Section 5: Recovering Risk Types and (Risk) Preferences

Jon Cohen

October 15, 2021

EFS Motivation

- Simple motivation: learn about preferences and types from choices and events
- Policy motivation: how bad is adverse selection and what should we do about it?

Big Picture Structure

1. Infer (distribution of) risk types from risk realizations
2. Infer (distribution of) risk preferences with the above + (distribution of) choices
3. Run counterfactuals using (joint distribution of) risk types + preferences

Big Picture Idea of Today's Recitation

- Review Cohen and Einav (2007) ECMA on car insurance choice
- Gain comfort with the idea of a model delivering

$$\text{choice}_i = f(\text{risk type}_i, \text{risk preference}_i)$$

- Gain comfort with behavioral and functional form assumptions to recover model parameters from (imperfect) data
- Point out tips for digesting structural papers as "structural signposts"

Picture of Big Picture Intuition

Notation: consumer i, choice j, prices p_j, indirect utilities μ_{ij}, choice regions A_{ij}

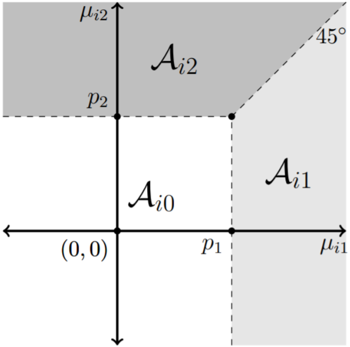


Figure 1: Choice regions for goods 0, 1, and 2.

Source: Berry and Haile (2021) WP

Outline

Setting: Car Insurance

- **Unobserved heterogeneity**: risk type and risk aversion
- **Realized risks**: accidents
- **Observed choices**: trading off premium (always pay) vs. deductible (pay only after accident) in menu of contracts

Unlike EFS, Cohen and Einav Have Price Variation!

$$d_{it} = min\{.5p_{it}, cap_t\}$$

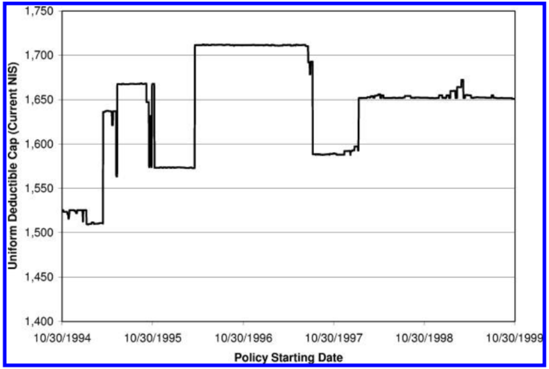


Figure 1. Variation in the Deductible Cap over Time

Firm says this was experimentation. Can avoid some assumptions w/ random variation...

Unlike EFS, Cohen and Einav Have Dynamics

- **Annuity guarantee**: Permanent decision at time of annuity purchase
- **Car insurance**: Temporary coverage as long as you continue paying premiums

Unlike EFS, Cohen and Einav Have Dynamics

- **Annuity guarantee**: Permanent decision at time of annuity purchase
- **Car insurance**: Temporary coverage as long as you continue paying premiums
  $\Rightarrow$ Consider "instantaneous" contract to isolate "static" demand
  1. **Tractable**: get simple closed form expressions for $\text{choice}_i = f(\text{risk type}_i, \text{risk preference}_i)$
  2. **Connected to research question** interested in risk preferences (not time preferences)
  3. **Realistic**: observe many cancellations in data

Outline

Model Overview

- Fully specify preferences to get $\text{choice}_i = f(\text{risk type}_i, \text{risk preference}_i)$
  - Identification problem: Choices driven by two unobserved dimensions
  - Identification solution: Make an assumption so one dimension is identified by something other than choice

Models Steps

1. Infer the distribution of risk types using observed risk realizations
   - What is the key assumption?
2. Given that, infer distribution of risk preferences from observed contract choices

Potential Problems

1. Moral hazard
2. Non-random attrition due to early cancellation
3. Unreported accidents
4. Two dimensions of unobserved risk: frequency and size

“Solutions"

(More or less) assume away!

1. Moral hazard

   $\rightarrow$ assume away!

2. Non-random attrition due to early cancellation

   $\rightarrow$ assume constant arrival rate and focus on per unit of time

3. Unreported accidents

   $\rightarrow$ assume threshold above which everything is reported

4. Two dimensions of unobserved risk: frequency and size

   $\rightarrow$ size is entirely idiosyncratic

Inferring Risk Type: Connection to EFS

- **Main idea**: infer riskiness of observable groups based on their risk realizations
- **Main assumption**: no moral hazard
- **Additional implementation challenges**: data censoring

Inferring Risk Type: Hypothetical with Perfect Data

- What type of data would we need to infer a **given person's risk type** under a given contract?

Inferring Risk Type: Hypothetical with Perfect Data

- What type of data would we need to infer a **given person's risk type** under a given contract?
  - In our sample, the number of accidents is a noisy measure of accident probability
  - Fundamental tension between needing a very long time series and allowing age effects
  - Impossible in EFS because we observe you dying only once

Inferring Risk Type: Hypothetical with Perfect Data

- What type of data would we need to infer a **given person's risk type** under a given contract?
  - In our sample, the number of accidents is a noisy measure of accident probability
  - Fundamental tension between needing a very long time series and allowing age effects
  - Impossible in EFS because we observe you dying only once
- What type of data would we need to infer whether a given person's **risk type changes** under different contracts?

Inferring Risk Type: Hypothetical with Perfect Data

- What type of data would we need to infer a **given person's risk type** under a given contract?
  - In our sample, the number of accidents is a noisy measure of accident probability
  - Fundamental tension between needing a very long time series and allowing age effects
  - Impossible in EFS because we observe you dying only once
- What type of data would we need to infer whether a given person's **risk type changes** under different contracts?
  - Need to additionally see people under multiple contracts
  - Ideally contracts would be randomly assigned

Inferring Risk Types: Structure with Actual Data

1. Model accidents as Poisson process w/ parameter λ
2. Parametrize based on observables: $ln(\lambda_i) = x_i'\beta + \varepsilon_i$, where $\varepsilon_i \sim N(0, \sigma_\lambda^2)$

Inferring Risk Types: Structure with Actual Data

1. Model accidents as Poisson process w/ parameter λ
   - λ_i captures i's "risk type"
   - Structural signpost # 1: Familiarize yourself with "go-to" distributions for contexts
2. Parametrize based on observables: $ln(\lambda_i) = x_i'\beta + \varepsilon_i$, where $\varepsilon_i \sim N(0, \sigma_\lambda^2)$
   - We don't get to observe λ_i, but we get its distribution from many people who look like i
   - Structural signpost # 2: $N()$ is computationally convenient and often a decent descriptor of the population characteristics
   - Structural signpost # 3: $\log()$ accommodates parameters with sign restrictions
   - Structural signpost # 4: Incorporate heterogeneity based on observables

Inferring Risk Type: Data

TABLE 2B—SUMMARY STATISTICS—CONTRACT CHOICES AND REALIZATIONS

Claims	Low	Regular	High	Very high	Total	Share
0	11,929 (0.193)	49,281 (0.796)	412 (0.007)	299 (0.005)	61,921 (1.00)	0.8034
1	3,124 (0.239)	9,867 (0.755)	47 (0.004)	35 (0.003)	13,073 (1.00)	0.1696
2	565 (0.308)	1,261 (0.688)	4 (0.002)	2 (0.001)	1,832 (1.00)	0.0238
3	71 (0.314)	154 (0.681)	1 (0.004)	0 (0.000)	226 (1.00)	0.0029
4	6 (0.353)	11 (0.647)	0 (0.000)	0 (0.000)	17 (1.00)	0.0002
5	1 (0.500)	1 (0.500)	0 (0.000)	0 (0.000)	2 (1.00)	0.00003

- Eyeball the positive correlation test
- Structural signpost # 5: Look for summary stats that drive the model (n.b. see Andrews, Gentzkow, Shapiro (2020) ECMA for a formal treatment)

Inferring Risk Type: Recap

- Assume no moral hazard so that realizations reveal type
- Observe only one realization $\rightarrow$ parametrize type based on observables
- Parameters: β, σ_λ

Inferring Risk Preferences: Connection to EFS

- **Main idea**: write down a model $\text{choice}_i = f(\text{risk type}_i, \text{risk preference}_i)$
- **Main assumptions**: choices are driven by inferred risk information and reveal underlying preferences
- **Additional implementation challenges**: discrete choices yield set identification rather than point identification

Inferring Risk Preferences: Cohen and Einav Overview

- Suppose we knew λ_i and choices are continuous

  $\Rightarrow$ can invert $\text{choice}_i = f(\text{risk type}_i, \text{risk preference}_i)$ to get exact risk preferences

- Discrete choices

  $\Rightarrow$ get bounds on risk preferences

- Observe λ_i's distribution rather than exact value

  $\Rightarrow$ get distribution of risk preferences

Inferring Risk Preferences: Mechanics

- Consider utility for coverage length t and ake $\lim t \to 0$ for “instantaneous contract"
- Derive indiff. condition btw contracts in terms of risk type/aversion

Inferring Risk Preferences: “Instantaneous Contract"

- Poisson:

$$P(k \text{ accidents over time } t) = \frac{(\lambda t)^k \exp(-\lambda t)}{k!}$$

  - $k > 1$ terms vanish as $t \to 0$

- EU from contract price p and deductible d over small interval t:

$$v(p,d) \approx \underbrace{(1-\lambda t)}_{P(\text{no accident})} u(w - pt) + \underbrace{(\lambda t)}_{P(\text{1 accident})} u(w - pt - d)$$

Inferring Risk Preferences: Indifference Condition

- Consider a high vs. low-deductible contract
  - *Notation check on deductibles*: $d^L < d^H \Rightarrow p^L > p^H$
- Indifference condition on contracts for the marginal type:

$$v(p^L, d^L) = v(p^H, d^H)$$

- Solve λ and take $t \to 0$ (see next slide for details):

$$\lambda = \frac{(p^L - p^H)u'(w)}{u(w - d^L) - u(w - d^H)}$$

- Rearrange to see MB vs. MC of low-deductible plan

Math Asides on Solving for λ

- Substitute for $v()$'s and divide through by t
- p^H, p^L disappear in terms with t only in $u(\cdot)$ since $\lim_{t\to 0} p^H t = \lim_{t\to 0} p^L t = 0$
- Need to express terms with t as derivative:

$$
\begin{aligned}
&\frac{1}{t}\left[u(w-p^H t) - u(w-p^L t)\right] \\
=&\frac{1}{t}\left[\Big(u(w-p^H t) - u(w)\Big) - \Big(u(w-p^L t) - u(w)\Big)\right] \\
=&p^H \frac{u(w-p^H t) - u(w)}{p^H t} - p^L \frac{u(w-p^L t) - u(w)}{p^L t} \\
=&(p^L - p^H)u'(w)
\end{aligned}
$$

Inferring Risk Preferences: Taking Stock

- Previously backed out (the distribution of) risk type λ
- Derived locus of risk type/aversion indifferent btw contracts:

$$\lambda = \frac{(p^L - p^H)u'(w)}{u(w - d^L) - u(w - d^H)}$$

- What can we do with an expression containing $u(\cdot)$, $u'(\cdot)$, and w?
  - Recall Baily-Chetty formula that mapped unobservable $u'(c)$ gap into observables

Inferring Risk Preferences: From hopeless $u(\cdot)$ to hopeful r

- Take 2nd order Taylor expansion of u around w in previous expression
  - $u(w - d^L) \approx u(w) - u'(w)d^L + \frac{1}{2}u''(w)[d^L]^2$
  - When is this exact?
  - Structural signpost # 6: Don't miss forest through trees. Goal is connecting model to data.
- Recall coefficient of **absolute** risk aversion $r(w) = -\frac{u''(w)}{u'(w)}$
  - EFS assumed CRRA because choice was over fraction of wealth
  - Cohen and Einav assume CARA because choice is over dollar amount
- Do algebra in the privacy of your own home:

$$r^*(\lambda) = \frac{\frac{p^L - p^H}{\lambda(d^H - d^L)} - 1}{\frac{1}{2}(d^H - d^L)}$$

Recap: What is this telling us?

Outline

Institutional Details

Model

Identification of Risk Preferences

Overview of Identification Section

- We made an behavioral assumption (no MH) and functional form assumption (Poisson parameters distributed lognormal) to *identify* risk types
- We derived a model to get $\text{choice}_i = f(\text{risk type}_i, \text{risk preference}_i)$
- I'll have a brief digression about model *identification*
- Then we will discuss assumptions that *identify* risk preferences

Big Picture of Model Identification

What does model identification mean?

Big Picture of Model Identification

What does model identification mean?

- A model is (set) identified if different (sets of) values of the parameters imply different distributions of observable data (Matzkin 2013 ARE)
- Identification is a binary property of a(n economic or econometric) model

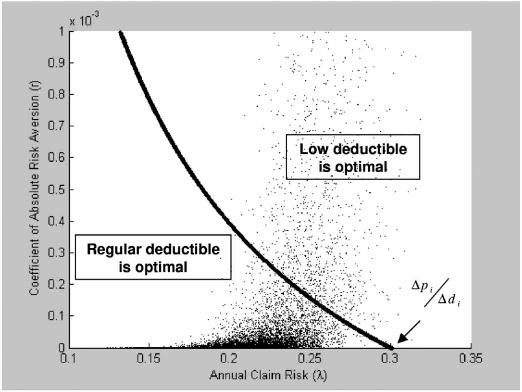
Recap on Model Identification: Examples

1. **Econometric model**: Additively linear in age, calendar time, and cohort
   - Age = calendar time - cohort
   - Different individual parameters consistent with same distribution of observable data

     $\Rightarrow$ not identified (Ameriks and Zeldes 2004)
2. **Economic model**: Equilibrium relationship between supply and demand
   - Supply (demand) shifters identify the demand (supply) curve

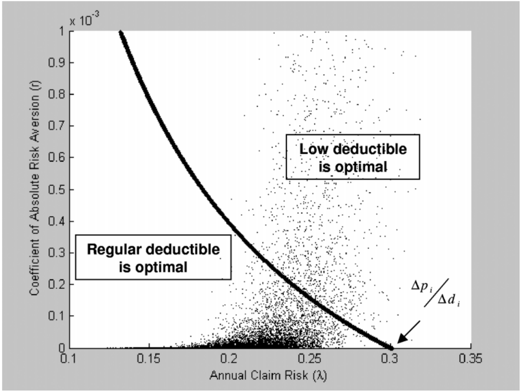
Identifying Risk Preferences: Illustration

Indifference condition from before: $r^*(\lambda) = 2\left[\frac{1}{\lambda}\frac{\Delta p}{(\Delta d)^2} - \frac{1}{\Delta d}\right]$



Identifying Risk Preferences: Illustration

Indifference condition from before: $r^*(\lambda) = 2\left[\frac{1}{\lambda}\frac{\Delta p}{(\Delta d)^2} - \frac{1}{\Delta d}\right]$



1. What does variation in Δp, Δd, and $\Delta p/\Delta d$ do? (www.desmos.com/calculator)
2. What variation and outcomes do we observe in the data?

Nonparametric Identification of r

- Given λ, choice identifies a **set** of possible r
- Random variation in prices and menus delivers many such sets

Nonparametric Identification of r

- Given λ, choice identifies a **set** of possible r
- Random variation in prices and menus delivers many such sets
- Sufficient variation identifies r without additional functional form assumptions
  - *Intuition*: Find out a number by repeatedly asking if it's $> x$ for many different x

Need for Parametric Assumptions

1. Cohen and Einav: Don't actually observe infinite price variation
2. EFS: Don't observe any price variation, so can get only identified sets without further assumptions
3. Structural signpost # 7: Keep track of what assumptions are required by setting vs. want of point identification/lack of infinite data vs. lack of random variation

Parametric Identification: Setup

- Recall parametric assumption on risk type:

$$\ln \lambda_i = x_i'\beta + \varepsilon_i$$

where $\varepsilon \sim N(0, \sigma_\lambda)$

- Additionally make parametric assumption on risk aversion:

$$\ln r_i = x_i'\gamma + v_i$$

where $v_i \sim N(0, \sigma_r)$

Parametric Identification: Setup

- Recall parametric assumption on risk type:

$$\ln \lambda_i = x_i'\beta + \varepsilon_i$$

where $\varepsilon \sim N(0, \sigma_\lambda)$

- Additionally make parametric assumption on risk aversion:

$$\ln r_i = x_i'\gamma + v_i$$

where $v_i \sim N(0, \sigma_r)$

- Note: We actually care about $\mu_\lambda \equiv E_i[\lambda_i]$ and $\mu_r \equiv E_i[r_i]$ rather than β and γ
- Allow $Cov(\varepsilon_i, v_i) = \rho\sigma_\lambda\sigma_r$ since motivation is joint distribution of unobservables

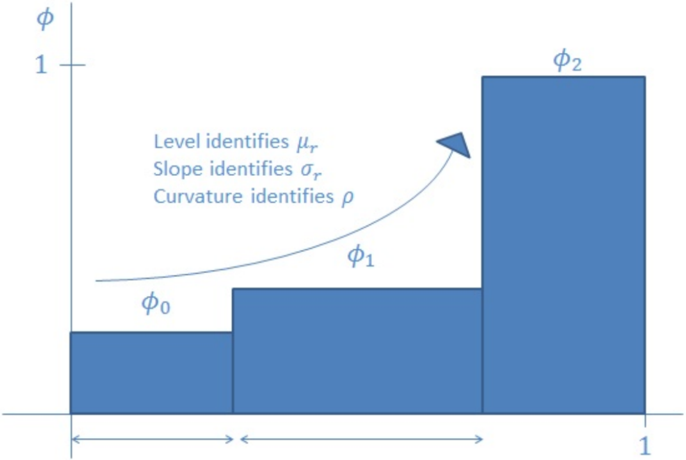
Parametric Identification: Intuition

- Model parameters: risk types $\mu_\lambda, \sigma_\lambda$; risk preferences μ_r, σ_r; and correlation ρ
- Rough intuition: Need 5 relevant moments to identify 5 parameters

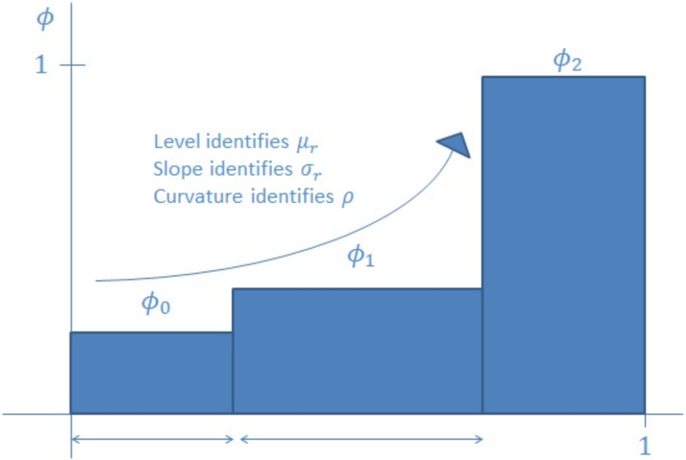
Parametric Identification: Intuition

- Fractions with k claim together identify $\mu_\lambda, \sigma_\lambda$
- Fractions choosing low deductible among those with 0,1, and 2 claims identify the remaining preference and correlation parameters
- Define $\phi_k \equiv$ fraction who chose the low deductible plan among those who realized k claims for $k \in \{0, 1, 2\}$

Parametric Identification: Graphical Intuition



Parametric Identification: Graphical Intuition



Sequential thought experiments:

0. We have $\mu_\lambda, \sigma_\lambda$
1. Suppose r constant (i.e. $\sigma_r = \rho = 0$) $\rightarrow$ bar height
2. Now suppose $\sigma_r > 0$ but $\rho = 0$ $\rightarrow$ bar slope
3. Now suppose $\sigma_r, \rho > 0 \rightarrow$ bar convexity

Parametric Identification: Estimation

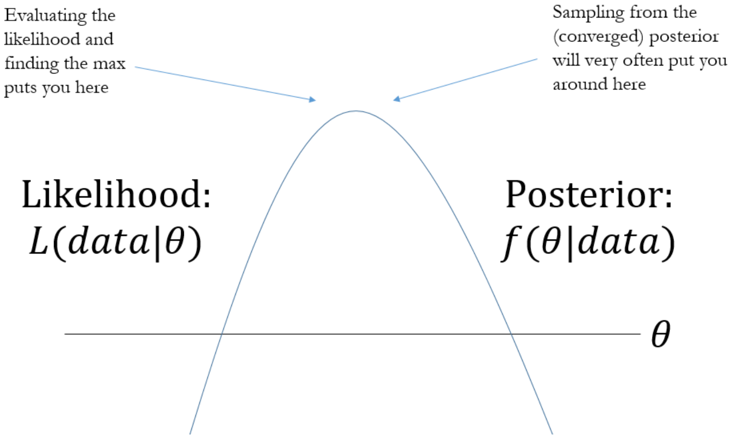
- Collect parameters: $\Theta = \{\beta, \sigma_\lambda, \gamma, \sigma_r, \rho\}$
- Write down likelihood of observed choices given a candidate Θ:

$$L(claims_i, choice_i|\Theta) = Pr(claims_i, choice_i|\lambda_i, r_i)Pr(\lambda_i, r_i|\Theta)$$

- Maximize likelihood?
  - Turns out this is computationally hard
  - (Evaluating likelihood *once* requires integrating over *both* λ_i and r_i for *every* i)
  - Gibbs Markov Chain Monte Carlo to the rescue!

Gibbs Sampling General Intuition

Both methods involve the data disciplining parameter estimation



Simulation methods: For when you/your computer is too dumb to evaluate something

Gibbs Sampling Explanation

- MLE $\approx$ frequentist, Gibbs MCMC $\approx$ Bayesian
- Gibbs procedure:
  1. Take draws of all parameters from priors
  2. Draw a single parameter from a posterior (given observables and other drawn parameters)
  3. Do previous step with a different parameter
  4. Continue iterating over all parameters many times

Crazy result: The sequence of posterior draws converges to the joint distribution[1]

Upshot: After a bunch of iterations, averaging over many subsequent draws delivers (mean) parameter estimate

Gibbs Sampling Application-Specific Intuition

- Given (parameters governing distribution of) r_i, observed choices tell you which (parameters governing distribution of) λ_i are likely
- Given those (parameters governing distribution of) λ_i, (parameters governing distribution of) r_i are likely
- Applying the discipline of observed choices many times eventually delivers parameters “close to the truth"

Conclusion

- Separately identify multidimensional unobserved types: risk type and preference
  - "Looking under the hood" of the positive correlation test
  - Common themes in "structural" insurance papers:
    1. Use ex post realizations to infer ex ante risk
    2. Make assumption on contract choice process
- Requires assumptions to see what objects in data can be mapped to unobservables
- Also requires a lot of structure. In my view, the paper transparently:
  1. Argues why assumptions and structure are necessary
  2. Shows where identification comes from
  3. Focuses on an interesting question without getting distracted