A Testing for Moral Hazard Survey Versus Claims Data

Presented to the
Australian Health Economics Society

Conference September 2010

D Rowell
Dr HS Nghiem
Prof. LB Connelly

Financial support for this research was provided by ACERH and the Brian Gray Scholarship for 2010 awarded by APRA and the RBA

ACERH
AUSTRALIAN CENTRE FOR ECONOMIC RESEARCH ON HEALTH

THE UNIVERSITY OF WESTERN AUSTRALIA

ANU
THE AUSTRALIAN NATIONAL UNIVERSITY

THE UNIVERSITY OF QUEENSLAND AUSTRALIA

Introduction

- Moral hazard refers to the tendency of insurance to reduce an individual's motive to prevent loss and therefore increase the probability of a claim.

- The aim is to present a test for moral hazard using conditional correlation with cross section data

- We argue that survey data is better suited to this task than claims data.

Outline

1. Empirical Literature
   i. Challenge
   ii. Literature

2. What is wrong with claims data?
   - Chiappori & Salanié (2000)

3. What is right with survey data?
   - Finkelstein & McGarry (2006)

4. A test for moral hazard

Empirical Challenge

Asymmetric information

$$Claim = f(Insurance \mid \mathrm{X})$$

Where **X** is a vector of covariates which reflect the insurer’s information set

Empirical challenge

At least 2 types of asymmetric information

1. Moral hazard –*hidden action*

$$Claim = \alpha_0 + \alpha_1 Insurance + \alpha_2 \boldsymbol{X} + \varepsilon_i$$

2. Adverse Selection –*hidden information*

$$Insurance = \beta_0 + \beta_1 Claim + \beta_2 \boldsymbol{X}_i + \eta_i$$

Chiappori & Salanié (2000)

- Test for asymmetric information using claims data
  - $$Insurnace = \alpha_o + \alpha_1 X + \varepsilon_i$$
  - $$Claim = \beta_o + \beta_1 X + \eta_i$$
- *X = 55 variables insurer's info. set*

Gender (1)	*Type of use (3)*	*Profession (7)*
Make of car (7)	*Area (4)*	*Age of car (11)*
Age of driver (8)	*Region (9)*	*Performance of car (5)*

- *Test for asymmetric information in a sub-sample of young drivers*
  - $$H_0 : \mathrm{cov}(\varepsilon_i, \eta_i)$$
  - $$H_0 : \rho = 0$$

Chiappori & Salanié (2000)

- Test for moral hazard using a ‘natural experiment’
  - Some young drivers ‘inherit’ their parent’s *bonus malus* discount if car ‘is’ jointly owned
  - *IBM* = 1 if young driver ‘inherits’ their parent’s maximum *bonus malus* discount
- Re-estimated bivariate probit model

$$Insurance = \alpha_o + \alpha_1 X + \alpha_2 IBM + \varepsilon_i \qquad (1)$$

$$Claim = \beta_o + \beta_1 X + \beta_2 IBM + \eta_i \qquad (2)$$

Chiappori & Salanié (2000)

Interpretation of β_2?

1. Pure familial correlation hypothesis: Low-risk parents produce low-risk children

2. Pure Moral Hazard hypothesis: The risk rating of premiums means that the cost of each successive claim is increasing. Preventative effort is increased with higher premiums

Three Options

A. Familial correlation & no moral hazard
   - *IBM* will be negatively correlated with RTC

B. No familial correlation & No moral hazard
   - *IBM* will not be correlated with RTC

C. No familial correlation & moral hazard
   - *IBM* will be positively correlated with RTC

β_2 is negative and significant, which rejects the moral hazard hypothesis' for the alternative hypothesis 'that the parents' performances are positively correlated with the child's.

Literature review

- Following Chiappori and Salanié (2000) the empirical literature has avoided analysis with cross sectional data
  - Abbring (2003) use longitudinal claims data to test for moral hazard using survival analysis
  - Dionne *et al* (2004, 2006 and 2007) use panel data to test for moral hazard using Granger causality.
  - Cohen (2005) test for asymmetric information using claims data but state that a test for moral hazard is beyond the scope of their paper.
  - Israel (2007) use longitudinal claims data to test for moral hazard using a difference in difference approach

What's wrong with Claims data?

- Chiappori & Salanié (2000) revisited
- A Natural experiment
  - occurs when some (often unintended) feature of the setup we are studying produces exogenous variation in an another otherwise endogenous explanatory variable" (Wooldridge 2002)
- But the *IBM* is
  - Observable to insurer
  - Used by insurers to risk rate some young policyholders
  - Correlated with driving ability
- Therefore
  - *IBM* should be in *X*
  - *X* should comprise 56 exogenous variables not 55
  - Claims data will invariably be constrained in this way

Chiappori & Salanié (2000) revisited

- But,
  - Moral hazard and a familial correlation could exist
  - There are therefore 4 possibilities!

	Nil Moral Hazard	Moral Hazard
Nil Familial Correlation	Nil Correlation	+ve correlation
Familial Correlation	-ve correlation	Ambiguous

What's right with Survey Data?

Finkelstein and McGarry (2006)

- Aim
  - To investigate asymmetric information in a market for long-term care insurance
- Data
  - Health and Retirement Study (HRS) uses survey data!
    - Approximate the insurer's information set *X*
    - Contains private information not observable to the insurer
      - Prob. nursing home placement
      - Preference for insurance

Finkelstein and McGarry (2006)

- Method

1. Provide evidence of private information by estimating 2 probit models

$$NHP = \alpha_o + \alpha_1 X + \alpha_2 Prob\_NHP + \varepsilon_i$$
$$LTCINS = \beta_o + \beta_1 X + \beta_2 Prob\_NHP + \eta_i$$

Results *SB*
$\alpha_2 > 0$
$B_2 > 0$

2. Traditional tests for asymmetric information fail to identify this private information

   i. Bivariate probit model: $\rho = 0$

$$NHP = \alpha_o + \alpha_1 X + \varepsilon_i$$
$$LTCINS = \beta_o + \beta_1 X + \eta_i$$

   ii. Probit model: $\alpha_2 = 0$

$$NHP = \alpha_o + \alpha_1 X + \alpha_2 LTCINS + \varepsilon_i$$

Finkelstein and McGarry (2006)

3. Hypothesise that other dimensions of private information e.g. preference for insurance (*PI**)

4. Re-estimate the probit models using seat-belt (*SB*) use as a proxy for *PI**

$$NHP = \alpha_o + \alpha_1 X + \alpha_2 Prob\_NHP + \alpha_3 SB + \varepsilon_i$$

$$LTCINS = \beta_o + \beta_1 X + \beta_2 Prob\_NHP + \beta_3 SB + \eta_i$$

Results *SB*
$\alpha_3 < 0$
$\beta_3 > 0$

"...the existence of multiple forms of private information in an insurance market and demonstrate how these factors can have offsetting effects on the correlation between insurance coverage and risk occurrence, thus invalidating the standard test of asymmetric information."

A Test for Moral Hazard

- Structural Model

$$NHP = \beta_0 + X\beta_1 + \beta_2 LTCINS + PI^* + RT^* + \mu_i$$

- Empirical test for Moral Hazard
  - **Bivariate probit with one instrumental variable and proxy**

$$NHP = \beta_0 + \beta_1 X + \beta_2 LTCINS\_R + \beta_3 SB + \varepsilon_i$$

$$LTCINS\_R = \alpha_0 + \alpha_1 X + \alpha_2 LTCINS\_S + \beta_3 SB + \eta_i$$

Selection of an IV

Insurance status of spouse (*LTCINS_S*) selected as IV for *LTCINS_R*

1. *LTCINS_S* correlated with *LTCINS_R ?*
   - Becker (1981) has argued that within couples matching occurs on many criteria.
   - Specifically intra-spousal correlation occurs in
     - Risk aversion (Spivey 2010)
     - Health status (Wilson 2002)
     - Health maintenance activities (Macken, Yates et al. 2000)

2. *LTCINS_S* is not correlated with error term
   - Once living arrangements of the spouse are controlled for the respondent's decision to enter a nursing home will not be correlated with spouse's insurance status

Results from Biprobit model

	NHP(1995-2000)		LTCINS (1995)	
	Coefficient	*p*-value	Coefficient	*p*-value
Respondent's long-term care insurance (1995)	-0.176	0.43	n.a.	n.a.
Spouse's long-term care insurance (1995)	n.a.	n.a.	1.942	<0.01
Wears seatbelt	0.009	0.89	0.145	0.08
Female	-2.745	<0.01	0.360	0.70
Coupled household	-1.435	0.01	-0.367	0.50
Spouse's age (=0 if unmarried)	0.017	0.01	-0.003	0.67
2nd quartile -Income	0.045	0.55	0.191	0.05
3rd quartile -Income	0.032	0.71	0.312	<0.01
4th quartile -Income	0.028	0.79	0.519	<0.01
2nd quartile -Assets	0.013	0.86	0.253	0.01
3rd quartile -Assets	0.036	0.67	0.363	<0.01
4th quartile -Assets	0.055	0.56	0.310	<0.01
ADL limitation: walking	-0.205	0.21	-0.132	0.61
ADL limitation: dressing	-0.142	0.33	-0.347	0.11
ADL limitation: bathing	0.144	0.36	0.111	0.64
ADL limitation: eating	-0.138	0.45	0.169	0.51
ADL limitation: toileting	-0.148	0.36	-0.078	0.74
IADL limitation: medication	0.091	0.60	0.151	0.52
IADL limitation: shopping	0.155	0.14	0.027	0.85
Use wheelchair	-0.094	0.58	0.264	0.28
Use walker	0.265	0.02	-0.180	0.33
Use crutch	-0.016	0.97	-4.514	1.00
Use oxygen	0.012	0.97	-5.768	1.00
Use cane	0.103	0.17	0.024	0.82
BMI > 30	0.108	0.14	-0.179	0.05
BMI < 20	0.037	0.68	0.051	0.62
Incontinent	-0.054	0.39	0.067	0.34

…continued

	NHP (1995-2000)		LTCINS (1995)	
	Coefficient	*p*-value	Coefficient	*p*-value
Takes medication	0.067	0.35	0.131	0.09
Smokes	0.000	1.00	0.111	0.25
Depressed	0.135	0.03	-0.171	0.03
ETOH	0.044	0.78	0.018	0.91
Cognitively impaired	-0.356	0.93	-56.998	0.19
Diabetic	0.022	0.80	-0.139	0.17
Insulin dependent diabetic	0.102	0.48	-0.148	0.43
Diabetic with renal Dx.	0.124	0.53	0.239	0.38
CVA	0.222	0.01	-0.021	0.85
Heart Dx.	0.101	0.17	-0.037	0.67
Heart Rx.	-0.196	0.02	0.057	0.57
AMI	0.063	0.65	-0.054	0.77
CCF	0.242	0.10	0.104	0.57
Hypertensive	-0.013	0.82	0.076	0.20
Hip replacement	0.206	0.06	0.056	0.69
Lung Dx.	-0.066	0.44	0.205	0.02
Cancer	0.012	0.86	-0.003	0.97
Psychiatric Dx.	0.149	0.03	0.104	0.19
Arthritis	-0.071	0.19	-0.034	0.55
Any days in nursing home (pre 1995)	0.469	0.01	0.291	0.18
Receives home help	0.129	0.22	-0.307	0.06
Injury from fall	0.101	0.06	0.135	0.02
# ADL's by age	0.003	0.10	0.000	0.89
# IADL's by age	-0.001	0.43	-0.001	0.42
Cognitive impairment by age	0.015	0.78	0.670	0.18
Sex by age	0.036	<0.01	-0.006	0.65
Sex by cognitive impairment	3.399	0.49	58.623	0.18
Sex by # ADL's	1.201	0.00	0.396	0.50
Sex by # IADL's	-0.330	0.47	-0.040	0.95
Sex by cognitive impairment by age	-0.054	0.38	-0.693	0.17
Sex by # ADL's by age	-0.017	<0.01	-0.004	0.56
Sex by # IADL's by age	0.005	0.43	0.001	0.85
Constant	-1.701	<0.01	-1.924	<0.01

Results from Biprobit model

	NHP (1995-2000)		LTCINS (1995)	
	Coeff.	*p*-value	Coeff.	*p*-value
LTCINS_R (1995)	-0.176	0.43	n.a.	n.a.
LTCINS_S (1995)	n.a.	n.a.	1.942	<0.01
Wears seatbelt	0.009	0.89	0.145	0.08
Constant	-1.701	<0.01	-1.924	<0.01

Biprobit model with a control for Spouse lives in nursing home

	NHP (1995-2000)		LTCINS (1995)	
	Coeff.	*p*-value	Coeff.	*p*-value
LTCINS_R (1995)	-0.146	0.52	n.a.	n.a.
LTCINS_S (1995)	n.a.	n.a.	1.944	<0.01
Wears seatbelt	0.009	0.89	0.147	0.08
Spouse lives in nursing home 1995-2000	0.77	<0.01	0.086	0.58
Constant	-1.673	<0.01	-1.924	<0.01

Quality of the IV

1. *LTCINS_S* $\notin X$

2. Correlated with endogenous variable *LTCINS_R*
   i. Coefficient for *LTCINS_S* = 1.942 (p-value < 0.01)
   ii. Test for weak instruments R^2 increased from 7.9% to 22.8%

$$LTCINS\_R = \alpha_0 + \alpha_1 X + \alpha_2 LTCINS\_S + \beta_3 SB + \eta_i$$

3. We claim the IV is uncorrelated with ε_i
   - This assertion usually must be taken on faith since *RT** is an unobserved component of ε_i
   - But the HRS contains *Prob_NHP* which is used as a proxy for *RT**

Quality of IV?

- Estimate this bivariate probit with an additional proxy for *RT** i.e. *Prob_NHP*

$$NHP = \beta_0 + \beta_1 X + \beta_2 LTCINS\_R + \beta_3 SB + \beta_4 Prob\_NHP + \varepsilon_i$$

$$LTCINS\_R = \alpha_0 + \alpha_1 X + \alpha_2 LTCINS\_S + \alpha_3 SB + \alpha_4 Prob\_NHP + \eta_i$$

Quality of IV?

-0.176 (*p*-value =0.43)

	NHP (1995-2000)		LTCINS (1995)	
	Coefficient	*p*-value	Coefficient	*p*-value
Respondent's long-term care insurance (1995)	-0.209	0.37	n.a.	n.a.
Spouse's long-term care insurance (1995)	n.a.	n.a.	1.937	< 0.01
Wears seatbelt	-0.030	0.69	0.181	0.04
Predicted probability of NHP	0.268	0.01	0.487	< 0.01
Constant	-1.734	< 0.01	-2.008	< 0.01

Conclusions

1. Empirical result: No *ex ante* moral hazard in insurance market for long-term care
   i. Plausible?
   ii. Moral hazard in this market may be *ex post*

2. Method: Suggest testing for moral hazard with an IV
   i. Claims data are the wrong data. They can not discriminate between the different dimensions of private information

   ii. The test for moral hazard used by Chiappori & Salanié (2000) was confected because of their reliance on claims data.

   iii. The analysis of survey data as outlined by Finkelstein and McGarry (2006) offers a new and prommissing lines of investigation. It suggests survey data not observable to the insurer may be used to discrimnionate between different types of asymetic information.

Bibliography

Abbring, J. H., J. J. Heckman, et al. (2003). "Adverse Selection and Moral Hazard In Insurance: Can Dynamic Data Help to Distinguish?" Journal of the European Economic Association 1(2-3): 512-521.

Becker, G. S. (1981). A treatise on the family. Cambridge, Massachusetts, Harvard University Press.

Berk, R. A. (1983). "An introduction to sample selection bias in sociological data." American Sociological Review 48(3): 386-398.

Brown III, H., J. Pagán, et al. (2005). "The impact of diabetes on employment: genetic IVs in a bivariate probit." Health Economics 14(5): 537-544.

Chiappori, P. A. and B. Salanié (2000). "Testing for Asymmetric Information in Insurance Markets." The Journal of Political Economy 108(1): 56-78.

Clarke, K. A. (2005). "The phantom menace: Omitted variable bias in econometric research." Conflict Management and Peace Science 22(4): 341-352.

Cohen, A. (2005). "Asymmetric Information and Learning: Evidence from the Automobile Insurance Market." Review of Economics and Statistics 87(2): 197-207.

Cohen, A. and P. Siegelman (2010). "Testing for Adverse Selection in Insurance Markets." Journal of Risk and Insurance 77(1): 39-84.

Dionne, G., C. Gouriéroux, et al. (2001). "Testing for Evidence of Adverse Selection in the Automobile Insurance Market: A Comment." Journal of Political Economy 109(2): 444-453.

Dionne, G., P.-C. Michaud, et al. (2004). "Separating Moral Hazard from Adverse Selection and Learning in Automobile Insurance: Longitudinal Evidence from France."

Dionne, G., P.-C. Michaud, et al. (2006). "Separating Moral Hazard from Adverse Selection and Learning in Automobile Insurance: Longitudinal Evidence from France."

Dionne, G., P.-C. Michaud, et al. (2007). "Separating Moral Hazard from Adverse Selection and Learning in Automobile Insurance: Longitudinal Evidence from France."

Durlauf, S. N. and L. E. Blume (2008). The new Palgrave dictionary of economics. Basingstoke, Hampshire ; New York, Palgrave Macmillan.

Finkelstein, A. and K. McGarry (2004). "Multiple Dimensions of Private Information: Evidence from the Long-Term Care Insurance Market." National Bureau of Economic Research

Finkelstein, A. and K. McGarry (2006). "Multiple Dimensions of Private Information: Evidence from the Long-Term Care Insurance Market." The American Economic Review 96(4): 938- 958.

Finkelstein, A. and J. Poterba (2004). "Adverse Selection in Insurance Markets: Policyholder Evidence from the U.K. Annuity Market." The Journal of Political Economy 112(1): 183-208.

Gauss, C. F. (1809). Theoria motus corporum coelestivium in sectionibus conicus solem ambientium. Hamburgi, Perthes et Besser.

Heckman, J. J. (1979). "Sample Selection Bias as a Specification Error." Econometrica 47(1): 153-161.

Israel, M. (2007). "Do We Drive More Safely When Accidents are More Expensive? Identifying Moral Hazard from Experience Rating Schemes." Under First Review at The Journal of Political Economy.

Macken, L., B. Yates, et al. (2000). "Concordance of risk factors in female spouses of male patients with coronary heart disease." Journal of Cardiopulmonary Rehabilitation and Prevention 20(6): 361.

Maddala, G. S. (1999). Limited-dependent and qualitative variables in econometrics. Cambridge [Cambridgeshire] ; New York, Cambridge University Press.

Puelz, R. and A. Snow (1994). "Evidence on Adverse Selection: Equilibrium Signaling and Cross-Subsidization in the Insurance Market." The Journal of Political Economy 102(2): 236-257.

Spivey, C. (2010). "Desperation or desire? The role of risk aversion in marriage." Economic Inquiry 48(2): 499-516.

Theil, H. (1958). Economic forecasts and policy. Amsterdam, North-Holland Pub. Co.

Wilson, S. (2002). "The health capital of families: an investigation of the inter-spousal correlation in health status." Social Science & Medicine 55(7): 1157-1172.

Wooldridge, J. M. (2002). Econometric analysis of cross section and panel data. Cambridge, Mass., MIT Press.

Working, E. J. (1927). "What Do Statistical "Demand Curves" Show?" The Quarterly Journal of Economics 41(2): 212-235.

Wright, P. G. (1928). The Tariff on Animal and Vegetable Oils. New York, Macmillan.

Zohoori, N. (1997). "Does Indignity Matter? A Comparison of Empirical Analysis with and without Control for Endogeneity." Annals of Epidemiology 7(4): 258-266.