Identifying and Estimating Models of Optimal Contracting: Part 1

Robert A. Miller

AES Summer School 2020

July 2020

- In an Arrow Debreu world with a Walrasian equilibrium, it doesn't matter whether an employee is paid the value of his marginal product less the amenity value with a certain wage or a piece rate.
- Both the employer and the employee can adjust their portfolio of financial assets at the competitive equilibrium rate to achieve the same resource allocation.
- For example if the uncertainty is idiosyncratic, both the employee or the employer could full insure at actuarially fair rates.
- These two lectures analyze compensation and labor supply when the contract form matters.
- This arises naturally in environments with asymmetric information.

A Pure Moral Hazard Model

Framework

- A risk neutral principal proposes a compensation plan to a risk averse agent, an explicit contract or an implicit agreement, which depends on the future realization of gross revenue to the principal.
- The agent accepts or rejects the principal's (implicit) offer.
- If he rejects the offer he receives a fixed utility from an outside option.
- If he accepts the offer, the agent chooses between pursuing the principal's objectives of value maximization (working), versus following objectives he would pursue if he was paid a fixed wage (shirking).
- The principal observes whether the offer is accepted, but not the agent's work routine.
- After revenue is realized, the agent receives compensation according to the explicit contract or implicit agreement, and the principal pockets the remainder as profit.

A Pure Moral Hazard Model

Choices of the agent

- Denote the workplace employment decision of the agent by an indicator $l_0 \in \{0, 1\}$, where $l_0 = 1$ means the agent rejects the principal's offer.
- Denote the effort level choices by $l_j \in \{0, 1\}$ for $j \in \{1, 2\}$, where diligence work is defined by setting $l_2 = 1$, and shirking is defined by setting $l_1 = 1$.
- Since taking the outside option, working diligently and shirking are mutually exclusive activities, $l_0 + l_1 + l_2 = 1$.

- Gross revenue to the principal is denoted by x, a random variable drawn from a probability distribution that is determined by the agent's work routine.
- After x is revealed the both the principal and the agent at the end of the period, the agent receives compensation according to the contract or implicit agreement.
- To reflect its potential dependence on (or measurability with respect to) x, we denote compensation by $w(x)$.
- The principal's profit is revenue less compensation, $x - w(x)$.

A Pure Moral Hazard Model

Marginal product of the agent

- Denote by $f(x)$ the probability density function for revenue conditional on the agent working, and let $f(x)g(x)$ denote the probability density function for revenue when the agent shirks.
- We assume:

$$E[xg(x)] \equiv \int xf(x)g(x)\,dx < \int xf(x)\,dx \equiv E[x]$$

- The inequality reflects the preference of principal for working over shirking.
- Since $f(x)$ and $f(x)g(x)$ are densities, $g(x)$, the ratio of the two densities, is a likelihood ratio.
- That is $g(x)$ is nonnegative for all x, bounded, and:

$$E[g(x)] \equiv \int g(x)f(x)\,dx = 1$$

- We assume the agent is an expected utility maximizer and utility is exponential in compensation, taking the form:

$$-I_0 - I_1 \alpha_1 E\left[e^{-\gamma w(x)} g(x)\right] - I_2 \alpha_2 E\left[e^{-\gamma w(x)}\right]$$

where without further loss of generality we normalize the utility of the outside option to negative one.

- Thus γ is the coefficient of absolute risk aversion, and α_j is a utility parameter with consumption equivalent $-\gamma^{-1}\log(\alpha_j)$ that measures the distaste from effort level $j \in \{1, 2\}$.
- We assume $\alpha_2 > \alpha_1$ meaning that shirking gives more utility to the agent, than working.
- A conflict of interest arises between the principal and the agent because he prefers shirking, meaning $\alpha_1 < \alpha_2$, yet the principal prefers working since $E[xg(x)] < E[x]$.

Solving the Pure Moral Hazard Model

Participation constraint

- To induce the agent to accept the principal's offer and engage in his preferred activity, shirking, it suffices to propose a contract that gives the agent an expected utility of at least minus one.
- In this case we require $w(x)$ to satisfy the inequality:

$$\alpha_1 E\left[e^{-\gamma w(x)} g(x)\right] \leq 1$$

Solving the Pure Moral Hazard Model

Participation and incentive compatibility constraints

- To elicit work from the agent, the principal must offer a contract that gives the agent a higher expected utility than the outside option, and a higher expected utility than shirking.
- In this case we require:

$$\alpha_2 E\left[e^{-\gamma w(x)}\right] \leq 1$$

and:

$$\alpha_2 E\left[e^{-\gamma w(x)}\right] \leq \alpha_1 E\left[e^{-\gamma w(x)} g(x)\right]$$

Solving the Pure Moral Hazard Model

Cost minimization inducing work

- Defining $v(x) \equiv \exp\left[-\gamma w(x)\right]$ note that:

$$-E\left[w(x)\right] = \gamma^{-1} E\left\{\log\left[v(x)\right]\right\}$$

the participation constraint can be expressed as:

$$\alpha_2 E\left[v(x)\right] \leq 1$$

and the incentive compatibility constraint becomes:

$$\alpha_2 E\left[v(x)\right] \leq \alpha_1 E\left[v(x) g(x)\right]$$

- In the transformed problem we maximize a strictly concave objective function with linear constraints. Applying the Kuhn Tucker theorem applies, we choose $v(x)$ for each x to maximize:

$$E\left\{\log\left[v(x)\right]\right\} + \eta_0 E\left[1 - \alpha_2 v(x)\right] + \eta_1 E\left[\alpha_1 g(x) v(x) - \alpha_2 v(x)\right]$$

Lemma (Margiotta and Miller, 2000)

To minimize the cost of inducing the agent to accept employment and work diligently the board offers the contract:

$$w^o(x) \equiv \gamma^{-1} \ln \alpha_2 + \gamma^{-1} \ln \left[1 + \eta \left(\frac{\alpha_2}{\alpha_1} \right) - \eta g(x) \right]$$

where η is the unique positive solution to the equation:

$$E\left[\frac{g(x)}{\alpha_2 + \eta[(\alpha_2/\alpha_1) - g(x)]} \right] = E\left[\frac{(\alpha_2/\alpha_1)}{\alpha_2 + \eta[(\alpha_2/\alpha_1) - g(x)]} \right]$$

- Differentiate the Lagrangian with respect $v(x)$ to obtain:

$$v(x)^{-1} = \eta_0 \alpha_2 + \eta_1 \alpha_2 - \eta_1 \alpha_1 g(x)$$

- We can show both constraints are met with equality, establishing the formula for η, and showing $\eta_0 = 1$, to yield:

$$v(x)^{-1} = \alpha_2 + \eta_1 \alpha_2 - \eta_1 \alpha_1 g(x)$$

Solving the Pure Moral Hazard Model

Intuition for cost minimizing contract

- There is no point exposing the manager to uncertainty in a shirking contract by tying compensation to revenue.
- Hence a agent paid to shirk is offered a fixed wage that just offsets his nonpecuniary benefits, $\gamma^{-1} \ln \alpha_1$.
- The certainty equivalent of the cost minimizing contract that induces diligent work is $\gamma^{-1} \ln \alpha_2$, higher than the optimal shirking contract to compensate for the lower nonpecuniary benefits because $\alpha_2 > \alpha_1$.
- Moreover the agent is paid a positive risk premium of $E\left[w^o(x)\right] - \gamma^{-1} \ln \alpha_2$.
- In this model of pure moral hazard these two factors, that working is less enjoyable than shirking, and more certainty in compensation is preferable, explains why compensating an agent to align his interests with the principal is more expensive than merely paying them enough to accept employment.

Measuring the Importance of Moral Hazard

Three measures

- Recall the optimal compensation with moral hazard is $w^o(x)$ and to meet the participation constraint, shareholders must pay $\gamma^{-1} \ln \alpha_2$.
- Therefore the maximal amount shareholders would pay to rid the firm of the moral hazard problem is:

$$\tau_1 \equiv E_t \left[w^o(x) - \gamma^{-1} \ln \alpha_2 \right] = \gamma^{-1} E \left\{ \ln \left[1 + \eta \left(\frac{\alpha_2}{\alpha_1} \right) - \eta g(x) \right] \right\}$$

- A second measure of moral hazard is the nonpecuniary benefits the manager obtains from shirking.
- This is the monetized utility loss from working versus shirking:

$$\tau_2 \equiv \gamma^{-1} \ln \alpha_1 - \gamma^{-1} \ln \alpha_2 = -\gamma^{-1} \ln \left(\alpha_2 / \alpha_1 \right)$$

- Third is the gross loss a firm incurs from the manager shirking instead of working:

$$\tau_3 \equiv E \left[x - xg(x) \right]$$

A Fully Parametric Specification

Truncated Normal distribution and Absolute Risk Aversion (CARA)

- Assume x is distributed truncated normal with lower truncation point ψ (representing bankruptcy or limited liability) with mean μ_w (μ_s) and variance σ^2 for parent normal if agent works (shirks):

$$\begin{aligned} f(x) &= \frac{1}{\sigma_w\sqrt{2\pi}}\Phi\left(\frac{\mu_w-\psi}{\sigma}\right)^{-1}\exp\left[\frac{-(x-\mu_w)^2}{2\sigma^2}\right] \\ \ln g(x) &= \ln\Phi\left[(\mu_s-\psi)/\sigma\right]-\ln\Phi\left[(\mu_w-\psi)/\sigma\right] \\ &\quad +\frac{\mu_w^2-\mu_s^2}{2\sigma^2}+\frac{(\mu_s-\mu_w)}{\sigma^2}x \end{aligned}$$

- Thus the model is parameterized by $(\psi,\mu_w,\sigma,\mu_s,\gamma,\alpha_1,\alpha_2)$.
- Suppose there are N observations on $(\widetilde{w}_n,x_n)$ where:

$$\widetilde{w}_n \equiv w_n+\epsilon_n \text{ and } E\left[\epsilon_n \,|x_n\right]=0.$$

A Fully Parametric Specification

Estimation

- Margiotta and Miller (2000) estimate:
  1. ψ with $\widehat{\psi} \equiv \min\{x_1, \ldots, x_N\}$. (Note $\widehat{\psi}$ converges to ψ at rate faster than $\sqrt{N}$ but is sensitive to measurement error.)
  2. (μ_w, σ) with LIML by forming likelihood for $f(x)$ with $\{x_1, \ldots, x_N\}$ under the assumption that $\widehat{\psi} = \psi$. (No first stage correction is necessary.)
  3. $(\mu_s, \gamma, \alpha_1, \alpha_2)$ with NLS based on

$$\widetilde{w}_n = \gamma^{-1} \ln \alpha_2 + \gamma^{-1} \ln \left[1 + \eta \left(\frac{\alpha_2}{\alpha_1} \right) - \eta g(x) \right] + \epsilon_n$$

     using an inner loop at each iteration to solve for η as a mapping of $(\alpha_2, \alpha_2, \mu_s)$ given $(\widehat{\psi}, \widehat{\mu}_w, \widehat{\sigma})$.
  4. Correct the standard errors for $(\mu_s, \gamma, \alpha_1, \alpha_2)$ in the third step induced by $(\widehat{\mu}_w, \widehat{\sigma})$ obtained from the second step.

A Fully Parametric Specification

Estimating the importance of moral hazard (Tables 1 and 8, Margiotta and Miller 2000)

- We used the Masson-Antle-Smith (MAS) data set (37 firms in aerospace, electronics, chemicals from 1944 - 1977).
- The annual cost of moral hazard pales in comparison to losses shareholders would make if managers were paid a fixed wage.

		Aerospace	Chemicals	Electronics
After-tax compensation	All	126,822	86,094	76,958
		(410,590)	(493,072)	(428,977)
	CEO	144,731	120,618	96,688
		(475,720)	(602,724)	(610,090)
	Non-CEO	118,211	69,283	64,522
		(375,687)	(428,969)	(257,306)
Pretax salary and bonus	All	136,408	121,786	82,223
		(61,319)	(63,111)	(34,787)
	CEO	175,965	154,324	106,522
		(58,025)	(70,338)	(38,316)
	Non-CEO	117,386	105,943	66,907
		(53,305)	(52,440)	(21,073)
After-tax value of options granted	All	16,821	11,759	16,947
		(58,726)	(46,206)	(65,051)
	CEO	19,463	14,525	19,721
		(61,331)	(55,171)	(77,674)
	Non-CEO	15,551	10,412	15,200
		(57,474)	(41,100)	(55,780)
Return on stock held	All	8,797	−5,601	(2,872)
		(294,955)	(477,451)	(384,681)
	CEO	6,790	5,835	(5,681)
		(306,196)	(531,743)	(564,701)
	Non-CEO	9,763	−11,169	(1,102)
		(289,817)	(400,172)	203,096
Pretax value of stock bonus	All	0	273	518
		0	(1,852)	(4,821)
	CEO	0	239	0
		0	(1,496)	0
	Non-CEO	0	290	845
		0	(2,003)	(6,139)
Return on options held	All	17,530	1,444	3,507
		(163,653)	(114,107)	(92,296)
	CEO	18,002	5,523	1,386
		(218,544)	(156,450)	(110,359)
	Non-CEO	17,303	−548	4,845
		(129,605)	(86,293)	79,080

Cross-Sectional Information on Executive Compensation and Cost of Moral Hazard in 1967 $

Measure	Industry	Executive	Cost
Δ_1	Aerospace	CEO	186,689
		Non-CEO	2,370
	Chemicals	CEO	232,966
		Non-CEO	2,680
	Electronics	CEO	173,643
		Non-CEO	2,327
Δ_2		CEO	259,181
		Non-CEO	3,272
Δ_3	Aerospace		263,283,500
	Chemicals		85,355,000
	Electronics		104,222,000

50 Years of Managerial Compensation

Changes in managerial compensation (Table 3, Gayle and Miller, 2009)

- We compare MAS data with data from:
  - S&P 500 COMPUSTAT CRSP (2,610 firms 1995 -2004, 2000 $US)
  - A subset formed from those firms in the three MAS sectors.

Rank	Sector	Old	New restricted	New all
All	All	528	4,121	2,319
		(1,243)	(19,283)	(12,121)
CEO	All	729	6,109	5,320
		(1,472)	(24,250)	(19,369)
Non-CEO	All	400	2,256	1,562
		(1,026)	(12,729)	(9,303)
All	Aerospace	744	6,407	
		(1,140)	(20,689)	
CEO	Aerospace	950	11,664	
		(1,292)	(19,416)	
Non-CEO	Aerospace	624	1,997	
		(695)	(18,563)	
All	Chemicals	543	2,802	
		(1,348)	(9,555)	
CEO	Chemicals	718	3,673	
		(1,527)	(7,072)	
Non-CEO	Chemicals	401	477	
		(241)	(23,390)	
All	Electronics	370	4,501	
		(1,057)	(22,118)	
CEO	Electronics	457	5,325	
		(1,407)	(24,576)	
Non-CEO	Electronics	108	1,635	
		(61)	(18,810)	

50 Years of Managerial Compensation

Changes in components of managerial compensation (Table 4, Gayle and Miller, 2009)

Variable	Rank	Old	New restricted	New all
Salary and bonus	All	219 (114)	838 (1,066)	667 (905)
	CEO	261 (115)	1,037 (1,365)	1,127 (1,282)
	Non-CEO	179 (97)	640 (576)	552 (738)
Value of options granted	All	79 (338)	2,401 (13,225)	903 (3,753)
	CEO	111 (439)	3,402 (18,172)	1,782 (7,169)
	Non-CEO	51 (198)	1,401 (4,237)	681 (2,106)
Value of restricted stock granted	All	11 (95)	187 (1,633)	152 (936)
	CEO	8 (72)	242 (2,021)	298 (1,464)
	Non-CEO	13 (112)	133 (1,118)	115 (743)
Change in wealth from options held	All	5 (134)	785 (14,636)	281 (8,710)
	CEO	7 (167)	1,667 (17,078)	1,474 (13,567)
	Non-CEO	3 (94)	−76 (11,706)	−18 (6,939)
Change in wealth from stock held	All	−3 (439)	−40 (5,681)	125 (4,350)
	CEO	0.434 (479)	−14 (6,712)	264 (6,791)
	Non-CEO	−7 (398)	−64 (4,496)	90 (3,473)

50 Years of Managerial Compensation

Changes in sample composition of firms (Table 2, Gayle and Miller, 2009)

Variable	Sector	Old	New restricted	New all
Sales	All	1,243 (2,250)	3,028 (6,830)	4,168 (109,000)
	Aerospace	1,886 (3,236)	11,500 (14,900)	
	Chemicals	1,246 (2,018)	2,252 (2,091)	
	Electronics	319 (536)	2,469 (6,223)	
Value of equity	All	589 (1,034)	1,273 (2,863)	1,868 (4,648)
	Aerospace	391 (680)	3,132 (3,826)	
	Chemicals	677 (1,107)	800 (869)	
	Electronics	159 (365)	1,283 (3,096)	
Number of firms	All	37	151	1,517
	Aerospace	5	11	
	Chemicals	25	40	
	Electronics	7	100	
Number of employees	All	27,370 (28,850)	12,208 (26,676)	18,341 (46,960)
	Aerospace	49,920 (34,335)	58,139 (69,452)	
	Chemicals	23,537 (25,268)	8,351 (9,323)	
	Electronics	10,485 (7,664)	9,195 (18,266)	
Total assets	All	525 (924)	3,035 (6,550)	9,926 (40,300)
	Aerospace	726 (130)	10,600 (12,900)	
	Chemicals	548 (851)	2,385 (2,380)	
	Electronics	146 (233)	2,551 (6,311)	
Observations	All	1,797	3,260	82,578
	Aerospace	355	233	
	Chemicals	1,092	935	
	Electronics	252	2,092	

50 Years of Managerial Compensation

What were the driving forces behind these changes?

- If managers in the COMPUSTAT population ran firms the same size as managers in MAS, their compensation would have increased by a factor of 2.3, the increase in national income per capita.
- After adjusting for the general increase in living standards over these years, the model attributes:
  - Hardly any of the increased managerial compensation to changes in $\gamma^{-1} \ln \alpha_2 / \alpha_0$, or the certainty equivalent wage
  - practically all the increase to changes the risk premium τ_1
- The factors driving the change in τ_1 were:
  - not risk preferences: managers in the MAS (COMPUSTAT) population were willing to \$240,670 (\$248,620) to avoid a gamble of winning or losing \$1 million.
  - not $\Delta f(x)$: the biggest $\Delta\tau_1$ in aerospace where the abnormal returns became less dispersed, which reduces the risk premium
  - the sharp increase in α_2 / α_1 mainly due to increased firm assets, which provides managers with more opportunities to shirk.

Identification

To what extent are these results an artifact of the functional forms imposed on the data?

- The model is defined by:
  - $f(x)$ the probability density function of x from working
  - $g(x)$ the likelihood ratio for shirking versus working
  - α_2 distaste for working relative to outside option
  - α_1 distaste for shirking relative to outside option
  - γ risk-aversion parameter.
- The panel data set is $\{x_{nt}, w_{nt}\}_{n=1,t=1}^{N,T}$ where $w(x) = E[w_{nt} | x_{nt}]$.
- Thus $f(x)$ and $w(x)$ are identified.
- This leaves only $g(x)$ plus $(\alpha_1, \alpha_2, \gamma_t)$ to identify.
- Our parameter estimates and welfare measures assume $f(x)$ and $f(x)g(x)$ are truncated normal?
- Are these reasonable assumptions?

Identification

A dynamic extension to the static model

- Each period t:
  - the agent chooses his consumption c_t.
  - the principal announces the compensation function $w_t(x_{t+1})$.
  - the agent chooses $l_{tj} \in \{0, 1\}$ for $j \in \{0, 1, 2\}$.
  - Output x_{t+1} occurs and he is paid
- For some $\beta \in (0, 1)$ his lifetime utility is:

$$-\sum_{t=0}^{\infty} \sum_{j=0}^{2} \beta^t \alpha_j l_{tj} \exp(-\gamma c_t)$$

  - $l_{t0} = 1$ means resigning, $l_{t2} = 1$ working and $l_{t1} = 1$ shirking
  - α_j measures the distaste from choosing $j \in \{0, 1, 2\}$
  - γ is the constant absolute level of risk aversion
  - $f(x)$ and $g(x)$ characterize production under working and shirking.
- In addition we assume:
  - $g(x) \to 0$ as $x \to \infty$. (Only work produces extraordinary outcomes.)
  - complete markets exist for all publicly disclosed events.
  - b_{t+1} is known at period t where b_t denote the bond price.

Lemma

The optimal long-term contract can be implemented by replicating optimal short-term contracts. If the agent, offered a contract of $w_t(x)$ *will retire in period* t *or* $t+1$*, he optimally chooses* (l_{t0}, l_{t1}, l_{t2}) *to minimize:*

$$l_{t0} + (\alpha_1 l_{t1} + \alpha_2 l_{t2})^{1/(b_t - 1)} E_t \left[\exp\left(-\frac{\gamma_t w_t(x)}{b_{t+1}} \right) [g_t(x) l_{t1} + l_{t2}] \right].$$

- This result is somewhat contentious:
  - There is no role for granting versus vesting options and stock, an institutional feature. (CEOs occasionally get punished.)
  - The median (average) tenure of a CEO is about 5 (7) years.
  - Complete markets is often questioned but . . .
    - The evidence against is spotty (Altug and Miller,1990)
    - Managers save and are financially savvy
    - Introducing an unmotivated inefficiency is problematic.

- Similar to the static model define:

$$v_t(x) \equiv \exp\left(-\gamma w_t(x) / b_{t+1}\right). \qquad (1)$$

- The participation and incentive-compatibility constraints also follow their static model analogues:

$$\alpha_2^{-1/(b_t-1)} \geq E\left[v_t(x)\right]$$
$$0 \geq E\left[\left(g(x) - (\alpha_2/\alpha_1)^{1/(b_t-1)}\right) v_t(x)\right].$$

- Minimizing expected compensation amounts to choosing v for each x to maximize:

$$\int \ln\left[v_t(x)\right] f(x) \mathrm{d}x$$

subject to the two constraints above.

- Note $\ln v$ is concave increasing in v, the constraints are linear, so the Kuhn Tucker theorem applies.

Identification

What if the risk parameter is known?

- The FOC for the Lagrangian is:

$$\begin{aligned} v_t(x)^{-1} &= \alpha_2 \left[1 + \theta_t \left(\alpha_2/\alpha_1\right) - \theta_t g(x)\right] \\ &= \overline{v}_t^{-1} - \alpha_2 \theta_t g(x) \end{aligned}$$

where θ_t is the ratio of the two Lagrange multipliers and:

$$\lim_{x\to\infty} \left[g(x)\right] = 0 \Rightarrow \lim_{x\to\infty} \left[v_t(x)^{-1}\right] = \alpha_2 \left[1 + \theta_t \left(\alpha_2/\alpha_1\right)\right] \equiv \overline{v}_t^{-1}$$

- This implies:

$$g(x) = \frac{\overline{v}_t^{-1} - v_t(x)^{-1}}{\alpha_2 \theta_t} = \frac{\overline{v}_t^{-1} - v_t(x)^{-1}}{\overline{v}_t^{-1} - E\left[v_t(x)^{-1}\right]}$$

because:

$$E\left[v_t(x)^{-1}\right] = \alpha_2 \left[1 + \theta_t \left(\alpha_2/\alpha_1\right) - \theta_t\right] \Rightarrow \overline{v}_t^{-1} - \alpha_2 \theta_t$$

- Also since both constraints bind:

$$\alpha_2^{-1/(b_t-1)} = E\left[v_t(x)\right]$$

$$0 = E\left[\left(g(x) - (\alpha_2/\alpha_1)^{1/(b_t-1)}\right) v_t(x)\right]$$

- Summarizing and making the dependence on γ explicit:

$$\overline{v}_t \equiv \lim_{x\to\infty}\left[\exp\left(-\gamma w_t(x)\,/\,b_{t+1}\right)\right] \equiv \exp\left(-\gamma \overline{w}_t(x)\,/\,b_{t+1}\right)$$

$$\alpha_2(\gamma) = E\left[v_t(x,\gamma)\right]^{1-b_t} \tag{2}$$

$$\alpha_1(\gamma) = \alpha_2(\gamma)\left\{\frac{\overline{v}_t(\gamma)^{-1} - E\left[v_t(x,\gamma)^{-1}\right]}{\overline{v}_t(\gamma)^{-1} - E\left[v_t(x,\gamma)\right]^{-1}}\right\}^{b_t-1}$$

$$g(x,\gamma) = \frac{\overline{v}_t(\gamma)^{-1} - v_t(x,\gamma)^{-1}}{\overline{v}_t(\gamma)^{-1} - E\left[v_t(x,\gamma)^{-1}\right]}$$

Identification

Observational equivalence

- Denote by γ^* the risk aversion parameter generating the data.
- The diagram below plots $\psi_t(\gamma) \equiv E[v_t(x,\gamma)]$ using the facts that:

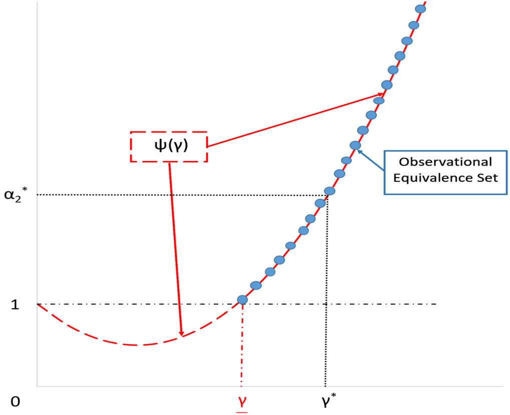
$$v_t(x,0) = 1, \qquad \frac{\partial}{\partial\gamma}\left[v_t(x,\gamma)\right]_{\gamma=0} = \frac{-w(x)}{b_{t+1}},$$
$$\frac{\partial^2}{\partial\gamma^2} v_t(x,\gamma) = \left(\frac{w(x)}{b_{t+1}}\right)^2 v_t(x,\gamma) > 0.$$

- We also assume $E[w_t(x)] \geq 0$, typically the case.
- This implies $\alpha_2^* \equiv \alpha_2(\gamma^*) > 1$.
- When the bond price is constant throughout the sample, observers cannot distinguish between agents with:
  1. A high risk tolerance and unpleasant working conditions
  2. Lower risk tolerance but more nonpecuniary benefits.
- Therefore α_2, α_1, and $g(x)$ are partially identified, or indexed by γ.

Identification

Illustrating the identified set with a constant bond prices

- Nevertheless the testable restriction $E\left[w_t(x)\right] \geq 0$ helps identify γ^*.
- It implies $\gamma^* \geq \underline{\gamma}$.
- The set of risk aversion parameters sharing the same data generating process are called *observationally equivalent.*



Identification

Do bond prices help to identify the risk aversion parameter?

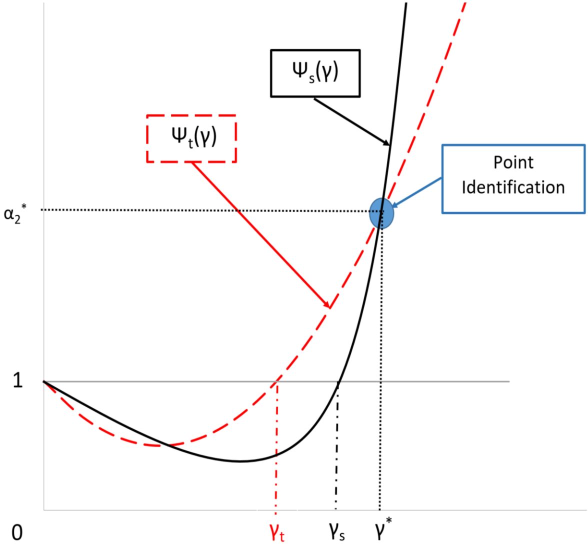
- Sample variation in the bond prices further shrinks the observationally equivalent set.
- Suppose that $b_s \neq b_t$ for periods s and t, and define:

$$q_1(\gamma) \equiv (1 - b_s) \log E\left[v_s(x, \gamma)\right] - (1 - b_t) \log E\left[v_t(x, \gamma)\right] \quad (3)$$

- Then $q_1(\gamma^*) = 0$ because $\log\left[\alpha_2(\gamma^*)\right] = (1 - b_t) \log E\left[v_t(x, \gamma^*)\right]$.
- How useful is this (nonlinear) equality in pinning down γ^*?
- Since this equation is nonlinear in γ, a unique solution is not guaranteed.

Identification

Illustrating how different bond prices help identify the risk aversion parameter



Identification

Does the profit maximization condition help to identify the risk aversion parameter?

- The profits from inducing the agent to work are $x - w^{o}(x)$.
- The profits from employing him to shirk are $xg(x) - \gamma^{-1}\log(\alpha_1)$.
- The principal incentivizes work if and only if:

$$E[x] - E[w^{o}(x)] - E[xg(x)] + \log(\alpha_1) \geq 0$$

- Substituting $\alpha_1(\gamma)$ for α_1 and $g(x,\gamma)$ for $g(x)$ into the LHS define:

$$\begin{aligned} q_2(\gamma) \equiv\ & cov\left(x, e^{\gamma w^{o}(x)}\right) \Big/ \left\{ e^{\gamma \overline{w}} - E\left[e^{\gamma w^{o}(x)}\right] \right\} \\ & + \gamma^{-1} \ln \left\{ \frac{1 - E\left[e^{\gamma w^{o}(x) - \gamma \overline{w}}\right]}{E\left[e^{-\gamma w^{o}(x)}\right] - e^{-\gamma \overline{w}}} \right\} - E[w^{o}(x)] \end{aligned} \qquad (4)$$

- Compensation for shirking does not depend on x (testable).
- Therefore if compensation depends on x then $q_2(\gamma^{*}) \geq 0$.

Identification
Sharp and tight bounds

- What is the largest observationally equivalent set that includes γ^*?
- Let $Q(\gamma)$ be a quadratic form summing elements like:
  - $[q_1(\gamma)]^2$ in the case of equalities
  - $[\max\{0, -q_2(\gamma)\}]^2$ in the case of positive inequalities
- Let $\Gamma \equiv \{\gamma : Q(\gamma) = 0\}$ be the set of risk aversion parameters satisfying the inequalities and equalities we have derived.
- Γ is sharp if every element in it is observationally equivalent to γ^*.
- Γ is tight if it contains every observationally equivalent element γ^*.
- Characterizing identification means defining a tight, sharp set.
- The equalities and inequalities described above are sharp and tight for this model. (See Gayle and Miller, 2015.)