Moral Hazard and the Optimality of Debt

Benjamin Hébert

Stanford University

November 2015

Moral Hazard and Debt

- Why is debt so prevalent?
  - Doesn't it cause excessive risk-taking?
  - "Heads I win, tails you lose."
- Example: securitization.
  - The "seller" makes loans, sells security to investors.
  - In aggregate, outside investors ("buyers") purchase a debt claim.
  - The seller retains a "first-loss", "horizontal," or "levered equity" tranche.
  - Common examples: mortgages, credit card debt, auto loans.
  - Fun example: Bowie bonds (asset: song royalties).

Moral Hazard and Securitization

Why is moral hazard and securitization interesting?

- Bad incentives in securitization prior to financial crisis.
- Many corporate finance theories of debt don't fit.
  - Taxes, control rights, free cash flow, costly state verification.
- Theory will apply to corporate finance and principal-agent problems.

What I argue: debt is optimal for moral hazard despite excess risk-taking.

Mini-outline:

1. Model setup.
2. Intuition for why debt is optimal.
3. Related literature.

Assets and Agents

- There are two key times, 0 and 1.
- The seller can create or modify an asset.
  - Example: mortgage originator (Countrywide, New Century, etc...).
- The seller is risk-neutral.
- The seller discounts time-1 cashflows to time-0 with factor β_s.
- The buyer is risk-neutral, discounts cashflows with factor $\beta_b > \beta_s$.
- Define the "gains from trade" as $\kappa = \frac{\beta_b - \beta_s}{\beta_s}$.

States

- Ω is the set of time-1 states of the model.
- Start with discrete models: $N+1$ possible states, $\Omega = \{0, 1, \dots N\}$.
- Asset value v_i depends on state $i \in \Omega$.
- Assume v_i is weakly increasing in i, $v_0 = 0$, $v_N > v_0$.
- Could have two states $i, j \in \Omega$ with same asset value, $v_i = v_j$.

Securities

- At time 0, the seller can sell a security to the buyer.
- The value of the security is s_i, can be contingent on the state.
  - Can contract on asset value v_i, also other information.
- Limited liability: seller can promise only the value of the asset.
- In this sense, the security is backed by the asset: $\forall i \in \Omega,\ s_i \leq v_i$.
- Limited liability for the buyer: $\forall i \in \Omega,\ s_i \geq 0$.

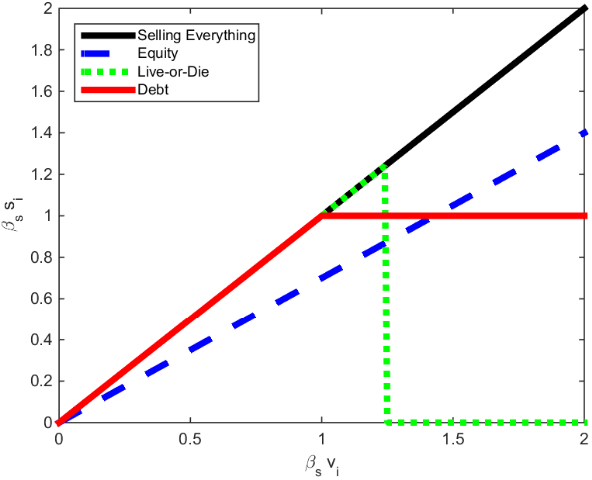
Securities

- At time 0, the seller can sell a security to the buyer.
- The value of the security is s_i, can be contingent on the state.
    - Can contract on asset value v_i, also other information.
- Limited liability: seller can promise only the value of the asset.
- In this sense, the security is backed by the asset: $\forall i \in \Omega,\ s_i \leq v_i$.
- Limited liability for the buyer: $\forall i \in \Omega,\ s_i \geq 0$.

- Space of allowed securities includes:
    - Selling all of the asset.
    - Selling none of the asset.
    - Selling a fractional claim on the asset.
    - Debt: for some $\bar{v} \in (0, v_N)$, $s_i = \min(v_i, \bar{v})$.

Security Designs

Figure: Important Security Designs



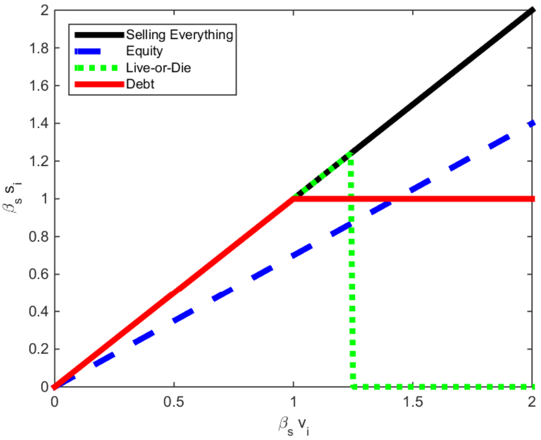
Classic Security Design Papers

- Definitions:
  - Effort: improving the mean asset value $E[v_i]$.
  - Risk shifting: changing the distribution of v_i without changing $E[v_i]$.
- Jensen and Meckling [1976]:
  - Debt good if only effort is possible.
  - Equity minimizes risk-shifting.
- Innes [1990]:
  - Only effort (no risk-shifting).
  - "Live-or-die" is optimal.
  - If restricted to monotone contracts, debt is optimal.
- This paper (and others): what is optimal when both effort and risk-shifting are possible? My answer: debt.

Why is Debt Optimal? Three Intuitions

1. Debt is "in between" the live-or-die security and equity.
2. Debt is the minimum variance security.
   - of all L.L. securities with the same E. V.
   - Reduced effort and risk-shifting costs summarized by security variance.
3. Constant payment, truncated for limited liability.

Figure: Important Security Designs



When is Debt Optimal?

- Static models, exact result depends on two assumptions:
  - The seller can choose any probability distribution ("non-parametric").
  - The cost of this choice is the Kullback-Leibler divergence.
- Static models, debt is approximately optimal:
  - Non-parametric models with invariant divergence cost functions.
- Dynamic models, exact result:
  - Continuous time model of effort, with quadratic costs.
- Dynamic models, approximate result:
  - Continuous time model of effort, convex costs.
- In the paper:
  - Static, parametric models.
  - Microfoundation based on rational inattention.

Static Security Design Papers

- Static models of moral hazard:
  - Seller can use multiple actions, not just effort ("parametric").
  - Unifying theme: debt not optimal (with or w/o monotonicity).
  - Mehran et al. [2013], Biais and Casamatta [1999], Edmans and Liu [2011], Fender and Mitchell [2009], Hellwig [2009].
- Security design with adverse selection:
  - Dang et al. [2011], DeMarzo and Duffie [1999], DeMarzo [2005], Gorton and Pennacchi [1990], Nachman and Noe [1994], Yang [2013].
  - This paper and "buyer-side" adv. sel. complementary.
  - Seller-side adv. sel. mitigation: non-bank lenders, etc.
- Empirical evidence consistent with moral hazard (mortgages):
  - Ashcraft et al. [2014], Begley and Purnanandam [2014], Demiroglu and James [2012], Elul [2015], Jiang et al. [2013], Keys et al. [2010], Krainer and Laderman [2014], Nadauld and Sherlund [2013], Purnanandam [2011], Rajan et al. [2010].
- Related theory:
  - Holmström and Milgrom [1987], Carroll [2015], Vanasco [2013].

Timing and Bargaining

- At time 0,
  1. The seller designs security ("D").
  2. The seller makes a take-it-or-leave-it offer ("O") at price K.
  3. The buyer accepts or rejects.
  4. If the buyer rejects, it is as-if $s_i = 0$ for all i, $K = 0$.
  5. The seller creates/modifies the asset ("M").
- At time 1:
  - The value of the asset is realized.
  - The security pays out its cashflows.
- Bargaining power irrelevant in "DOM" timing.
  - Only the price changes.

Timing Doesn't Matter

- Timing examples:
  - Royalties altered by Bowie's actions after security sold: "DOM".
  - For mortgages, create asset first, then sell security: "MDO".
  - "Shelf registration" timing (DeMarzo and Duffie [1999]): "DMO".

Theorem

If there is a unique sub-game perfect equilibrium with the "DOM" timing, the same price/security design/moral hazard actions are the strong proper equilibria of the "MDO" and "DMO" timings.

- Relies on "forward induction" intuition.
- The seller won't offer a strictly dominated security.
- Some technical caveats.

Moral Hazard

- After security design, the seller creates or modifies the asset.
- The probability that state i occurs is p^i.
- The seller chooses some $p \in M$ (set of prob. dists.).
  - Assume non-parametric: M is entire probability simplex.
- Each p has a cost (next slide).
- Cannot contract on p.
- Identical to Holmström and Milgrom [1987].
- Flexibility related to Yang [2013].

Moral Hazard Cont'd

- The seller pays a cost $\theta D(p||q)$, where D is a divergence.
  - Definition of divergence: $D(p||q) > 0$ for all $p \neq q$, $D(q||q) = 0$.
  - Extra assumptions: D is convex and smooth in p.
  - θ is a positive constant, and q is the minimum-cost distribution.
  - Assume that q has full support: $\forall i,\ q^i > 0$.
- Almost without loss of generality. Example
- Additional assumptions:
  - Units: $\beta_s \sum_i \hat{p}^i v_i = 1$, where $\hat{p}$ is "sell-nothing" distribution.
  - Largest possible value: $v_N > \sum_i q^i v_i + \beta_b^{-1} \kappa\theta$.

The Moral Hazard Sub-Problem

- Assume the security has been designed, and a price accepted.
- Define the seller's retained tranche, $\eta_i = \beta_s(v_i - s_i)$.
- Define the cost function $\psi(p) = \theta D(p||q)$.
- The seller's moral hazard problem and indirect utility function:

$$\phi(\eta) = \sup_{p \in M} \{\sum_i p^i \eta_i - \psi(p)\}.$$

- This leads to optimal policy $p(\eta)$.
  - Parametric: not necessarily unique. Non-parametric: unique.
  - Non-parametric, always interior: $\phi(\eta)$ is convex conjugate of $\psi(p)$.

The Security Design Problem

- At the security design stage, the seller solves

$$U(\eta^*) = \max_{\eta}\{\beta_b \sum_i p^i(\eta) s_i(\eta) + \phi(\eta)\},$$

subject to the limited liability constraint that $\eta_i \in [0, \beta_s v_i]$.

- Definition: $s_i(\eta) = v_i - \beta_s^{-1}\eta_i$.
- $p(\eta)$ and $\phi(\eta)$ from the moral hazard problem.

The Kullback-Leibler Divergence

- The cost function is the Kullback-Leibler divergence (relative entropy):

$$D_{KL}(p||q) = \sum_{i=0}^{N} p^i \ln(\frac{p^i}{q^i}).$$

- The KL divergence guarantees interior p.
  - In MH problem, unique p for each η, unique η for each p.
  - Holmström and Milgrom [1987] first emphasized duality of η and p.
- This divergence has many applications:
  - In Hansen and Sargent [2008] robustness theory.
  - In information theory and rational inattention (Sims [2003]).
  - Variety of uses in econometrics and statistics.
- Examples of $p(\eta)$: Example

Security Design FOC

- The security design FOC is

$$\kappa(p^i(\eta^*) - \lambda^i + \omega^i) - \beta_b \sum_{j>0} \frac{\partial p^i(\eta)}{\partial \eta_j}|_{\eta=\eta^*} s_j(\eta^*) = 0.$$

- λ and ω are limited liability multipliers (scaled by κ for convenience).
- Two important effects of changing s_i:
  - Change in gains from trade, $\kappa p^i(\eta^*)$.
  - Change in security value due to changing incentives.
- Key object: $\frac{\partial p^i(\eta)}{\partial \eta_j}$.

Moral Hazard FOC

- Moral hazard FOC is

$$\eta_i = \frac{\partial \psi(p)}{\partial p^i}|_{p=p^*(\eta)}.$$

- Differentiate with respect to η_j:

$$\frac{\partial p^i(\eta)}{\partial \eta_j} = [\partial_i \partial_j \psi(p)]^{-1}.$$

- Holds for any convex cost function with interior solution for p.

The KL Divergence

- The Hessian of the KL divergence is the Fisher info. matrix, $g_{ij}(p)$:

$$\partial_i \partial_j \psi(p) = \theta g_{ij}(p).$$

- $N = 2$ example (non-parametric):

$$g_{ij}(p) = \begin{bmatrix} \frac{1}{p^1} & 0 \\ 0 & \frac{1}{p^2} \end{bmatrix} + \begin{bmatrix} \frac{1}{p^0} & \frac{1}{p^0} \\ \frac{1}{p^0} & \frac{1}{p^0} \end{bmatrix}.$$

- Inverse Fisher information metric ($N = 2$):

$$[g_{ij}(p)]^{-1} = g^{ij}(p) = \begin{bmatrix} p^1 & 0 \\ 0 & p^2 \end{bmatrix} - \begin{bmatrix} p^1 \\ p^2 \end{bmatrix} \begin{bmatrix} p^1 & p^2 \end{bmatrix}.$$

- The Cramér-Rao bound is an equality: $V^p[s_i] = \sum_i \sum_j s_i s_j g^{ij}(p)$.

Debt is Optimal

- Security design FOC with KL divergence:

$$\kappa(p^{*i} - \lambda^i + \omega^i) - \beta_b \theta^{-1} \sum_{j>0} g^{ij}(p^*) s_j(\eta^*) = 0.$$

- Equivalent problem:

$$\max_s \kappa E^{p^*}[s_i] - \frac{1}{2}\theta^{-1}\beta_b V^{p^*}[s_i] + \kappa \sum_i \lambda^i (v_i - s_i) + \kappa \sum_i \omega^i s_i.$$

- The FOC solves a mean-variance optimization.

Theorem

The optimal security design is a debt, $s_j = \min(v_j, \bar{v})$, for all Ω and q.

Proof

Intuition

- With the KL divergence cost function, we trade off mean and variance.
- The higher mean-value of the security, the more the agents trade.
- Higher variance of the security payoff has multiple effects:
    - It can reduce the seller's incentive to improve $E[v_i]$ ("effort").
    - It creates incentives for the seller to shift cashflows.
- Consider small change in the retained tranche, ε_i.
    - ΔU from moral hazard costs (effort + risk-shifting): $-Cov^{p^*}(\varepsilon_i, \eta_i)$.
    - ΔU from discounted asset value: $\beta_s Cov^{p^*}(\varepsilon_i, v_i)$.
    - Combined effect: $\beta_s Cov^{p^*}(\varepsilon_i, s_i)$, change in security variance.
- Key assumption: risk-shifting and effort are both costly.

Relaxing Assumptions

- Timing and bargaining assumptions not essential:
  - Results hold if the buyer had some bargaining power.
  - Results hold if the security designed after asset created.
- Buyer can be risk averse, any increasing, non-convex utility function.
- Don't need $v_0 = 0$. Also don't need $s_i \geq 0$ (buyer LL).
- Redundant states are "extra" information.
  - The "asset" is actually a pool of many assets.
  - The redundant states have information about each asset's value.
  - Result: optimal security design depends only on total value.

Comparative Statics

- How risky is the optimal debt contract? Value of the "put option" is

$$\beta_b \bar{v} - \beta_b E^{p^*}[s_i] = \kappa\theta.$$

- The smaller the moral hazard (bigger θ), the riskier the debt.
- The bigger the gains from trade (bigger κ), the riskier the debt.
- The probability distributions q and p^* do not change the option value.
  - They do change the option-value/debt-level ($\bar{v}$) relationship.
- Empirical literature (Nadauld and Weisbach [2012]): $\kappa \approx 0.85\%$.

Two possible calibration strategies:

1. Estimate θ based on $\hat{p}$ (sell-nothing) vs p^*. Discussion
2. Infer θ from the security design: $\theta \approx 2$. Calculations

Is Relative Entropy Special?

- Consider the class of (smooth) f-divergences:

$$D_f(p||q) = \sum_{i=0}^{N} q^i f(\frac{p^i}{q^i})$$

where $f(u)$ is a smooth, convex function, $f(1) = 0$, $f'(1) = 0$, $f''(1) = 1$.

- Example (KL divergence):

$$f(u) = u \ln u - u + 1$$

- Others: χ^2 divergence, Hellinger distance, reverse KL.
- I assume $p(\eta)$ is interior for all η.

Theorem

If the optimal security design is debt for all sample spaces Ω and zero-cost probability distributions q, then that f-divergence is the Kullback-Leibler divergence.

Approximately Optimal Debt Contracts

- Approximation: moral hazard θ^{-1} and gains from trade κ near 0.
  - Possible motivation: time period is short.
  - Legal or reputational constraints on moral hazard.
  - The limit is degenerate (no moral hazard or gains from trade).
  - Approximation results are for θ^{-1}, κ small but positive.
  - Key assumption: higher order terms negligible.
- "First-order terms dominate higher order terms" is not the same as "problem is economically small."

Mean-Variance Tradeoffs

- Chentsov [1982]: For all f-divergences (invariant divergences),

$$\partial_i \partial_j D(p||q)|_{p=q} = c g_{ij}(p)|_{p=q},$$

for some constant $c > 0$.

- All f-divergences look like KL divergence when $p \approx q$.

Theorem

The difference in utilities achieved by an arbitrary security s and the sell-nothing security is

$$U(s;\theta^{-1},\kappa) - U(0;\theta^{-1},\kappa) = \kappa E^q[\beta_s s_i] - \frac{1}{2}\theta^{-1} V^q[\beta_s s_i] + O(\theta^{-2} + \theta^{-1}\kappa).$$

Debt is Approximately Optimal

- Define s_{debt} as the mean-variance optimal limited liability security:

$$s_{debt}(\theta^{-1}, \kappa) = \arg\max_{s} \kappa E^q[\beta_s s_i] - \frac{1}{2}\theta^{-1} V^q[\beta_s s_i]$$

over the set of LL securities.

Corollary

Let $s^*(\theta^{-1}, \kappa)$ be the optimal security given θ^{-1} and κ. Then

1. $\lim_{\theta^{-1}\to 0^+} s^*(\theta^{-1}, \bar{\kappa}\theta^{-1}) = s_{debt}(1, \bar{\kappa})$.
2. $U(s^*(\theta^{-1}, \kappa); \theta^{-1}, \kappa) - U(s_{debt}(\theta^{-1}, \kappa); \theta^{-1}, \kappa) = O(\theta^{-2} + \theta^{-1}\kappa)$.

- Numerical example: Example

Second Order Approximations

Corollary

Chentsov [1982]: for all invariant divergences, $\exists \alpha \in \mathbb{R}$ s.t.

$$\partial_i \partial_j \partial_k D(p||q)|_{p=q} = (\frac{3+\alpha}{2}) \partial_i g_{jk}(p)|_{p=q}.$$

- KL: $\alpha = -1$. Hellinger, χ^2, reverse KL: $\alpha = 0, -3, 1$.

Corollary

The second order-optimal security, for $\alpha < 1 + \frac{2}{\kappa}$,

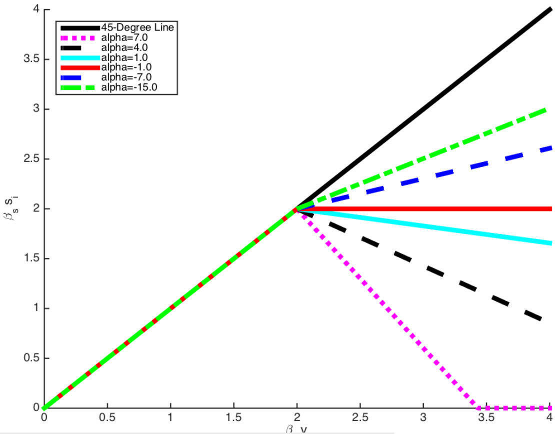
$$s_i^{d-eq}(\theta^{-1}, \kappa) = \begin{cases} v_i & \text{if } v_i < \bar{v} \\ (\frac{-\kappa(1+\alpha)}{2+\kappa(1-\alpha)}(v_i - \bar{v}) + \bar{v})_+ & \text{if } v_i \geq \bar{v} \end{cases}$$

for some $\bar{v} \in (0, v_N)$.

Second-Order Optimal Securities

- α controls which problem is most severe: effort or risk-shifting.
- The optimal contract goes from "Live-or-Die" (largest α) to equity (smallest α). Debt security is "in between."
- There is a "pecking order." First order: debt. Second order: debt and equity. Higher order: custom contract.

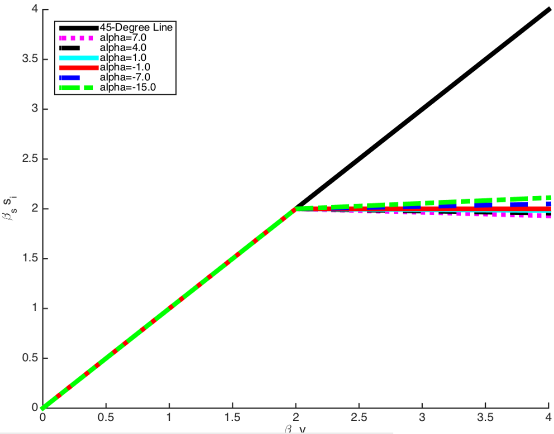
Figure: Second Order Security Designs ($\kappa = 17\%$)



Second-Order Optimal Securities

- For common α values (Hellinger, χ^2, reverse KL) and empirically relevant κ, second order optimal contracts close to debt.
- In example with $\alpha = -7$, debt achieves 99.96% of the gains of the optimal contract, vs. selling everything.
- How to estimate/calibrate α?

Figure: Second Order Security Designs ($\kappa = 0.85\%$)



Dynamic Models Outline

Outline:

1. CT model setup.
2. How CT and static models are related.
3. With quadratic effort costs, debt is optimal.
4. With convex efforts costs, debt is approx. optimal.

Most related literature:

- Chassang [2013], Cvitanić et al. [2009], DeMarzo and Sannikov [2006], Hartman-Glaser et al. [2012], Holmström and Milgrom [1987], Malamud et al. [2013], Sannikov [2013], Schaettler and Sung [1993].

Model Setup

- At time zero, seller and buyer trade security.
- Between times zero and one, seller applies effort.
- At time one, payoffs are received.
- V_t denotes the asset value at time t.
- V denotes entire path of V_t, for $t \in [0, 1]$.
- The security design $s(V)$ must be limited liability, $s(V) \in [0, V_1]$.
- The security is $\mathcal{F}_1^V$-measurable (filtration generated by V_t).
- Key idea: V_1 is what matters. Path is "redundant state."

The Asset Value Process

- V_t follows a controlled diffusion process:

$$dV_t = b(V_t, t)dt + u_t\sigma(V_t, t)dt + \sigma(V_t, t)dW_t.$$

- The initial value $V_0 > 0$ is known by buyer and seller at time zero.
- The functions $b(V_t, t)$ and $\sigma(V_t, t) > 0$ are given.
  - I assume restrictions for existence and integrability, discussed later.
- The drift u_t ("effort") is controlled by the seller.
  - Effort always increases the expected time-1 asset value.
- W_t is a BM on the canonical filtered space, $(\Omega, \mathcal{F}, \{\mathcal{F}_t^W\}_{t=0}^1, \tilde{P})$.
- Moral hazard: W_t not observable/contractible.

Effort Strategies and Costs

- Flow cost of effort function, $g(t, V_t, u_t)$.
  - $g(\cdot)$ is continuously twice differentiable, convex in u_t, weakly positive.
  - $g(t, V_t, 0) = 0$ for all t and V_t.
- The admissible set of strategies $\mathscr{U}$ is the set of $\mathcal{F}_t^V$-adapted, square-integrable feedback control strategies u_t such that

$$Z_t = \exp(\int_0^t u_s dB_s - \frac{1}{2}\int_0^t u_s^2 ds)$$

  satisfies $E^{\tilde{P}}[Z_t^4] < \infty$ for all t.
- This implies Novikov's condition holds (Cvitanić et al. [2009]).

Seller's CT Problem

- Retained tranche $\eta(V) = \beta_s(V_1 - s(V))$, as before.
- The moral hazard sub-problem is

$$\begin{aligned}\phi_{CT}(\eta) &= \sup_{\{u_t\}\in\mathscr{U}} \phi_{CT}(\eta; \{u_t\}) \\ &= \sup_{\{u_t\}\in\mathscr{U}} \{E^{\tilde{P}}[\eta(V)] - E^{\tilde{P}}[\int_0^1 g(t, V_t, u_t)dt]\}.\end{aligned}$$

- Let S be the set of limited-liability, $\mathcal{F}_1^V$-measurable security designs.
- The security design problem is

$$U_{CT}(s^*) = \sup_{s\in S}\{\beta_b E^{\tilde{P}}[s(V)] + \phi_{CT}(\eta(V))\}.$$

Weak vs. Strong Formulation

- "Weak" formulation: replace V with X, $\tilde{P}$ with P.
  - Cvitanić et al. [2009], Schaettler and Sung [1993].
  - Equivalent for our purposes (security design).
- Consider a process

$$dX_t = b(X_t, t)dt + \sigma(X_t, t)dB_t,$$

where B is a Brownian motion on the probability space $(\Omega, \mathcal{F}, Q)$.

- For any $\{u_t\} \in \mathscr{U}$, $Z_t = \exp(\int_0^t u_s dB_s - \frac{1}{2}\int_0^t u_s^2 ds)$ is a martingale.
- By Girsanov's theorem, define a measure P by $\frac{dP}{dQ} = Z_1$.
- Under measure P, X has the same law that V does under measure $\tilde{P}$.
- Key point: we defined measure Q and Brownian motion B_t.

Choosing a Distribution

- Each effort strategy $\{u_t\} \in \mathscr{U}$ defines a $\frac{dP}{dQ}$.
- The reverse is also true. Let M be the set of measures on $(\Omega, \mathcal{F})$ s.t.
  - All $P \in M$ are absolutely continuous with respect to Q.
  - $E^Q[(\frac{dP}{dQ})^4] < \infty$.
- For each $P \in M$, there exists a unique (up to an evanescence) $\{u_t\} \in \mathscr{U}$ such that

$$\frac{dP}{dQ} = \exp(\int_0^1 u_s dB_s - \frac{1}{2}\int_0^1 u_s^2 ds).$$

The Moral Hazard Problem Revisited

- The moral hazard problem can be rewritten as

$$\phi_{CT}(\eta) = \sup_{P \in M} \{E^Q[\frac{dP}{dQ}\eta(X)] - D_g(P||Q)\},$$

where $D_g(P||Q)$ is a divergence, defined as

$$D_g(P||Q) = E^P[\int_0^1 g(t, X_t, u_t)dt],$$

where $\{u_t\} \in \mathscr{U}$ is the control strategy that generates $\frac{dP}{dQ}$.

- Two questions:
  1. Is there a $g(\cdot)$ corresponding to the KL divergence?
  2. Is there a $g(\cdot)$ s.t. $D_g(\cdot)$ has the Fisher “matrix” as its “Hessian?”

Quadratic Costs

- Consider flow cost fuction $g(t, X_t, u_t) = \frac{\theta}{2} u_t^2$.
- Define $D_{KL}(P||Q) = E^Q[\frac{dP}{dQ} \ln(\frac{dP}{dQ})]$.
  - Bierkens and Kappen [2014] show $\theta D_{KL}(P||Q) = D_g(P||Q) = E_0^Q[\frac{\theta}{2} \int_0^1 (u_t)^2 dt]$. Algebra
- Integrability condition: $E^Q[\exp(4\theta^{-1} X_1)] < \infty$.

Theorem

In the continuous time model with quadratic costs, for all functions $b(V_t, t)$ and $\sigma(V_t, t)$ such that the integrability conditions are satisfied and a solution to the SDE exists, the optimal security design is a debt contract:

$$s_j(\eta^*) = \min(v_j, \bar{v}).$$

Discussion

- Surprising equivalence of static and dynamic contracting problems.
- In the dynamic effort problem, seller can "choose a distribution."
- The cost is the KL divergence, and the mean-variance intuition applies.
- Key distinction: final outcomes V_1 vs. paths V.
  - Recall: in static problem, Ω could contain redundant values.
  - Same idea: inefficient to pay seller different when $V_1(\omega) = V_1(\omega')$.
- Comparison with Holmström and Milgrom [1987]:
  - In HM, process is arithmetic BM, "constant" contract optimal.
  - In this model, because of limited liability, debt is optimal.
- $E^Q[\cdot]$ could also be expectations under common risk-neutral measure.
- Extension: payment at time $T > 1$. Extension

Writedowns (Renegotiation)

- Seller has sold optimal debt security to the buyer.
- Suppose that at time $t = 0.5$, seller can offer buyer a new security.
  - but cannot offer payments. LL binds.
  - Assume no gains from trade at time 0.5.
- If $v_{0.5}$ is very low, seller has weak incentives for effort.
- Can seller and buyer find a Pareto-improving writedown? Yes, if:

$$\beta_b E_0^P[s] > \theta \Longleftrightarrow \frac{\bar{v} - E_0^P[s]}{E_0^P[s]} < \kappa.$$

- If moral hazard is large enough relative to security value.

Convex Cost Functions

- Now consider convex cost functions $g(t, V_t, u_t) = \theta\psi(u_t)$.
  - For technical reasons, require $|u_t| \leq \bar{u}$.
  - $\psi(\cdot)$ convex, continuously twice differentiable, minimized at $u_t = 0$.
  - For all $|u_t| \leq \bar{u}$, $\psi''(u) \in [K_1, K_2]$ for $0 < K_1 < 1 < K_2$.
- Integrability: for all feasible $\{u_t\}$, $E^P[X_1^4] < \infty$.
- Consider first-order approximation in κ and θ^{-1} (same as earlier).
- Goal: show that debt is first-order optimal.

Connection to Fisher "Matrix" (Heuristic)

- Let $\frac{dP}{dQ}(\gamma,\tau) = \exp(\int_0^1 (\gamma u_s + \tau v_s) dB_s - \frac{1}{2}\int_0^1 (\gamma u_s + \tau v_s)^2 ds)$.
  - γ, τ are parameters. u_s and v_s are "directions" (Cameron-Martin).

Connection to Fisher "Matrix" (Heuristic)

- Let $\frac{dP}{dQ}(\gamma,\tau) = \exp(\int_0^1 (\gamma u_s + \tau v_s) dB_s - \frac{1}{2}\int_0^1 (\gamma u_s + \tau v_s)^2 ds)$.
  - γ, τ are parameters. u_s and v_s are "directions" (Cameron-Martin).

- Fisher information in directions u, v:

$$\begin{aligned} I(u,v) &= E^{P(\gamma,\tau)}[(\frac{\partial}{\partial\gamma}\ln(\frac{dP}{dQ}(\gamma,\tau)))(\frac{\partial}{\partial\tau}\ln(\frac{dP}{dQ}(\gamma,\tau)))]|_{\gamma=\tau=0} \\ &= E^Q[\int_0^1 u_s v_s ds]. \end{aligned}$$

Connection to Fisher "Matrix" (Heuristic)

- Let $\frac{dP}{dQ}(\gamma,\tau) = \exp(\int_0^1 (\gamma u_s + \tau v_s) dB_s - \frac{1}{2}\int_0^1 (\gamma u_s + \tau v_s)^2 ds)$.
  - γ, τ are parameters. u_s and v_s are "directions" (Cameron-Martin).
- Fisher information in directions u, v:

$$\begin{aligned} I(u,v) &= E^{P(\gamma,\tau)}[(\frac{\partial}{\partial\gamma}\ln(\frac{dP}{dQ}(\gamma,\tau)))(\frac{\partial}{\partial\tau}\ln(\frac{dP}{dQ}(\gamma,\tau)))]|_{\gamma=\tau=0} \\ &= E^{Q}[\int_0^1 u_s v_s ds]. \end{aligned}$$

- Second variation ("Hessian") of $D_\psi(P||Q)$:

$$\frac{\partial}{\partial\tau}\frac{\partial}{\partial\gamma} D_\psi(P(\gamma,\tau)||Q)|_{\gamma=\tau=0} =$$

$$\theta\frac{\partial}{\partial\tau}\frac{\partial}{\partial\gamma} E^{P(\gamma,\tau)}[\int_0^1 \psi(\gamma u_s + \tau v_s) ds]|_{\gamma=\tau=0} = \theta E^{Q}[\int_0^1 u_s v_s ds].$$

Mean-Variance

Theorem

In the continuous time problem, with convex flow cost function $\psi(\cdot)$, for all functions $b(V_t, t)$ and $\sigma(V_t, t)$ such that the integrability conditions are satisfied and a solution to the SDE exists, the difference in utilities achieved by an arbitrary security s and the sell-nothing security is

$$U(s;\theta^{-1},\kappa) - U(0;\theta^{-1},\kappa) = \kappa E^Q[\beta_s s] - \frac{1}{2}\theta^{-1} V^Q[\beta_s s] + O(\theta^{-2} + \theta^{-1}\kappa).$$

- Debt securities are first-order optimal.
- Open question: is this related to invariant divergences?

Conclusions

- Debt optimally balances concerns about effort and risk-shifting.
  - This results goes against the intuition from previous papers.
- The exact result applies with KL divergence cost functions.
  - These can be motivated with rational inattention (see paper).
  - The KL divergence can be justified from a continuous time problem.
- The result holds approximately in many static and dynamic models.
- The paper use new techniques to explain the prevalence of debt.

Proof

- Using definition of Fisher information metric, rewrite FOC:

$$\beta_b s_j = \beta_b \sum_k p^k s_k + \kappa\theta - \kappa\left(\frac{\lambda^j - \omega^j}{p^j}\right)$$

- Define an endogenous constant, $\bar{v} = \sum_k p^k s_k + \frac{\kappa}{\beta_b}\theta$.
- $\omega^j > 0$ never possible.
- If $\lambda^j = 0$, so $s_j \leq v_j$, then $s_j = \bar{v}$.
- If $\lambda^j > 0$, then $s_j = v_j$, and by the formula above, $v_j < \bar{v}$.
- Therefore, $s_j = \min(v_j, \bar{v})$, which is debt (can show that $\bar{v} \in (0, v_N)$).

Back

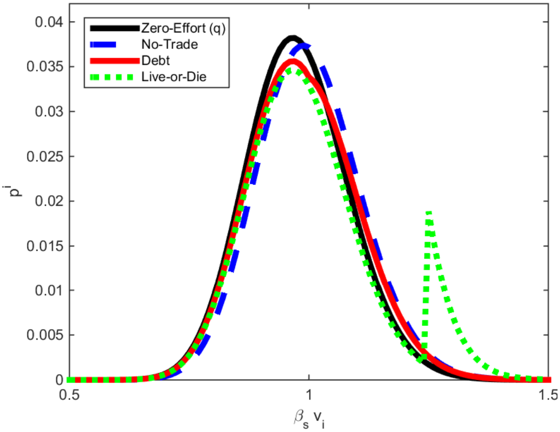
Generality of Divergences

- Example: seller chooses log-mean (effort) e of log-normal distribution.
  - Convex cost $c(e - e_q)$, minimized at $c(0) = 0$.
- In this framework:
  - Set M is set of log-normal distributions with log-variance σ^2.
  - Assume $p \in M$, calculate e: $e = \int_\Omega p(v) \ln(v) dv$.
  - $\theta D(p||q) = c(\int_\Omega \ln(v)[p(v) - q(v)]dv)$.

Back

Moral Hazard Solutions

Figure: Optimal $p(\eta)$



Back

Calibration from Moral Hazard

Assuming KL divergence cost function:

$$\theta^{-1} \approx E^{\hat{p}}[v_i] \cdot \frac{E^{\hat{p}}[v_i] - E^{p^*}[v_i]}{Cov^{p^*}(v_i, s_i^*)}$$

- Empirical work estimates ex-post differences.
  - Ex-post effect size vs. ex-ante effect size.
  - Elul [2015], Keys et al. [2010], Krainer and Laderman [2014], Purnanandam [2011]
- Hard to estimate $Cov^{p^*}(v_i, s_i^*)$.
  - How much risk did sellers think they retained?
  - Evidence of some risk retention in Demiroglu and James [2012], Gorton [2009].

Back

Inference from Security Design

Rearrange optimal security design equation:

$$\underbrace{\frac{\beta_b \bar{v} - \beta_b E^{p^*}[s_i]}{\beta_b E^{p^*}[s_i]}}_{Spread} \underbrace{\frac{E^{p^*}[s_i]}{E^{p^*}[v_i]}}_{Share} \underbrace{(1 - \frac{E^{\hat{p}}[v_i] - E^{p^*}[v_i]}{E^{\hat{p}}[v_i]})}_{Moral\ Hazard} \kappa^{-1} = \theta$$

- Spread: the initial yield on securitized assets, above discount rate.
  - Initial spread on 06-2 ABX (90/10 weights on AAA and BBB): 34*bps*
- Share: the fraction of the asset value that is sold to investors. ≈ 1.
- Moral Hazard: percentage diff. between ex-ante securitized/non-securitized asset value.
  - Nearly 1 (consistent w/ empirical evidence).
- Nadauld and Weisbach [2012] estimate $\kappa \approx 17 bps$
- Estimate θ of roughly 2. Back

Example Asymptotic Utility

Figure: Utility vs. Sell-Everything

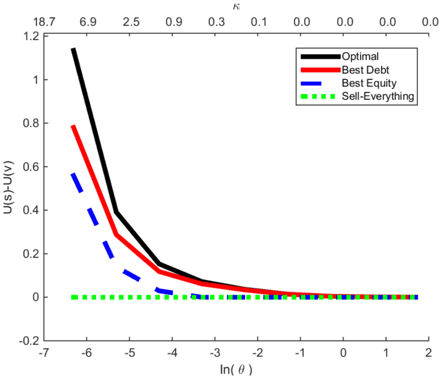
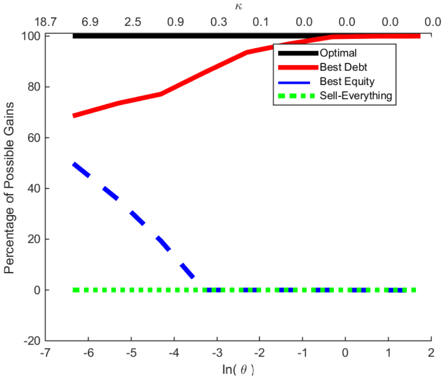


Figure: Percentage of Possible Gains



- Back

KL and Quadratic Costs

- By Girsanov's theorem, $B_t^P = B_t - \int_0^t u_s ds$ is a BM under measure P.

$$\begin{aligned} D_{KL}(P||Q) &= E^P[\ln(\frac{dP}{dQ})] \\ &= E^P[\int_0^t u(X,s)dB_s - \frac{1}{2}\int_0^t u(X,s)^2 ds] \\ &= E^P[\int_0^t u(X,s)dB_s^P + \frac{1}{2}\int_0^t u(X,s)^2 ds] \\ &= \frac{1}{2}E^P[\int_0^t u(X,s)^2 ds]. \end{aligned}$$

Back

Extension

- Originator controls V_t for $t \in [0,1]$, but security pays at time $T > 1$.
- Suppose that $\frac{dP}{dQ}$ is $\mathcal{F}_1$-measurable. Then

$$E_0^P[\ln(\frac{dP}{dQ})] = E_0^P[\ln(E_1^P[\frac{dP}{dQ}])].$$

- The KL divergence evaluated at time 1 and T will be the same.
- Seller controls $\frac{dP}{dQ}$, subject to measurability constraint.
- Just like a parametric problem (see paper).

Back

Adam B Ashcraft, Kunal Gooriah, and Amir Kermani. Does skin-in-the-game affect security performance? evidence from the conduit cmbs market. 2014.

Taylor A Begley and Amiyatosh K Purnanandam. Design of financial securities: Empirical evidence from private-label rmbs deals. *Available at SSRN 2022380*, 2014.

B. Biais and C. Casamatta. Optimal leverage and aggregate investment. *The Journal of Finance*, 54(4):1291–1323, 1999.

Joris Bierkens and Hilbert J Kappen. Explicit solution of relative entropy weighted control. *Systems & Control Letters*, 72:36–43, 2014.

G. Carroll. Robustness and linear contracts. *The American Economic Review*, 105(2):536–563, 2015.

S. Chassang. Calibrated incentive contracts. *Econometrica*, 81(5): 1935–1971, 2013.

N.N. Chentsov. Statistical decision rules and optimal inference. transl. math. *Monographs, American Mathematical Society, Providence, RI*, 1982.

J. Cvitanić, X. Wan, and J. Zhang. Optimal compensation with hidden action and lump-sum payment in a continuous-time model. *Applied Mathematics and Optimization*, 59(1):99–146, 2009.

T.V. Dang, G. Gorton, and B. Holmström. Ignorance and the optimality of debt for liquidity provision. Technical report, Working Paper, Yale University, 2011.

P. DeMarzo and D. Duffie. A liquidity-based model of security design. *Econometrica*, 67(1):65–99, 1999.

P. M. DeMarzo. The pooling and tranching of securities: A model of informed intermediation. *Review of Financial Studies*, 18(1):1–35, 2005.

P. M. DeMarzo and Y. Sannikov. Optimal security design and dynamic capital structure in a continuous-time agency model. *The Journal of Finance*, 61(6):2681–2724, 2006.

C. Demiroglu and C. James. How important is having skin in the game? originator-sponsor affiliation and losses on mortgage-backed securities. *Review of Financial Studies*, 25(11):3217–3258, 2012.

A. Edmans and Q. Liu. Inside debt. *Review of Finance*, 15(1):75–102, 2011.

R. Elul. Securitization and mortgage default. *Journal of Financial Services Research*, pages 1–29, 2015.

I. Fender and J. Mitchell. Incentives and tranche retention in securitisation: a screening model. *CEPR Discussion Paper No. DP7483*, 2009.

G. Gorton. The subprime panic. *European Financial Management*, 15(1): 10–46, 2009.

G. Gorton and G. Pennacchi. Financial intermediaries and liquidity creation. *Journal of Finance*, pages 49–71, 1990.

L.P. Hansen and T.J. Sargent. *Robustness*. Princeton Univ Pr, 2008.

B. Hartman-Glaser, T. Piskorski, and A. Tchistyi. Optimal securitization with moral hazard. *Journal of Financial Economics*, 104(1):186–202, 2012.

M.F. Hellwig. A reconsideration of the jensen-meckling model of outside finance. *Journal of Financial Intermediation*, 18(4):495–525, 2009.

B. Holmström and P. Milgrom. Aggregation and linearity in the provision of intertemporal incentives. *Econometrica*, pages 303–328, 1987.

R.D. Innes. Limited liability and incentive contracting with ex-ante action choices. *Journal of Economic Theory*, 52(1):45–67, 1990.

M.C. Jensen and W. Meckling. Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, 3 (4):305–360, 1976.

W. Jiang, A. A. Nelson, and E. Vytlacil. Securitization and loan performance: Ex ante and ex post relations in the mortgage market. *Review of Financial Studies*, page hht073, 2013.

B.J. Keys, T. Mukherjee, A. Seru, and V. Vig. Did securitization lead to lax screening? evidence from subprime loans. *The Quarterly Journal of Economics*, 125(1):307–362, 2010.

J. Krainer and E. Laderman. Mortgage loan securitization and relative loan performance. *Journal of Financial Services Research*, 45(1):39–66, 2014.

S. Malamud, H. Rui, and A. Whinston. Optimal incentives and securitization of defaultable assets. *Journal of Financial Economics*, 107 (1):111–135, 2013.

H. Mehran, V. V. Acharya, and A. V. Thakor. Caught between scylla and charybdis? regulating bank leverage when there is rent seeking and risk shifting. *Regulating Bank Leverage When There is Rent Seeking and Risk Shifting (February 2013). ECGI-Finance Working Paper*, (365), 2013.

D. C. Nachman and T. H. Noe. Optimal design of securities under asymmetric information. *Review of Financial Studies*, 7(1):1–44, 1994.

T. D. Nadauld and S. M. Sherlund. The impact of securitization on the expansion of subprime credit. *Journal of Financial Economics*, 107(2): 454–476, 2013.

T. D. Nadauld and M. S. Weisbach. Did securitization affect the cost of corporate debt? *Journal of financial economics*, 105(2):332–352, 2012.

A. Purnanandam. Originate-to-distribute model and the subprime mortgage crisis. *Review of Financial Studies*, 24(6):1881–1915, 2011.

U. Rajan, A. Seru, and V. Vig. The failure of models that predict failure: distance, incentives and defaults. *Chicago GSB Research Paper*, (08-19), 2010.

Y. Sannikov. Moral hazard and long-run incentives, 2013.

H. Schaettler and J. Sung. The first-order approach to the continuous-time principal–agent problem with exponential utility. *Journal of Economic Theory*, 61(2):331–371, 1993.

C.A. Sims. Implications of rational inattention. *Journal of Monetary Economics*, 50(3):665–690, 2003.

V. M. Vanasco. Information acquisition vs. liquidity in financial markets. Technical report, 2013.

M. Yang. Optimality of debt under flexible information acquisition. *Available at SSRN 2103971*, 2013.