Bargaining with Optimism: Medical Malpractice Lawsuits in Florida

Antonio Merlo and Xun Tang
Department of Economics
Rice University

March 2016

Introduction

- We study a structural bargaining model where both sides are optimistic about the disagreement outcome
- Empirical context:
  - patients and doctors bargain for settlement of medical malpractice lawsuits in Florida, U.S.
  - Florida law requires a mandatory settlement conference mediated by court officials and attended by attorneys
- Motivation
  - rationalize failure to reach settlement
  - infer beliefs from both sides
  - address policy questions (e.g., impact of tort reforms)

Preview of Results

- Identification of structural elements despite data deficiency
  - Florida malpractice cases offer a typical empirical environment
  - data reports transfer paid by defendant and settlement probability
  - timing of settlement not reported in data
- Estimate the model using Maximum Simulated Likelihood (MSL)
  - severity has significant impact on settlement and compensation at least at certain levels
  - some correlation between severity and jury decisions
  - patient and doctor beliefs partially captures such patterns
  - two sides' beliefs are negatively correlated

Related Literature

- Theory:
  - Merlo and Wilson (1994): Stochastic sequential bargaining
  - Yildiz (2004, 2011): Bargaining with uncommon prior
- Empirical
  - medical malpractice lawsuits: Sieg (2000); Watanabe (2009)
  - plea bargaining: de Silveira (2012)
- Models with unobserved heterogeneity
  - Hu (2008), Hu and Schennach (2008);
  - An, Hu and Shum (2010), Hu, McAdams and Shum (2013)

The Model

- Two sides: p (plaintiff) and d (defendant)
- Potential compensation C: "surplus" or "cake" to be divided.
- Dates in the legal process
  - T_0 : filing of the lawsuit with a county court;
  - T_s : settlement conference (mandatory by the state law of Florida);
  - T_c : court trial and jury decision (if necessary).
- Wait-time: $T \equiv T_c - T_s$.

The Model

- At T_s the defendant offers to settle by paying S
- The conference concludes
  - either with a settlement $(A = 1)$ and transfer S to the plaintiff;
  - or with no settlement $(A = 0)$ and a trial takes place at T_c.

The Model

- If a trial is needed (at T_c),
  - either the court rules in favor of plaintiff ($D = 1$) and the defendant pays C to the plaintiff
  - or the charges against the defendant are dropped ($D = 0$).
- Beliefs at the settlement conference:
  - $\mu_p \in [0, 1]$: plaintiff belief for $D = 1$;
  - $\mu_d \in [0, 1]$: defendant belief for $D = 0$;
  - optimism: $\mu_p + \mu_d > 1$.
- δ : time discount factor

Nash Equilibrium

- Plaintiff accepts an offer iff $S \geq \delta^T \mu_p C$.
- The defendant offers $S = \delta^T \mu_p C$.
- Settlement occurs iff

$$\delta^T \mu_p C \leq \delta^T (1 - \mu_d) C + \sum_{r=1}^{T} \delta^r K,$$

where K is the defendant's litigation cost per period.

The Data

- Data reports A (dummy for settlement) and
  - S (if $A = 1$); or
  - D (if $A = 0$) and C (if $A = 0$, $D = 1$).
- Recall $A = 1$ iff $YC \leq \phi(T)K$, where $Y \equiv \mu_p + \mu_d - 1$ and $\phi(t) \equiv \sum_{r=1}^{t} \delta^{r-t}$ is increasing in t.
- Case-level: county, severity, income, T_0
- Individual-level: age, gender, doctor information

Identification

- Suppose cases are homogenous
- Structural element: distribution of (μ_p, μ_d) and C
- Orthogonality conditions
  - T, D, K and (μ_p, μ_d, C) mutually independent
- Link between observed outcome and model elements:
  - (i) settlement occurs ($A = 0$) iff $YC \leq \phi(T)K$; and accepted offer S is $\delta^T \mu_p C$;
  - (ii) C reported when there is no settlement ($A = 0$) and plaintiff wins ($D = 1$).

Identification

- As we will show, the distribution of $(S, A, C) \mid T, K$ is identified if wait-time T can be conditioned on using data.

- However, there is no reliable measure of T in data
  - T_c not reported for cases settled outside the court
  - T_s not reported for cases ruled in the court.
  - reported measures of T_s and T_c are imprecise

- We treat T as unobserved heterogeneity and recover outcome distribution given T, K first.
- Assume: lawsuits filed with the same county in the same period (month) have the same T.
- This is plausible given how settlement conferences are scheduled in practice.

Step 1: Recover conditional outcome distribution

- A "cluster": a collection of cases filed at the same county court in the same month.
  - reasonable to maintain that these cases share the same unreported T.
  - beliefs, costs, and potential compensation are independent across cases
- i, j, l: cases filed in the same cluster.
- $\mathcal{E}_{i,l}$: the event "$A_i = 0, D_i = 1$ and $A_l = 1$".

- By the law of total probability,

$$f_{C_i,S_l}(c,s,A_j=1|\mathcal{E}_{i,l},\mathbf{k})$$
$$=\sum_t\left(\begin{array}{c} f_{C_i}(c|A_i=0,D_i=1,T=t,k_i)\times \\ \mathbb{E}(A_j|T=t,k_j)\times \\ f_{S_l}\left(s,T=t|\mathcal{E}_{i,l},k_i,k_l\right) \end{array}\right)$$

- We have used:
  - independence between cases within a cluster
  - orthogonality conditions maintained
  - definition of $\mathcal{E}_{i,l}$.

- First, we recover $\mathbb{E}\left(A \mid t,k\right)$ and $f_S\left(s \mid A=1,t,k\right)$ $\forall t,k$.

- Notation (fix a vector of $\mathbf{k} \equiv (k_i, k_j, k_l)$):
  - $\mathcal{B}_M, \mathcal{D}_M$: partition of support of S and C into M intervals (each denoted b_m, d_m);
  - L_{C_i,S_l}: M-by-M with entry being the prob $C_i \in d_m$ and $S_l \in b_{m'} \mid \mathcal{E}_{i,l}, \mathbf{k}$;
  - Λ_{C_i,S_l} : M-by-M with entry $f(C_i \in d_m, A_j = 1, S_l \in b_{m'} \mid \mathcal{E}_{i,l}, \mathbf{k})$;
  - $L_{C_i|T}$: M-by-$|\mathcal{T}|$ with entry $\Pr(C_i \in d_m|(1 - A_i)D_i = 1, t, k_i)$;
  - Δ_j : $|\mathcal{T}|$-by-$|\mathcal{T}|$ diagonal with entry $\mathbb{E}(A_j|k_j, t)$;
  - L_{T,S_l} : $|\mathcal{T}|$-by-M with entry $\Pr(T = t, S_l \in b_m|\mathcal{E}_{i,l}, k_i, k_l)$

- In matrix notation

$$\Lambda_{C_i,S_l} = L_{C_i|T}\Delta_j L_{T,S_l} \text{ and } L_{C_i,S_l} = L_{C_i|T} L_{T,S_l}.$$

- Assume for any $(\mu_p, \mu_d) \in \mathcal{M}$, C is continuously distributed over $\mathcal{C} \equiv (0, \overline{c})$; K is continuously distributed over $\mathcal{K} \equiv (0, \bar{k})$.

- Lemma. There exists $\mathcal{B}_{|\mathcal{T}|}$ and $\mathcal{D}_{|\mathcal{T}|}$ such that L_{C_i,S_l} has full-rank.

Intuition

- There is sufficient variation in the conditional distributions of C and S as T changes.
- Specifically, the conditional supports are nested.
- Under the cluster structure and maintained orthogonality conditions, these two sources of variation interact and induce substantial correlation between observed transfers even after T is integrated out.

- Proposition. Conditional outcome distribution are identified.
  - Intuition: $\Lambda_{C_i,S_l}(L_{C_i,S_l})^{-1} = L_{C_i|T}\Delta_j \left(L_{C_i|T}\right)^{-1}$.
  - Scale of $L_{C_i|T}$ and label of Δ_j are pinned down using equilibrium implication.
  - By symmetric argument, $L_{S_l|T}$ is identified.
- An important note: several key conditions that are necessary for eigen-value decomposition arise intrinsically from the model structure (given the exogeneity conditions)

- Next, recover the conditional density of S_i over full support
  - l_s: $|\mathcal{T}|$-vector with m-th coordinate being $f_{S_i}(s, C_l \in d_m | \mathcal{E}_{i,l}, k_i, k_l)$.
  - $l_s = (L_{T,C_i})' \lambda_s$, w/ λ_s a $|\mathcal{T}|$-vector with the t-th component $f_{S_i}(s | A_l = 1, t, k_l)$.
  - L_{T,C_i}: squared matrix $\Pr(T = t, C_i \in d_m | \mathcal{E}_{i,l}, k_i, k_l)$ that is identified as $(L_{S_l|T})^{-1} L_{S_l,C_i}$.
  - Thus λ_s is identified $\forall s$.

Step 2: Belief distribution

- Assume $(\mu_p, \mu_d) \perp C$. For simplicity, suppose $\phi(t)k < \bar{c}$.
- Define $\psi(k, t, c)$ as

$$
\begin{aligned}
\Pr(C_i &\leq c | A_i = 0, D_i = 1, T = t, K_i = k) \\
&\times \Pr(A_i = 0 | T = t, K_i = k)
\end{aligned}
$$

which is identified from previous steps.

- By our maintained independence,

$$
\begin{aligned}
\psi(k, t, c) = & \quad \Pr(Y_i C_i > \phi(t)k, C_i \leq c) \\
\Rightarrow \frac{\partial}{\partial c}\psi(k, t, c) = & \quad \Pr\left(Y_i > \phi(t)k/c\right) f_C(c)
\end{aligned}
$$

- Fix any $c_0 \in \mathcal{C}$. For any $c \in \mathcal{C}$, find a triple (t, k_0, k) with $k_0/k = c_0/c$ and $\phi(t)k < c$.

- By construction, for any $t \in \mathcal{T}$,

$$\frac{\partial\psi(k, t, c)/\partial c}{\partial\psi(k_0, t, c_0)/\partial c} = \frac{f_C(c)}{f_C(c_0)}.$$

- Thus the density of C is (over-)identified.

- For any $\alpha \in (0,1)$, there exist t_α, k_α and $c_\alpha \in \mathcal{C}$ such that $\phi(t_\alpha)k_\alpha/c_\alpha = \alpha$.

- Then

$$\begin{aligned} \partial\psi(k_\alpha, t_\alpha, c_\alpha)/\partial c = &\quad \Pr(Y_i > \phi(t_\alpha)k_\alpha/c_\alpha)f_C(c_\alpha) \\ \Rightarrow \Pr(Y_i > \alpha) = &\quad \frac{\partial\psi(k_\alpha, t_\alpha, c_\alpha)/\partial c}{f_C(c_\alpha)}. \end{aligned}$$

- This identifies the marginal distribution of Y.

- Define $\varphi(k, t, s)$ as

$$\Pr(S \leq s, A = 1 | k, t) = \Pr(\mu_p C \leq s/\delta^t, YC \leq \phi(t)k),$$

  which is identified for all t, k and $s \in \mathcal{S}$.

- Assume $\phi(|\mathcal{T}|)\bar{k} \geq \bar{c}$.
  - When total defense costs is high, settlement probability approaches one.
  - This condition is empirically supported by data.

- By logarithm transform, $\varphi(k, t, s)$ is

$$\Pr(V_1 + W \leq \log s - t \log \delta, V_2 + W \leq \log \phi(t) + \log k),$$

where $V_1 \equiv \log \mu_p$, $V_2 \equiv \log Y$ and $W \equiv \log C$.

- Joint distribution of V_1, V_2 is uniquely recovered.

Estimation: Distribution of C

- Panel structure:
  - each cluster (month-county pairs) is indexed by n
  - cases within a cluster are indexed by i.
- Define $Z_{n,i} \equiv S_{n,i}$ if $A_{n,i} = 1$; $Z_{n,i} \equiv C_{n,i}$ if $A_{n,i} = 0$ and $D_{n,i} = 1$; and $Z_{n,i} \equiv 0$ otherwise.
- $x_{n,i}, w_{n,i}$: case characteristics that affect C and (μ_p, μ_d) respectively
- First Step: estimate the distribution of C given $x_{n,i}$
  - $C \mid x$: truncated exponential with a rate $\lambda(x_{n,i}; \beta) \equiv \exp\{x_{n,i}\beta\}$
  - MLE: $\hat{\beta} \equiv \max_\beta \sum_{n,i} d_{n,i}(1 - a_{n,i}) \left[x_{n,i}\beta - \exp\{x_{n,i}\beta\}c_{n,i}\right]$

Estimation: Belief Distribution

- Parametrization: $\mu_p = 1 - \tilde{Y}$ & $\mu_d = Y + \tilde{Y}$, where
  - $(\tilde{Y}, Y, 1 - \tilde{Y} - Y)$: Dirichlet with parameters $\alpha_{j,n,i} \equiv \exp\{w_{n,i}\rho_j\}$ for $j = 1, 2, 3$
  - optimism: $Y = \mu_p + \mu_d - 1$
  - $\tilde{Y} \mid w_{n,i} : Beta(\alpha_{1,n,i}, \alpha_{2,n,i} + \alpha_{3,n,i})$
  - $Y \mid \tilde{Y} = \tau, w_{n,i} : (1 - \tau) Beta(\alpha_{2,n,i}, \alpha_{3,n,i})$
- The log-likelihood $L_N(\rho, \beta, \theta)$ equals:

$$\sum_{n=1}^{N} \ln \left[\sum_t h_n(t;\theta) \prod_i f_{n,i}(t;\rho,\beta) \right]$$

where $f_{n,i}(t;\rho,\beta)$ is the density of $Z_{n,i}, A_{n,i}, D_{n,i}$ given $T_n = t$, $w_{n,i}$; and $h_n(t;\theta)$ is the density of T_n.

MSL Estimation

- Simulated Maximum Likelihood Estimator:

$$(\hat{\rho}, \hat{\theta}) \equiv \arg\max_{\rho,\theta} \hat{L}_N(\rho, \theta, \hat{\beta}).$$

  where $\hat{L}_N(\rho, \theta, \beta)$ approximates $L_N(\rho, \theta, \beta)$ using simulated draws.

- $(\hat{\rho}, \hat{\theta})$: converge at root-n rate to zero-mean normal under standard conditions with $S \to \infty$ and $\sqrt{N}/S \to \infty$.

- Limiting covariance can be consistently estimated based on the analog principle using simulated observations.

Florida Medical Malpractice Lawsuits

- Data Source: Florida Department of Financial Service
  - same as that used in Sieg (2000) and Watanabe (2009)
  - 8,765 cases filed with 66 county courts between 1978-1998
  - either settled outside the court or resolved through court trials
- Data records:
  - patient age, gender; level of severity
  - doctor's board certification, education background
  - date of initial filing and "final date of disposition"
  - outcome from settlement conference or court trial, including transfer from the defendant

Table 1: Settlement Outcome (Unit: $1k)

Board	Severity	# obs	$\hat{p}_{A=1}$		$\hat{\mu}_{S\|A=1}$	
Yes	low	987	0.712	(0.014)	41.948	(2.160)
	med	1572	0.792	(0.010)	104.948	(3.814)
	high	1867	0.835	(0.009)	278.616	(8.763)
	death	1642	0.834	(0.009)	195.332	(6.680)
No	low	711	0.812	(0.015)	36.681	(2.705)
	med	679	0.851	(0.014)	82.668	(4.289)
	high	589	0.844	(0.015)	295.610	(23.80)
	death	718	0.872	(0.013)	183.646	(9.042)

Table 2: Mean Compensation Ruled by Court
(Unit: $1k; # of obs: 251)

severity	age<18	18<age<30	30<age<60	age>60
low	76.319	111.362	79.480	11.285
	(32.791)	(20.714)	(21.721)	(9.318)
med	178.760	237.077	274.842	283.982
	(48.561)	(41.528)	(163.311)	(123.511)
high	397.492	560.042	570.572	486.831
	(114.898)	(91.707)	(225.287)	(155.053)
death	359.431	340.535	268.804	656.517
	(75.706)	(63.721)	(67.235)	(176.139)

Table 3. Logit Estimates for Settlement (A)
(# of obs: 8,765)

	(1)	(2)	(3)
constant	1.012 (0.523)*	1.086 (0.530)**	0.2391 (1.852)
sev	0.253 (0.156)*	0.248 (0.157)	0.7654 (0.684)
age	0.004 (0.008)	-0.002 (0.009)	-2.204e-07 (0.009)
gender	-0.217 (0.056)***	-0.171 (0.057)***	-0.2152 (0.057)***
board	-0.309 (0.063)***	-0.325 (0.064)***	-0.3084 (0.064)***
graduate	0.162 (0.067)***	0.229 (0.068)***	0.1699 (0.068)***
income	0.001 (0.019)	4.684e-07 (0.019)	0.04617 (0.146)
costs	0.013 (0.027)	0.016 (0.028)	0.03391 (0.055)
age*inc	4.472e-05 (2.791e-04)	1.283e-05 (2.843e-04)	-5.153e-06 (2.872e-04)
age*costs	-5.915e-04 (3.246e-04)*	-5.506e-04 (3.293e-04)*	-2.818e-06 (3.306e-04)
sev*inc	-5.131e-04 (0.006)	-3.096e-06 (0.006)	-0.028 (0.054)
sev*costs	-5.244e-04 (0.008)	-0.007 (0.008)	0.004 (0.019)
age^2		9.024e-05 (5.443e-05)*	1.127e-05 (6.302e-05)
inc^2			-3.125e-04 (0.003)
$costs^2$			0.002 (0.003)
$sev*age^2$			1.123e-05 (1.546e-05)
$sev*inc^2$			2.832e-04 (0.001)
$sev*costs^2$			-5.683e-04 (0.001)
Log L	-4119.189	-4118.391	-4117.147
p-value	$<$0.001	$<$0.001	$<$0.001

Notes: p-val is for LRT test of joint significance; income and costs are in units of $1k.

Table 4. Logit Estimates for Court Decisions (D)
(# of obs: 1,161)

	(1)	(2)	(3)
constant	-3.815 (1.294)***	-3.815 (1.304)***	0.019 (2.089)
severity	1.089 (0.376)***	1.089 (0.376)***	0.921 (0.365)***
age	0.005 (0.020)	0.013 (0.022)	1.796e-05 (0.021)
gender	-0.428 (0.141)***	-0.428 (0.141)***	-0.430 (0.141)***
board	-0.195 (0.154)	-0.196 (0.154)	-0.182 (0.154)
graduate	-0.236 (0.175)	-0.236 (0.175)	-0.273 (0.177)
income	0.081 (0.046)*	0.074 (0.046)	-0.208 (0.146)
sev*inc	-0.037 (0.014)***	-0.037 (0.014)***	-0.031 (0.013)***
age*inc	-6.163e-05 (7.048e-04)	7.353e-05 (7.261e-04)	3.500e-04 (6.928e-04)
age^2		-1.498e-04 (1.390e-04)	-7.890e-04 (1.369e-04)
$income^2$			0.005 (0.003)*
log lkh.	-689.226	-688.621	-688.164
p-value	0.004	0.005	0.006

Notes: p-val is for LRT test of joint significance; income and costs are in units of $1k.

Table 5. Distribution of Total Compensation
(Unit: $1k; # of obs: 251)

	(1)	(2)
constant	-4.136 (1.217)***	-1.286 (8.247)
severity	-0.566 (0.073)***	-0.632 (1.378)
age	0.097 (0.090)	-0.961 (2.035)
gender	0.116 (0.161)	0.092 (0.167)
board	-0.620 (0.200)***	-0.638 (0.204)***
graduate	-0.116 (0.198)	-0.122 (0.225)
log income	-0.003 (0.338)	-0.775 (2.419)
sev*log income		0.022 (0.400)
age*log income		0.201 (0.594)
age^2		0.084 (0.097)
sev*age^2		-2.426e-04 (0.020)
Log likelihood	-1662.21	-1661.06
Pseudo-R^2	0.512	0.541
p-value for LRT	$<$0.001	$<$0.001

Notes: income are in units of $1k; Bootstrap standard errors in parentheses;
***: sig at 1%.

Table 6. Distribution of Total Compensation

	(1)	(2)
severity	197.705 (35.875)***	196.078 (36.536)***
age	-33.999 (31.489)	-87.530 (57.490)*
log income	0.965 (117.926)	12.564 (174.246)

Notes: income are in units of $1k; Bootstrap standard errors in parentheses; ***: sig at 1%; *: sig at 15%.

Table 7. Distribution of Total Compensation
(# of obs: 8,765)

Severity	Low		Med		High		Death	
	μ_p	μ_d	μ_p	μ_d	μ_p	μ_d	μ_p	μ_d
Mean	0.457	0.576	0.558	0.466	0.884	0.161	0.905	0.122
	(0.087)	(0.086)	(0.091)	(0.092)	(0.056)	(0.054)	(0.028)	(0.045)
Median	0.452	0.586	0.567	0.461	0.891	0.129	0.918	0.090
	(0.101)	(0.098)	(0.106)	(0.105)	(0.062)	(0.060)	(0.020)	(0.037)
Skewness	0.115	-0.215	-0.172	0.099	-0.462	0.442	-1.015	0.976
	(0.252)	(0.249)	(0.275)	(0.268)	(0.775)	(0.725)	(0.161)	(0.124)

Conclusion

- This paper rationalizes failure for settlement under optimism.
- Robust identification (which only requires some weak exogeneity conditions).
- MSL estimates provide evidence for optimism.
- Plaintiff and defendant beliefs negatively correlated.
- In progress: counterfactual tort reform

Linking parameters to belief distribution

	μ_p	μ_d
Marg. distribution	$Beta(\alpha_2+\alpha_3, \alpha_1)$	$Beta(\alpha_1+\alpha_2, \alpha_3)$
Mean	$\frac{\alpha_2+\alpha_3}{\alpha_0}$	$\frac{\alpha_1+\alpha_2}{\alpha_0}$
Variance	$\frac{\alpha_1(\alpha_2+\alpha_3)}{\alpha_0^2(\alpha_0+1)}$	$\frac{\alpha_3(\alpha_1+\alpha_2)}{\alpha_0^2(\alpha_0+1)}$
Skewness	$\frac{2(\alpha_1-\alpha_2-\alpha_3)\sqrt{\alpha_0+1}}{(\alpha_0+2)\sqrt{\alpha_1(\alpha_2+\alpha_3)}}$	$\frac{2(\alpha_3-\alpha_1-\alpha_2)\sqrt{\alpha_0+1}}{(\alpha_0+2)\sqrt{\alpha_3(\alpha_1+\alpha_2)}}$
Mode	$\frac{\alpha_2+\alpha_3-1}{\alpha_0-2}$	$\frac{\alpha_1+\alpha_2-1}{\alpha_0-2}$