Identifying the Contract Optimality without Knowing Functions: Moral Hazard and Statistic Inference

Rongzhu Ke
rogerke@mit.edu

Department of Economics, MIT

June 21st, 2008
Outline

- Motivations
- Modeling Environments
- Test Procedures
- An Empirical Example: Piece Rate Contract
- Model Extensions and Discussions
- Conclusions and Implications
Motivations

- To know how well people do in contract design;
- To put moral hazard theory to work;
- To provide a theoretical foundation for empirical contract analyses.
Motivations (cont...)

Suppose an economist observes a sequence of data \( \{(x_i, w_i)\}_{i=1}^{n} \) (and possibly only this data), where \( x_i \) and \( w_i \) are the output and wage payment of observation \( i \), respectively. What can he tell about the underlying contract \( w_i = s(x_i) \) for the agent without completely knowing the agent’s utility function, cost of efforts, and production function, especially with the existence of moral hazard?
Research Questions

Q1: does there exist any function tuple \( p = (u, c, f) \in (U \times C \times F) \) to rationalize the observed data?

Q2: If no, can we bound the loss of profit due to sub-optimality?

Q3a: If yes, is it unique (and in what sense) and how to find it based on the data?

Q3b: If yes, can we identify the parameters?

Q4: How do we identify the mechanism of incentive versus those of other mechanisms?
Related Topics

- GARP: (Varian, 1984; Blundell, 2003, 2007);
  — Not address contract related issues.

- Econometrics in Adverse Selection: (Wilson, 1993; Athey and Haile, 2005; d’Haultfoeuille & Fevrier, 2007);
  — Not address moral hazard issues.

- Review Essay: (Chiappori and Salanie, 2003)
  — Not address the contract optimality itself.
Modeling Environments

- The principal’s problem is:

\[
\max_{\{a, s(x)\}} \int v(x - s(x)) f(x, a) \, dx
\]

s.t.

\[
\int [u(s(x)) - c(a)] f(x, a) \, dx \geq U \quad \text{(IR)}
\]

\[
\int [u(s(x)) - c(a)] f(x, a) \, dx \geq \int [u(s(x)) - c(a')] f(x, a') \, dx, \forall a, a' \in \mathcal{A} \quad \text{(IC)}
\]

- \(v(.)\): principal’s utility; \(u(.)\): agent’s monetary utility;
- \(X\): output, \(f(x, a)\): p.d.f. of output, given the agent’s unobservable effort \(a > 0\);
- \(s(x) \in S\): a take-it-or-leave-it contract offered to the agent;
- \(c(a)\): agent’s disutility of effort, separable from the monetary utility.
Modeling Environments: Regularity Conditions

- A1: Agent is risk-averse and the principal is risk neutral.
- A2: disutility of effort is increasing and weakly convex in effort;
- A3: expected output is increasing and weakly concave in effort;
- A4: the score is well-defined everywhere, namely \( \frac{\partial}{\partial a} \log f(x, a) > -\infty \) for any \( x \in \mathcal{X} \);
- A4': or the payment is uniformly bounded from below, namely, \( s(x) \geq s > -\infty \) for any \( x \in \mathcal{X} \).
Modeling Environments: Solutions

Assuming the first order approach (FOA) is valid, then:

$$\frac{v'(x - s(x))}{u'(s(x))} = \lambda + \mu \frac{f_a(x, a)}{f(x, a)} \quad \text{(solves } s(x) \text{)} \quad (1)$$

$$\int u(s(x)) f_a(x, a) dx - c'(a) = 0 \quad (2)$$

$$\int v(x - s(x)) f_a(x, a) dx + \mu \left( \int u(s(x)) f_{aa}(x, a) dx - c''(a) \right) = 0 \quad (3)$$

$$\int [u(s(x)) - c(a)] f(x, a) dx \geq U \quad \text{(IR)}$$

and the second order condition (SOC):

$$\int u(s(x)) f_{aa}(x, a) dx - c''(a) < 0 \quad (4)$$
Test Procedures—Benchmark Case

- $u(.)$ and $f(x,a)$ is known, $c(a)$ is unknown.
- Denote $\frac{1}{u'(s)} = h(s)$ and $l_a(x,a) = \frac{\partial}{\partial a} \log f(x,a) = \frac{f_a(x,a)}{f(x,a)}$. 
Test Procedures—Benchmark Case (cont...)

**Theorem 1:** Assume data set is i.i.d., then \( \frac{1}{n} \sum h(w_i) \) is the efficient estimate of \( \lambda \), i.e.

\[
\rho \equiv \frac{\text{Cov}(h(w_i), l_a(x_i,a))}{\sqrt{\text{Var}(h(w_i)) \text{Var}(l_a(x_i,a))}} = 1,
\]

if and only if the first order condition (1) holds (with probability 1).
Test Procedures—Benchmark Case (cont...)

- Step 1: Estimate effort parameter in $f(x, a)$;
  $$\hat{a} = \arg \max_a \frac{1}{n} \sum_{i=1}^{n} l(x_i, a)$$

- Step 2: Examine the significance of moral hazard and the validity of FOA;
  To justify $\mu > 0$, it suffices to test:
  $$H_0 : \int u(w)f_a(x, a) dx > 0$$
  To justify SOC, it suffices to test
  $$H_0 : -\int u(w)f_{aa}(x, a) dx > 0$$
Test Procedures—Benchmark Case (cont...)

Step 3: Test the first order condition (1);
Let

\[ \hat{Q} = \frac{1}{n-1} \sum_{i=1}^{n} [h(w_i) - \bar{h}]^2 = \frac{1}{n-1} \sum_{i=1}^{n} [h(w_i) - \hat{E} h(w_i)]^2, \tag{6} \]

\[ \hat{J} = \hat{E} h(w_i) l_a(x_i, \hat{a}) \tag{7} \]

be the sample analogue of \( \text{Var}(h(w_i)) \) and \( \text{Cov}(h(w_i), l_a(x_i, a)) \) respectively, and let

\[ \hat{\rho} = \frac{\hat{J}}{\sqrt{\hat{Z}} \hat{Q}} \tag{8} \]

be the estimator of correlation coefficient between \( h(w_i) \) and \( l_a(x_i, a) \).
**Theorem 2**: To test the first order condition (1), is to test $\hat{\rho} \to^p 1$, where

$$\hat{\rho} = \frac{j}{\sqrt{\hat{Z} \hat{Q}}}$$

has the following asymptotic distribution:

(i) if $\rho < 1$,

$$\sqrt{n}(\hat{\rho} - \rho) \to^d \mathcal{N} [0, n\text{AVar}(\hat{\rho})].$$

(For convenience, in the final expression of asymptotic objects, we suppress the arguments $w$ or $x$ if there is no confusion).

(ii) If $\rho = 1$, under the null hypothesis:

$$\frac{n\hat{Q}(1 - \hat{\rho})^2}{2\hat{\lambda}^2} \to^d \chi_1^2.$$

We can either use sample analogue or bootstrapping to compute AVar.
Step 4: Estimate the loss of profit.

\textbf{Definition}

A contract \( w = s(x) \) is called the conditional constrained optimum (CCO), if and only if there is no other incentive compatible contract \( \tilde{s}(x) \neq s(x) \) with positive probability such that (i)

\[
\int v(x - s(x)) f(x, a^*) dx < \int v(x - \tilde{s}(x)) f(x, a^*) dx \text{ given the same effort level } a^* \text{ being implemented},
\]

and (ii) the agent has the same utility

\[
\int u(s(x)) f(x, a^*) dx - c(a^*) = \int u(\tilde{s}(x)) f(x, a^*) dx - c(a^*).
\]
Compared with the CCO, (i) the profit loss of the observed contract is estimated by

$$\Delta \Pi = \Pi(\hat{s}^*(x)) - \Pi(s(x)) = \hat{E}w_i - \hat{E}h^{-1}[\hat{\lambda}^* + \hat{\mu}^*l_{\hat{a}^*}(x_i, \hat{a}^*)]$$

where $\hat{a}^* = \hat{a}$ is the consistent estimator of the effort, $(\hat{\lambda}^*, \hat{\mu}^*, \hat{s}^*(x_i))$ solves the following equations

$$h^{-1}[\hat{\lambda}^* + \hat{\mu}^*l_{\hat{a}^*}(x_i, \hat{a}^*)] = \hat{s}^*(x_i) \text{ for almost every } x_i$$

$$\int u(\hat{s}^*(x_i))l_{\hat{a}}(x_i, \hat{a})f(x, \hat{a})dx = \hat{E}u(w_i)l_{\hat{a}}(x_i, \hat{a})$$

$$\int u(\hat{s}^*(x))f(x, \hat{a})dx = \hat{E}u(w_i);$$
Test Procedures—Benchmark Case (A Simulated Example)

- Knowing $u(w) = 2\sqrt{w}$, $f(x, a) = \frac{1}{a}e^{-\frac{x}{a}}$ and observing 100 observation.
- Contract A is the optimal one, while contract B is piece-rate (sub-optimal).

(Insert table 1 here)
Table 1: A simulated example

<table>
<thead>
<tr>
<th></th>
<th>Contract A</th>
<th></th>
<th>Contract B</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>True</td>
<td>Estimate</td>
<td>z-value</td>
<td>True</td>
</tr>
<tr>
<td></td>
<td>(S.D.)</td>
<td>(S.D.)</td>
<td></td>
<td>(S.D.)</td>
</tr>
<tr>
<td>$h$</td>
<td>10</td>
<td>10.0330</td>
<td>10.1196</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>(8.1631)</td>
<td>(7.8098)</td>
<td></td>
<td>(0.8163)</td>
</tr>
<tr>
<td>$\lambda \rightarrow \mathbb{E} h$</td>
<td>1</td>
<td>2.0053</td>
<td>1.9994</td>
<td>1.9329</td>
</tr>
<tr>
<td></td>
<td>(0.8163)</td>
<td>(0.9172)</td>
<td></td>
<td>(0.9172)</td>
</tr>
<tr>
<td>$\mu \rightarrow \frac{\mathrm{Cov}(h, y)}{2d}$</td>
<td>10</td>
<td>10.1062</td>
<td>0.9971</td>
<td>9.3075</td>
</tr>
<tr>
<td></td>
<td>(0.0011)</td>
<td>(0.0011)</td>
<td></td>
<td>(0.0011)</td>
</tr>
<tr>
<td>$\vartheta \rightarrow \mathbb{E} I^2$</td>
<td>0.01</td>
<td>0.6664</td>
<td>1.0923</td>
<td>0.8431</td>
</tr>
<tr>
<td></td>
<td>(0.0011)</td>
<td>(0.0011)</td>
<td></td>
<td>(0.0011)</td>
</tr>
<tr>
<td>$\rho - \frac{\bar{z}}{\bar{y}}$</td>
<td>1</td>
<td>0.9900</td>
<td>-0.2228</td>
<td>0.9565</td>
</tr>
<tr>
<td></td>
<td>(0.4488)</td>
<td>(0.3593)</td>
<td></td>
<td>(0.3593)</td>
</tr>
<tr>
<td>$\frac{\delta(x_1, x_2)}{2x}$</td>
<td>0</td>
<td>8.29×10^{-4}</td>
<td>0.0516</td>
<td>0.0250</td>
</tr>
<tr>
<td>$\Delta \Pi$ per capita</td>
<td>0</td>
<td>0.0131</td>
<td>0.4000</td>
<td>0.09</td>
</tr>
<tr>
<td>$A C$ per capita</td>
<td>1</td>
<td>0.6567</td>
<td>0.3275</td>
<td>0.7563</td>
</tr>
<tr>
<td>$T L$ per capita</td>
<td>0</td>
<td>0.6698</td>
<td>0.4000</td>
<td>0.09</td>
</tr>
<tr>
<td>$\bar{M}(\bar{d})$</td>
<td>0.2</td>
<td>0.1319</td>
<td>14.0854***</td>
<td>0.1999</td>
</tr>
<tr>
<td></td>
<td>(9.3643×10^{-2})</td>
<td>(0.1044)</td>
<td></td>
<td>(0.1044)</td>
</tr>
<tr>
<td>$\bar{K}(\bar{d})$</td>
<td>0</td>
<td>0.0168</td>
<td>11.5654***</td>
<td>0.0100</td>
</tr>
<tr>
<td></td>
<td>(1.4526×10^{-2})</td>
<td>(1.4526×10^{-2})</td>
<td></td>
<td>(1.4526×10^{-2})</td>
</tr>
<tr>
<td>$\bar{M}^*(\bar{d})$</td>
<td>0.2</td>
<td>0.1292</td>
<td>0.1999</td>
<td>0.1625</td>
</tr>
<tr>
<td>$\bar{K}^*(\bar{d})$</td>
<td>0</td>
<td>0.0302</td>
<td>0.0302</td>
<td>0.0279</td>
</tr>
<tr>
<td>profit per capita</td>
<td>0</td>
<td>5.3721</td>
<td>4.99</td>
<td>4.5507</td>
</tr>
<tr>
<td></td>
<td>(4.3602)</td>
<td>(3.9172)</td>
<td></td>
<td>(3.9172)</td>
</tr>
<tr>
<td>average payment</td>
<td>5</td>
<td>4.6809</td>
<td>5.01</td>
<td>4.5689</td>
</tr>
<tr>
<td></td>
<td>(3.9368)</td>
<td>(3.9368)</td>
<td></td>
<td>(3.9368)</td>
</tr>
<tr>
<td>#obs</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

Note: the number in ( ) is standard deviation or asymptotic standard deviation; the number in [ ] is z-value by bootstrapping; *** indicates significance at 99.9% level, ** indicates significance at 99% level; * indicates significance at 90% level.
Test Procedures II—*Functionally knowing $u(\cdot)$ & $f(\cdot, a)$*

- Test procedure for given agent;
- Test procedure for selection of agent;
- Test procedure for matching technology and contract
Test Procedures II—*Functionally knowing* $u(.)$ & $f(., a)$

**Given Agent**

- Step 1: estimate parameters in $f(x, a, \theta)$;
- Step 2: find the most favorable parameters in $u(s(x), \gamma)$ by minimizing some criterion function.
- Step 3: test the null hypothesis $H_0: \rho(\gamma_0) = 1$ and justify the SOC.
- Step 4: bound the loss of profit due to sub-optimality for given agent.
Test Procedures II—*Functionally knowing* $u(.)$ & $f(., a)$ (cont...)

Selection of agent:

- Let $V(\gamma)$ be the solution to problem (P1), then if the principal recognizes the profitability of selecting different agent, he should solve the further optimization problem:

$$P2: \max_{\gamma \in V} V(\gamma) \text{ s.t. IC and IR.}$$

- The solution to P2 means the principal "choosing the right agent to offer the right contract".

- The first order condition for saddle point is:

$$\frac{\partial u(s, \gamma)}{\partial \gamma} \int_{< s} (\lambda + \mu l_a(x, a^*)) f(x, a) dx + \int_{\geq s} \frac{\partial}{\partial \gamma} u(s^*, \gamma) f(x, a^*) dx = 0$$

(9)

- The profit loss is:

$$\hat{\Delta \Pi} = \frac{1}{n} \sum [w_i - \hat{s}^*(x_i)] + \frac{1}{n} \sum [\hat{s}^*(x_i) - \hat{s}^{**}(x_i)]$$

incentive error

selection error
Test Procedures II—*Functionally knowing* $u(.)$ & $f(., a)$ (cont...)

**Matching Technology and Contract:**

- Let $V(\theta)$ be the solution to problem P1 when there is parameter(s) in distribution function $f(x, ., .)$, the further optimization problem:

  $$P3 : \max_{\theta \in \hat{\theta}} V(\theta) \text{s.t. IC and IR}$$

- The solution to P3 means the principal "choosing the right technology to match the right contract".

Condition to test the $H_{0m}: L^*_\theta = 0$, where can be approximated by

$$\hat{L}^*_\theta = \hat{E}(x_i - s^*(x_i))l_\theta(x_i, \hat{\theta}) + \hat{\lambda}\hat{E}u(s^*(x_i), \hat{\gamma}_0)l_\theta(x_i, \hat{\theta})$$

$$+ \hat{\mu}\hat{E}u(s^*(x_i), \hat{\gamma}_0)[l_{\theta a}(x_i, \hat{\theta}) + l_a(x_i, \hat{\theta})]l_\theta(x_i, \hat{\theta}).$$
Test Procedures II—An Empirical Example

Summary Statistic

Table 2: Summary Statistics (Cotton Spinner in Jiutian Textile Mill, Zhejiang, China)

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Piece Rate</td>
<td>0.1867</td>
</tr>
<tr>
<td>Obs#</td>
<td>100</td>
</tr>
<tr>
<td>Max</td>
<td>302.4</td>
</tr>
<tr>
<td>Min</td>
<td>105.0</td>
</tr>
<tr>
<td>Median</td>
<td>194.9250</td>
</tr>
<tr>
<td>Mean</td>
<td>194.4075</td>
</tr>
<tr>
<td>S.D.</td>
<td>34.4156</td>
</tr>
<tr>
<td>Chi2 Stat</td>
<td>5.9063</td>
</tr>
<tr>
<td>p-value</td>
<td>0.2063</td>
</tr>
</tbody>
</table>
### Table 4. Analysis of Piece Rate (Cotton Spinner in Zhejiang, China)

<table>
<thead>
<tr>
<th>Criterion Function</th>
<th>$L(\lambda, \mu, \gamma)$</th>
<th>Stat. (S.D.)</th>
<th>Z-value</th>
<th>$V(\gamma^\star)$</th>
<th>Stat. (S.D.)</th>
<th>Z-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{a}$</td>
<td></td>
<td>$\hat{a}$</td>
<td>5.2542</td>
<td>(0.1798)</td>
<td>5.2542</td>
<td>(0.1798)</td>
</tr>
<tr>
<td>$\hat{Z} \rightarrow E\lambda^2(x_i)$</td>
<td>30.9185</td>
<td>30.9185</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$1 - \gamma_0$</td>
<td>0.8650</td>
<td>(4.2426e-002)</td>
<td>(N.E.)</td>
<td>5</td>
<td>(N.E.)</td>
<td></td>
</tr>
<tr>
<td>Trunc. $x_0$</td>
<td>22.3100</td>
<td>8.4868e-007</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{\lambda} \rightarrow E\lambda$</td>
<td>0.6038</td>
<td>1.2638e-007</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{\mu} \rightarrow \frac{\text{Cov}(h, \lambda)}{\text{E}^2(h)}$</td>
<td>11.6816</td>
<td>7.5675e-0015</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{\rho} \rightarrow \text{Var}(h)$</td>
<td>0.9649</td>
<td>0.0275</td>
<td>0.8079</td>
<td>(0.1458)</td>
<td>12.9007***</td>
<td></td>
</tr>
<tr>
<td>$\chi^2_\mu$</td>
<td>3.6572e-004***</td>
<td>(8.4352)</td>
<td></td>
<td></td>
<td>(1.1225)</td>
<td></td>
</tr>
<tr>
<td>$\Delta \Pi_{1b}$</td>
<td>0.0804</td>
<td>0.7160</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$100% \text{Profit}$</td>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta \Pi_{2b}$</td>
<td>33.7677</td>
<td>33.7677</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$100% \text{Profit}$</td>
<td>21.3600</td>
<td>21.3600</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{M}(\alpha)$</td>
<td>1.6142</td>
<td>1.7144***</td>
<td>9.6841e-007</td>
<td>24.0638***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{E}(\alpha)$</td>
<td>(9.4157)</td>
<td>(4.0243e-007)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{\lambda}^*$</td>
<td>12.5000</td>
<td>0.1617</td>
<td>7.1778e-006</td>
<td>9.7687**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{\mu}^*$</td>
<td>(5.3692e+002)</td>
<td>(7.3477e-006)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\gamma_{\min}$</td>
<td>22.3075</td>
<td>101.9600</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\gamma_{\max}$</td>
<td>0.5929</td>
<td>0.1041</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\bar{M}$</td>
<td>0.8652</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\gamma$</td>
<td>-4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta \bar{C}$</td>
<td>0.4811</td>
<td>0.0946</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$L_{\theta}$</td>
<td>-98.6413</td>
<td>337.0450</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>profit per capita</td>
<td>158.12601 (28.3169)</td>
<td>191.9460</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>average payment</td>
<td>36.2815 (61.111)</td>
<td>2.4453</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: *** indicates significance at 99.9% level; ** indicates significance at 95% level.
Test Procedures III
— Functionally knowing $u(.)$, partially knowing $f(x, a)$

- $u(.)$ is parameterized by some unknown parameters.
- $f(x, a)$ is semi-parametrically known, i.e.

$$f(x, a) = \frac{e^{\omega(x, a)} v(x)}{\int e^{\omega(x, a)} v(x) dx} \left( \int e^{\omega(x, a)} v(x) dx < 0 \right)$$

where real valued function $v(x) > 0$ is unknown, $\omega(x, a)$ but is parameterized by some unknown parameters.

- In this case, analyses will be similar.
Extentions and Discussions

- Non-parametric test procedures;
- Heterogeneous data-generating processes;
- Unobservable shocks.
Conclusions and Implications

- What we can do based on a single sequence of data:
  - Test the existence of moral hazard;
  - Test the optimality of contract;
  - Bound the profit loss comparing to the potential optimal contract (given agent and/or selected agent);
  - Suggest a more plausible contract based on the data;

- Implications: Labor contract, public finance, Empirical contract study, Corporate finance...
Comments and Questions?

Thank you!