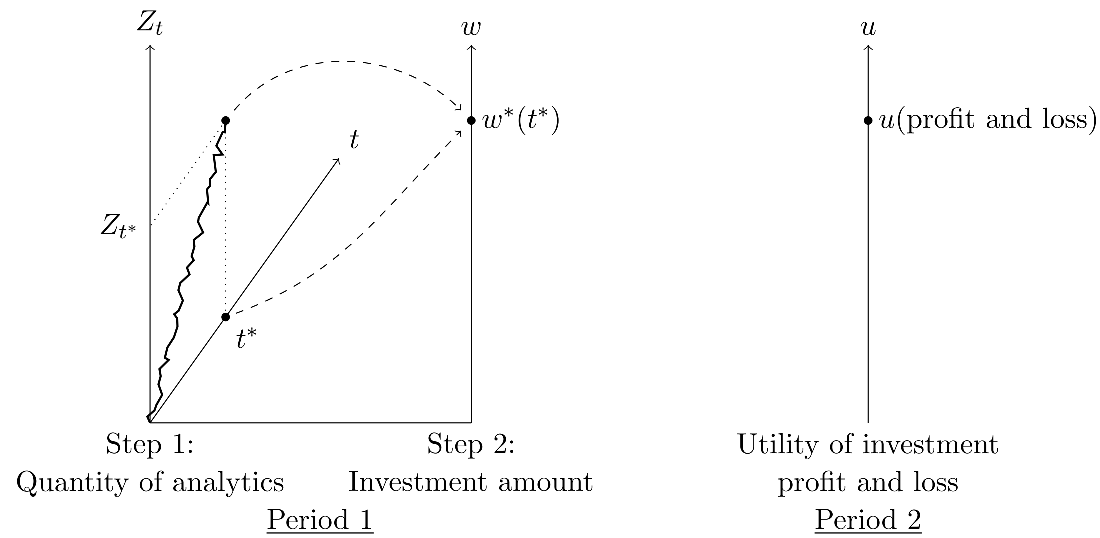


Figure 2: **Model structure.** In period 1, the DM performs two steps. First, the DM optimizes the quantity of data analytics t^*. This corresponds to observing the noisy signal Z_t from 0 to t^*. Second, the DM chooses the investment amount $w^*(t^*)$ based on the realized signal from 0 to t^*. In period 2, the investment profit and loss is realized, which gives utility $u(\text{profit and loss}) = u\left(w^*(t^*)[\theta - 1] + [\kappa - ct^* - w^*(t^*)][e^r - 1]\right)$, see (1).

of the payoff and the DM is risk averse. Thus, if the cost of data analytics is not too high, the DM uses analytics as a risk management device.

3.1 Optimal investment

Recall that we consider a two-period model, where the DM has CARA utility function in (1) with the coefficient of absolute risk aversion γ and initial wealth κ. We first give the optimal investment level of the DM for a given analytics quantity t in the following lemma.

Lemma 2 (Optimal investment)

Given analytics quantity t and CARA utility function with $\gamma > 0$, the optimal investment level of the risky investment is given by

$$w^*(t) = \max\left\{\frac{\hat{\theta}_t - e^r}{\gamma s_t}, 0\right\}, \tag{2}$$

where r is the borrowing and lending rate, $\hat{\theta}_t$ and s_t are the posterior expected payoff and variance.

Proof. From (1) we get the maximization problem:

$$\begin{aligned}&\sup_{w\geq 0} E\left[-\exp(-\gamma[w(\theta-1)+(\kappa-ct-w)(e^r-1)])|\mathcal{F}_t\right]\\&=\sup_{w\geq 0}-\exp\left(-\gamma w(\hat{\theta}_t-1)-\gamma(\kappa-ct-w)(e^r-1)+\frac{\gamma^2 w^2}{2}s_t\right),\end{aligned}$$

where the expected utility is strictly concave in w. The first order condition gives

$$-\gamma(\hat{\theta}_t-1)+\gamma(e^r-1)+\gamma^2 w s_t=0$$

and, therefore, $w=\frac{\hat{\theta}_t-e^r}{\gamma s_t}$. By the strict concavity of the objective function and the constraint on w, we get (2). □

By (2), the higher the expected payoff $\hat{\theta}_t$, the more the DM invests in the risky opportunity. Furthermore, the higher the uncertainty s_t, the less the DM invests. A more risk averse investor invests less, as expected. When t rises, s_t decreases and, unconditionally on the signal outcome, the DM invests more. So, the more the DM uses analytics on the risky opportunity, the lower the risk and the more she invests on average.

Note also that, as e.g. in Merton (1969) and Grossman and Stiglitz (1980), due to the CARA utility the optimal investment (2) is independent of wealth κ, which means that the DM's leverage level ($\max\{w^*(t)+ct-\kappa,0\}$) is given by κ. We address this issue in section 6, where we consider the case of financial constraint. That is, in that section the leverage is bounded and we show that some of our results are driven by leverage.

3.2 Value and marginal value of data analytics

Substituting the optimal investment into the utility function in (1) gives the expected utility for any quantity t of data analytics. This is given in the following lemma and the proof is in Appendix C.2.

Lemma 3 (Expected utility)

Given the quantity of analytics t, the expected utility of the optimal investment before observing the *outcome of the analytics is given by*

$$
\begin{aligned}
u^*(t) =& E\left[-\exp\left(-\gamma\left[w^*(t)(\theta-1)+(\kappa-ct-w^*(t))(e^r-1)\right]\right)\right] \\
=& -e^{\gamma(\kappa-ct)(1-e^r)}\Phi(d(t)) - \frac{\Phi\left(\frac{-d(t)}{\sqrt{s_0t+1}}\right)\exp\left(-\frac{\hat{\theta}_0^2+e^{2r}-2\hat{\theta}_0e^r+2\gamma s_0(\kappa-ct)(e^r-1)}{2s_0}\right)}{\sqrt{s_0t+1}},
\end{aligned}
$$

where $\Phi(\cdot)$ is cumulative standard normal distribution and $d(t) := \left(e^r-\hat{\theta}_0\right)\sqrt{\frac{s_0t+1}{s_0^2t}}$. The expectation is with respect to $\mathcal{F}_0$, so unconditional on the signal outcome.

This result allows us to analyze how the DM's expected utility changes with respect to t. The expected utility is not only a function of the quantity t; it depends also on risk aversion γ, initial wealth κ, borrowing and lending rate r, cost of data analytics c, prior mean $\hat{\theta}_0$ and variance s_0. For simplicity, in Lemma 3 we write $u(t)$, but we will analyze later how the model parameters affect the expected utility, and then we might write, e.g., $u(t;\gamma)$. Note that, by (1), we use the unconditional expectation in Lemma 3 because that is needed in deciding the quantity of analytics.

Next we introduce the value of data analytics which is the difference of the unconditional expected utilities with a positive quantity of data analytics and zero quantity of data analytics. The following proposition gives the value of data analytics and the proof is in Appendix C.3.

Proposition 1 (Value of data analytics)
The value of data analytics $v(t) := u^*(t) - u^*(0)$ is given by

$$
\begin{cases}
-e^{\gamma(\kappa-ct)(1-e^r)}\Phi(d(t)) + \left(1-\frac{e^{\gamma ct(e^r-1)}}{\sqrt{s_0t+1}}\Phi\left(\frac{\hat{\theta}_0-e^r}{s_0\sqrt{t}}\right)\right)\exp\left(-\frac{\hat{\theta}_0^2+e^{2r}-2\hat{\theta}_0e^r+2\gamma s_0\kappa(e^r-1)}{2s_0}\right) & \text{if } \hat{\theta}_0 > e^r \\
e^{\gamma\kappa(1-e^r)}\left(1-\frac{e^{\gamma ct(e^r-1)}}{2}\left(1+\frac{1}{\sqrt{s_0t+1}}\right)\right) & \text{if } \hat{\theta}_0 = e^r \\
e^{\gamma\kappa(1-e^r)}\left(1-e^{\gamma ct(e^r-1)}\Phi(d(t))\right) - \frac{\Phi\left(\frac{\hat{\theta}_0-e^r}{s_0\sqrt{t}}\right)\exp\left(-\frac{\hat{\theta}_0^2+e^{2r}-2\hat{\theta}_0e^r+2\gamma s_0(\kappa-ct)(e^r-1)}{2s_0}\right)}{\sqrt{s_0t+1}} & \text{if } \hat{\theta}_0 < e^r
\end{cases}
$$

So, Lemma 3 gives $u^*(t)$ and $u^*(0)$ which then give the result. We calculate the marginal value of data analytics from Proposition 1. This helps us solve for the optimal quantity of data analytics.

Lemma 4 (Marginal value of data analytics)
When $\hat{\theta}_0 \neq e^r$, the marginal value of data analytics is given by

$$
\begin{aligned}
v'(t) =& \frac{e^{\gamma(e^r-1)(ct-\kappa)}}{2(s_0t+1)^{3/2}}\Bigg[-2c\gamma\left(e^r-1\right)\left(s_0t+1\right)^{3/2}\Phi\left(d(t)\right) \\
& -2c\gamma\left(e^r-1\right)\left(s_0t+1\right)e^{-\frac{(e^r-\hat{\theta}_0)^2}{2s_0}}\Phi\left(-\frac{d(t)}{\sqrt{s_0t+1}}\right) + s_0e^{-\frac{(e^r-\hat{\theta}_0)^2}{2s_0}}\Phi\left(-\frac{d(t)}{\sqrt{s_0t+1}}\right)\Bigg]
\end{aligned}
$$

for all $t \geq 0$. *When* $\hat{\theta}_0 = e^r$, *the marginal value is given by*

$$v'(t) = \frac{e^{\gamma(e^r-1)(ct-\kappa)}}{2(s_0t+1)^{3/2}}\Big[-c\gamma\left(e^r-1\right)\left(s_0t+1\right)^{3/2} - c\gamma\left(e^r-1\right)\left(s_0t+1\right) + \frac{s_0}{2}\Big]$$

for all $t \geq 0$.

Proof. $\partial_t v(t) = \partial_t(u^*(t) - u^*(0)) = \partial_t u^*(t)$. The results follows from Lemma 3 by taking the derivative. □

From Lemma 4 we get directly the following result.

Corollary 1 (Negative marginal value at zero)
$\lim_{t\downarrow 0} v'(t)$ *is given by*

$$\begin{cases} \frac{1}{2}\exp\left(\kappa\gamma(1-e^r) - \frac{(e^r-\hat{\theta}_0)^2}{2s_0}\right)\left[s_0 - 2c\gamma(e^r-1)\right] < 0 & \text{if } \hat{\theta}_0 > e^r \text{ and } s_0 < 2c\gamma(e^r-1) \\ \frac{1}{4}\exp\left(\kappa\gamma(1-e^r)\right)\left[s_0 - 4c\gamma(e^r-1)\right] < 0 & \text{if } \hat{\theta}_0 = e^r \text{ and } s_0 < 4c\gamma(e^r-1) \\ (1-e^r)c\gamma\exp(\kappa\gamma(1-e^r)) < 0 & \text{if } \hat{\theta}_0 < e^r \end{cases}$$

where r, c, *and* γ *are strictly positive.*

Note that, by Corollary 1, we have $\lim_{t\downarrow 0, \hat{\theta}_0\uparrow e^r} v' \neq \lim_{t\downarrow 0, \hat{\theta}_0\downarrow e^r} v' \neq \lim_{t\downarrow 0, \hat{\theta}_0 = e^r} v'$. Further, for instance, the condition in the last line of Corollary 1 can be written as $s_0/(2\gamma) < c(e^r-1)$, which means that the marginal value at $t = 0$ is negative if lending out money equal to the cost of one unit of data analytics is greater than $s_0/(2\gamma)$. Note that $s_0/(2\gamma)$ is high if uncertainty (s_0) is high and the DM has a low level of risk aversion (γ) and, therefore, she would prefer to invest in the risky opportunity and data analytics, and this way decrease the investment uncertainty. However, the condition $s_0/(2\gamma) < c(e^r-1)$ says that it is better to lend the money out; hence, the marginal value of data analytics is negative at $t = 0$.

Let us define choke-off cost c_{CH} as the highest cost c such that the value of information is positive. That is, if $c < c_{CH}$ then $v(t) > 0$ for some $t > 0$.

Lemma 5 (Choke-off cost)
Let $v(t) = v(t;c)$. *There exists choke-off cost* $c_{CH} > 0$ *such that if* $c < c_{CH}$ *then* $v(t;c) > 0$ *for some* $t > 0$. *The choke-off cost* $c_{CH} = \frac{s_0}{2\gamma(e^r-1)}$ *if* $\hat{\theta}_0 > e^r$, $c_{CH} = \frac{s_0}{4\gamma(e^r-1)}$ *if* $\hat{\theta}_0 = e^r$, *and* $c_{CH} \in (0, \frac{s_0}{2\gamma(e^r-1)})$ *if* $\hat{\theta}_0 < e^r$.

Proof. We consider three cases:

(i) $\hat{\theta}_0 > e^r$: By Corollary 1, $v'(0;c) < 0$ if $c > \frac{s_0}{2\gamma(e^r-1)}$. Since $v(t;c)$ is strictly concave in t and $v(0;c) = 0$, $v(t;c) < 0$ for all $t > 0$ if $c > \frac{s_0}{2\gamma(e^r-1)}$.

(ii) $\hat{\theta}_0 = e^r$: This is proved similarly as *(i)*.

(iii) $\hat{\theta}_0 < e^r$: By Lemma 4, $v'(t;c) < 0$ for all $t > 0$ if $c = \frac{s_0}{2\gamma(e^r-1)}$. Since $v'(t;c)$ is continuous with respect to c, there exists $\epsilon > 0$ such that $v'(t;c) < 0$ for all $c \in \left(\frac{s_0}{2\gamma(e^r-1)} - \epsilon, \frac{s_0}{2\gamma(e^r-1)} + \epsilon\right)$ and $t > 0$. Because $v(0;c) = 0$, the value of data analytics is negative for all $t > 0$, and $c = \frac{s_0}{2\gamma(e^r-1)}$ cannot be c_{CH}. Since $v'(t;0) > 0$, the Intermediate Value Theorem gives that $c_{CH} \in (0, \frac{s_0}{2\gamma(e^r-1)})$. □

By Lemma 5, there is a positive value of data analytics when $c < c_{CH}$. Thus, in this case the analytics raises the DM's expected utility. The choke-off cost for $\hat{\theta}_0 \leq e^r$ is lower than the choke-off cost for $\hat{\theta}_0 > e^r$, i.e., the DM requires a lower marginal cost to get value from analytics if the expected payoff is low. This is because when the unconditional expected risky asset payoff is low, the DM expects to lend her money and, thus, she requires more incentive, in the form of lower cost of analytics, to use data analytics on the risky investment opportunity.

The next proposition gives the shape of the value of data analytics with respect to the quantity of data analytics, and the proof is in Appendix C.4.

Proposition 2 (Shape of the value)
Let $c < c_{CH}$, so that $v(t) > 0$ for some $t > 0$. The value of data analytics $v(\cdot) : \mathbb{R}_+ \to \mathbb{R}$ satisfies:

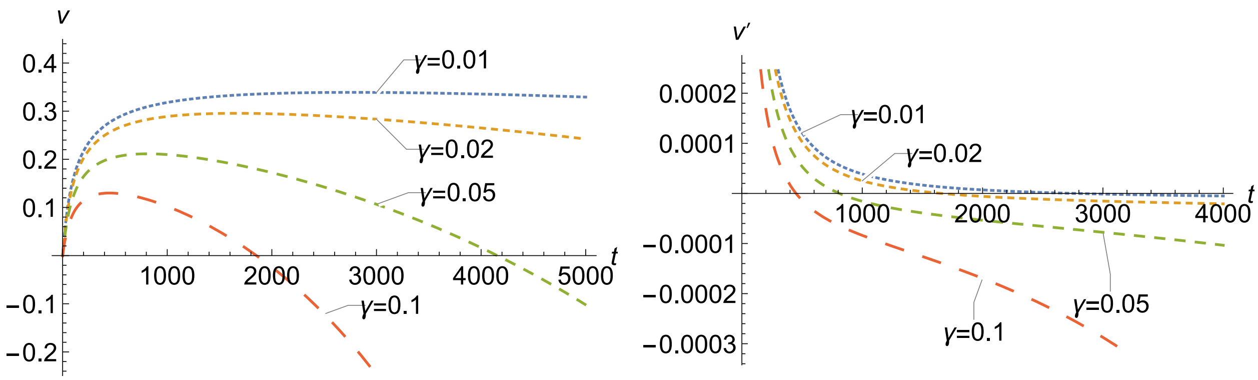
(i) Strictly concave if $\hat{\theta}_0 \geq e^r$.

(ii) Not globally concave if $\hat{\theta}_0 < e^r$.

By Proposition 2, the optimization of the quantity of data analytics is simpler when $\hat{\theta}_0 \geq e^r$.

Corollary 2 (Risk aversion)
Let $v(t) = v(t;\gamma)$. When $c < c_{CH}$ then $v(t;\gamma)$ falls in risk aversion γ for all t such that $v(t;\gamma) > 0$.

Proof. From Proposition 1 we get $\frac{\partial v}{\partial \gamma} = (1 - e^r)(\kappa v(t) - ctu^*(t))$ which is negative when $v(t) > 0$ as $u^*(t)$ is always negative (see Lemma 3). □



(a) Value of analytics $v(t)$

(b) Marginal value of data analytics $v'(t)$

Figure 3: **Value and marginal value of data analytics with different risk aversion** γ**.** Parameter values: $r = 0.05$, $s_0 = 0.01$, $\hat{\theta}_0 = 1.15$, $\kappa = 100$, and $c = 0.1$.

Thus, the value of data analytics falls in risk aversion. This is because, by Lemma 2, the higher the risk aversion the less the DM invests in the risky opportunity and, therefore, information on the investment is less valuable to her.

We next plot the value and marginal value of data analytics with respect to the quantity of data analytics t under different parameter values. Since the value of data analytics depends on $\hat{\theta}_0$ (Proposition 1), we consider two cases, $\hat{\theta}_0 \geq e^r$ and $\hat{\theta}_0 < e^r$.

Case $\hat{\theta}_0 \geq e^r$***: Expected payoff higher than the lending return***

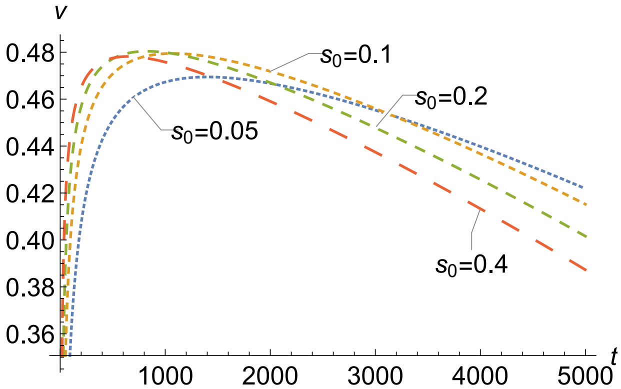
Figure 3 plots the value and marginal value of data analytics against analytics quantity for four different values of risk aversion, $\gamma = 0.01, 0.02, 0.05$, and 0.1. Consistent with Corollary 2, Figure 3a shows that the value of data analytics falls in risk aversion. Figure 3b shows that the marginal value falls with risk aversion as well.

Figure 4 plots the value and marginal value of data analytics against analytics quantity for four different values of uncertainty, $s_0 = 0.05$, 0.1, 0.2, and 0.4. The figure shows that the marginal value of data analytics falls in the uncertainty.

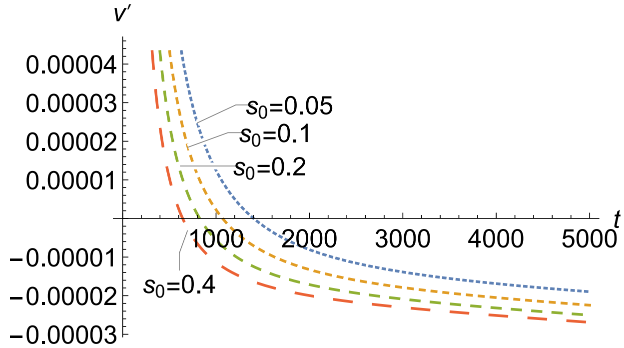
Figures 3a and 4a show that when $\hat{\theta}_0 \geq e^r$, the value of data analytics is strictly concave with respect to t, and this is proved in Proposition 2.

Case $\hat{\theta}_0 < e^r$***: Expected payoff lower than the lending return***

Figure 5 plots the value and marginal value of data analytics against analytics quantity for four different values of risk aversion, $\gamma = 0.01, 0.02, 0.05$, and 0.1. Consistent with Corollary 2, the

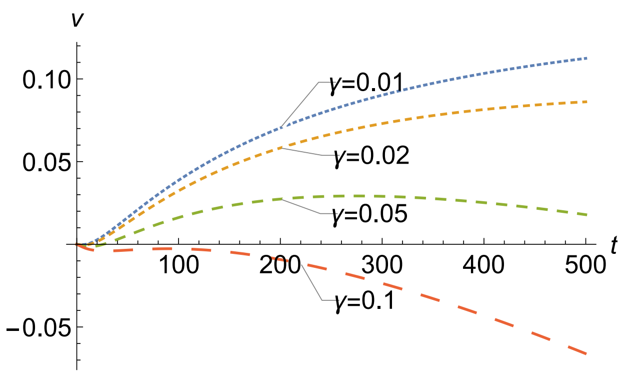


(a) Value of data analytics $v(t)$

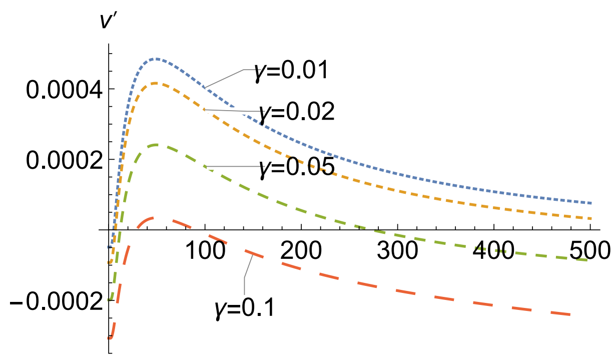


(b) Marginal value of data analytics $v'(t)$

Figure 4: **Value and marginal value of data analytics with different initial uncertainty level s_0.** Parameter values: $r = 0.05$, $\gamma = 0.01$, $\hat{\theta}_0 = 1.15$, $\kappa = 100$, and $c = 0.1$.



(a) Value of data analytics $v(t)$



(b) Marginal value of analytics $v'(t)$

Figure 5: **Value and marginal value of data analytics with different risk aversion γ.** Parameter values: $r = 0.05$, $s_0 = 0.01$, $\hat{\theta}_0 = 1$, $\kappa = 100$, and $c = 0.1$.

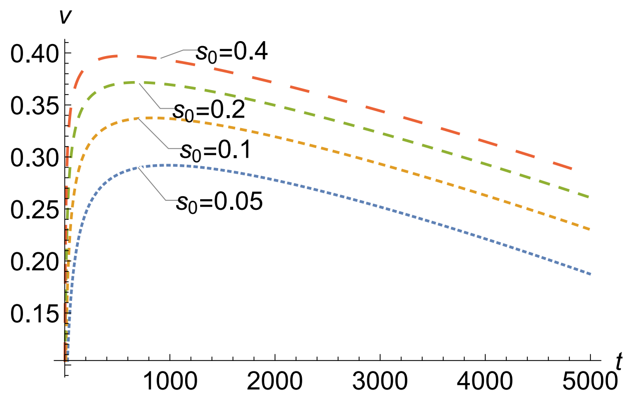
value and marginal value of data analytics falls in risk aversion.

Figure 6 plots the value and marginal value of data analytics against analytics quantity for four different values of uncertainty, $s_0 = 0.05$, 0.1, 0.2, and 0.4. The value of data analytics rises with uncertainty, but the relationship between the marginal value of analytics and the uncertainty is more complicated.

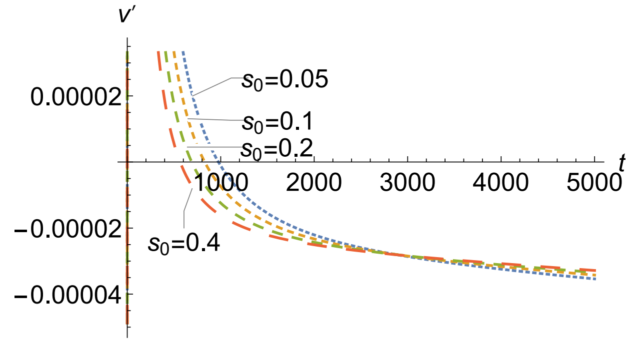
Finally, the following proposition analyzes the value of data analytics with respect to the expected payoff of the investment (the proof is in Appendix C.5).

Proposition 3 (Expected payoff)

The value of data analytics $v(t) = v(t; \hat{\theta}_0)$ behaves as follows with respect to the expected payoff $\hat{\theta}_0$:



(a) Value of data analytics $v(t)$



(b) Marginal Value of data analytics $v'(t)$

Figure 6: **Value and marginal value of data analytics with different initial uncertainty level s_0.** Parameter values: $r = 0.05$, $\gamma = 0.01$, $\hat{\theta}_0 = 1$, $\kappa = 100$, and $c = 0.1$.

(i) *If $c < c_{CH}$ then for each $t > 0$, where $v(t; \hat{\theta}_0) > 0$, there exists $\varsigma > e^r$ such that $v(t; \hat{\theta}_0)$ falls in $\hat{\theta}_0$ for all $\hat{\theta}_0 \in (\varsigma, \infty)$.*

(ii) *If $\hat{\theta}_0 < e^r$ then $v(t; \hat{\theta}_0)$ rises in $\hat{\theta}_0$ for all $t > 0$.*

(iii) $\lim_{\hat{\theta}_0 \to \infty} v(t; \hat{\theta}_0) = \lim_{\hat{\theta}_0 \to \infty} \partial_t v(t; \hat{\theta}_0) = 0$ *for all $t \geq 0$.*

By Proposition 3, when the expected payoff is high then the value of data analytics falls with respect to $\hat{\theta}_0$. Hence, the value is low for investments with high expected payoff. On the other hand, the value of data analytics rises in $\hat{\theta}_0$ when the expected payoff is less than e^r. Therefore, the value of data analytics is also low for investments that have low expected payoff. The highest value of data analytics is hence achieved when $\hat{\theta}_0 > e^r$, but not for a very high expected payoff. That is, data analytics gives highest value for mediocre investments. Figure 7 illustrates this.

4 Optimal quantity of data analytics

In this section, we consider the optimal quantity of data analytics. We consider two cases. First, in subsection 4.1 we analyze the case where the DM can take leverage, and then in section 6 we study the case where the DM cannot take leverage.

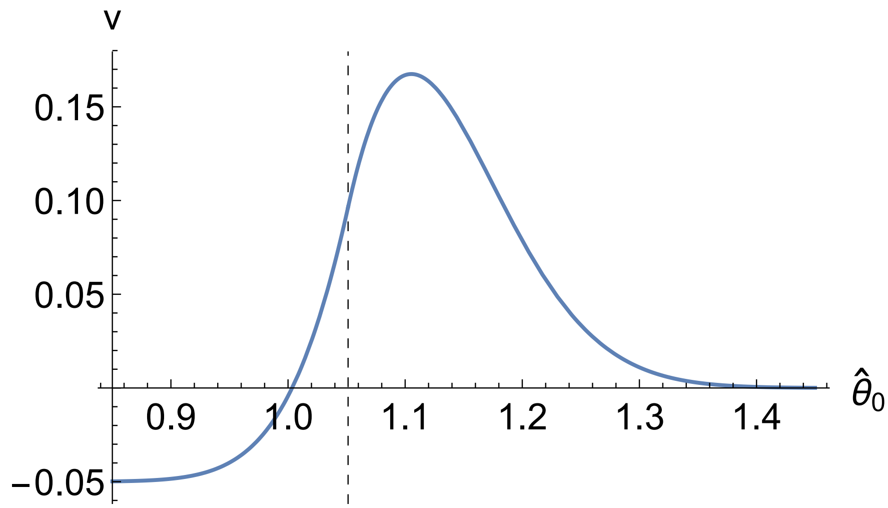


Figure 7: **Value of data analytics $v(t)$ with respect to $\hat{\theta}_0$.** Dotted line represents $\hat{\theta}_0 = e^r$. The peak value occurs close to e^r (see Proposition 3). Parameter values: $r = 0.05$, $\gamma = 0.01$, $s_0 = 0.01$, $\kappa = 100$, $c = 1$, and $t = 100$.

4.1 Optimal quantity with leverage

The optimal quantity of data analytics is defined as follows

$$t^* := \arg\min\{\ t \mid v(t) \text{ is a global maximum}\}.$$

The following result gives the existence of t^*.

Theorem 1 (Optimal quantity of data analytics)

Let $c > 0$. There exists bounded and nonnegative t^* such that $\sup_{t \geq 0} v(t) = v(t^*)$.

Proof. We consider two cases:

(*i*) $\hat{\theta}_0 \geq e^r$: In this case, $v(t)$ is strictly concave and, hence, admits a unique global maximum.

(*ii*) $\hat{\theta}_0 < e^r$: By definition, $v(0) = 0$ and $\lim_{t \to \infty} v(t) = -\infty$ since $c > 0$. Thus, $\exists M \in \mathbb{R}$ such that $\forall t > M$, $v(t) < 0$. Set $t_\infty \in (M, \infty)$ and consider the interval $[0, t_\infty]$. Since $v(t)$ is continuous on the closed and bounded interval $[0, t_\infty]$, by the Weierstrass Extreme Value Theorem, $v(t)$ achieves maximum and minimum on $[0, t_\infty]$ at least once. Set t^* as the minimum value of t where $v(t)$ achieves the maximum. $\square$

The following corollary gives the quantity of data analytics under zero cost.

Corollary 3 (Zero cost)

When $c = 0$ then $t^* = \infty$.

Proof. Let $c = 0$. If $\hat{\theta}_0 \neq e^r$ then

$$v'(t) = \frac{s_0}{2(s_0 t + 1)^{3/2}} \exp\left(-\frac{\left(e^r - \hat{\theta}_0\right)^2 + 2s_0\kappa\gamma\left(e^r - 1\right)}{2s_0}\right)\Phi\left(-\frac{d(t)}{\sqrt{s_0 t + 1}}\right) > 0,$$

and $v'(t) = \frac{s_0 e^{\gamma\kappa(1-e^r)}}{4(s_0 t+1)^{3/2}} > 0$ if $\hat{\theta}_0 = e^r$. Thus, in both the cases each extra unit of data analytics raises the expected utility of the DM. □

This is an intuitive result since if there is no cost of using data analytics then the DM should use infinite quantity and learn the payoff perfectly.

The next corollary gives conditions when the DM acquires either a high quantity of data analytics or not at all.

Corollary 4 (Small quantity of analytics)

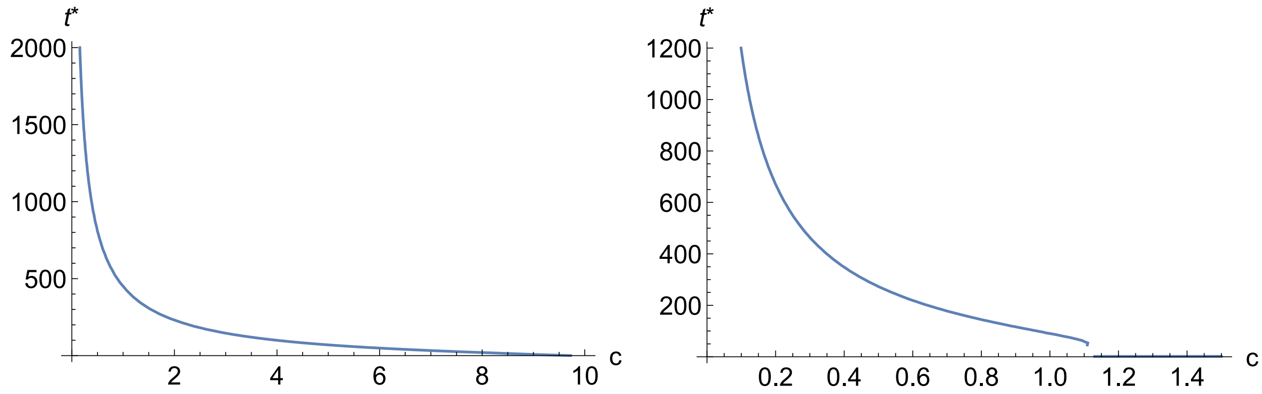
When $c \in (0, c_{CH})$ and $\hat{\theta}_0 < e^r$, the DM does not acquire small quantities of analytics.

Proof. By Corollary 1 and the conditions of this corollary, we have $\lim_{t\downarrow 0} v'(t) < 0$. Since $v(0) = u^*(0) - u^*(0) = 0$ and $v(t)$ is continuous, $v'(0) < 0$ implies that there exists $\epsilon > 0$ such that the value of data analytics $v(t) < 0$ for all $t \in (0, \epsilon)$. That is, under the above conditions, small quantities of analytics reduce the expected utility of the DM. □

Corollary 4 says that the DM does not buy small quantities of data analytics when the expected payoff is lower than the lending and borrowing rate.

The optimal quantity can be solved by using the first order condition, $v'(t) = 0$. By Lemma 4, this condition can be written as

$$\begin{aligned}&-2c\gamma\left(e^r - 1\right)\left(s_0 t + 1\right)^{3/2}\Phi\left(d(t)\right) - 2c\gamma\left(e^r - 1\right)\left(s_0 t + 1\right)e^{-\frac{\left(e^r - \hat{\theta}_0\right)^2}{2s_0}}\Phi\left(-\frac{d(t)}{\sqrt{s_0 t + 1}}\right)\\&+ s_0 e^{-\frac{\left(e^r - \hat{\theta}_0\right)^2}{2s_0}}\Phi\left(-\frac{d(t)}{\sqrt{s_0 t + 1}}\right) = 0\end{aligned} \tag{3}$$



(a) $\hat{\theta}_0 = 1.15 > e^r$ (b) $\hat{\theta}_0 = 1 < e^r$

Figure 8: **Optimal quantity of data analytics t^* with respect to c.** Parameter values: $r = 0.05$, $\gamma = 0.01$, $s_0 = 0.01$, and $\kappa = 100$.

for all $\hat{\theta}_0 \neq e^r$, and

$$-c\gamma\left(e^r - 1\right)\left(s_0 t + 1\right)^{3/2} - c\gamma\left(e^r - 1\right)\left(s_0 t + 1\right) + \frac{s_0}{2} = 0 \tag{4}$$

for $\hat{\theta}_0 = e^r$.

Next we give the properties of the optimal quantity of data analytics (the proof is in Appendix C.6).

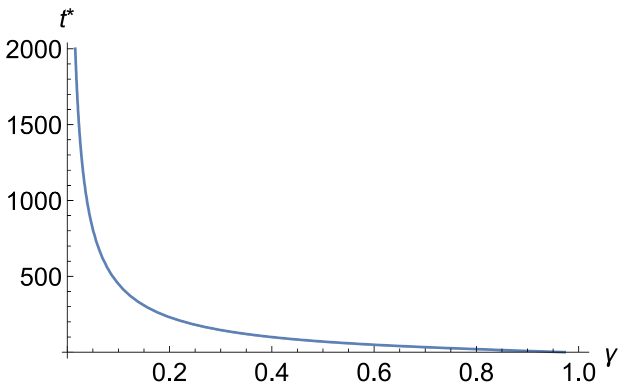
Theorem 2 (Properties of the optimal quantity)

We have:

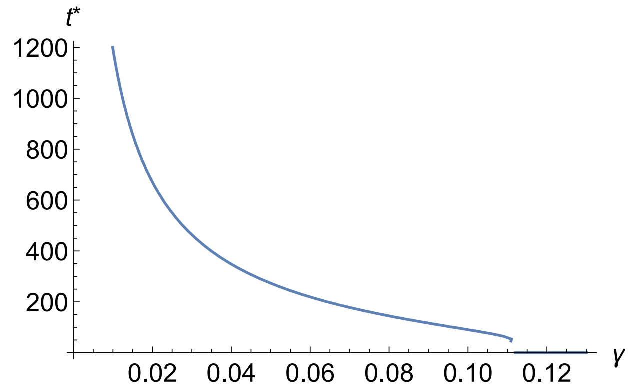
(i) Optimal quantity of analytics t^* is independent of initial wealth κ.

(ii) If $c < c_{CH}$ and $\hat{\theta}_0 \geq e^r$ then optimal quantity t^* falls in cost c and risk aversion γ.

By Theorem 2, the optimal quantity of analytics rises when the cost of analytics or the DM's risk aversion falls. Further, the optimal quantity is independent of the initial wealth κ, because, by the CARA utility, the first order condition in (3) and (4) are independent of κ. Figures 8 – 11 illustrate the optimal quantity of data analytics with respect to different model parameters. Figure 8 shows that the optimal quantity of data analytics falls in the marginal cost, and it gives the demand curve of data analytics. Hence, the DM uses more analytics the less costly it is. The discontinuity in Figure 8b is due to the choke-off cost; consistent with Corollary 4 in this case the DM never acquires small quantities of data analytics. When $\hat{\theta}_0 > e^r$, there is no discontinuity as

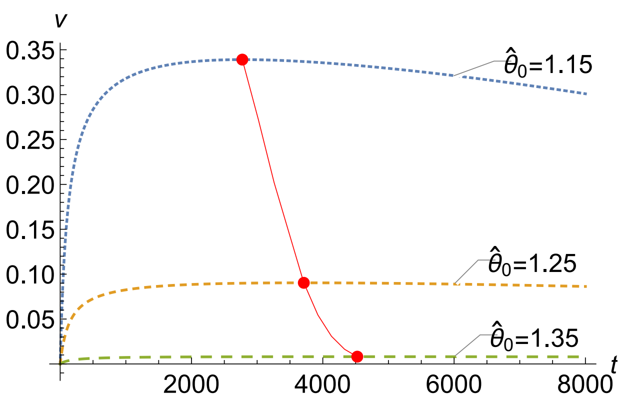


(a) $\hat{\theta}_0 = 1.15 > e^r$

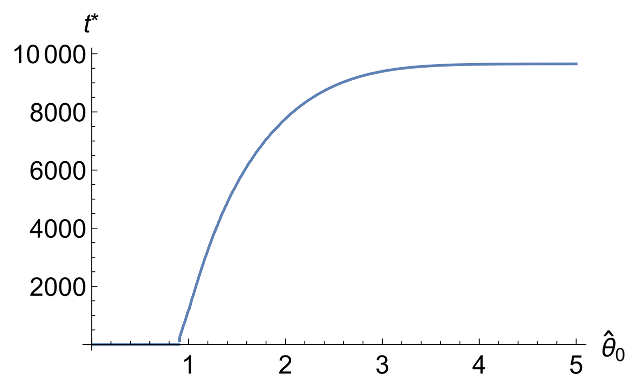


(b) $\hat{\theta}_0 = 1 < e^r$

Figure 9: **Optimal quantity of data analytics t^* with respect to γ.** Parameter values: $r = 0.05$, $c = 0.1$, $s_0 = 0.01$, and $\kappa = 100$.



(a) Value of data analytics with different initial expected payoffs $\hat{\theta}_0$; red solid line gives the optimal quantity t^* for the different initial expected payoffs

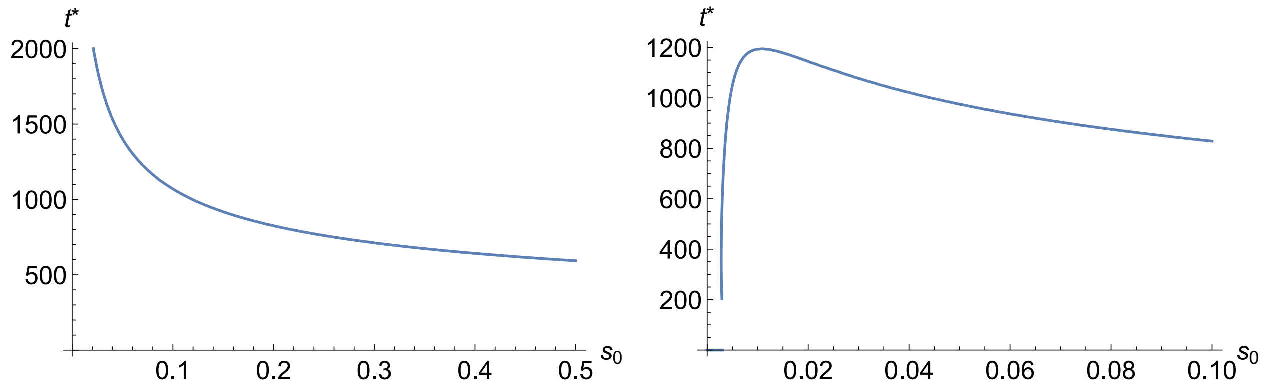


(b) Optimal quantity against $\hat{\theta}_0$

Figure 10: **Optimal quantity of data analytics t^* with respect to $\hat{\theta}_0$.** Parameter values: $r = 0.05$, $c = 0.1$, $s_0 = 0.01$, $\kappa = 100$, and $\gamma = 0.01$.

the value function is continuous and strictly concave. Figure 9 shows that the optimal quantity of data analytics falls in risk aversion. This is because, by Lemma 2, a highly risk averse DM invests less in the risky investment opportunity than a less risk averse DM, ceteris paribus, and therefore, the highly risk averse DM has less incentive to learn and use data analytics. The discontinuity in Figure 9b is again due to the choke-off cost. Figure 10 shows that the optimal quantity of data analytics is non-decreasing in the expected payoff. This is surprising since, by Proposition 3 and Figures 7 and 10a, the value of data analytics is highest for mediocre investment opportunities.[2] We conjecture that Figure 10b is driven by the amount of leverage. That is, leverage and, therefore,

[2]Kihlstrom (1974) and Keppo et al. (2008) show that demand for information is hill-shaped in beliefs, and greatest when the DM is most uncertain, so that very high or very low expectations lead to low demand.



(a) $\hat{\theta}_0 = 1.15 > e^r$

(b) $\hat{\theta}_0 = 1 < e^r$

Figure 11: **Optimal quantity of data analytics t^* with respect to s_0.** Parameter values: $r = 0.05$, $\gamma = 0.01$, $c = 0.1$, and $\kappa = 100$.

also the size of the risky investment rise in the expected payoff, which increases the DM's incentive to acquire more analytics. For robustness we sample million times parameter values $(\hat{\theta}_0, s_0, \gamma) \in [e^r, 5] \times [0.01, 0.05] \times [0.01, 0.05]$, and the other parameter values are fixed and given in Figure 10; we found that the non-decreasing property in Figure 10 holds for each of them. Further, when we do the same robustness test without leverage (see Section 6), the non-decreasing property vanishes. Thus, these tests indicate that our conjecture on the leverage is true.

Figure 11 shows that the effect of uncertainty level s_0 on the optimal quantity of data analytics is similar to the effect of risk aversion γ in Figure 9. This is because, by Lemma 2, when s_0 or γ rises, the DM invests less in the risky opportunity. However, if $\hat{\theta}_0 < e^r$ and s_0 approaches zero then the optimal quantity also approaches zero because in this case the DM learns the risky investment is less profitable than the lending. Note that in Figure (11b) there is discontinuity close to zero due to the choke-off cost.

5 Expected optimal investment

In this section we solve for the unconditional expected investment amount and leverage before the DM observes the noisy signal on the investment payoff. We first give the result for the expected optimal investment when the DM is allowed to take leverage (the proof is Appendix C.7), and in the next section we consider the case without leverage.

Proposition 4 (Expected optimal investment with leverage)

The expected optimal investment is given by

(*i*) $c \leq c_{CH}$*: Optimal quantity* $t^* > 0$ *and the expected optimal investment is given by*

$$
\begin{aligned}
E\left[w^*(t^*)\right] &= E\left[\max\left\{\frac{\hat{\theta}_{t^*} - e^r}{\gamma s_{t^*}}, 0\right\}\right] \\
&= \frac{(\hat{\theta}_0 - e^r)(1 + t^* s_0)}{s_0 \gamma} \Phi\left(\frac{(\hat{\theta}_0 - e^r)\sqrt{\frac{1}{t^*} + s_0}}{s_0}\right) + \frac{\sqrt{t^* + s_0 (t^*)^2}}{\gamma} \phi\left(\frac{(e^r - \hat{\theta}_0)\sqrt{\frac{1}{t^*} + s_0}}{s_0}\right),
\end{aligned}
$$

where the expectation is with respect to $\mathcal{F}_0$*, so unconditional on the signal outcome.*

(*ii*) $c > c_{CH}$*: Optimal quantity* $t^* = 0$ *and the optimal investment is given by*

$$
E\left[w^*(0)\right] = w^*(0) = \begin{cases} \frac{\hat{\theta}_0 - e^r}{s_0 \gamma} & \text{if } \hat{\theta}_0 > e^r \\ 0 & \text{if } \hat{\theta}_0 \leq e^r \end{cases}.
$$

The expected leverage is given by

$$
E\left[\max\left\{w^*(t^*) + ct^* - \kappa, 0\right\}\right] = \left(\frac{(\hat{\theta}_0 - e^r)(1 + t^* s_0)}{s_0 \gamma} + ct^* - \kappa\right) \Phi\left(-b^*\right) + \frac{\sqrt{t^* + s_0 (t^*)^2}}{\gamma} \phi\left(b^*\right),
$$

where $b^* = \left(\frac{(\kappa - ct)\gamma s_0}{1 + s_0 t^*} + e^r - \hat{\theta}_0\right)\sqrt{\frac{s_0 t^* + 1}{s_0^2 t^*}}$.

To analyze the relationship between leverage and the cost of data analytics, by Proposition 4, we first calculate the derivative of the expected leverage with respect to the cost of analytics:

$$
\begin{aligned}
&\frac{\partial E[\max\{w^*(t^*) + ct^* - \kappa, 0\}]}{\partial c} \\
&= \begin{cases} 0 & \text{if } c > c_{CH} \\ \left(\frac{\hat{\theta}_0 - e^r}{\gamma} + c\right) \Phi\left(-b^*\right) \frac{\partial t^*}{\partial c} + \phi\left(b^*\right) \frac{1 + 2 s_0 t^*}{2\gamma\sqrt{t^* + s_0 (t^*)^2}} \frac{\partial t^*}{\partial c} + t^* \Phi\left(-b^*\right) & \text{otherwise} \end{cases}
\end{aligned} \tag{5}
$$

where $b^* = \left(\frac{(\kappa - ct^*)\gamma s_0}{1 + s_0 t^*} + e^r - \hat{\theta}_0\right)\sqrt{\frac{s_0 t^* + 1}{s_0^2 t^*}}$. Since this derivative depends on the optimal quantity of data analytics t^* that we cannot solve analytically, we next numerically analyze (5).

The DM takes leverage if the expected payoff of the investment is greater than the lending payoff, and her investment is greater than her initial wealth. Hence, we numerically consider a case where $\hat{\theta}_0 \gg e^r$. We sample 100,000 parameter values $(\kappa, s_0, \gamma) \in [100, 1000] \times [0.01, 0.05] \times [0.01, 0.05]$, while the other parameters are fixed ($\hat{\theta}_0 = 1.15$, $r = 0.05$, and $c = 1$). In each of these cases derivative (5) is negative, which supports our conjecture on the falling leverage with respect to the cost of data analytics.

If there is no financial constraint then due to the increased leverage, the falling cost of data analytics in Figure 1 may lead to higher losses during the crises. This is illustrated in Figure 12a, where the expected losses are higher under the low cost of analytics. As mentioned before, this is driven by leverage: Under the low cost, the DM acquires more data analytics, which makes the investment opportunity less risky and, therefore, she takes more leverage when investing. In Figure 12b we analyze further the effect of cost of data analytics on investment losses. Here we assume that the realized payoff θ is a certain number of standard deviations below the expected payoff. Note that the higher the cost of data analytics, the higher the standard deviation is. As can be seen, the losses are again higher under the low cost of analytics due to the higher leverage.

To further analyze the effect of leverage, let us perform the analysis in Figure 12 without leverage; in this case the maximum quantity of data analytics is κ/c and the maximum risky investment is $\kappa - ct^* \geq 0$ (see Section 6). To do that, we sample 10,000 parameter values $(s_0, \gamma) \in [0.01, 0.05] \times [0.01, 0.05]$, while the other parameters are fixed ($\hat{\theta}_0 = 1.06$, $\kappa = 1000$, and $r = 0.05$). For each of these cases we compare the losses under the costs of data analytics $c = 0.01$ and $c = 1$, respectively. Without leverage the losses corresponding to Figure 12a are again higher under the lower cost of data analytics, but not for the losses corresponding to Figure 12b. This shows that leverage constraints are effective in reducing losses during crises, which is important, for instance, in banking regulation (see e.g. Peura and Keppo 2006, Hart and Zingales 2011).

Consistent with Figure 12, let us define the following two negative shocks.

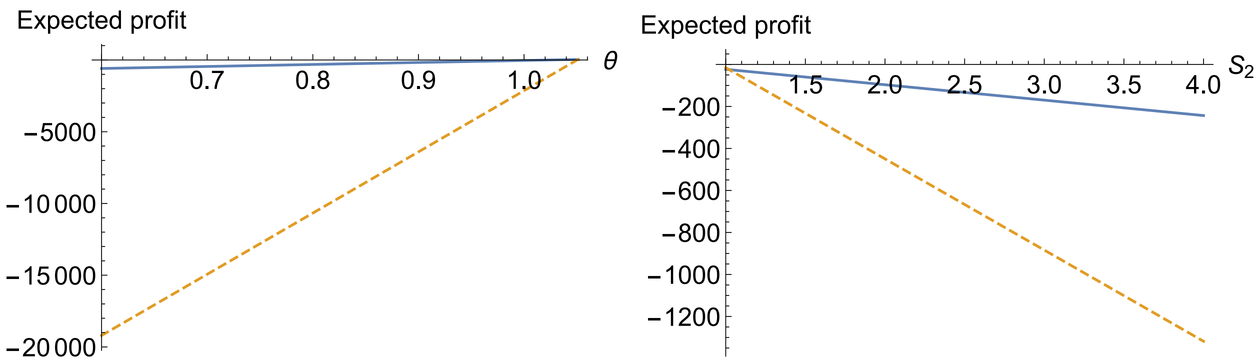
Definition 1 (Shocks)

We consider the following two negative shocks:

(*i*) *The payoff of the risky investment* $\theta = S_1$, *where* S_1 *is the shock level and* $S_1 \in (-\infty, e^r)$.

(*ii*) *The payoff* $\theta = \hat{\theta}_0 - S_2\sqrt{s_{t^*}}$, *where* S_2 *is the shock level and* $S_2 \in (0, \infty)$.

Now we can state the following result (the proof is in Appendix C.8) and it is illustrated in Figure 12.



(a) **Expected profit and loss with respect to the realized payoff θ.** This is shock (i) in Definition 1. The blue solid line is with high cost ($c = 1$) and the orange dotted line is with low cost ($c = 0.01$).

(b) **Expected profit and loss with respect to the number of standard deviations below the expected payoff.** This is shock (ii) in Definition 1. Realized payoff $\theta = \hat{\theta}_0 - S_2\sqrt{s_{t^*}}$. The blue solid line is with high cost ($c = 1$) and the orange dotted line is with low cost ($c = 0.01$).

Figure 12: **Expected profit and loss of an investment under different analytics costs.** Other parameter values: $\hat{\theta}_0 = 1.06$ $s_0 = 0.01$, $r = 0.05$, $\gamma = 0.01$, and $\kappa = 1000$. The expected profit and loss is given by $E[w^*(t^*)](\theta - 1) + (\kappa - ct^* - E[w^*(t^*)])(e^r - 1)$.

Theorem 3 (Higher losses)

Consider a DM who is able to borrow. Let $\hat{\theta}_0 > e^r$ and consider two costs of data analytics, c_h and c_l, where $c_{CH} > c_h > c_l > 0$. We have:

(i) *There exists bounded $\bar{S}_1$ such that the loss under c_l and shock S_1 is higher than the loss under c_h and shock S_1 for all $S_1 < \bar{S}_1$, where the shock is given by (i) in Definition 1.*

(ii) *There exist bounded $\bar{S}_2$ such that the loss under c_l and shock S_2 is higher than the loss under c_h and shock S_2 for all $S_2 > \bar{S}_2$, where the shock is given by (ii) in Definition 1.*

To understand Theorem 3, let us write down the expected profit and loss under negative shock (ii) in Definition 1 (shock (i) can be handled in a similar way). By Lemma 2, in this case the expected profit and loss is given by

$$
\begin{aligned}
&E[w^*(t^*)](\theta - 1) + (\kappa - ct^* - E[w^*(t^*)])(e^r - 1) \\
&= E\left[\max\left\{\frac{\hat{\theta}_{t^*} - e^r}{\gamma s_{t^*}}, 0\right\}\right]\left(\hat{\theta}_0 - S_2\sqrt{s_{t^*}} - 1\right) + \left(\kappa - ct^* - E\left[\max\left\{\frac{\hat{\theta}_{t^*} - e^r}{\gamma s_{t^*}}, 0\right\}\right]\right)(e^r - 1).
\end{aligned}
$$

This gives the derivative of the expected profit and loss with respect to shock S_2:

$$-E\left[\max\left\{\frac{\hat{\theta}_{t^*}-e^r}{\gamma s_{t^*}},0\right\}\right]\sqrt{s_{t^*}} = -E\left[\max\left\{\frac{\hat{\theta}_{t^*}-e^r}{\gamma\sqrt{s_{t^*}}},0\right\}\right] \leq 0, \tag{6}$$

which is negative and rises in the posterior variance of the investment payoff, s_{t^*}. That is, the lower the variance s_{t^*} is, the more the expected profit and loss fall due to the negative shock S_2. Note that, by Lemma 1 and Theorem 2, the posterior variance s_{t^*} rises in the cost of analytics c. Therefore, by (6), when c falls, the investment position gets riskier. This is surprising since one could guess that the lower posterior variance of the investment payoff leads to a less risky investment. However, as discussed after (5), the falling posterior variance raises the expected investment and leverage, which is a stronger effect on the investment risk than the lower posterior variance.

6 Financial constraint

In this section, we consider a financial constraint so that the DM is not able to take leverage. Hence, the DM's investment in the risky opportunity is bounded by $\kappa - ct \geq 0$, and the quantity of data analytics is bounded by κ/c. The optimal risky investment in this case is given by the following lemma.

Lemma 6 (Optimal investment without leverage)
Given analytics quantity $t \in [0, \frac{\kappa}{c}]$ and the CARA utility function with $\gamma > 0$, the optimal level of risky investment without leverage is given by

$$w^*_{NL}(t) = \min\left\{\max\left\{\frac{\hat{\theta}_t - e^r}{\gamma s_t},0\right\}, \kappa - ct\right\}.$$

Proof. By Lemma 2, the optimization problem without leverage:

$$\max_{w\in[0,\kappa-ct]} E\left[-\exp\left(-\gamma\left[w(\theta-1)+(\kappa-ct-w)(e^r-1)\right]\right)|\mathcal{F}_t\right].$$

By the concavity of the objective function, if $w^* \in (0, \kappa - ct)$ in Lemma 2 then $w^*_{NL} = w^*$; if $w^* = 0$ then $w^*_{NL} = 0$; if $w^* > \kappa - ct$ then $w^*_{NL} = \kappa - ct$. □

As in Section 3, we first give the expected utility of the optimal investment, and the proof is in Appendix C.9.

Lemma 7 (Expected utility without leverage)

The expected utility of the optimal investment under a given quantity of analytics t*:*

$$
\begin{aligned}
u^*_{NL}(t) = &- \frac{\exp\left(-\frac{\left(\hat{\theta}_0 - e^r\right)^2 + 2\gamma(\kappa - ct)s_0(e^r-1)}{2s_0}\right)\left(\Phi\left(\frac{e^r-\hat{\theta}_0}{s_0\sqrt{t}} + \frac{\gamma(\kappa-ct)}{\sqrt{t}}\right) - \Phi\left(\frac{e^r-\hat{\theta}_0}{s_0\sqrt{t}}\right)\right)}{\sqrt{s_0 t + 1}} \\
&- e^{\frac{\gamma^2(\kappa-ct)^2 s_0}{2} + \gamma(\kappa-ct)\left(1-\hat{\theta}_0\right)}\Phi\left(\frac{\sqrt{s_0 t+1}\left(\hat{\theta}_0 - e^r\right)}{s_0\sqrt{t}} - \frac{\gamma(\kappa-ct)\sqrt{s_0 t + 1}}{\sqrt{t}}\right) \\
&- e^{\gamma(\kappa-ct)(1-e^r)}\Phi\left(\frac{\sqrt{s_0 t+1}\left(e^r - \hat{\theta}_0\right)}{s_0\sqrt{t}}\right),
\end{aligned}
$$

where the expectation is with respect to $\mathcal{F}_0$*, so unconditional on the signal outcome.*

Comparing Lemmas 3 and 7, we get the following corollary (the proof is Appendix C.10) and it is illustrated in Figure 13.

Corollary 5 (Expected utility difference)

The difference between the expected utilities with and without leverage:

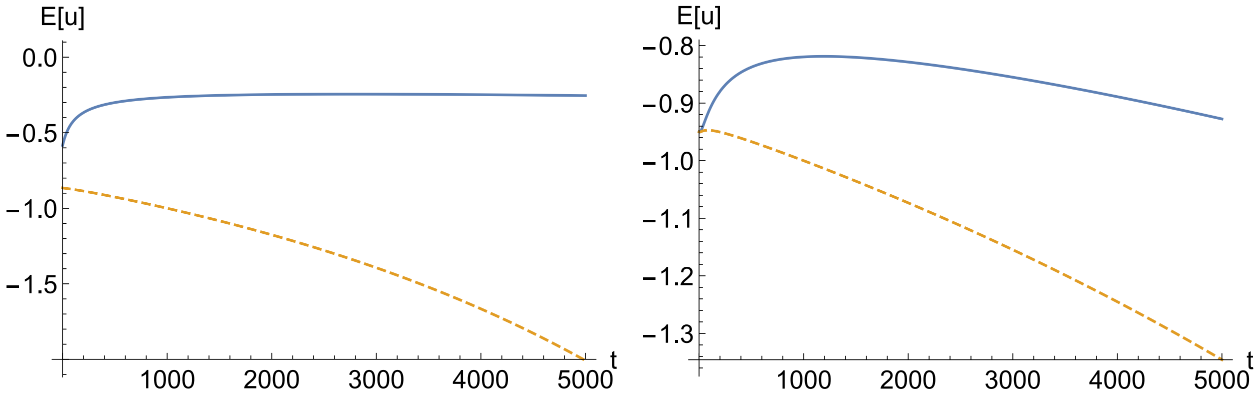
(i) *If the quantity of data analytics* $t = 0$ *and if the leverage constraint in Lemma 6 is not binding then the expected utility without leverage equals the expected utility with leverage.*

(ii) *If* $t > 0$ *then the expected utility without leverage is strictly lower than the expected utility with leverage.*

We next calculate the value of data analytics. For that, let us define $\Theta := e^r + \kappa\gamma s_0$. Then if quantity of analytics $t = 0$ and expected payoff $\hat{\theta}_0 > \Theta$ then the DM is bounded by her leverage constraint, and thus, $w^*(0) = \kappa$.

Proposition 5 (Value of data analytics without leverage)

The value of data analytics $v_{NL}(t) := u^*_{NL}(t) - u^*_{NL}(0)$ *is given as follows:*



(a) $\hat{\theta}_0 = 1.15 > e^r$ (b) $\hat{\theta}_0 = 1 < e^r$

Figure 13: **Expected utility of data analytics with and without leverage.** The blue solid line is with leverage and the orange dotted line is without leverage. Parameter values: $r = 0.05$, $\gamma = 0.01$, $s_0 = 0.01$, $c = 0.1$, and $\kappa = 100$.

(i) $\hat{\theta}_0 > \Theta$*: The leverage constraint is active,* $w^*(0) = \kappa$*, and* $v_{NL}(t)$ *is given by*

$$
\begin{aligned}
&-\frac{\exp\left(-\frac{\left(\hat{\theta}_0-e^r\right)^2+2\gamma(\kappa-ct)s_0(e^r-1)}{2s_0}\right)\left(\Phi\left(\frac{e^r-\hat{\theta}_0}{s_0\sqrt{t}}+\frac{\gamma(\kappa-ct)}{\sqrt{t}}\right)-\Phi\left(\frac{e^r-\hat{\theta}_0}{s_0\sqrt{t}}\right)\right)}{\sqrt{s_0t+1}} \\
&+e^{\frac{\gamma^2\kappa^2 s_0}{2}+\gamma\kappa\left(1-\hat{\theta}_0\right)}-e^{\frac{\gamma^2(\kappa-ct)^2 s_0}{2}+\gamma(\kappa-ct)\left(1-\hat{\theta}_0\right)}\Phi\left(\frac{\sqrt{s_0t+1}\left(\hat{\theta}_0-e^r\right)}{s_0\sqrt{t}}-\frac{\gamma(\kappa-ct)\sqrt{s_0t+1}}{\sqrt{t}}\right) \\
&-e^{\gamma(\kappa-ct)(1-e^r)}\Phi\left(\frac{\sqrt{s_0t+1}\left(e^r-\hat{\theta}_0\right)}{s_0\sqrt{t}}\right).
\end{aligned}
$$

(ii) $\hat{\theta}_0 \in (e^r, \Theta]$*: No constraints are active and* $v_{NL}(t)$ *is given by*

$$
\begin{aligned}
&\exp\left(-\frac{\left(\hat{\theta}_0-e^r\right)^2+2\gamma\kappa s_0\left(e^r-1\right)}{2s_0}\right)\left(1-e^{\gamma ct(e^r-1)}\frac{\left(\Phi\left(\frac{e^r-\hat{\theta}_0}{s_0\sqrt{t}}+\frac{\gamma(\kappa-ct)}{\sqrt{t}}\right)-\Phi\left(\frac{e^r-\hat{\theta}_0}{s_0\sqrt{t}}\right)\right)}{\sqrt{s_0t+1}}\right) \\
&-e^{\frac{\gamma^2(\kappa-ct)^2 s_0}{2}+\gamma(\kappa-ct)\left(1-\hat{\theta}_0\right)}\Phi\left(\frac{\sqrt{s_0t+1}\left(\hat{\theta}_0-e^r\right)}{s_0\sqrt{t}}-\frac{\gamma(\kappa-ct)\sqrt{s_0t+1}}{\sqrt{t}}\right) \\
&-e^{\gamma(\kappa-ct)\left(1-e^{rT}\right)}\Phi\left(\frac{\sqrt{s_0t+1}\left(e^r-\hat{\theta}_0\right)}{s_0\sqrt{t}}\right).
\end{aligned}
$$

(iii) $\hat{\theta}_0 \le e^r$*: The investment constraint is active,* $w^*(0) = 0$*, and* $v_{NL}(t)$ *is given by*

$$-\frac{\exp\left(-\frac{\left(\hat{\theta}_0 - e^r\right)^2 + 2\gamma(\kappa - ct)s_0(e^r - 1)}{2s_0}\right)\left(\Phi\left(\frac{e^r - \hat{\theta}_0}{s_0\sqrt{t}} + \frac{\gamma(\kappa - ct)}{\sqrt{t}}\right) - \Phi\left(\frac{e^r - \hat{\theta}_0}{s_0\sqrt{t}}\right)\right)}{\sqrt{s_0 t + 1}}$$
$$- e^{\frac{\gamma^2(\kappa - ct)^2 s_0}{2} + \gamma(\kappa - ct)\left(1 - \hat{\theta}_0\right)}\Phi\left(\frac{\sqrt{s_0 t + 1}\left(\hat{\theta}_0 - e^r\right)}{s_0\sqrt{t}} - \frac{\gamma(\kappa - ct)\sqrt{s_0 t + 1}}{\sqrt{t}}\right)$$
$$+ e^{\gamma\kappa(1 - e^r)}\left(1 - e^{\gamma ct(e^r - 1)}\Phi\left(\frac{\sqrt{s_0 t + 1}\left(e^r - \hat{\theta}_0\right)}{s_0\sqrt{t}}\right)\right).$$

The key difference between Propositions 1 and 5 is the one additional region in Proposition 5. This is due to the financial constraint in Lemma 6 that prevents leverage and requires us to consider the region where the DM's budget constraint is not tight ($e^r < \hat{\theta}_0 \le \Theta$).

The marginal value of data analytics is complicated, but we are able to analyze that when $\hat{\theta}_0 > \Theta$ and $s_0 - 2c\gamma(e^r - 1) > 0$. The first condition means that with zero data analytics ($t = 0$), the DM would prefer to invest more than her initial wealth and, therefore, her investment is bounded by the financial constraint. The second condition $s_0 - 2c\gamma(e^r - 1) > 0$ means that without the financial constraint the marginal value is positive at $t = 0$ (see Corollary 1), or equivalently in this case, $c < c_{CH}$. The following corollary compares the marginal values of data analytics with and without leverage under these two conditions (the proof is in Appendix C.11), and Figure 14 illustrates this result.

Corollary 6 (Marginal values with and without leverage)
If $\hat{\theta}_0 > \Theta$ *and* $c < c_{CH}$ *then the marginal value of data analytics without leverage is strictly lower than the marginal value of data analytics with leverage for all* $t \in \left(0, \frac{s_0 - 2c\gamma(e^r - 1)}{2s_0 c\gamma(e^r - 1)}\right)$.

From Corollary 6 we get the following result.

Proposition 6 (Lower optimal quantity)
If $\hat{\theta}_0 > \Theta$ *and* $c < c_{CH}$ *then the optimal quantity* t^* *is lower for a financially constrained DM.*

Proof. Let $v'(t)$ and $v'_{NL}(t)$ be the marginal values of data analytics with and without leverage, and t^* and t^*_{NL} be the respective optimal quantities. Then $v'(t)$ is strictly decreasing in t since the value of data analytics with leverage is strictly concave when $\hat{\theta}_0 > \Theta$. Suppose that $t^* \le t^*_{NL}$. Then $v'(t^*_{NL}) \le v'(t^*) = 0 = v'_{NL}(t^*_{NL})$, but this contradicts Corollary 6. □

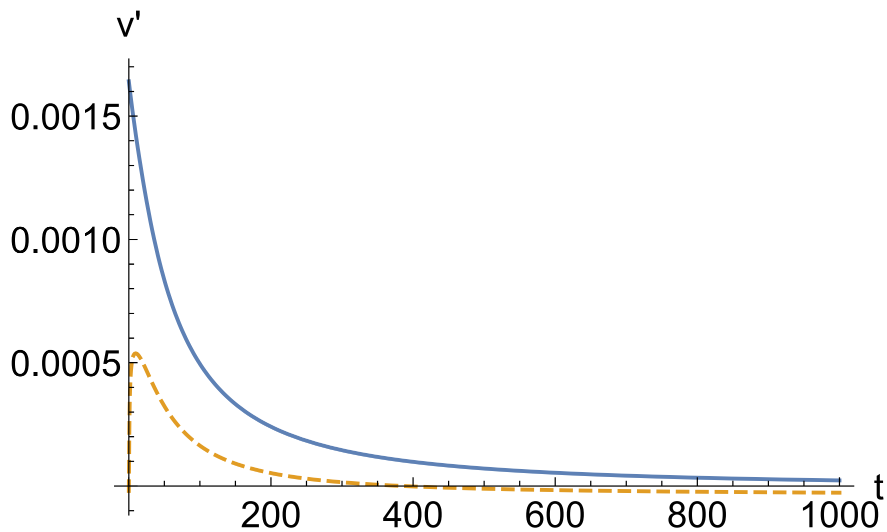


Figure 14: **The marginal value of data analytics with and without leverage.** The blue solid line is with leverage and the orange dotted line is without leverage. Parameter values: $r = 0.05$, $\gamma = 0.01$, $s_0 = 0.01$, $\kappa = 1000$, $c = 0.1$, and $\hat{\theta}_0 = 1.16$.

The value of data analytics for a financially constrained DM is bounded above by the corresponding value of data analytics without the financial constraint. The next result shows when the value under the financial constraint is strictly lower than the value without the constraint.

Corollary 7 (Value of data analytics with and without leverage)
The value of data analytics without leverage is strictly lower than the value of data analytics with leverage if one of the following conditions holds:

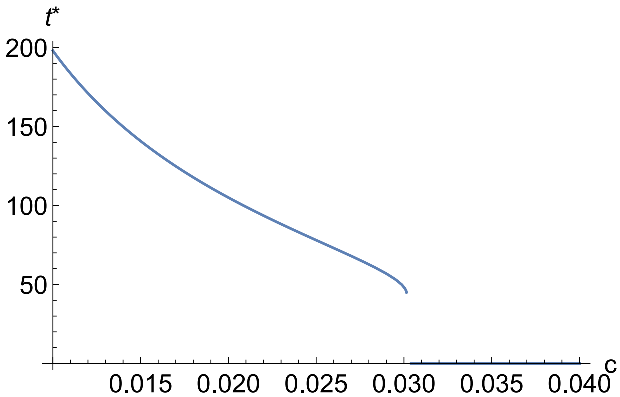
(i) $\hat{\theta}_0 > \Theta$, $c < c_{CH}$, *and* $t \in \left(0, \frac{s_0 - 2c\gamma(e^r - 1)}{2s_0 c\gamma(e^r - 1)}\right)$,

(ii) $\hat{\theta}_0 \leq \Theta$.

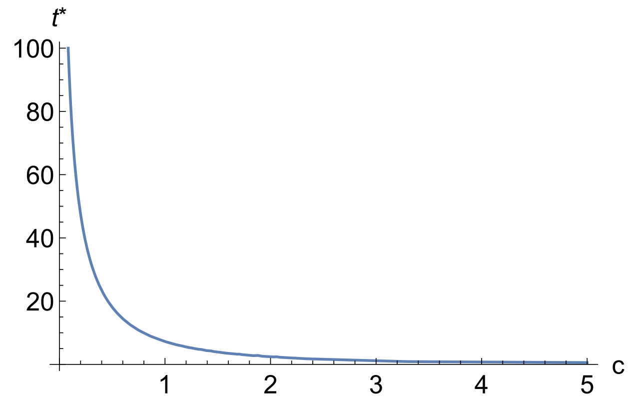
Proof. We have two cases:

(i) $\hat{\theta}_0 > \Theta$: We have $u^*(0) \neq u^*_{NL}(0)$. However, $v(0) = v_{NL}(0) = 0$ since $v(0) = u^*(0) - u^*(0) = 0$ and $v_{NL}(0) = u^*_{NL}(0) - u^*_{NL}(0) = 0$. Define $V(t) := v(t) - v_{NL}(t)$. Then $V(0) = 0$, $V(t)$ is continuous, and, by Corollary 6, $V'(t) = v'(t) - v'_{NL}(t) > 0$ for all $0 < t < \frac{s_0 - 2c\gamma(e^r - 1)}{2s_0 c\gamma(e^r - 1)}$. Since $V(0) = 0$ and $V'(t) > 0$, we have $v(t) > v_{NL}(t)$ for all $0 < t < \frac{s_0 - 2c\gamma(e^r - 1)}{2s_0 c\gamma(e^r - 1)}$.

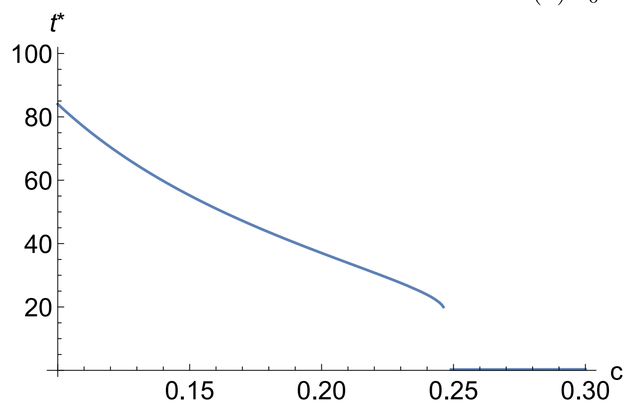
(ii) $\hat{\theta}_0 \leq \Theta$: We have $u^*(0) = u^*_{NL}(0)$. By Corollary 5, $u^*(t) > u^*_{NL}(t)$ $\forall t > 0$. Therefore, $v(t) = u^*(t) - u^*(0) > u^*_{NL}(t) - u^*_{NL}(0) = v_{NL}(t)$. $\square$



(a) $\hat{\theta}_0 = 1.15 > \Theta$



(b) $\hat{\theta}_0 = 1.06 < \Theta$

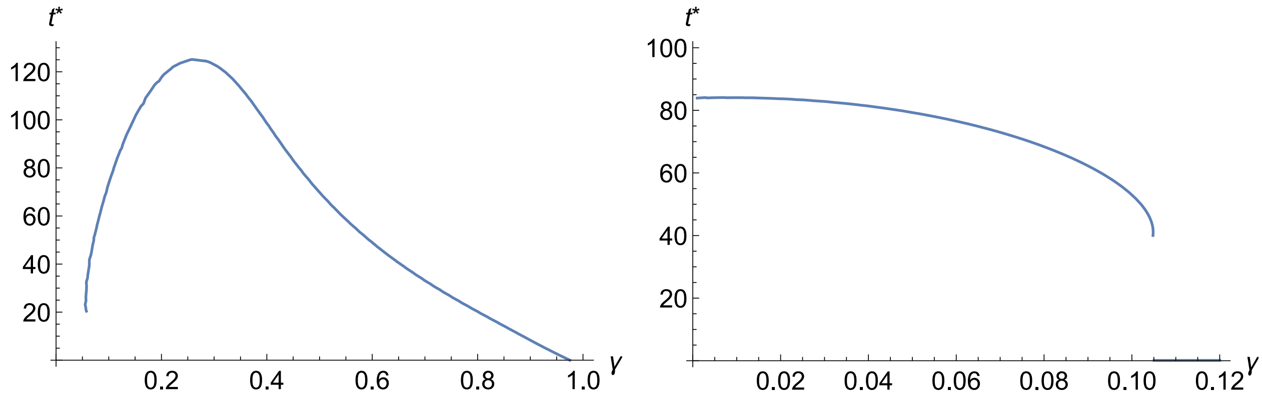


(c) $\hat{\theta}_0 = 1 < e^r$

Figure 15: **Optimal quantity of data analytics t^* without leverage and with respect to c.** Parameter values: $r = 0.05$, $\gamma = 0.01$, $s_0 = 0.01$, and $\kappa = 100$.

We illustrate the optimal quantity of data analytics without leverage in Figures 15 and 16. Figure 15 shows that the optimal quantity falls in cost, and it gives the demand curve for data analytics. Hence, the DM uses more analytics the less costly it is. The discontinuities in Figures 15a and 15c are due to the choke-off cost.

Figure 16 shows that the optimal quantity of data analytics without leverage is first rising and then falling in the risk aversion. The rising part is due to the leverage constraint. By Lemma 6, increasing risk aversion at a low level does not change the investment, because the leverage constraint is tight, but it raises the DM's incentive to decrease the uncertainty by data analytics. However, when the risk aversion is high enough, the investment amount falls because the leverage constraint is not active anymore (see Lemma 6), and this decreases the incentive to use data analytics. There are only two regions plotted in Figure 16 because Θ rises in γ, which means that in Figure 16a we consider both $\Theta > \hat{\theta}_0 > e^r$ and $\hat{\theta}_0 > \Theta$.



(a) $\hat{\theta}_0 = 1.15 > e^r$ (b) $\hat{\theta}_0 = 1 < e^r$

Figure 16: **Optimal quantity of data analytics t^* without leverage and with respect to γ.** Parameter values: $r = 0.05$, $c = 0.1$, $s_0 = 0.01$, and $\kappa = 100$.

Next we derive the expected optimal investment when the DM is not allowed to take leverage (the proof is in Appendix C.12).

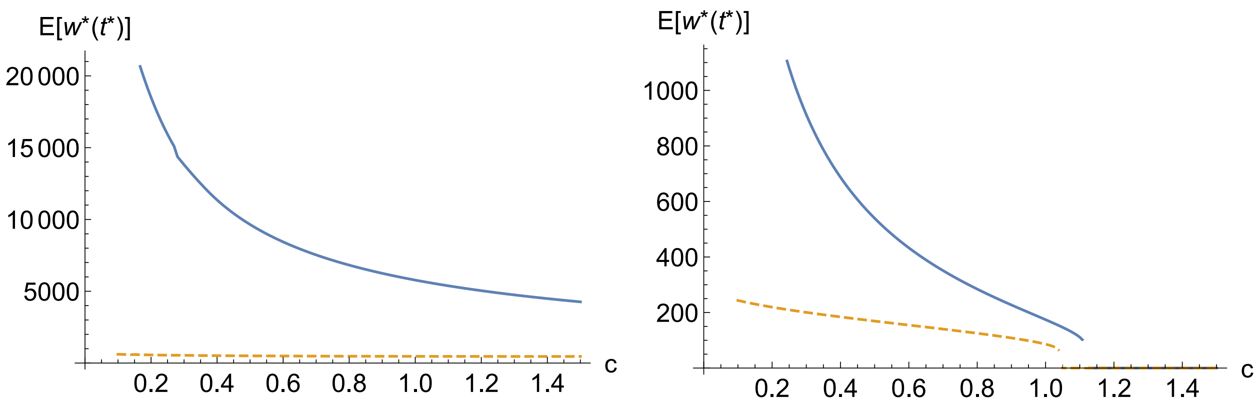
Proposition 7 (Expected optimal investment without leverage)

The expected optimal investment is given by

(i) $c \leq c_{CH}$*:*

$$\begin{aligned}
E\left[w^*_{NL}(t^*)\right] =& E\left[\min\left\{\max\left\{\frac{\hat{\theta}_{t^*} - e^r}{\gamma s_{t^*}}, 0\right\}, \kappa - ct\right\}\right] \\
=& \frac{(\hat{\theta}_0 - e^r)(1 + t^* s_0)}{s_0 \gamma}\left(\Phi\left(d^* + \frac{(\kappa - ct)\gamma}{\sqrt{t^*}\sqrt{1 + s_0 t^*}}\right) - \Phi\left(d^*\right)\right) \\
&+ \frac{\sqrt{t^* + s_0 (t^*)^2}}{\gamma}\left(\phi\left(d^*\right) - \phi\left(d^* + \frac{(\kappa - ct)\gamma}{\sqrt{t^*}\sqrt{1 + s_0 t^*}}\right)\right) \\
&+ (\kappa - ct)\left(1 - \Phi\left(d^* + \frac{(\kappa - ct)\gamma}{\sqrt{t^*}\sqrt{1 + s_0 t^*}}\right)\right),
\end{aligned}$$

where $d^* = \frac{(e^r - \hat{\theta}_0)\sqrt{\frac{1}{t^*} + s_0}}{s_0}$ *and the expectation is with respect to $\mathcal{F}_0$, so unconditional on the signal outcome.*



(a) $\hat{\theta}_0 = 1.15 > e^r$

(b) $\hat{\theta}_0 = 1 < e^r$

Figure 17: **Expected optimal investment with respect to** c**.** The blue solid line is with leverage and the orange dotted line is without leverage. Parameter values: $r = 0.05$, $\gamma = 0.01$, $s_0 = 0.01$, and $\kappa = 1000$.

(ii) $c > c_{CH}$*:*

$$E\left[w^*_{NL}(0)\right] = \begin{cases} \kappa & \textit{if } \hat{\theta}_0 > \Theta \\ \frac{\hat{\theta}_0 - e^r}{s_0 \gamma} & \textit{if } e^r < \hat{\theta}_0 \leq \Theta \\ 0 & \textit{if } \hat{\theta}_0 \leq e^r \end{cases}$$

The expected optimal investment is lower for a financially constrained DM, because her investment is bounded. We are not able to compare directly the expected optimal investments in Propositions 4 and 7 because their optimal quantities of data analytics are different and we do not have explicit formulas for t^*. Therefore, we compare them numerically in Figure 17, where we plot the expected optimal investments with and without leverage against the cost of data analytics.

We see from Figure 17 that the expected optimal investment with leverage falls in the cost of data analytics c. Therefore, lower cost of data analytics raises the expected size of the investment. That is, when c is low, the DM acquires more data analytics which then lowers the risk and raises the investment amount as shown in Figure 17. The rising investment size means higher leverage and, thus, leverage falls in the cost of data analytics. This is consistent with Maksimovic et al. (1999) who show that in high-growth industries levered firms acquire more information than all-equity financed firms.

7 Conclusion

We have derived a new investment model where a risk averse decision maker faces a joint decision of the level of data analytics and investment amount. Data analytics is costly due to the price of datasets and efforts in collecting and processing the data. The more the investor buys and analyzes data, the better she can forecast the investment payoff, which decreases the riskiness of the investment opportunity. This indicates that data analytics is a risk management device. After observing the results of the data analytics, the decision maker decides how much to invest and, if needed, she uses leverage in the investment.

We showed that the more data analytics the decision maker acquires, the more she invests on average and, therefore, the more she uses leverage. Further, the lower the cost of data analytics, naturally the more the decision maker acquires analytics, which decreases the riskiness of the investment opportunity and this way raises the investment amount and leverage. We showed that the higher leverage, driven by the falling cost of data analytics, leads to higher losses during the crises. We also showed that decision makers with low risk aversion and without leverage constraints acquire more data analytics because they invest more and, therefore, they have a higher incentive to collect new information on the investments. We also showed that the demand of data analytics is highest with high expected return opportunities, as decision makers possessing such opportunities use a high level of leverage and, therefore, decide to analyze the investment opportunities thoroughly.

A Nonlinear cost

We use the same model as in our main analysis, except the cost of data analytics is quadratic, ct^2. We get directly the following result.

Lemma 8 (Expected utility)
Given the quantity of analytics t, the expected utility of the optimal investment before observing the outcome of the analytics is given by

$$\begin{aligned} u^*(t) =& E\left[-\exp(-\gamma(w^*(t)(\theta-1)+(\kappa-ct^2-w^*(t))(e^r-1)))\right] \\ =& -e^{\gamma(\kappa-ct^2)(1-e^r)}\Phi(d(t)) - \frac{\Phi\left(\frac{-d(t)}{\sqrt{s_0t+1}}\right)\exp\left(-\frac{\hat{\theta}_0^2+e^{2r}-2\hat{\theta}_0e^r+2\gamma s_0(\kappa-ct^2)(e^r-1)}{2s_0}\right)}{\sqrt{s_0t+1}}, \end{aligned}$$

where $\Phi(\cdot)$ is cumulative standard normal distribution and $d(t) := \left(e^r-\hat{\theta}_0\right)\sqrt{\frac{s_0t+1}{s_0^2t}}$. The expectation is with respect to $\mathcal{F}_0$, so unconditional on the signal outcome.

The optimal investment remains the same as in the main analysis. Next we give the marginal value of data analytics under the quadratic cost:

Lemma 9 (Marginal value of data analytics)
For $t>0$, the marginal value of data analytics is given by

$$\begin{aligned} v'(t) =& \frac{e^{\gamma(e^r-1)\left(ct^2-\kappa\right)}}{2(s_0t+1)^{3/2}}\Bigg[-4ct\gamma\left(e^r-1\right)\left(s_0t+1\right)^{3/2}\Phi\left(d(t)\right) \\ & -4ct\gamma\left(e^r-1\right)\left(s_0t+1\right)e^{-\frac{\left(e^r-\hat{\theta}_0\right)^2}{2s_0}}\Phi\left(-\frac{d(t)}{\sqrt{s_0t+1}}\right)+s_0e^{-\frac{\left(e^r-\hat{\theta}_0\right)^2}{2s_0}}\Phi\left(-\frac{d(t)}{\sqrt{s_0t+1}}\right)\Bigg]. \end{aligned}$$

As with the linear cost, the value function is strictly concave for $\hat{\theta}_0 > e^r$.

Proposition 8 (Shape of the value of data analytics)
The value of data analytics function is strictly concave if $\hat{\theta}_0 > e^r$.

Proof. The second derivative of the value function

$$\begin{aligned} v'' =& \frac{e^{\gamma(e^r-1)\left(ct^2-\kappa\right)}}{8t^2(s_0t+1)^{7/2}}\Bigg(-16ct^2\gamma\left(e^r-1\right)\left(s_0t+1\right)^{7/2}\left(2c\gamma\left(e^r-1\right)t^2+1\right)\Phi\left(\left(e^r-\hat{\theta}_0\right)\sqrt{\frac{s_0t+1}{s_0^2t}}\right) \\ & -2t^2(s_0t+1)e^{-\frac{\left(e^r-\hat{\theta}_0\right)^2}{2s_0}}\left(8c\gamma\left(e^r-1\right)\left(s_0t+1\right)\left(2ct^2\gamma\left(e^r-1\right)\left(s_0t+1\right)+1\right)+3s_0^2\right)\Phi\left(\frac{\left(\hat{\theta}_0-e^r\right)}{s_0\sqrt{t}}\right) \\ & +\sqrt{\frac{2t}{\pi}}\left(e^r-\hat{\theta}_0\right)\left(s_0t+1\right)^2e^{-\frac{\left(e^r-\hat{\theta}_0\right)^2(s_0t+1)}{2s_0^2t}}\Bigg) \end{aligned}$$

is strictly negative if $\hat{\theta}_0 > e^r$. $\square$

So when $\hat{\theta}_0 > e^r$, we have a unique global maximum given by the first order condition $v'(t) = 0$. The first order condition can be written as follows

$$
\begin{aligned}
& - 4ct\gamma \left(e^r - 1\right) \left(s_0 t + 1\right)^{3/2} \Phi\left(d(t)\right) - 4ct\gamma \left(e^r - 1\right) (s_0 t + 1) e^{-\frac{\left(e^r - \hat{\theta}_0\right)^2}{2 s_0}} \Phi\left(-\frac{d(t)}{\sqrt{s_0 t + 1}}\right) \\
& + s_0 e^{-\frac{\left(e^r - \hat{\theta}_0\right)^2}{2 s_0}} \Phi\left(-\frac{d(t)}{\sqrt{s_0 t + 1}}\right) = 0. \qquad (7)
\end{aligned}
$$

As in the main analysis, we have the following results.

Theorem 4 (Properties of the optimal quantity)

We have:

(i) *The optimal quantity t^* is independent of initial wealth κ.*

(ii) *If $\hat{\theta}_0 > e^r$, the optimal quantity t^* falls in cost c.*

Proof. We can see from (7) that the first order condition that it is independent of κ, thus the optimal quantity is also independent of κ. Let the left hand side of (7) be $f(t^*)$. We have:

$$
\begin{aligned}
\frac{\partial f}{\partial t^*} = & \frac{\left(e^r - \hat{\theta}_0\right) \phi\left(\left(e^r - \hat{\theta}_0\right) \sqrt{\frac{s_0 t + 1}{s_0^2 t}}\right)}{2 t^{3/2}} - 4c\gamma \left(e^r - 1\right) (2 s_0 t + 1) e^{-\frac{\left(e^r - \hat{\theta}_0\right)^2}{2 s_0}} \Phi\left(\frac{\hat{\theta}_0 - e^r}{s_0 \sqrt{t}}\right) \\
& - 2c\gamma \left(e^r - 1\right) (5 s_0 t + 2) \sqrt{s_0 t + 1} \Phi\left(\left(e^r - \hat{\theta}_0\right) \sqrt{\frac{s_0 t + 1}{s_0^2 t}}\right) \\
\frac{\partial f}{\partial c} = & - 4t\gamma \left(e^r - 1\right) (s_0 t + 1) \left(e^{-\frac{\left(e^r - \hat{\theta}_0\right)^2}{2 s_0}} \Phi\left(\frac{\hat{\theta}_0 - e^r}{s_0 \sqrt{t}}\right) + \sqrt{s_0 t + 1} \Phi\left(\left(e^r - \hat{\theta}_0\right) \sqrt{\frac{s_0 t + 1}{s_0^2 t}}\right)\right)
\end{aligned}
$$

and by the Implicit Function Theorem, we get $\frac{\partial t^*}{\partial c} = -\frac{\frac{\partial f}{\partial c}}{\frac{\partial f}{\partial t^*}}$ which is strictly negative when $\hat{\theta}_0 > e^r$. $\square$

Since the optimal investment formula is the same as in the main analysis, the expected optimal investment function remains the same. Then, by Theorem 4, Theorem 3 holds also for the quadratic cost of data analytics.

B Multiple Investments

We extend the results to n independent investment opportunities, with the parameters of the DM remaining the same. For each investment $i \in \{1, 2, ..., n\}$, we have the unknown payoff θ_i, and the prior expected payoff $\hat{\theta}_{0,i}$ and variance $s_{0,i}$. The cost of acquiring quantity t_i of data analytics for investment i is c_i. Hence, the utility of the DM in the next period is given by

$$\begin{aligned}
&-\exp\left(-\gamma\left[\sum_{i=1}^{n} w_i(\theta_i - 1) + (\kappa - \sum_{i=1}^{n} c_i t_i - \sum_{i=1}^{n} w_i)(e^r - 1)\right]\right) \\
&= -\exp(-\gamma[\boldsymbol{w} \cdot (\boldsymbol{\theta} - \mathbf{1}) + (\kappa - \boldsymbol{c} \cdot \boldsymbol{t} - \boldsymbol{w} \cdot \mathbf{1})(e^r - 1)]).
\end{aligned}$$

We assume that signals on different investments are independent. As in the main analysis, let $(\Omega_i, \mathcal{F}_i, \mathbb{P}_i)$ be a probability space for investment $i \in \{1, 2, ..., n\}$. The quantity of data analytics t_i corresponds to observing a noisy signal $Z_{t_i,i}$ on payoff θ_i from 0 to t_i. The noisy signal follows $dZ_{t_i,i} = \theta_i dt_i + dB_{t_i,i}$, where $B_{t_i,i}$ is a standard Wiener process independent of the other uncertainties. $\mathcal{F}_{t_i,i}$ is the information level from observing the noisy signal $Z_{t_i,i}$ up to quantity t_i. Thus, the DM can learn about the payoff θ_i by obtaining quantity $t_i \geq 0$ of data analytics at a marginal cost $c_i > 0$.

Lemmas 10 and 11 are derived directly from their counterparts in the single investment case and their proofs are exactly the same, except the variables are indexed by $i \in \{1, 2, ..., n\}$.

Lemma 10 (Unconditional and conditional payoffs for investment i)
For each investment $i \in \{1, 2, ..., n\}$, unconditionally on the signal outcome, we have

$$\theta_i = \hat{\theta}_{0,i} + \sqrt{\frac{s_{0,i}^2 t_i}{1 + s_{0,i} t_i}} \epsilon_{t,i} + \sqrt{s_{t,i}} \epsilon_i$$

and conditional on the signal outcome i.e., with respect to information $\mathcal{F}_{t,i}$, we have

$$\theta_i = \hat{\theta}_{t,i} + \sqrt{s_{t,i}} \epsilon_t,$$

where $\hat{\theta}_{0,i}$ and $s_{0,i}$ are the DM's prior mean and variance for investment i, $s_{t,i} = \frac{s_{0,i}}{1+t_i s_{0,i}}$ is the posterior variance of investment i, $\epsilon_{t,i}$ is $\mathcal{F}_{t,i}$-measurable standard normal random variable, ϵ_i is a standard normal variable with respect to informations $\mathcal{F}_{0,i}$ and $\mathcal{F}_{t,i}$, ϵ_i and $\epsilon_{t,i}$ are independent with respect to information $\mathcal{F}_{0,i}$.

Proof. See Lemma 1. □

Lemma 11 (Optimal investment for multiple investments)
Given analytics quantity $\boldsymbol{t} = (t_1, t_2, \ldots, t_n)$ and CARA utility function with $\gamma > 0$, the optimal level

of risky investment is given by $\boldsymbol{w^*}(\boldsymbol{t}) = (w_1^*(t_1), w_2^*(t_2), \ldots, w_n^*(t_n))$, *where*

$$w_i^*(t_i) = \max\left\{\frac{\hat{\theta}_{t_i,i} - e^r}{\gamma s_{t_i,i}}, 0\right\}. \tag{8}$$

Proof. See Lemma 2. □

Lemma 12 (Expected utility)

The expected utility of the optimal investment for multiple investments

$$\begin{aligned} u^*(\boldsymbol{t}) =& E\left[-\exp(-\gamma[\boldsymbol{w^*(t)} \cdot (\boldsymbol{\theta} - \mathbf{1}) + (\kappa - \boldsymbol{c} \cdot \boldsymbol{t} - \boldsymbol{w^*(t)} \cdot \mathbf{1})(e^r - 1)])\right] \\ =& -e^{\gamma(\kappa - \sum_{i=1}^n c_i t_i)(1-e^r)} \prod_{i=1}^n \left(\Phi(d_i(t_i)) + \frac{\Phi\left(\frac{-d_i(t_i)}{\sqrt{s_{0,i}t_i+1}}\right)\exp\left(-\frac{(\hat{\theta}_{0,i}-e^r)^2}{2s_{0,i}}\right)}{\sqrt{s_{0,i}t_i + 1}}\right), \end{aligned}$$

where $\Phi(\cdot)$ *is cumulative standard normal distribution and* $d_i(t_i) := \left(e^r - \hat{\theta}_{0,i}\right)\sqrt{\frac{s_{0,i}t_i+1}{s_{0,i}^2 t_i}}$.

Proof. Similar to the single investment case, we find the expected utility by integrating the utility in the next period over $\mathbb{R}^{2n}$ with respect to the $2n$ normal PDFs and replacing the ϵ_i's and $\epsilon_{t,i}$'s by variables x_i and x_{i+1} for investment i:

$$\int_{\mathbb{R}^{2n}} \cdots \int -\exp(-\gamma[\boldsymbol{w^*(t)} \cdot (\boldsymbol{\theta} - \mathbf{1}) + (\kappa - \boldsymbol{c} \cdot \boldsymbol{t} - \boldsymbol{w^*(t)} \cdot \mathbf{1})(e^r - 1)]) \frac{1}{(2\pi)^n} e^{-\frac{1}{2}\sum_{i=1}^{2n} x_i} \prod_{i=1}^{2n} dx_i.$$

We then integrate with respect to the dummy variables associated with $\epsilon_{1,i}$ to get

$$\frac{1}{(\sqrt{2\pi})^n} \int_{\mathbb{R}^n} \cdots \int -e^{\gamma(\kappa - \sum_{i=1}^n c_i t_i)(1-e^r)} \prod_{i=1}^n \exp\left(\gamma w_i^* e^r + \frac{\gamma^2 s_{0,i} w_i^{*2}}{2s_{0,i}t_i + 2} - \gamma w_i^* x_i \sqrt{\frac{s_{0,i}^2 t_i}{s_{0,i}t_i + 1}} - \gamma\hat{\theta}_{0,i} w_i^* - \frac{x_i^2}{2}\right) \prod_{i=1}^n dx_i.$$

Observing that each i-th term in the product is similar to that in the single investment case, and by the independence of the normal variables, we get:

$$-e^{\gamma(\kappa - \sum_{i=1}^n c_i t_i)(1-e^r)} \prod_{i=1}^n \left(\Phi(d_i(t_i)) + \frac{\Phi\left(\frac{-d_i(t_i)}{\sqrt{s_{0,i}t_i+1}}\right)\exp\left(-\frac{(\hat{\theta}_{0,i}-e^r)^2}{2s_{0,i}}\right)}{\sqrt{s_{0,i}t_i + 1}}\right).$$

□

In the next proposition, we introduce new index notation i_j that we formally define in the proof. This index represents investments that have $\hat{\theta}_{0,i_j} = \hat{\theta}_{0,k} > e^r$ where $k \in \{1,2,\ldots,n\}$. Thus, the set $\{i_1, i_2, \ldots, i_m\}$ is a subset of the set of n independent investments $\{1,2,\ldots,n\}$.

Proposition 9 (Value of data analytics)
Let i_j be the index for investments such that $\hat{\theta}_{0,i_j} > e^r$ and assume there are $0 \le m \le n$ such investments. Then the value of data analytics $v(\boldsymbol{t}) := u^*(\boldsymbol{t}) - u^*(\boldsymbol{0})$ is given by

$$
\begin{aligned}
&e^{\gamma\kappa(1-e^r)} \prod_{i_j=i_1}^{i_m} \exp\left(-\frac{(\hat{\theta}_{0,i_j} - e^r)^2}{2s_{0,i_j}}\right) \\
&-e^{\gamma(\kappa - \sum_{i=1}^n c_i t_i)(1-e^r)} \prod_{i=1}^{n} \left(\Phi(d_i(t_i)) + \frac{\Phi\left(\frac{-d_i(t_i)}{\sqrt{s_{0,i}t_i+1}}\right) \exp\left(-\frac{(\hat{\theta}_{0,i}-e^r)^2}{2s_{0,i}}\right)}{\sqrt{s_{0,i}t_i+1}} \right).
\end{aligned}
$$

Note that if $m = 0$ then $\prod_{i_j=i_1}^{i_m} \exp\left(-\frac{(\hat{\theta}_{0,i_j}-e^r)^2}{2s_{0,i_j}}\right) = 1$.

Proof. Observe that

$$
\lim_{t_i \downarrow 0} d_i(t_i) = \lim_{t_i \downarrow 0} \left(e^r - \hat{\theta}_{0,i}\right) \sqrt{\frac{s_{0,i}t_i + 1}{s_{0,i}^2 t_i}} = \begin{cases} -\infty & \text{if } \hat{\theta}_{0,i} > e^r \\ 0 & \hat{\theta}_{0,i} = e^r \\ +\infty & \text{if } \hat{\theta}_{0,i} < e^r \end{cases}
$$

and

$$
\lim_{t_i \downarrow 0} \frac{-d_i(t_i)}{\sqrt{s_{0,i}t_i + 1}} = \lim_{t_i \downarrow 0} \left(e^r - \hat{\theta}_{0,i}\right) \frac{-1}{\sqrt{s_{0,i}^2 t_i}} = \begin{cases} +\infty & \text{if } \hat{\theta}_{0,i} > e^r \\ 0 & \text{if } \hat{\theta}_{0,i} = e^r \\ -\infty & \text{if } \hat{\theta}_{0,i} < e^r \end{cases}
$$

Define $J := \{\hat{\theta}_{0,i} \mid \hat{\theta}_{0,i} > e^r, i \in \{1,2,\ldots,n\}\}$ and $m := |J|$. If $J \neq \emptyset$, i.e., $m \neq 0$, then define index $j \in \{1,2,\ldots,m\}$ so that $\hat{\theta}_{0,i_j} \in J$ for all $j \in \{1,2,\ldots,m\}$. To find $\lim_{\boldsymbol{t}\downarrow\boldsymbol{0}} u^*(\boldsymbol{t})$, we have

$$
e^{\gamma(\kappa - \sum_{i=1}^n c_i t_i)(1-e^r)} \to e^{\gamma\kappa(1-e^r)}
$$

and

$$\prod_{i=1}^{n}\left(\Phi(d_i(t_i))+\frac{\Phi\left(\frac{-d_i(t_i)}{\sqrt{s_{0,i}t_i+1}}\right)\exp\left(-\frac{(\hat{\theta}_{0,i}-e^r)^2}{2s_{0,i}}\right)}{\sqrt{s_{0,i}t_i+1}}\right)\to\begin{cases}\prod_{i_j=i_1}^{i_m}\exp\left(-\frac{(\hat{\theta}_{0,i_j}-e^r)^2}{2s_{0,i_j}}\right) & \text{if } m\neq 0\\ 1 & \text{if } m=0\end{cases}$$

when $\boldsymbol{t}\downarrow 0$. We have this because the $n-m$ investments with $\hat{\theta}_{0,i}\leq e^r$ have their corresponding limits as one in the product, and the terms involving the m investments with $\hat{\theta}_{0,i_j}>e^r$ have corresponding limits $\exp\left(-\frac{(\hat{\theta}_{0,i_j}-e^r)^2}{2s_{0,i_j}}\right)$. □

To avoid confusion in the sum and product terms, in the next corollary we use z for index for investments.

Corollary 8 (Marginal value of data analytics)
The marginal value of data analytics for the z-th investment is given by

$$\frac{\partial v(\boldsymbol{t})}{\partial t_z}=\frac{e^{\gamma(e^r-1)(c_zt_z-\kappa)}}{2(s_{0,z}t_z+1)^{3/2}}\Bigg[-2c_z\gamma\left(e^r-1\right)\left(s_{0,z}t_z+1\right)^{3/2}\Phi\left(d(t_z)\right)$$
$$-2c_z\gamma\left(e^r-1\right)\left(s_{0,z}t_z+1\right)e^{-\frac{\left(e^r-\hat{\theta}_{0,z}\right)^2}{2s_{0,z}}}\Phi\left(-\frac{d(t_z)}{\sqrt{s_{0,z}t_z+1}}\right)+s_{0,z}e^{-\frac{\left(e^r-\hat{\theta}_{0,z}\right)^2}{2s_{0,z}}}\Phi\left(-\frac{d(t_z)}{\sqrt{s_{0,z}t_z+1}}\right)\Bigg]\cdot$$
$$e^{\gamma\left(\sum\limits_{i=1,i\neq z}^{n}c_it_i\right)(e^r-1)}\prod_{i=1,i\neq z}^{n}\left(\Phi(d_i(t_i))+\frac{\Phi\left(\frac{-d_i(t_i)}{\sqrt{s_{0,i}t_i+1}}\right)\exp\left(-\frac{(\hat{\theta}_{0,i}-e^r)^2}{2s_{0,i}}\right)}{\sqrt{s_{0,i}t_i+1}}\right).$$

Proof. Since we take derivative with respect to only the z-th investment, we can treat the other variables as constants. Thus, the result is obtained directly from the single investment case in Lemma 4, but multiplied by the other variables associated with the other investments. □

Our results with a single investment case hold for each investment in the multiple investment case because the marginal value of data analytics for each investment is just a positive scalar multiplied by the marginal value of data analytics in the single investment case.

Hence, following the same argument as in the single investment case, the optimal quantity of data analytics for investment j is given by the first order condition $\frac{\partial v(\boldsymbol{t})}{\partial t_j}=c_j$. To solve for $\boldsymbol{t}^*$

requires solving the system of non-linear equations

$$\begin{cases} \frac{\partial v(\boldsymbol{t})}{\partial t_1} & = 0 \\ \frac{\partial v(\boldsymbol{t})}{\partial t_2} & = 0 \\ & \vdots \\ \frac{\partial v(\boldsymbol{t})}{\partial t_n} & = 0 \end{cases} \Rightarrow \begin{cases} f_1(t_1) & = 0 \\ f_2(t_2) & = 0 \\ & \vdots \\ f_n(t_n) & = 0 \end{cases}$$

where

$$\begin{aligned} f_i(t_i) = & -2c_i\gamma\left(e^r-1\right)\left(s_{0,i}t_i+1\right)^{3/2}\Phi\left(d(t_i)\right) \\ & -2c_i\gamma\left(e^r-1\right)\left(s_{0,i}t_i+1\right)e^{-\frac{\left(e^r-\hat{\theta}_{0,i}\right)^2}{2s_{0,i}}}\Phi\left(-\frac{d(t_i)}{\sqrt{s_{0,i}t_i+1}}\right) \\ & + s_{0,i}e^{-\frac{\left(e^r-\hat{\theta}_{0,i}\right)^2}{2s_{0,i}}}\Phi\left(-\frac{d(t_i)}{\sqrt{s_{0,i}t_i+1}}\right). \end{aligned}$$

Note each equation for investment i is independent of the other $n-1$ equations, and $f_i(t_i)$ is just the indexed version of (3) in the single investment case. Hence, Theorem 2 holds for each investment i.

Furthermore, the optimal level of investment i is given by Lemma 11 which has the same formula as in Lemma 2. Hence, the corresponding expected optimal investment i with leverage is given by Proposition 4. Then the total expected optimal investment is given by

$$E[\boldsymbol{w}^*(\boldsymbol{t}^*)] = \sum_{i=1}^{n} E[w_i^*(t_i^*)].$$

As in the single investment case, the falling cost of data analytics raises leverage, which causes higher losses during the crises. This holds for each investment $i \in \{1, 2, \ldots, n\}$.

C Omitted Proofs

C.1 Proof of Lemma 1

The payoff θ is constant but unknown, and the DM learns that from noisy observations $dZ_t = \theta dt + dB_t$. By the one dimensional Kalman-Bucy filter (see e.g. Øksendal (2003)), belief $\hat{\theta}_t = E[\theta|\mathcal{F}_t]$ follows

$$d\hat{\theta}_t = -s_t\hat{\theta}_t dt + s_t dZ_t, \tag{9}$$

where the DM's prior mean $\hat{\theta}_0 = E[\theta|\mathcal{F}_0]$, and variance $s_t = E[(\theta - \hat{\theta}_t)^2|\mathcal{F}_t]$ satisfies the (deterministic) Riccati equation $\frac{ds_t}{dt} = -s_t^2$, where s_0 is the prior variance. Solving the Riccati equation gives $s_t = \left(s_0^{-1} + \int_0^t 1\, ds\right)^{-1} = \frac{s_0}{1+ts_0}$. By the dynamics of Z_t and (9), we have

$$d\hat{\theta}_t = s_t(dZ_t - \hat{\theta}_t dt) = s_t(\theta - \hat{\theta}_t)dt + s_t dB_t,$$

which gives $\hat{\theta}_t = \hat{\theta}_0 + \int_0^t s_y(\theta - \hat{\theta}_y)dy + \int_0^t s_y\, dB_y$. Therefore, we have

$$\begin{aligned}
E[\hat{\theta}_t] &= \hat{\theta}_0 + E\left[\int_0^t s_y(\theta - \hat{\theta}_y)dy\right] = \hat{\theta}_0 + \int_0^t E\left[s_y(\theta - \hat{\theta}_y)\right] dy \\
&= \hat{\theta}_0 + \int_0^t E\left[E\left[s_y(\theta - \hat{\theta}_y)|\mathcal{F}_y\right]\right] dy = \hat{\theta}_0 \text{ (since } E[\theta|\mathcal{F}_y] = \hat{\theta}_y).
\end{aligned}$$

Next we solve for the variance of $\hat{\theta}_t$. By Ito's lemma,

$$d\hat{\theta}_t^2 = 2\hat{\theta}_t d\hat{\theta}_t + \frac{1}{2}(2)(d\hat{\theta}_t)^2 = 2\hat{\theta}_t s_t(\theta - \hat{\theta}_t)dt + 2s_t\hat{\theta}_t dB_t + s_t^2 dt,$$

which gives $\hat{\theta}_t^2 = \hat{\theta}_0^2 + \int_0^t 2\hat{\theta}_y s_y(\theta - \hat{\theta}_y)dy + \int_0^t 2s_y\hat{\theta}_y dB_y + \int_0^t s_y^2 dy$. Then

$$\begin{aligned}
E(\hat{\theta}_t^2) &= \hat{\theta}_0^2 + E\left[\int_0^t 2\hat{\theta}_y s_y(\theta - \hat{\theta}_y)dy\right] + E\left[\int_0^t s_y^2\, dy\right] \\
&= \hat{\theta}_0^2 + \int_0^t E\left[2\hat{\theta}_y s_y(\theta - \hat{\theta}_y)\right] dy + E\left[\int_0^t s_y^2\, dy\right] \\
&= \hat{\theta}_0^2 + \int_0^t E\left[E\left[2\hat{\theta}_y s_y(\theta - \hat{\theta}_y)|\mathcal{F}_y\right]\right] dy + \frac{s_0^2 t}{1+ts_0} \\
&= \hat{\theta}_0^2 + \int_0^t E\left[2\hat{\theta}_y s_y E\left[(\theta - \hat{\theta}_y)|\mathcal{F}_y\right]\right] dy + \frac{s_0^2 t}{1+ts_0} = \hat{\theta}_0^2 + \frac{s_0^2 t}{1+ts_0}.
\end{aligned}$$

Therefore, $Var(\hat{\theta}_t) = \frac{s_0^2 t}{1+ts_0}$ and, unconditional on the signal outcome, $\hat{\theta}_t \sim N(\hat{\theta}_0, \frac{s_0^2 t}{1+ts_0})$. Hence, $\theta = \hat{\theta}_0 + \sqrt{\frac{s_0^2 t}{1+s_0 t}}\epsilon_t + \sqrt{s_t}\epsilon$, where $s_t = \frac{s_0}{1+ts_0}$, ϵ_t is $\mathcal{F}_t$-measurable standard normal random variable, ϵ is a standard normal variable with respect to informations $\mathcal{F}_0$ and $\mathcal{F}_t$, and ϵ and ϵ_t are independent with respect to information $\mathcal{F}_0$. $\square$

C.2 Proof of Lemma 3

By Lemma 1, $\theta = \hat{\theta}_0 + \sqrt{\frac{s_0^2 t}{1+s_0 t}}\epsilon_t + \sqrt{s_t}\epsilon$, where $\epsilon, \epsilon_t \sim N(0,1)$ with respect to information $\mathcal{F}_0$, $\epsilon \perp \epsilon_t$, $s_t = \frac{s_0}{1+ts_0}$, and $\hat{\theta}_t = \hat{\theta}_0 + \sqrt{\frac{s_0^2 t}{1+s_0 t}}\epsilon_t$. Let z and x be the realized values of ϵ and ϵ_t, respectively. Then we have with quantity t of analytics $\hat{\theta}_t = \hat{\theta}_0 + x\sqrt{\frac{s_0^2 t}{s_0 t+1}}$ and $\theta = \hat{\theta}_0 + x\sqrt{\frac{s_0^2 t}{s_0 t+1}} + z\sqrt{\frac{s_0}{s_0 t+1}}$. Further, the utility in the next period:

$$\begin{aligned}&-\exp\left[-\gamma\left(w^*\left(\theta-1\right)+\left(\kappa-ct-w^*\right)\left(e^r-1\right)\right)\right]\\&=-\exp\left(-\gamma\left[w^*\left(\hat{\theta}_0-1+x\sqrt{\frac{s_0^2 t}{s_0 t+1}}+z\sqrt{\frac{s_0}{s_0 t+1}}\right)+\left(e^r-1\right)\left(\kappa-ct-w^*\right)\right]\right).\end{aligned}$$

To find the expected utility, we take the double integral over the entire domain $\mathbb{R}^2$ with respect to the two normal PDFs:

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} -e^{-\frac{x^2+z^2}{2}}\exp\left(-\gamma\left[\left(e^r-1\right)\left(\kappa-ct-w^*\right)+w^*\left(\hat{\theta}_0-1+x\sqrt{\frac{s_0^2 t}{s_0 t+1}}+z\sqrt{\frac{s_0}{s_0 t+1}}\right)\right]\right)dzdx. \tag{10}$$

We first evaluate the integral with respect to z:

$$\begin{aligned}&\int_{-\infty}^{\infty} -e^{-\frac{x^2+z^2}{2}}\exp\left(-\gamma\left[\left(e^r-1\right)\left(\kappa-ct-w^*\right)+w^*\left(\hat{\theta}_0-1+x\sqrt{\frac{s_0^2 t}{s_0 t+1}}+z\sqrt{\frac{s_0}{s_0 t+1}}\right)\right]\right)dz\\&=-\sqrt{2\pi}e^{-\frac{x^2}{2}}\exp\left(-\gamma\left[\left(e^r-1\right)\left(\kappa-ct-w^*\right)+w^*\left(\hat{\theta}_0-1+x\sqrt{\frac{s_0^2 t}{s_0 t+1}}\right)\right]\right)\int_{-\infty}^{\infty}\frac{e^{-\frac{z^2+2\gamma w^* z\sqrt{\frac{s_0}{s_0 t+1}}}{2}}}{\sqrt{2\pi}}dz\\&=-\sqrt{2\pi}\exp\left(\gamma w^* e^r-\gamma(\kappa-ct)e^r+\frac{\gamma^2 s_0 w^{*2}}{2s_0 t+2}-\gamma w^* x\sqrt{\frac{s_0^2 t}{s_0 t+1}}-\gamma\hat{\theta}_0 w^*+\gamma(\kappa-ct)-\frac{x^2}{2}\right).\end{aligned} \tag{11}$$

To evaluate the integral with respect to x, we determine the domain where the optimal investment is positive. The optimal investment is given by

$$w^* = \max\left\{\frac{\hat{\theta}_t - e^r}{\gamma s_t}, 0\right\} = \max\left\{\frac{(s_0 t+1)\left(\hat{\theta}_0 - e^r + x\sqrt{\frac{s_0^2 t}{s_0 t+1}}\right)}{\gamma s_0}, 0\right\}.$$

For w^* to be positive, we need to have $\hat{\theta}_0 - e^r + x\sqrt{\frac{s_0^2 t}{s_0 t+1}} > 0$, which gives $x > \left(e^r - \hat{\theta}_0\right)\sqrt{\frac{s_0 t+1}{s_0^2 t}} :=$ $d(t)$. Hence, from (11) we get

$$\begin{aligned}
&\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} -\exp\left(\gamma w^* e^r - \gamma(\kappa - ct)e^r + \frac{\gamma^2 s_0 w^{*2}}{2s_0 t+2} - \gamma w^* x\sqrt{\frac{s_0^2 t}{s_0 t+1}} - \gamma\hat{\theta}_0 w^* + \gamma(\kappa - ct) - \frac{x^2}{2}\right) dx \\
=&\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{d(t)} -\exp\left(\gamma w^* e^r - \gamma(\kappa - ct)e^r + \frac{\gamma^2 s_0 w^{*2}}{2s_0 t+2} - \gamma w^* x\sqrt{\frac{s_0^2 t}{s_0 t+1}} - \gamma\hat{\theta}_0 w^* + \gamma(\kappa - ct) - \frac{x^2}{2}\right) dx \\
&+ \frac{1}{\sqrt{2\pi}}\int_{d(t)}^{\infty} -\exp\left(\gamma w^* e^r - \gamma(\kappa - ct)e^r + \frac{\gamma^2 s_0 w^{*2}}{2s_0 t+2} - \gamma w^* x\sqrt{\frac{s_0^2 t}{s_0 t+1}} - \gamma\hat{\theta}_0 w^* + \gamma(\kappa - ct) - \frac{x^2}{2}\right) dx.
\end{aligned}$$

The first part $\left(\int_{-\infty}^{d(t)}\right)$ is evaluate as follows:

$$\begin{aligned}
&\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{d(t)} -\exp\left(\gamma w^* e^r - \gamma(\kappa - ct)e^r + \frac{\gamma^2 s_0 w^{*2}}{2s_0 t+2} - \gamma w^* x\sqrt{\frac{s_0^2 t}{s_0 t+1}} - \gamma\hat{\theta}_0 w^* + \gamma(\kappa - ct) - \frac{x^2}{2}\right) dx \\
=&\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{d(t)} -e^{-\gamma(\kappa-ct)e^r + \gamma(\kappa-ct) - \frac{x^2}{2}}\, dx = -e^{\gamma(\kappa-ct)(1-e^r)}\Phi(d(t)).
\end{aligned}$$

The second part $\left(\int_{d(t)}^{\infty}\right)$ is evaluated by using $w^* = \frac{1}{\gamma s_0}\left(s_0 t+1\right)\left(\hat{\theta}_0 - e^r + x\sqrt{\frac{s_0^2 t}{s_0 t+1}}\right)$, and using two substitutions, $y = x\sqrt{1+s_0 t}$ and $Y = y + (\hat{\theta}_0 - e^r)\sqrt{t}$, and standard techniques of evaluating expectations of normal distributions. These give

$$\begin{aligned}
&\frac{1}{\sqrt{2\pi}}\int_{d(t)}^{\infty} -\exp\left(\gamma w^* e^r - \gamma(\kappa - ct)e^r + \frac{\gamma^2 s_0 w^{*2}}{2s_0 t+2} - \gamma w^* x\sqrt{\frac{s_0^2 t}{s_0 t+1}} - \gamma\hat{\theta}_0 w^* + \gamma(\kappa - ct) - \frac{x^2}{2}\right) dx \\
=& -\frac{\Phi\left(\frac{-d(t)}{\sqrt{s_0 t+1}}\right)\exp\left(-\frac{\hat{\theta}_0^2 + e^{2r} - 2\hat{\theta}_0 e^r + 2\gamma s_0(\kappa-ct)(e^r-1)}{2s_0}\right)}{\sqrt{s_0 t+1}}.
\end{aligned}$$

Hence, the expected utility is given by

$$u^*(t) = -e^{\gamma(\kappa-ct)(1-e^r)}\Phi(d(t)) - \frac{\Phi\left(\frac{-d(t)}{\sqrt{s_0 t+1}}\right)\exp\left(-\frac{\hat{\theta}_0^2 + e^{2r} - 2\hat{\theta}_0 e^r + 2\gamma s_0(\kappa-ct)(e^r-1)}{2s_0}\right)}{\sqrt{s_0 t+1}},$$

where $d(t) = \left(e^r - \hat{\theta}_0\right)\sqrt{\frac{s_0 t+1}{s_0^2 t}}$. □

C.3 Proof of Proposition 1

This proof has two parts. In the first part, we derive $u^*(0)$ directly from the definition. In the second part, we show that $u^*(t)$ is continuous at $t = 0$.

(i) Let $t = 0$, so that $\hat{\theta}_0 = \hat{\theta}_0$, $s_0 = s_0$, and $\theta = \hat{\theta}_0 + \sqrt{s_0}z$, where z is the outcome of the standard normal variable. This gives the utility in the next period as $-\exp\left(-\gamma\left[w^*\left(\hat{\theta}_0 - 1 + z\sqrt{s_0}\right) + (e^r - 1)(\kappa - w\right.\right.$ where $w^* = \max\left\{\frac{\hat{\theta}_0 - e^r}{\gamma s_0}, 0\right\}$. To calculate the expected utility, we integrate over the standard normal PDF:

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} -e^{-\frac{z^2}{2}} \exp\left(-\gamma\left[(e^r - 1)(\kappa - w^*) + w^*\left(\hat{\theta}_0 - 1 + z\sqrt{s_0}\right)\right]\right) dz$$

$$= -e^{-\gamma(w^*\hat{\theta}_0 + \kappa e^r - \kappa - w^* e^r) + \frac{\gamma^2 w^{*2} s_0}{2}} = \begin{cases} -\exp\left(-\frac{\hat{\theta}_0^2 + e^{2r} - 2\hat{\theta}_0 e^r + 2\gamma s_0 \kappa (e^r - 1)}{2s_0}\right) & \text{if } \hat{\theta}_0 > e^r \\ -e^{-\gamma\kappa(e^r - 1)} & \text{if } \hat{\theta}_0 \leq e^r \end{cases}$$

(ii) Let $t \downarrow 0$ in the expected utility $u^*(t)$. Observe that, by Lemma 3, in $u^*(t)$ we have:

$$\lim_{t\downarrow 0} d(t) = \lim_{t\downarrow 0}\left(e^r - \hat{\theta}_0\right)\sqrt{\frac{s_0 t + 1}{s_0^2 t}} = \begin{cases} -\infty & \text{if } \hat{\theta}_0 > e^r \\ 0 & \text{if } \hat{\theta}_0 = e^r \\ +\infty & \text{if } \hat{\theta}_0 < e^r \end{cases}$$

and

$$\lim_{t\downarrow 0} \frac{-d(t)}{\sqrt{s_0 t + 1}} = \lim_{t\downarrow 0}\left(e^r - \hat{\theta}_0\right)\frac{-1}{\sqrt{s_0^2 t}} = \begin{cases} +\infty & \text{if } \hat{\theta}_0 > e^r \\ 0 & \text{if } \hat{\theta}_0 = e^r \\ -\infty & \text{if } \hat{\theta}_0 < e^r \end{cases}$$

Then we have:

$$\lim_{t\downarrow 0} u^*(t) = \lim_{t\downarrow 0} -e^{\gamma(\kappa - ct)(1 - e^r)}\Phi(d(t)) - \lim_{t\downarrow 0} \frac{\Phi\left(\frac{-d(t)}{\sqrt{s_0 t + 1}}\right)\exp\left(-\frac{\hat{\theta}_0^2 + e^{2r} - 2\hat{\theta}_0 e^r + 2\gamma s_0(\kappa - ct)(e^r - 1)}{2s_0}\right)}{\sqrt{s_0 t + 1}}$$

$$= \begin{cases} -e^{\gamma\kappa(1 - e^r)}\Phi(-\infty) - \Phi(+\infty)\exp\left(-\frac{\hat{\theta}_0^2 + e^{2r} - 2\hat{\theta}_0 e^r + 2\gamma s_0 \kappa (e^r - 1)}{2s_0}\right) & \text{if } \hat{\theta}_0 > e^r \\ -e^{\gamma\kappa(1 - e^r)}\Phi(0) - \Phi(0)\exp\left(-\gamma\kappa\left(e^r - 1\right)\right) & \text{if } \hat{\theta}_0 = e^r \\ -e^{\gamma\kappa(1 - e^r)}\Phi(+\infty) - \Phi(-\infty)\exp\left(-\frac{\hat{\theta}_0^2 + e^{2r} - 2\hat{\theta}_0 e^r + 2\gamma s_0 \kappa (e^r - 1)}{2s_0}\right) & \text{if } \hat{\theta}_0 < e^r \end{cases}$$

$$= \begin{cases} -\exp\left(-\frac{\hat{\theta}_0^2 + e^{2r} - 2\hat{\theta}_0 e^r + 2\gamma s_0 \kappa (e^r - 1)}{2s_0}\right) & \text{if } \hat{\theta}_0 > e^r \\ -e^{-\gamma\kappa(e^r - 1)} & \text{if } \hat{\theta}_0 \leq e^r \end{cases}$$

Therefore, the value of data analytics $u^*(t) - u^*(0)$ is given by

$$\begin{cases} -e^{\gamma(\kappa-ct)(1-e^r)}\Phi(d(t)) + \left(1 - \frac{e^{\gamma ct(e^r-1)}}{\sqrt{s_0t+1}}\Phi\left(\frac{\hat{\theta}_0-e^r}{s_0\sqrt{t}}\right)\right)\exp\left(-\frac{\hat{\theta}_0^2+e^{2r}-2\hat{\theta}_0e^r+2\gamma s_0\kappa(e^r-1)}{2s_0}\right) & \text{if } \hat{\theta}_0 > e^r \\ e^{\gamma\kappa(1-e^r)}\left(1 - \frac{e^{\gamma ct(e^r-1)}}{2}\left(1 + \frac{1}{\sqrt{s_0t+1}}\right)\right) & \text{if } \hat{\theta}_0 = e^r \\ e^{\gamma\kappa(1-e^r)}\left(1 - e^{\gamma ct(e^r-1)}\Phi(d(t))\right) - \frac{\Phi\left(\frac{\hat{\theta}_0-e^r}{s_0\sqrt{t}}\right)\exp\left(-\frac{\hat{\theta}_0^2+e^{2r}-2\hat{\theta}_0e^r+2\gamma s_0(\kappa-ct)(e^r-1)}{2s_0}\right)}{\sqrt{s_0t+1}} & \text{if } \hat{\theta}_0 < e^r \end{cases} \quad \square$$

C.4 Proof of Proposition 2

Here we consider three cases: The no negative investment constraint is not binding, the marginal value v' is not continuous in $\hat{\theta}_0$ (by Corollary 1, there is discontinuity when $t = 0$ and $\hat{\theta}_0 = e^r$), and the no negative investment constraint is binding.

(*i*) $\hat{\theta}_0 > e^r$: The second derivative of the value function is as follows:

$$\begin{aligned} v'' = &\frac{e^{\gamma(e^r-1)(ct-\kappa)}}{8t^2\left(s_0t+1\right)^{7/2}}\Bigg(-8c^2t^2\gamma^2\left(e^r-1\right)^2\left(s_0t+1\right)^{7/2}\Phi\left(\left(e^r-\hat{\theta}_0\right)\sqrt{\frac{s_0t+1}{s_0^2t}}\right) \\ &- 2s_0^2t^2(s_0t+1)e^{-\frac{\left(e^r-\hat{\theta}_0\right)^2}{2s_0}}\Phi\left(\frac{\left(\hat{\theta}_0-e^r\right)}{s_0\sqrt{t}}\right)\left(\left[2c\gamma\left(e^r-1\right)\left(\frac{s_0t+1}{s_0}\right)-1\right]^2+2\right) \\ &+ \sqrt{\frac{2t}{\pi}}\left(e^r-\hat{\theta}_0\right)\left(s_0t+1\right)^2e^{-\frac{\left(e^r-\hat{\theta}_0\right)^2(s_0t+1)}{2s_0^2t}}\Bigg). \end{aligned}$$

Hence, $v'' < 0$ for all $t > 0$.

(*ii*) $\hat{\theta}_0 = e^r$: Take the second derivative of the value function as follows:

$$\begin{aligned} v'' = &\frac{e^{\gamma(e^r-1)(ct-\kappa)}}{8(s_0t+1)^{5/2}}\Bigg(-4c^2\gamma^2\left(e^r-1\right)^2\left(s_0t+1\right)^{5/2} \\ &- \left[2c\gamma\left(e^r-1\right)+s_0\left(2c\gamma\left(e^r-1\right)t-1\right)\right]^2-2s_0^2\Bigg). \end{aligned}$$

Again, $v'' < 0$ for all $t > 0$.

(*iii*) $\hat{\theta}_0 < e^r$: $v(0) = 0$ and $v'(0) < 0$. This means that there exists an interval $(0, \epsilon)$ such that for all $t_\epsilon \in (0, \epsilon)$, $v(t_\epsilon) < 0$. If $c < c_{CH}$, then $v(t) > 0$ in some region, and by continuity, there exists a t_0 such that $v(t_0) = 0$ and $t_0 > t_\epsilon$. By the Mean Value Theorem, there exists a $d \in (t_\epsilon, t_0)$ such that $v'(d) = \frac{v(t_0)-v(t_\epsilon)}{t_0-t_\epsilon} = \frac{-v(t_\epsilon)}{t_0-t_\epsilon} > 0$. Applying the Mean Value Theorem again, there exists an $e \in (0, d)$ such that $v''(e) = \frac{v'(d)-v'(0)}{d-0} = \frac{v'(d)-v'(0)}{d} > 0$. Hence, $v(t)$ is not globally concave if $c < c_{CH}$ and $\hat{\theta}_0 < e^r$. $\square$

C.5 Proof of Proposition 3

The value of data analytics $v(t;\hat{\theta}_0)$ is nonzero if $t>0$.

(*i*) $\hat{\theta}_0 > e^r$: We have:

$$
\begin{aligned}
\frac{\partial v}{\partial \hat{\theta}_0} = \quad & e^{\gamma(\kappa-ct)(1-e^r)}\phi(d(t))\sqrt{\frac{s_0t+1}{s_0^2t}} \\
& -\frac{e^{\gamma ct(e^r-1)}}{s_0\sqrt{t}\sqrt{s_0t+1}}\phi\left(\frac{\hat{\theta}_0-e^r}{s_0\sqrt{t}}\right)\exp\left(-\frac{(\hat{\theta}_0-e^r)^2}{2s_0}-\gamma\kappa(e^r-1)\right) \\
& +\frac{e^r-\hat{\theta}_0}{s_0}\left(1-\frac{e^{\gamma ct(e^r-1)}}{\sqrt{s_0t+1}}\Phi\left(\frac{\hat{\theta}_0-e^r}{s_0\sqrt{t}}\right)\right)\exp\left(-\frac{(\hat{\theta}_0-e^r)^2}{2s_0}-\gamma\kappa(e^r-1)\right).
\end{aligned}
\tag{12}
$$

Let $c <_{CH}$. Then for $v(t;\hat{\theta}_0)\geq 0$, we need $1-\frac{e^{\gamma ct(e^r-1)}}{\sqrt{s_0t+1}}\Phi\left(\frac{\hat{\theta}_0-e^r}{s_0\sqrt{t}}\right)>0$, and $s_0>2c\gamma(e^r-1)$, otherwise $v(t;\hat{\theta}_0)$ is negative by concavity (see Corollary 1 and Proposition 2). Let us consider the first and third terms in (12):

$$
\begin{aligned}
& e^{\gamma(\kappa-ct)(1-e^r)}\phi(d(t))\sqrt{\frac{s_0t+1}{s_0^2t}}+\frac{e^r-\hat{\theta}_0}{s_0}\left(1-\frac{e^{\gamma ct(e^r-1)}}{\sqrt{s_0t+1}}\Phi\left(\frac{\hat{\theta}_0-e^r}{s_0\sqrt{t}}\right)\right)e^{-\frac{(\hat{\theta}_0-e^r)^2}{2s_0}-\gamma\kappa(e^r-1)} \\
=& e^{-\frac{(\hat{\theta}_0-e^r)^2}{2s_0}-\gamma\kappa(e^r-1)}\left(\frac{\sqrt{s_0t+1}}{s_0\sqrt{2\pi t}}e^{-\frac{(\hat{\theta}_0-e^r)^2}{2s_0^2t}+\gamma ct(e^r-1)}+\frac{e^r-\hat{\theta}_0}{s_0}\left[1-\frac{e^{\gamma ct(e^r-1)}}{\sqrt{s_0t+1}}\Phi\left(\frac{\hat{\theta}_0-e^r}{s_0\sqrt{t}}\right)\right]\right) \\
=& e^{-\frac{(\hat{\theta}_0-e^r)^2}{2s_0}-\gamma\kappa(e^r-1)}G(\hat{\theta}_0),
\end{aligned}
$$

which implicitly defines term $G(\hat{\theta}_0)$. Then the sign of the sum of the first and third terms in (12) is the sign of $G(\hat{\theta}_0)$, $G(e^r)>0$, $\lim_{\hat{\theta}_0\to\infty}G(\hat{\theta}_0)<0$, and $G(\hat{\theta}_0)$ is continuous for $\hat{\theta}_0>e^r$. By the Intermediate Value Theorem, $G(\hat{\theta}_0)=0$ at least once. Let $\hat{\theta}_0^i$ be such that $G(\hat{\theta}_0^i)=0$, where $i\in\{1,2,\dots\}$. Taking derivative with respect to $\hat{\theta}_0$, we have:

$$
\begin{aligned}
\frac{\partial G(\hat{\theta}_0)}{\partial\hat{\theta}_0} =& \left(\frac{e^r-\hat{\theta}_0}{s_0^2t}\right)\frac{\sqrt{s_0t+1}}{s_0\sqrt{2\pi t}}e^{-\frac{(\hat{\theta}_0-e^r)^2}{2s_0^2t}+\gamma ct(e^r-1)}-\frac{1}{s_0}\left(1-\frac{e^{\gamma ct(e^r-1)}}{\sqrt{s_0t+1}}\Phi\left(\frac{\hat{\theta}_0-e^r}{s_0\sqrt{t}}\right)\right) \\
& -\frac{e^r-\hat{\theta}_0}{s_0^2\sqrt{t}}\frac{e^{\gamma ct(e^r-1)}}{\sqrt{s_0t+1}}\phi\left(\frac{\hat{\theta}_0-e^r}{s_0\sqrt{t}}\right) \\
\frac{\partial G(\hat{\theta}_0^i)}{\partial\hat{\theta}_0} =& \left(\frac{e^r-\hat{\theta}_0^i}{s_0^2t}\right)\frac{\sqrt{s_0t+1}}{s_0\sqrt{2\pi t}}e^{-\frac{(\hat{\theta}_0^i-e^r)^2}{2s_0^2t}+\gamma ct(e^r-1)}-\frac{1}{s_0}\left(1-\frac{e^{\gamma ct(e^r-1)}}{\sqrt{s_0t+1}}\Phi\left(\frac{\hat{\theta}_0^i-e^r}{s_0\sqrt{t}}\right)\right) \\
& -\frac{e^r-\hat{\theta}_0^i}{s_0^2\sqrt{t}}\frac{e^{\gamma ct(e^r-1)}}{\sqrt{s_0t+1}}\phi\left(\frac{\hat{\theta}_0^i-e^r}{s_0\sqrt{t}}\right) \\
=& \frac{1}{\sqrt{2\pi}}e^{-\frac{(\hat{\theta}_0^i-e^r)^2}{2s_0^2t}+\gamma ct(e^r-1)}\left(\frac{e^r-\hat{\theta}_0^i}{s_0^2\sqrt{t}}\right)\left(\frac{1}{s_0t\sqrt{s_0t+1}}\right)-\frac{1}{s_0}\left(1-\frac{e^{\gamma ct(e^r-1)}}{\sqrt{s_0t+1}}\Phi\left(\frac{\hat{\theta}_0^i-e^r}{s_0\sqrt{t}}\right)\right).
\end{aligned}
$$

Hence, $\frac{\partial G(\hat{\theta}_0^i)}{\partial\hat{\theta}_0}<0$. Since $G(e^r)>0$ and $\lim_{\hat{\theta}_0\to\infty}G(\hat{\theta}_0)=-\infty$, then there is only one $\hat{\theta}_0^i$ with $G(\hat{\theta}_0^i)=0$, and let $\varsigma=\hat{\theta}_0^1$. Then when $\hat{\theta}_0>\varsigma$, $G(\hat{\theta}_0)<0$ and, therefore, $\frac{\partial v}{\partial\hat{\theta}_0}<0$ in (12).

(*ii*) $\hat{\theta}_0 < e^r$: We have:

$\frac{\partial v}{\partial \hat{\theta}_0} = \frac{\sqrt{t}}{\sqrt{1+s_0 t}} \exp\left(-\frac{(\hat{\theta}_0 - e^r)^2}{2s_0} - \gamma(\kappa - ct)(e^r - 1)\right) \left[\phi\left(\frac{e^r - \hat{\theta}_0}{s_0\sqrt{t}}\right) - \frac{e^r - \hat{\theta}_0}{s_0\sqrt{t}} \Phi\left(\frac{\hat{\theta}_0 - e^r}{s_0\sqrt{t}}\right)\right]$. The sign of $\frac{\partial v}{\partial \hat{\theta}_0}$ depends on the term in the square brackets. Let $X = \frac{\hat{\theta}_0 - e^r}{s_0\sqrt{t}}$ and $F(X) = \phi(X) + X\Phi(X)$. F is smooth, $F(0) = 0$, and $\frac{dF}{dX} = \Phi(X) > 0$. So, $F(X)$ is increasing in X for all $X \geq 0$, which implies $\frac{\partial v}{\partial \hat{\theta}_0} > 0$.

(*iii*) It is easy to see that $\lim_{\hat{\theta}_0 \to \infty} v'(t; \hat{\theta}_0) = 0$ for all $t \geq 0$. Since $v(0; \hat{\theta}_0) = 0$, we have:

$$\lim_{\hat{\theta}_0 \to \infty} v(t; \hat{\theta}_0) = \lim_{\hat{\theta}_0 \to \infty} v(0; \hat{\theta}_0) + \lim_{\hat{\theta}_0 \to \infty} \int_0^t v'(y; \hat{\theta}_0) dy = 0. \quad \square$$

C.6 Proof of Theorem 2

(*i*) (3) and (4) are independent of κ. Hence t^* is independent of κ.

(*ii*) Let $c < c_{CH}$ so that $t^* > 0$.

(*a*) $\hat{\theta}_0 > e^r$: Let $f(t)$ be the left hand side of (3) and $t^* \in (0, \infty)$ such that $f(t^*) = 0$. We have:

$$\begin{aligned}
\frac{\partial f}{\partial t^*} =& \frac{\left(e^r - \hat{\theta}_0\right) \phi\left(\left(e^r - \hat{\theta}_0\right)\sqrt{\frac{(s_0 t + 1)}{s_0^2 t}}\right)}{2t^{3/2}} - 2c\gamma\left(e^r - 1\right) s_0 e^{-\frac{\left(e^r - \hat{\theta}_0\right)^2}{2s_0}} \Phi\left(\frac{\hat{\theta}_0 - e^r}{s_0\sqrt{t}}\right) \\
&- 3c\gamma\left(e^r - 1\right) s_0 \sqrt{s_0 t + 1} \Phi\left(\left(e^r - \hat{\theta}_0\right)\sqrt{\frac{s_0 t + 1}{s_0^2 t}}\right) \\
\frac{\partial f}{\partial c} =& - 2\gamma\left(e^r - 1\right)\left(1 + s_0 t\right)\left[\left(s_0 t + 1\right)^{1/2} \Phi\left(\left(e^r - \hat{\theta}_0\right)\sqrt{\frac{s_0 t + 1}{s_0^2 t}}\right)\right. \\
&\left. + e^{-\frac{\left(e^r - \hat{\theta}_0\right)^2}{2s_0}} \Phi\left(\frac{\hat{\theta}_0 - e^r}{s_0\sqrt{t}}\right)\right] \\
\frac{\partial f}{\partial \gamma} =& - 2c\left(e^r - 1\right)\left(1 + s_0 t\right)\left[\left(s_0 t + 1\right)^{1/2} \Phi\left(\left(e^r - \hat{\theta}_0\right)\sqrt{\frac{s_0 t + 1}{s_0^2 t}}\right)\right. \\
&\left. + e^{-\frac{\left(e^r - \hat{\theta}_0\right)^2}{2s_0}} \Phi\left(\frac{\hat{\theta}_0 - e^r}{s_0\sqrt{t}}\right)\right].
\end{aligned}$$

(*b*) $\hat{\theta}_0 = e^r$: Let $f(t)$ be the left hand side of (4) and $t^* \in (0, \infty)$ such that $f(t^*) = 0$. Then we have:

$$\begin{aligned}
\frac{\partial f}{\partial t^*} =& -c\gamma\left(e^r - 1\right)s_0 - \frac{3}{2}c\gamma\left(e^r - 1\right)s_0\sqrt{s_0 t + 1}\\
\frac{\partial f}{\partial c} =& -\gamma\left(e^r - 1\right)\left(s_0 t + 1\right)^{3/2} - \gamma\left(e^r - 1\right)\left(s_0 t + 1\right)\\
\frac{\partial f}{\partial \gamma} =& -c\left(e^r - 1\right)\left(s_0 t + 1\right)^{3/2} - c\left(e^r - 1\right)\left(s_0 t + 1\right).
\end{aligned}$$

Hence, in both cases (*a*) and (*b*), $\frac{\partial f}{\partial c} < 0$, $\frac{\partial f}{\partial \gamma} < 0$, $\frac{\partial f}{\partial t^*} < 0$, and $v(t)$ is strictly concave, thus admits an unique maximum. By the Implicit Function Theorem, we have $\frac{\partial t^*}{\partial c} = -\frac{\frac{\partial f}{\partial c}}{\frac{\partial f}{\partial t^*}} < 0$ and $\frac{\partial t^*}{\partial \gamma} = -\frac{\frac{\partial f}{\partial \gamma}}{\frac{\partial f}{\partial t^*}} < 0$. □

C.7 Proof of Proposition 4

We consider two cases:

(*i*) $c \le c_{CH}$: From (2) we get

$$\begin{aligned}
E\left[w^*(t^*)\right] =& E\left[\max\left\{\frac{\hat{\theta}_{t^*} - e^r}{\gamma s_{t^*}}, 0\right\}\right] = E\left[\max\left\{\frac{(\hat{\theta}_0 + \epsilon_t\sqrt{\frac{s_0^2 t^*}{s_0 t^* + 1}} - e^r)(1 + s_0 t^*)}{\gamma s_0}, 0\right\}\right]\\
=& \int_{-\infty}^{\infty} \max\left\{\frac{(\hat{\theta}_0 + x\sqrt{\frac{s_0^2 t^*}{s_0 t^* + 1}} - e^r)(1 + s_0 t^*)}{\gamma s_0}, 0\right\}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\, dx.
\end{aligned}$$

Condition $w^*(t^*) > 0$ corresponds to $\hat{\theta}_0 - e^r + x\sqrt{\frac{s_0^2 t^*}{s_0 t^* + 1}} > 0$, which can be written as $x > \left(e^r - \hat{\theta}_0\right)\sqrt{\frac{s_0 t^* + 1}{s_0^2 t^*}} := d^*$. Hence,

$$\begin{aligned}
&\int_{-\infty}^{\infty} \max\left\{\frac{(\hat{\theta}_0 + x\sqrt{\frac{s_0^2 t^*}{s_0 t^* + 1}} - e^r)(1 + s_0 t^*)}{\gamma s_0}, 0\right\}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\, dx\\
=& \int_{d^*}^{\infty} \frac{(\hat{\theta}_0 + x\sqrt{\frac{s_0^2 t^*}{s_0 t^* + 1}} - e^r)(1 + s_0 t^*)}{\gamma s_0}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\, dx\\
=& \frac{(\hat{\theta}_0 - e^r)(1 + t^* s_0)}{\gamma s_0}\Phi(-d^*) + \frac{\sqrt{t^*}\sqrt{1 + s_0 t^*}}{\gamma}\int_{d^*}^{\infty}\frac{x}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\, dx\\
=& \frac{(\hat{\theta}_0 - e^r)(1 + t^* s_0)}{\gamma s_0}\Phi(-d^*) + \frac{\sqrt{t^*}\sqrt{1 + s_0 t^*}}{\gamma}\left[-\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\right]_{d^*}^{\infty}\\
=& \frac{(\hat{\theta}_0 - e^r)(1 + t^* s_0)}{\gamma s_0}\Phi(-d^*) + \frac{\sqrt{t^*}\sqrt{1 + s_0 t^*}}{\gamma}\phi(d^*).
\end{aligned}$$

(ii) $c > c_{CH}$: $t^* = 0$, hence $w^* = \max\left\{\frac{\hat{\theta}_0 - e^r}{\gamma s_0}, 0\right\} = \frac{\hat{\theta}_0 - e^r}{\gamma s_0}$ if $\hat{\theta}_0 > e^r$, and otherwise $w^* = 0$.

To find the expected leverage, we consider the values of x such that

$$\frac{1}{\gamma s_0}\left(\hat{\theta}_0 + x\sqrt{\frac{s_0^2 t^*}{s_0 t^* + 1}} - e^r\right)(1 + s_0 t^*) > \kappa - ct^*,$$

which gives $x > \left(\frac{(\kappa - ct^*)\gamma s_0}{1 + s_0 t^*} + e^r - \hat{\theta}_0\right)\sqrt{\frac{s_0 t^* + 1}{s_0^2 t^*}} := b^*$. Hence, the expected leverage:

$$\begin{aligned}
&E[\max\{w^*(t^*) + ct^* - \kappa, 0\}] \\
=&E\left[\max\left\{\max\left\{\frac{(\hat{\theta}_0 + \epsilon_2\sqrt{\frac{s_0^2 t^*}{s_0 t^* + 1}} - e^r)(1 + s_0 t^*)}{\gamma s_0}, 0\right\} + ct^* - \kappa, 0\right\}\right] \\
=&\int_{b^*}^{\infty}\left(\frac{(\hat{\theta}_0 + x\sqrt{\frac{s_0^2 t^*}{s_0 t^* + 1}} - e^r)(1 + s_0 t^*)}{\gamma s_0} + ct^* - \kappa\right)\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\,dx \\
=&\left(\frac{(\hat{\theta}_0 - e^r)(1 + t^* s_0)}{s_0\gamma} + ct^* - \kappa\right)\Phi\left(-b^*\right) + \frac{\sqrt{t^* + s_0(t^*)^2}}{\gamma}\phi\left(b^*\right). \quad \square
\end{aligned}$$

C.8 Proof of Theorem 3

(i) S_1: The expected profit and loss is given by

$$E[w^*(t^*)](S_1 - 1) + (\kappa - ct^* - E[w^*(t^*)])(e^r - 1).$$

By taking derivative with respect to S_1, we get $\frac{\partial E[Profit]}{\partial S_1} = E[w^*(t^*)]$. This is the slope of the profit and loss line in Figure 12a. Note that the slope is positive. By Proposition 4, $\frac{\partial E[w^*(t^*)]}{\partial c}$ is given by

$$\frac{\hat{\theta}_0 - e^r}{\gamma}\Phi\left(\frac{(\hat{\theta}_0 - e^r)\sqrt{\frac{1}{t^*} + s_0}}{s_0}\right)\frac{\partial t^*}{\partial c} + \phi\left(\frac{(e^r - \hat{\theta}_0)\sqrt{\frac{1}{t^*} + s_0}}{s_0}\right)\frac{1 + 2s_0 t^*}{2\gamma\sqrt{t^* + s_0(t^*)^2}}\frac{\partial t^*}{\partial c}$$

for $c < c_{CH}$. Since $\frac{\partial E[w^*(t^*)]}{\partial c}$ is negative for $\hat{\theta}_0 > e^r$, the slope of the expected profit and loss line increases with falling cost of data analytics. So, the expected profit and loss lines with different costs are not parallel. Let $\bar{S}_1$ be the shock value, where the two lines corresponding to high cost c_h and low cost c_l intersect. $\bar{S}_1$ exists and is unique because non-parallel lines intersect at one point. Hence, decreasing cost from c_h to c_l results in an anti-clockwise rotation of the expected profit and loss line, and for all $S_1 < \bar{S}_1$, the lower cost of analytics results in higher losses.

(ii) S_2: The expected profit and loss is given by

$$E[w^*(t^*)](\hat{\theta}_0 - S_2\sqrt{s_{t^*}} - 1) + (\kappa - ct^* - E[w^*(t^*)])(e^r - 1).$$

Taking derivative with respect to S_2, we have

$$\frac{\partial E[Profit]}{\partial S_2} = -E[w^*(t^*)]\sqrt{s_{t^*}} = -E[w^*(t^*)]\frac{\sqrt{s_0}}{\sqrt{1+s_0 t^*}}.$$

This is again the slope of the profit and loss line in Figure 12b. Note that the slope is negative. Taking derivative with respect to c we get

$$-\frac{(\hat{\theta}_0 - e^r)\sqrt{s_0}}{2\gamma\sqrt{1+s_0 t^*}}\Phi\left(\frac{(\hat{\theta}_0 - e^r)\sqrt{\frac{1}{t^*}+s_0}}{s_0}\right)\frac{\partial t^*}{\partial c} - \phi\left(\frac{(e^r - \hat{\theta}_0)\sqrt{\frac{1}{t^*}+s_0}}{s_0}\right)\frac{\sqrt{s_0}}{2\sqrt{t^*}\gamma}\frac{\partial t^*}{\partial c} \tag{13}$$

for $c < c_{CH}$. Since (13) is positive for $\hat{\theta}_0 > e^r$, the slope of the expected profit and loss line decreases with falling cost of data analytics. So, the expected profit and loss lines with different costs are not parallel. Let $\bar{S}_2$ be the shock value where the two lines corresponding to high cost c_h and low cost c_l intersect. $\bar{S}_2$ exists and is unique because non-parallel lines intersect at one point. Hence, decreasing cost from c_h to c_l results in a clockwise rotation of the expected profit and loss line, and for all $S_2 > \bar{S}_2$, lower cost of analytics results in high losses. □

C.9 Proof of Lemma 7

The proof is similar to that of Lemma 3. The expected utility without leverage is

$$u^*_{NL}(t) = E\left[-\exp(-\gamma(w^*_{NL}(\theta-1) + (\kappa - ct - w^*_{NL})(e^r-1)))\right] =$$
$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} -e^{-\frac{x^2+z^2}{2}}\exp\left(-\gamma\left[(e^r-1)(\kappa - ct - w^*_{NL}) + w^*_{NL}\left(\hat{\theta}_0 - 1 + x\sqrt{\frac{s_0^2 t}{s_0 t+1}} + z\sqrt{\frac{s_0}{s_0 t+1}}\right)\right]\right) dz dx. \tag{14}$$

We first evaluate the integral with respect to z, which gives the same form as in (11) in the proof of Lemma 3:

$$
\begin{aligned}
&\int_{-\infty}^{\infty} -e^{-\frac{x^2+z^2}{2}} \exp\left(-\gamma\left[(e^r-1)(\kappa-ct-w^*_{NL}) + w^*_{NL}\left(\hat{\theta}_0 - 1 + x\sqrt{\frac{s_0^2 t}{s_0 t+1}} + z\sqrt{\frac{s_0}{s_0 t+1}}\right)\right]\right) dz \\
=& -\sqrt{2\pi}e^{-\frac{x^2}{2}} \exp\left(-\gamma\left[(e^r-1)(\kappa-ct-w^*_{NL}) + w^*_{NL}\left(\hat{\theta}_0 - 1 + x\sqrt{\frac{s_0^2 t}{s_0 t+1}}\right)\right]\right) \int_{-\infty}^{\infty} \frac{e^{-\frac{z^2+2\gamma w^*_{NL} z\sqrt{\frac{s_0}{s_0t+1}}}{2}}}{\sqrt{2\pi}} dz \\
=& -\sqrt{2\pi}\exp\left(\gamma w^*_{NL} e^r - \gamma(\kappa-ct)e^r + \frac{\gamma^2 s_0 {w^*_{NL}}^2}{2s_0 t+2} - \gamma w^*_{NL} x\sqrt{\frac{s_0^2 t}{s_0 t+1}} - \gamma\hat{\theta}_0 w^*_{NL} + \gamma(\kappa-ct) - \frac{x^2}{2}\right). \qquad (15)
\end{aligned}
$$

To evaluate the integral with respect to x, we determine the domain where the optimal investment without leverage is positive, as well as where the leverage constraint is binding. The optimal investment without leverage is given by

$$
\begin{aligned}
w^*_{NL} &= \min\left\{\max\left\{\frac{\hat{\theta}_t - e^r}{\gamma s_t}, 0\right\}, \kappa - ct\right\} \\
&= \min\left\{\max\left\{\frac{(s_0 t+1)\left(\hat{\theta}_0 - e^r + x\sqrt{\frac{s_0^2 t}{s_0 t+1}}\right)}{\gamma s_0}, 0\right\}, \kappa - ct\right\}.
\end{aligned}
$$

Define $d(t) := \frac{\sqrt{s_0t+1}(e^r-\hat{\theta}_0)}{s_0\sqrt{t}}$ and $D(t) := d(t) + \frac{(\kappa-ct)\gamma}{\sqrt{t}\sqrt{s_0t+1}}$. If $x < d(t)$ then $w^*_{NL} = 0$, and if $x > D(t)$ then $w^*_{NL} = \kappa - ct$. Hence, dividing (15) by 2π (because earlier we moved $1/2\pi$ out of the integrals in (10) and (14)) and integrating over the domain give

$$
\begin{aligned}
&\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} -\exp\left(\gamma w^*_{NL} e^r - \gamma(\kappa-ct)e^r + \frac{\gamma^2 s_0 {w^*_{NL}}^2}{2s_0t+2} - \gamma w^*_{NL} x\sqrt{\frac{s_0^2 t}{s_0 t+1}} - \gamma\hat{\theta}_0 w^*_{NL} + \gamma(\kappa-ct) - \frac{x^2}{2}\right) dx \\
=&\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{d} -\exp\left(-\gamma(\kappa-ct)e^r + \gamma(\kappa-ct) - \frac{x^2}{2}\right) dx \\
&+ \frac{1}{\sqrt{2\pi}}\int_{d}^{D} -\exp\left(\gamma w^*_{NL} e^r - \gamma(\kappa-ct)e^r + \frac{\gamma^2 s_0 {w^*_{NL}}^2}{2s_0t+2} - \gamma w^*_{NL} x\sqrt{\frac{s_0^2 t}{s_0 t+1}} - \gamma\hat{\theta}_0 w^*_{NL} + \gamma(\kappa-ct) - \frac{x^2}{2}\right) dx \\
&+ \frac{1}{\sqrt{2\pi}}\int_{D}^{\infty} -\exp\left(\frac{\gamma^2 s_0(\kappa-ct)^2}{2s_0t+2} - \gamma(\kappa-ct)x\sqrt{\frac{s_0^2 t}{s_0 t+1}} + \gamma(\kappa-ct)(1-\hat{\theta}_0) - \frac{x^2}{2}\right) dx. \qquad (16)
\end{aligned}
$$

Simplifying the integrals by using the normal CDF gives the result. □

C.10 Proof of Corollary 5

(i) From the derivation of the expected utility without leverage in Lemma 7, we know that the integration the expected utilities with and without leverage are the same for the first iterated integral given by (11), except with w^* replaced by w^*_{NL}. If $t = 0$, then the expected utilities are given by this first iterated integral. Further, in this case if the leverage constraint is not binding, then $w^*_{NL} = w^*$ and the expected utilities are equal.

(ii) Continuing from (i), if $t > 0$ then integrating with respect to x in the region $\int_{-\infty}^{D(t)} \cdot dx$ for (11) and (15) gives the same result and we only need to compare the region $\int_{D(t)}^{\infty} \cdot dx$, where $D(t) = \frac{\sqrt{s_0 t+1}(e^r - \hat{\theta}_0)}{s_0\sqrt{t}} + \frac{(\kappa - ct)\gamma}{\sqrt{t}\sqrt{s_0 t+1}}$. In this region, $w^* \geq \kappa - ct$ and $w^*_{NL} = \kappa - ct$ due to the financial constraint. Note that $w^* = \frac{1}{\gamma s_0}(s_0 t+1)\left(\hat{\theta}_0 - e^r + x\sqrt{\frac{s_0^2 t}{s_0 t+1}}\right)$ is increasing with respect to x and
$w^*_{NL} = \min\left\{\frac{1}{\gamma s_0}(s_0 t+1)\left(\hat{\theta}_0 - e^r + x\sqrt{\frac{s_0^2 t}{s_0 t+1}}\right), \kappa - ct\right\}$. The integrand is

$$-\frac{1}{\sqrt{2\pi}}\exp\left(\gamma w e^r - \gamma(\kappa - ct)e^r + \frac{\gamma^2 s_0 w^2}{2s_0 t+2} - \gamma w x\sqrt{\frac{s_0^2 t}{s_0 t+1}} - \gamma\hat{\theta}_0 w + \gamma(\kappa - ct) - \frac{x^2}{2}\right).$$

and consider the terms involving w in the exponent:

$$\gamma w e^r + \frac{\gamma^2 s_0 w^2}{2s_0 t+2} - \gamma w x\sqrt{\frac{s_0^2 t}{s_0 t+1}} - \gamma\hat{\theta}_0 w = \frac{\gamma^2 s_0 w^2}{2s_0 t+2} - \gamma w\left(\hat{\theta}_0 - e^r + x\sqrt{\frac{s_0^2 t}{s_0 t+1}}\right)$$
$$= \frac{s_0\gamma^2}{2(s_0 t+1)}\left[\left(w - \frac{(s_0 t+1)\left(\hat{\theta}_0 - e^r + x\sqrt{\frac{s_0^2 t}{s_0 t+1}}\right)}{\gamma s_0}\right)^2 - \left(\frac{(s_0 t+1)\left(\hat{\theta}_0 - e^r + x\sqrt{\frac{s_0^2 t}{s_0 t+1}}\right)}{\gamma s_0}\right)^2\right].$$

which is always larger for $w = w^*_{NL} = \kappa - ct$ than for $w = w^* =$
$\frac{1}{\gamma s_0}(s_0 t+1)\left(\hat{\theta}_0 - e^r + x\sqrt{\frac{s_0^2 t}{s_0 t+1}}\right)$. Hence,

$$\int_D^{\infty} -\exp\left(\gamma w^*_{NL} e^r - \gamma(\kappa - ct)e^r + \frac{\gamma^2 s_0 {w^*_{NL}}^2}{2s_0 t+2} - \gamma w^*_{NL} x\sqrt{\frac{s_0^2 t}{s_0 t+1}} - \gamma\hat{\theta}_0 w^*_{NL} + \gamma(\kappa - ct) - \frac{x^2}{2}\right) dx$$
$$< \int_D^{\infty} -\exp\left(\gamma w^* e^r - \gamma(\kappa - ct)e^r + \frac{\gamma^2 s_0 {w^*}^2}{2s_0 t+2} - \gamma w^* x\sqrt{\frac{s_0^2 t}{s_0 t+1}} - \gamma\hat{\theta}_0 w^* + \gamma(\kappa - ct) - \frac{x^2}{2}\right) dx,$$

and the expected utility without leverage is strictly lower than the expected utility with leverage. □

C.11 Proof of Corollary 6

Since $v(t) = u^*(t) - u^*(0)$ and $v_{NL}(t) = u^*_{NL}(t) - u^*_{NL}(0)$, we only need to consider $u^*(t)$ and $u^*_{NL}(t)$ when differentiating $v(t)$ with respect to t. Further, comparing $u^*(t)$ with $u^*_{NL}(t)$ and removing terms that are equal, for $u^*_{NL}(t)$ we need to focus only on the derivatives of

$$-\frac{\exp\left(-\frac{\left(\hat{\theta}_0-e^r\right)^2+2\gamma(\kappa-ct)s_0(e^r-1)}{2s_0}\right)\Phi\left(\frac{e^r-\hat{\theta}_0}{s_0\sqrt{t}}+\frac{\gamma(\kappa-ct)}{\sqrt{t}}\right)}{\sqrt{s_0t+1}}$$

and

$$-e^{\frac{\gamma^2(\kappa-ct)^2s_0}{2}+\gamma(\kappa-ct)\left(1-\hat{\theta}_0\right)}\Phi\left(\frac{\sqrt{s_0t+1}\left(\hat{\theta}_0-e^r\right)}{s_0\sqrt{t}}-\frac{\gamma(\kappa-ct)\sqrt{s_0t+1}}{\sqrt{t}}\right);$$

and for $u^*(t)$ the derivative of

$$-\frac{\exp\left(-\frac{\left(\hat{\theta}_0-e^r\right)^2+2\gamma(\kappa-ct)s_0(e^r-1)}{2s_0}\right)}{\sqrt{s_0t+1}}.$$

Let us consider the term for $u^*(t)$ minus the two terms for $u^*_{NL}(t)$, and take derivative with respect to t. This gives

$$\begin{aligned}&\frac{\left(s_0-2c\gamma\left(e^r-1\right)\left(1+s_0t\right)\right)\left(1-\Phi\left(\frac{e^r-\hat{\theta}_0+\gamma s_0(\kappa-ct)}{s_0\sqrt{t}}\right)\right)\exp\left(\gamma\left(e^r-1\right)\left(ct-\kappa\right)-\frac{\left(e^r-\hat{\theta}_0\right)^2}{2s_0}\right)}{2(s_0t+1)^{3/2}}\\&+\frac{c\gamma s_0\sqrt{t}\exp\left(-\frac{\left(\hat{\theta}_0-e^r\right)^2+2\gamma(\kappa-ct)s_0(e^r-1)}{2s_0}\right)}{\sqrt{s_0t+1}}\phi\left(\frac{e^r-\hat{\theta}_0+s_0\gamma(\kappa-ct)}{s_0\sqrt{t}}\right)\\&+c\gamma(\hat{\theta}_0-1-\gamma s_0(\kappa-ct))e^{\frac{\gamma^2(\kappa-ct)^2s_0}{2}+\gamma(\kappa-ct)\left(1-\hat{\theta}_0\right)}\Phi\left(\frac{\sqrt{s_0t+1}\left(\hat{\theta}_0-e^r-\gamma s_0(\kappa-ct)\right)}{s_0\sqrt{t}}\right).\end{aligned}\tag{17}$$

If $\hat{\theta}_0>\Theta$ then the term $(\hat{\theta}_0-1-\gamma s_0(\kappa-ct))$ in (17) is postive. If $c<c_{CH}$, then $s_0>2c\gamma(e^r-1)$ and the marginal value of data analytics with leverage in Lemma 4 and the term $s_0-2c\gamma(e^r-1)(s_0t+1)$ in (17) are positive for all $t\in\left(0,\frac{s_0-2c\gamma(e^r-1)}{2s_0c\gamma(e^r-1)}\right)$. Hence, (17) is positive. □

C.12 Proof of Proposition 7

The proof is similar to Proposition 4, with one addition region in the integration. That is, to find the expectation, the integral $\int_{-\infty}^{\infty}$ is divided to $\int_{-\infty}^{d^*}+\int_{d^*}^{D^*}+\int_{D^*}^{\infty}$, where $d^*:=\frac{(e^r-\hat{\theta}_0)\sqrt{\frac{1}{t^*}+s_0}}{s_0}$ and $D^*:=d^*+\frac{(\kappa-ct)\gamma}{\sqrt{t^*}\sqrt{s_0t^*+1}}$. The result is obtained by direct integration. □

References

Afèche P, Ata B (2013) Bayesian dynamic pricing in queueing systems with unknown delay cost characteristics. *Manufacturing & Service Operations Management* 15(2):292–304.

Chade H, Schlee E (2002) Another look at the radner–stiglitz nonconcavity in the value of information. *Journal of Economic Theory* 107(2):421–452.

Cheung WC, Simchi-Levi D, Wang H (2017) Dynamic pricing and demand learning with limited price experimentation. *Operations Research* 65(6):1722–1731.

Chick S, Forster M, Pertile P (2017) A bayesian decision theoretic model of sequential experimentation with delayed response. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 79(5):1439–1462.

Chick SE, Hasija S, Nasiry J (2016) Information elicitation and influenza vaccine production. *Operations Research* 65(1):75–96.

Dugast J, Foucault T (2017) Data abundance and asset price informativeness. *https://ssrn.com/abstract=2398904* .

Goldfarb A, Tucker C (2017) Digital economics. Technical report, National Bureau of Economic Research.

Grossman SJ, Stiglitz JE (1980) On the impossibility of informationally efficient markets. *The American Economic Review* 70(3):393–408.

Guéant O, Pu J (2017) Portfolio choice, portfolio liquidation, and portfolio transition under drift uncertainty. *arXiv preprint arXiv:1611.07843* .

Hart O, Zingales L (2011) A new capital regulation for large financial institutions. *American Law and Economics Review* 13(2):453–490.

Karatzas I, Zhao X (2001) *Bayesian Adaptive Portfolio Optimization*, 632669 (Cambridge University Press), URL http://dx.doi.org/10.1017/CBO9780511569708.018.

Keppo J, Moscarini G, Smith L (2008) The demand for information: More heat than light. *Journal of Economic Theory* 138(1):21–50.

Kihlstrom R (1974) A bayesian model of demand for information about product quality. *International Economic Review* 99–118.

Lakner P (1998) Optimal trading strategy for an investor: the case of partial information. *Stochastic Processes and their Applications* 76(1):77–97.

Maksimovic V, Stomper A, Zechner J (1999) Capital structure, information acquisition and investment decisions in an industry framework. *Review of Finance* 2(3):251–271.

McCallum J (2017) Memory prices (1957-2017). `www.jcmit.net/memoryprice.htm`, accessed: 2017-05-03.

Merton RC (1969) Lifetime portfolio selection under uncertainty: The continuous-time case. *The Review of Economics and Statistics* 247–257.

Moscarini G, Smith L (2001) The optimal level of experimentation. *Econometrica* 69(6):1629–1644.

Moscarini G, Smith L (2002) The law of large demand for information. *Econometrica* 70(6):2351–2366.

Negoescu DM, Bimpikis K, Brandeau ML, Iancu DA (2017) Dynamic learning of patient response types: An application to treating chronic diseases. *Management Science* .

Øksendal B (2003) *Stochastic Differential Equations: An Introduction with Applications*. Hochschultext / Universitext (Springer), ISBN 9783540047582, URL `https://books.google.com.sg/books?id=kXw9hB4EEpUC`.

Peress J (2004) Wealth, information acquisition, and portfolio choice. *Review of Financial Studies* 17(3):879–914.

Peura S, Keppo J (2006) Optimal bank capital with costly recapitalization. *The Journal of Business* 79(4):2163–2201.

Radner R, Stiglitz J (1984) A nonconcavity in the value of information. *Bayesian models in economic theory* 5(3).

Verrecchia RE (1982) Information acquisition in a noisy rational expectations economy. *Econometrica: Journal of the Econometric Society* 1415–1430.

Xie J, Frazier PI, Chick SE (2016) Bayesian optimization via simulation with pairwise sampling and correlated prior beliefs. *Operations Research* 64(2):542–559.

Yu M, Debo L, Kapuscinski R (2015) Strategic waiting for consumer-generated quality information: Dynamic pricing of new experience goods. *Management Science* 62(2):410–435.