219B – Final Exam – Spring 2007

Question #1

In this first Question we consider the deductible choice in the home insurance industry, as in Sydnor (2006)'s paper. A home insurance contract is characterized by a premium P and a deductible level D. The home insurance contract offers two possibilities (to simplify): a high-deductible contract (P_{Hi}, D_{Hi}) and a low-deductible contract (P_{Lo}, D_{Lo}), with $D_{Hi} > D_{Lo}$ and $P_{Hi} < P_{Lo}$. The agent has wealth W and a utility function $U(C)$, where C is the amount of wealth left over after paying the premium and the (eventual) losses after the deductible. Finally, the probability of an accident is π and the loss in case of an accident is $L > D_{Hi}$.

a) Assuming expected utility, derive the condition under which the agent prefers the low-deductible to the high-deductible contract (assume that the probability of accident and the loss are independent of the deductible chosen).

b) Linearize now the utility function around W using the first-order Taylor approximation $U(C) = U(W) + U'(W)(C - W) + o(C - W)$. Neglect the term $o(C - W)$. Show that this implies that the agent chooses the low-deductible contract if (ad only if)

$$\pi (D_{Hi} - D_{Lo}) \geq P_{Lo} - P_{Hi}. \qquad (1)$$

c) Consider Table 2a from Sydnor (2006). Consider the consumers that choose a \$500 deductible (D_{Lo}) over the \$1,000 deductible (D_{Hi}) (first row) [Neglect for now the existence of the \$250 deductible] For them, (1) must hold. Fill in the average observed values for $\pi, D_{Hi} - D_{Lo}$ and $P_{Lo} - P_{Hi}$ in equation (1). Is equation (1) satisfied? Argue that you reject the null hypothesis of approximate risk-neutrality.

Table 2. Potential Savings with the $1000 Deductible

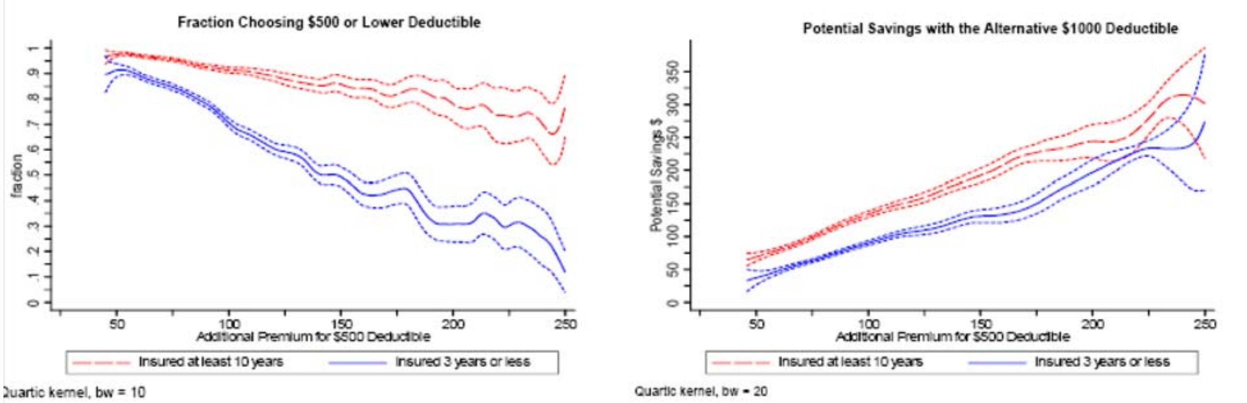
2a. Full Sample

Chosen Deductible	Number of claims per policy	Increase in out-of-pocket payments *per claim* with a $1000 deductible	Increase in out-of-pocket payments *per policy* with a $1000 deductible	Reduction in yearly premium per policy with $1000 deductible	Savings per policy with $1000 deductible
$500	0.043	469.86	19.93	99.85	79.93
N=23,782 (47.6%)	(.0014)	(2.91)	(0.67)	(0.26)	(0.71)
$250	0.049	651.61	31.98	158.93	126.95
N=17,536 (35.1%)	(.0018)	(6.59)	(1.20)	(0.45)	(1.28)

d) You are in a room full of economists. A first economist looks over Table 2a from Sydnor (2006) and says: "Well, of course, you implicitely assumed risk-neutrality. This puzzle is easily explained if we allow for risk-aversion." Debate this assersion.

e) A second economist says: "Well, maybe people did not have the right prior about the probability of an accident. Surely, the phenomenon is not there for people that have been there long enough." Debate this assertion. Refer to Figure 4 in your discussion.

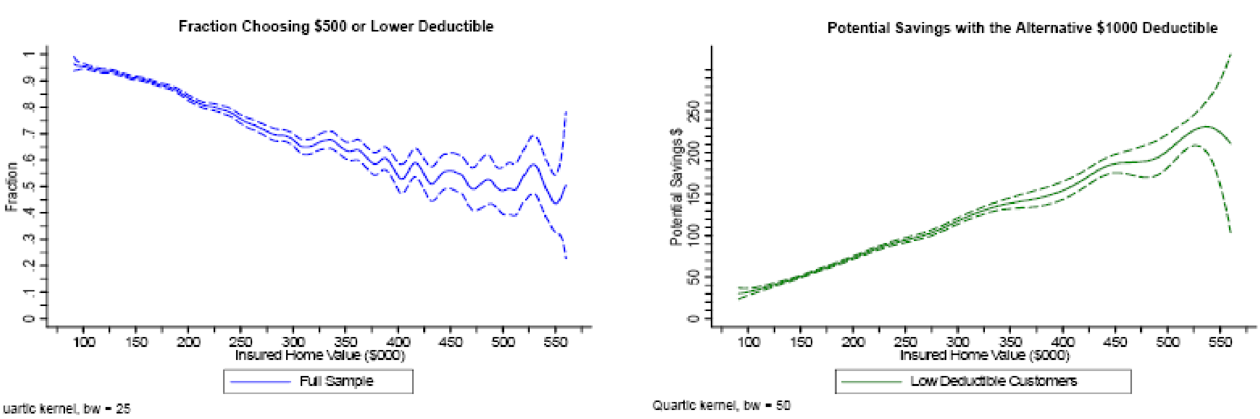
Figure 4. Deductible Choice as a Function of Additional Premium: Split by Years Insured with Company



f) What do you think is the 'right' explanation for why the puzzle is even larger for consumers that have been insured for at least 10 years?

g) A third economist says: "Surely this is due to credit constraints." Debate this assertion. If it helps, you can refer also to Figure 2.

Figure 2. Deductible Choice as a Function of Insured Home Value



h) We now consider a reference-dependent utility model of the above decision (Read on to the next point to have a full picture). Assume that the value of an insurance contract $V(P, D, \pi)$ is given by $v(-P) + w(\pi) v(-D)$ where $w(\pi)$ is the probability weighting function, and $v(x)$ is the value function according to prospect theory. Assume a piece-wise linear value function:

$$v(x) = \begin{cases} x & \text{if} \quad x \geq 0 \\ \lambda x & \text{if} \quad x < 0 \end{cases}$$

(the reference point here is zero) Write the condition under which the agent prefers the low-deductible to the high-deductible contract and simplify it. Given what you know about λ and $w(\pi)$, does this formulation of reference-dependent preferences explain partially or totally the finding in Table 2a?

i) As you are explaining this to the crowd of economists, the lone behavioral economist in the audience interrupts you and says "You really should not suffer loss-aversion for paying a premium, since you expect to pay a premium; you suffer loss aversion only from unexpected losses if an accident occurs". Discuss the reasonableness of this assumption in light of what you know of models of reference points as rational expectations (Koszegi and Rabin, 2006) Follow this suggestion, and repeat the steps in point (h). Does this version explain partially or totally the finding in Table 2a?

j) What identification strategy does this paper follow in terms of the taxonomy we introduced in class? Discuss briefly this strategy and one more paper in which it is used.

Question #2

Studies of peer effects (how interacting with others affects your decisions) are very common. But rarely can one separate the multiple components of the possible peer effects: (i) common shocks; (ii) exchange of information; (iii) social pressure, and so on. Here we consider the contribution of Mas and Moretti (2006). Mas and Moretti examine peer effects in productivity at check-out counters in a super-market chain.

a) Consider Column 1 in Table 2 in Mas and Moretti (2006). (Reading the notes is a good idea here). Describe the empirical specification and interpret quantitatively the finding.

Table 2: The effect of changes of average co-worker permanent productivity on reference person current productivity

	(1)	(2)	(3)	(4)	(5)	(6)
Δ Co-worker permanent Productivity	0.176 (0.023)	0.159 (0.023)			0.160 (0.026)	0.261 (0.033)
Δ Co-worker permanent productivity$_{t+1}$					-0.010 (0.026)	
Δ Co-worker permanent prod. × Above average worker						-0.214 (0.046)
Entry of above average productivity worker			0.011 (0.001)			
Exit of an above average productivity worker			-0.005 (0.001)			
Shift entry of above average productivity worker				0.006 (0.002)		
Shift exit of an above average productivity worker				-0.006 (0.002)		
Observations	1,734,140	1,734,140	1,734,164	1,734,164	1,356,643	1,734,140
Controls	No	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors clustered by store × calendar date × checker are in parentheses. The units of observation are checker × ten minute cells. The dependent variable is the change in the log productivity of a checker across ten minute periods. Individual productivity is defined as the number of items scanned per second over a ten minute period. The change in co-worker permanent productivity is computed as the simple average of co-worker permanent productivity components estimated by fitting equation (4) to the data. Controls are the change in the number of workers on duty in ten minute intervals (except columns 3-4), and 588 10 minute time interval by day of week dummies. In column (3)-(4), we also include in the models dummies for whether there was any entry or exit into or out of the shift, irrespective of the productivity of the worker entering or departing. Entry is 1 if in a given day a checker was on duty at time *t* but not on duty at time *t-1*, and 0 otherwise. Exit is 1 if in a given day a checker was on duty at time *t* but not on duty at time *t+1*, and 0 otherwise. Shift entry is 1 if an observation is a checker's first in a calendar date, and 0 otherwise. Shift exit is 1 if an observation is a checker's last in a calendar date, and 0 otherwise.

b) Why do the authors use the Co-Worker Permanent Productivity? Why not use the Co-Worker contemporaneous productivity?

c) In Column 2 the authors add a set of controls. Why is it helpful to know that adding controls does not change the coefficient much relative to Column 1?

d) Comment on how the standard errors are clustered. Give an example on a type of correlation that these standard errors allow and a type that they rule out.

e) How do you interpret the asymmetric effect in Column 3?

f) Interpret the finding in Column 5 (the ΔCo-Worker Permanent Productivity$_{t+1}$ refers to the change in Productivity occurring next period)

g) Outline two possible explanations of the findings in Table 2.

h) Table 6 exploits the geographical location of co-workers. Workers 'behind' are workers that are facing worker i and hence can see his productivity. How does the result in Table 6 help separate explanations?

Table 6: The effect of changes of average co-worker permanent productivity on reference person current productivity; Models by spatial orientation and proximity

	(1)	(2)	(3)	(4)	(5)
Δ Co-worker permanent productivity behind	0.233 (0.019)				
Δ Co-worker permanent productivity in front	0.007 (0.018)				
Observations	1,660,312	1,734,164	1,501,555	1,734,164	1,734,164

Notes: Standard errors clustered by store × calendar date × checker are in parentheses. The units of observation are checker × ten minute cells. The dependent variable is the change in the log productivity of a checker across ten minute periods. See note for Table 2 for a description of the dependent variable. All models include controls for the change in the number of workers on duty in a ten minute interval, and 588 ten minute time interval by day of week dummies. "Behind" refers to workers that are facing worker *i*. "In front" refers to workers that *i* is facing. "Closer" refers to workers that are one or two positions away from *i*. "Farther" denotes workers that are three or four positions away from *i*. The permanent productivity averages are taken over the indicated sub-groups. For example, Δ co-worker permanent productivity in front & closer denotes the change in the average permanent productivity of co-workers that worker *i* is facing are who are one or two registers away from *i*. In this example, if there are no workers that are positioned one or two positions in front of *i* in both *t-1* and *t*, then this change is coded as 0. The variable "Δ Worker" behind denotes the change in the presence of a worker for whom *i* is in their line-of-sight. In column (5), "Average FE" denotes the average permanent productivity of the workers who have entered or exited the checkout stands.

i) What identification strategy does this paper follow in terms of the taxonomy we introduced in class? Discuss briefly this strategy and one more paper in which it is used,

Question #3

In this Question we consider a simple model of inattention and its application to some of the papers we discussed in class. Asume that a commodity's price is given by $P + p$, where p is the part of the price that is less salient (think of the tax or of the shipping costs). An inattentive agent perceives the price to be $P + \theta p$, where $\theta \in [0, 1]$ captures the degree of inattention.

a) Consider first the Hossain and Morgan (2007) paper on eBay auction with varying levels of the shipping cost p. (Neglect the role of the reserve price) Assume that bidders have independent private values. Consider first fully attentive bidders ($\theta = 1$) Since eBay is essentially a second-price auction, argue that an attentive bidder with value v should bid $b^* = v - p$ (the bid does not include the shipping cost). What is the revenue raised by the seller (the revenue includes the shipping cost)?

b) Re-compute the optimal bidding and the revenue raised for the case of inattention ($\theta \in [0, 1]$).

c) Table 3 in Hossain and Morgan (2007) presents the revenue raised for treatment A (zero shipping cost, $p = 0$) and treatment B (high shipping cost $p = 3.99$). Using the information on the average revenue raised, provide an estimate for $\hat{\theta}$.(exclude the unsold item).

Table 3. Revenues from Low Reserve Treatments

CD Title	Revenues under Treatment A	Revenues under Treatment B	B - A	Percent Difference
Music	5.50	7.24	1.74	32%
Ooops! I Did it Again	6.50	7.74	1.24	19%
Serendipity	8.50	10.49	1.99	23%
O Brother Where Art Thou?	12.50	11.99	-0.51	-4%
Greatest Hits - Tim McGraw	11.00	15.99	4.99	45%
A Day Without Rain	13.50	14.99	1.49	11%
Automatic for the People	**0.00**	9.99	9.99	
Everyday	7.28	9.49	2.21	30%
Joshua Tree	6.07	8.25	2.18	36%
Unplugged in New York	4.50	5.24	0.74	16%
Average	*7.54*	*10.14*	*2.61*	*35%*
Average excluding unsold	*8.37*	*10.16*	*1.79*	*21%*

d) Table 4 in Hossain and Morgan (2007) presents the revenue raised for treatment C (shipping cost $p = 2$) and treatment D (high shipping cost. $p = 6$). Using the information on the average revenue raised, provide an estimate for $\hat{\theta}$.(exclude the unsold item). Provide an explanation for why $\hat{\theta}$ may be lower in this case (other than because of sampling error)

Table 4. Revenues from High Reserve Treatments

CD Title	Revenues under Treatment C	Revenues under Treatment D	D - C	Percent Difference
Music	9.00	8.00	-1.00	-11%
Ooops! I Did it Again	**0.00**	**0.00**	0.00	
Serendipity	12.50	13.50	1.00	8%
O Brother Where Art Thou?	11.52	11.00	-0.52	-5%
Greatest Hits - Tim McGraw	18.00	17.00	-1.00	-6%
A Day Without Rain	15.50	16.00	0.50	3%
Automatic for the People	**0.00**	**0.00**	0.00	
Everyday	10.50	13.50	3.00	29%
Joshua Tree	8.00	11.10	3.10	39%
Unplugged in New York	8.00	**0.00**	-8.00	-100%
Average	*9.30*	*9.01*	*-0.29*	*-3%*
Average excluding unsold	*12.15*	*12.87*	*0.73*	*6%*

e) What identification strategy does this paper follow in terms of the taxonomy we introduced in class? Discuss briefly this strategy and one more paper in which it is used.

Question #4

Consider the problem of 401(k) enrollment. This problem is based on Problem Set 1—As such, I am providing less guidance to the solution. Compared to the alternative activity, which has payoff 0, enrolling in a 401(k) has payoff $-k < 0$ at time t (the present) and payoff $b > 0$ for all periods from $t + 1$ on. ($t + 1$ included). The individual has to choose when to undertake the investment activity, that is, at t, at $t + 1$, at $t + 2$, etc. (The individual can also decide not to do it, which we define as doing it at $t = \infty$) Assume that both k and b are deterministic.

a) Solve for the investment decision for a time-consistent individual ($\beta = \hat{\beta} = 1$).

b) Solve for the investment decision for a fully naive present-biased individual ($\beta < \hat{\beta} = 1$).

c) Consider employees of a company with annual income of \$40,000 and a company match of 50 percent up to 6 percent contribution. Calibrate the model for the time-consistent and naive present-biased individual. Make all the assumptions you need about the parameters.

d) Compare these calibrations to the findings in Madrian and Shea (2001) as per the next Figure. Discuss.