14.472 Public Finance II

Asymmetric Information: Testing

Amy Finkelstein

Fall 2024

Empirical analysis of selection

- Theory: adverse selection can impair efficient operation of insurance markets and create scope for welfare improving government intervention
- Lays out an empirical agenda:
  - **[Today]** Does selection exist in a particular market
  - What are the efficiency costs of any detected selection
  - What are the welfare costs of alternative potential government interventions?

Lecture(s) road map

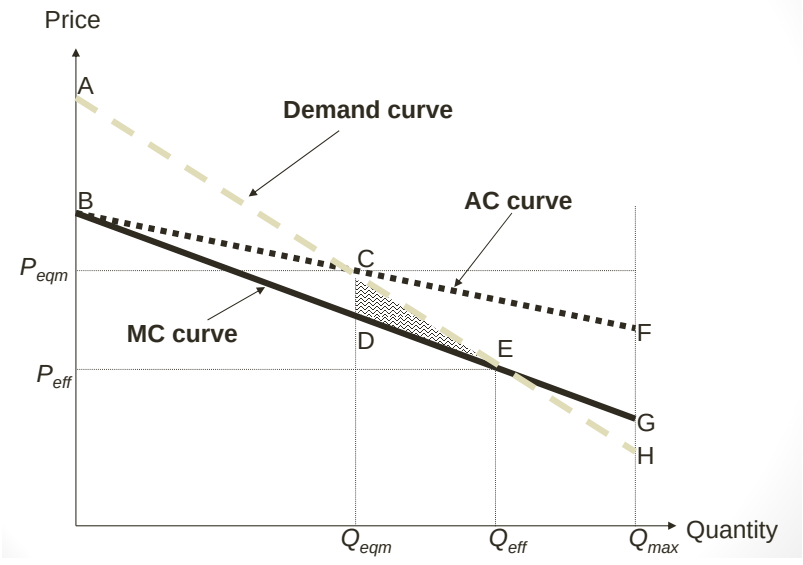
- Testing for selection
- Empirical welfare analysis I: Using data on choices and claims
- Empirical welfare analysis II: What happens when you can't use choice data
  - Don't trust revealed preference
  - Markets don't exist

Testing for Selection

Some perspective

- Seminal theoretical work on adverse selection in 1970s
  - Very influential in policy
  - 2001 Nobel prize
- Empirical testing only emerges several decades later!

Adverse selection: downward sloping marginal cost curve



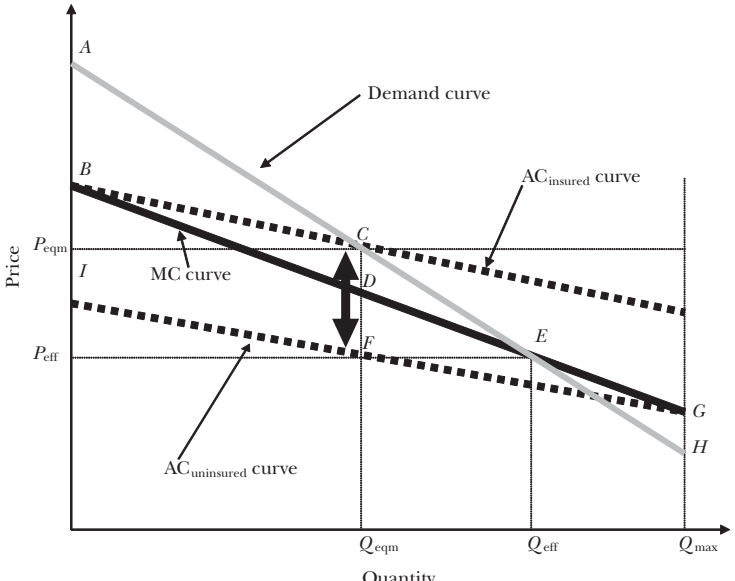
"Positive correlation test"

- Testing for adverse selection essentially requires testing whether MC curve downward sloping
- Making inferences about marginal individuals can be difficult
- Early empirical approaches developed strategies that could focus on averages
- 'Positive correlation' or 'bivariate probit' test (Chiappori and Salanie, JPE 2000)
  - 'Early' for emprical literautre on adverse selection in insurance markets
  - 'Late' relative to the theory (1970s!)

“Positive correlation test”

- Reject null of symmetric information if there is a positive correlation between insurance coverage and ex-post risk occurrence
- Are average costs of insured higher than average costs of uninsured?
- At any given price, and in particular at the equilibrium price, adverse selection implies that average cost of insured individuals is higher than average costs of uninsured individuals

Positive correlation test: graphical illustration



Example: Annuitant vs. Population Mortality (US 2007)

	Annuitant Mortality		Population Mortality	
	Men	Women	Men	Women
65	1.02%	0.57%	1.72%	1.16%
75	2.98	1.61	4.29	2.98
85	8.06	5.08	11.35	8.54

Positive correlation test: regression version

$$Coverage_i = X_i\beta + \epsilon_i$$
$$Accident_i = X_i\gamma + \mu_i$$

- Simultaneously estimate above two equations (e.g. bivariate probit)
  - Under the null of symmetric information, residuals should be uncorrelated
  - Statistically significant positive correlation between two implies rejection of the null hypothesis
- Note: Crucial to condition on what is priced by insurer.
  - Test is among individuals treated symmetrically by insurance company
  - As a result, many applications use data from a single insurance company and compare claims across individuals who are offered same contracts at same price but choose more vs. less coverage

Spawned Cottage Industry of Papers

- Acute health insurance and annuities
  - Higher risk people tend to have more insurance
- Life insurance, long term care insurance, Medigap:
  - Higher risk people have no more (and sometimes less) insurance.
- Auto insurance – evidence more mixed
- Life insurance vs annuities . . .
  - Annuity: payments as long as you survive
  - Insure opposite risks. And find selection in one not in another . . . ?

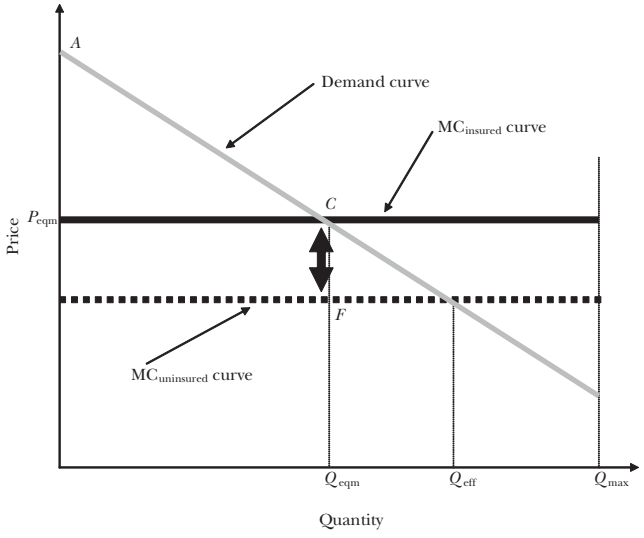
Two Important Limitations to Positive Correlation Test

1. Does not distinguish between adverse selection and moral hazard
2. Not robust to allowing for unobserved preference heterogeneity in addition to risk types

Moral hazard also generates positive correlation

- Adverse selection: those with private information they are high expected cost self select into insurance market
- Moral hazard: individuals identical before purchasing insurance; those with greater coverage have less incentive to take actions to reduce their expected costs ex post

Moral hazard also generates positive correlation



Positive correlation test is a joint test of adverse selection and moral hazard

- Conceptually very different: ex ante vs. ex post private information
- Policy implications different: government tends not to have comparative advantage with moral hazard
- So really want to know which you have detected

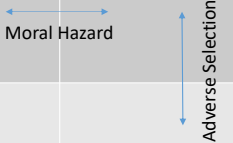
Distinguishing selection from moral hazard

- Key point: need exogenous variation in contracts
  - Basic problem: distinguishing treatment (moral hazard) from selection (selection!)
- Variety of sources of variation
  - Quasi-Exogenous variation (examples)
    - border design / geographic variation in health insurance premiums (Panhans 2019)
    - policy reforms with differential effects on flood insurance prices by house type (Wagner 2021)
  - Randomized experiments (examples)
    - health insurance in rural Pakistan (Fischer et al. 2018)
    - consumer loans in South Africa (Karlan and Zinman 2009)

Karlan and Zinman 2009

- Setting: Consumer lender (South Africa)
- Randomized offer interest rate and contract rate on loan
- Selection: compare repayment rate of those offered different rates (but receiving same rate)
- MH: compare repayment rates of those responding to same high offer rate but facing different contract rates

	High Contract Rate	Low Contract Rate
High Offer Rate		
Low Offer Rate	N/A	



Interpreting results of positive correlation test

- Positive correlation may reflect adverse selection, moral hazard, or both
- Lack of positive correlation may reflect:
  - No asymmetric information
  - Offsetting advantageous selection and moral hazard?

Unobserved heterogeneity in preferences

- Standard theory models: individuals may potentially differ on only one unobserved dimension: risk type
- In the presence of multiple dimensions of heterogeneity, inability to detect a positive correlation does not preclude presence of selection
  - If risk aversion is negatively correlated with risk, lower-risk individuals can purchase more insurance in equilibrium even when there is private information about risk type (e.g. Finkelstein and McGarry 2006; Fang et al. 2008)

Addressing Limitations of the Positive Correlation Test

- Two key issues:
  - Not robust to preference heterogeneity
  - Joint test of moral hazard and selection
- "Cost curve test" of selection (Einav, Finkelstein and Cullen, 2010)
  - Addresses both these issues
  - But higher data hurdle (need quasi-random variation in prices)

Cost curve test (Einav, Finkelstein and Cullen 2010)

- Idea: Slope of MC curve provides a direct test of existence and nature of selection
- Reject null of no selection if reject null of constant MC curve
- Slope of cost curve indicates if selection is adverse or advantageous

Cost curve test implementation

- Estimate average cost curve on sample who (endogenously) purchase the same contract:

$$c_i = \gamma + \delta p_i + \epsilon_i$$

  where c_i is insurable costs (claims), and p_i is price of insruance
- Estimating how costs change for the endogenously selected sample of those who stay insured as price (exogenously) varies
  - This is key idea of selection
- Data requirements higher than for positive correlation test:
  - As with positive correlation test, need to know insurance coverage (since sample based on that)
  - Additional data hurdle: need exogenous variation in prices

Aside: Selecting on endogenous outcome

- Useful if you want to understand the characteristics of those who respond to the intervention
- Other examples:
  - What type of DI applicants deterred from hassles (Deshpande and Li 2019)?
  - Who is the marginal child when abortion is legalized (Gruber, Levine and Staiger 1999)?
- More generally: "characterizing the compliers" (Abadie 2002).

Example: Colorado Health Insurance Exchange

- Panhans (2019)
- Colorado Health Insurance Exchange 2014
  - Created by Affordable Care Act (ACA)
  - Subsidized for low income individuals
- Statewide data on premiums, claims, insurance coverage (exchange coverage vs. not) .
- Source of premium variation: geographic discontinuities in insurance premiums at boundaries of "rating areas" established by law
  - Premiums change discretely at "artificial" boundaries of rating areas
  - Compare costs of those enrolled on either side of the border (fixed effect for each zip code pair $\varphi_{g(k)}$)

$$c_i = \gamma + \delta p_{ik} + \varphi_{g(k)} + \epsilon_i$$

Border Design

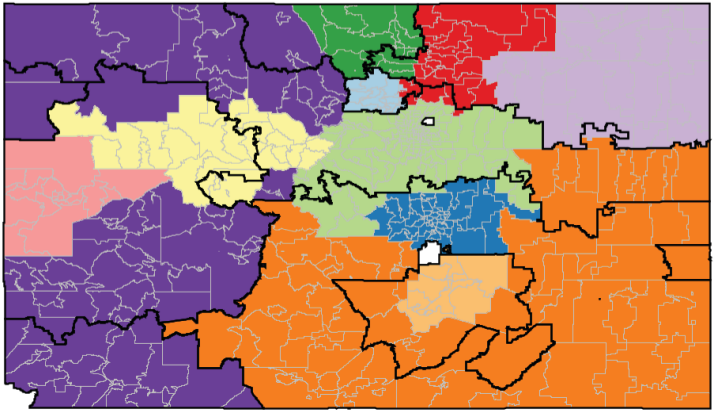


FIGURE 2. 2014 RATING AREAS IN COLORADO

Notes: Five-digit zip codes are shown grouped into rating areas based on color. The outlines designate the grouping of zip codes into medical markets, here defined as the Hospital Referral Region (HRR).

Border-induced premium variation

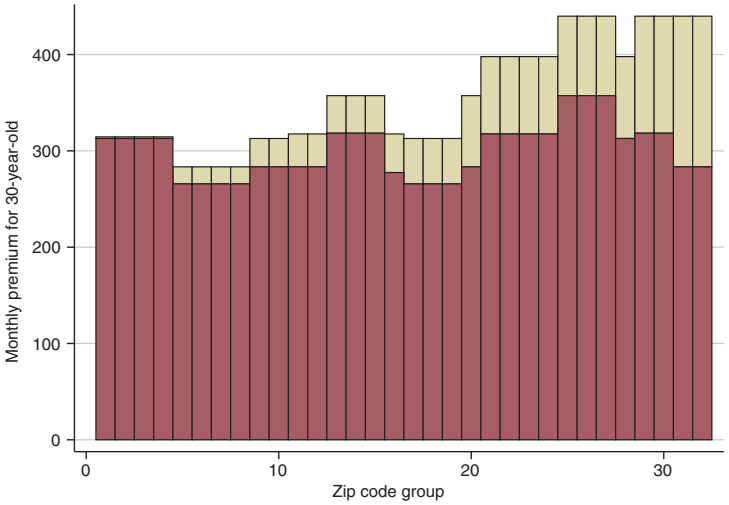


FIGURE 3. CHANGE IN PREMIUM ACROSS RATING AREA BOUNDARY

Cost curve indicates selection

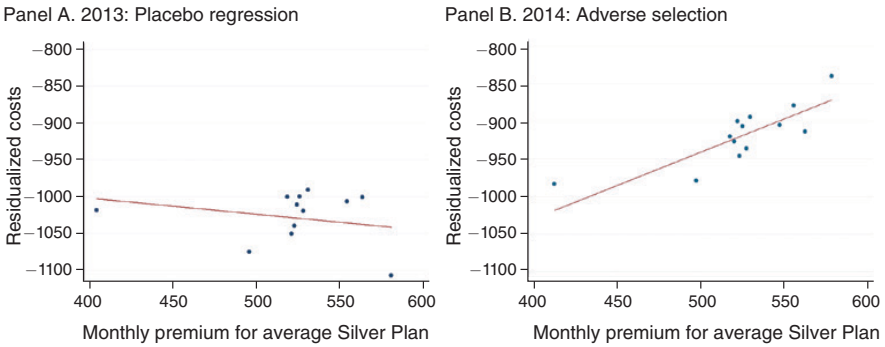


FIGURE 5. BINNED SCATTERPLOT OF SELECTION REGRESSION

Notes: Panel A presents graphically the results from the placebo regression in column 1 of Table 4. Panel B presents the results from the main OLS results in panel A, column 1 of Table 3, which indicate adverse selection. The sample means of premiums have been added back in to the premium residuals before plotting.

Cost curve test: comments

- Can detect adverse or advantageous selection (slope of curve)
- Not affects by existence (or lack thereof) of moral hazard
  - Estimate cost curve on sample in which coverage is fixed
  - OF course slope of cost curve may reflect selection based on different expected responsiveness to incentive effects
  - "Selection on moral hazard" (Einav, Finkelstein, Ryan, Schrimpf, Cullen AER 2013)

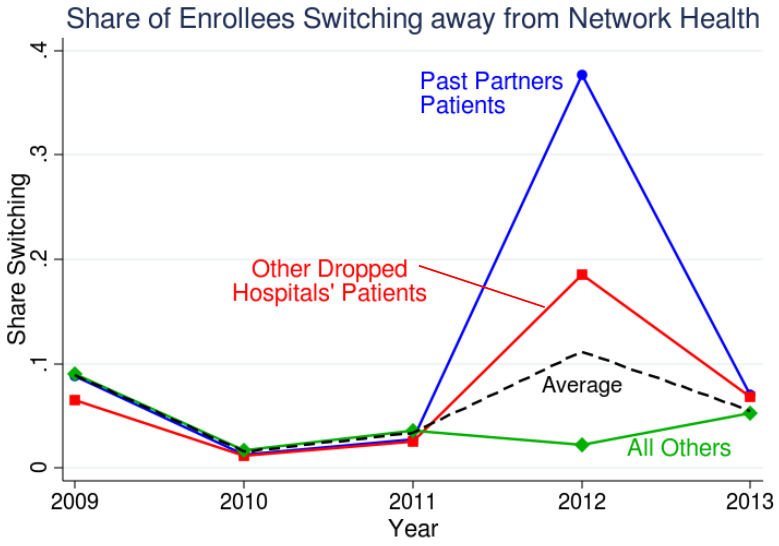
Other consequences of selection

- Most existing work looks at impact of adverse selection on (mis-) pricing and insurance coverage
- Selection may also give insurers incentives to distort plan benefits (Rothschild-Stiglitz 1976)
- Relatively little existing work looking at impact of selection on contract / benefit design
  - Cost-sharing to discourage high cost enrollees
    - e.g. high cost-sharing for HIV drugs in Part D when new drug development makes existing diagnosis-based risk adjustment inadequate (Carey 2017 AEJ:Policy)
  - Bundling gym memberships with health insurance plans attracts lower cost enrollees (Cooper and Trivedi 2012 NEJM)
  - Broader networks attract higher cost enrollees (Shepard, AER 2022)

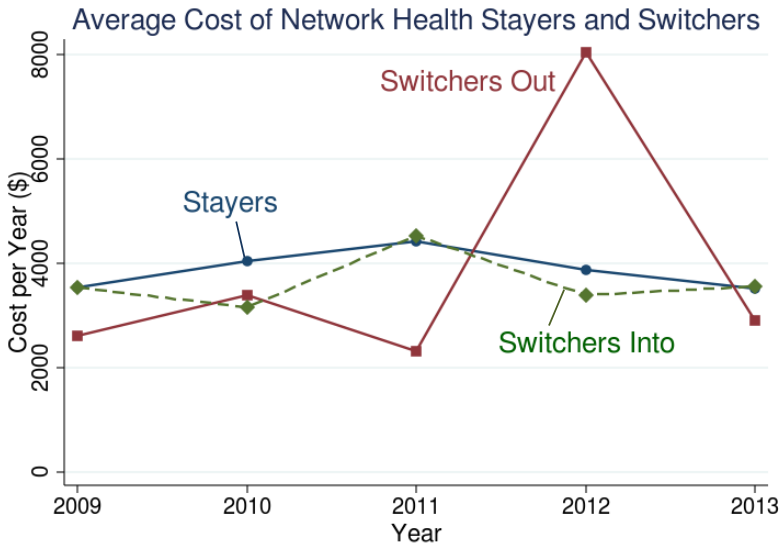
Shepard (2022): Selection and Benefit Design

- Massachusetts pioneer health insurance exchange for low income individuals (Commcare)
- In 2012, one plan (Network Health Plan) substantially limited its network of providers
  - Dropped Partners ('star hospitals' - MGH, BWH)

Limiting plan's network produces switching out of plan



High cost enrollees in particular switch out of limited network plan



Implications

- (Rare) evidence of how selection may affect benefit design
  - Broader networks attract higher cost enrollees
  - See also Kreider, Layton, Shepard, Wallace et al (JHE 2024): covering a top cancer hospital also creates adverse selection
- Interestingly, this adverse selection into broader network plans occurs despite 'state of the art' risk adjustment
  - To counteract adverse selection: plans are reimbursed more for patients with higher predicted costs ('risk scores'), predicted based on demographics and past diagnoses
- Why might risk scoring be inadequate?
  - Statistical challenge (need better data, better algorithms)
  - Economic / conceptual challenge: selection on levels (expected health costs) but also on slope (responsiveness of behavior to plan)

Some open testing questions

- Impact of selection on contract design
- Many markets where selection has had relatively little examination
  - e.g. Disability insurance. There's a lot of public policy (and research on that public policy) but not on potential underlying market failure
- Why don't insurance companies price on more observable characteristics?
- Does benefit design respond endogenously to risk distribution (a la Rothschild and Stiglitz)?