(M)oral Hazard?*

Erik Grönqvist**
Stockholm School of Economics

November, 2006

SSE/EFI Working Paper Series in Economics and Finance No. 642

Abstract

Would you go to the dentist more often if it were free? Observational data is here used to analyze the impact of full-coverage insurance on dental care utilization using different identification strategies. The challenge of assessing the bite of moral hazard without an experimental study design is to separate it from adverse selection, as agents act on private and generally unobservable information. By utilizing a quasi-experimental feature of the insurance scheme the moral hazard effect is identified on observables, and by having access to an instrument the effect is identified with IV. Moral hazard is assessed using both difference-in-differences and cross-sectional estimations.

Keywords: Asymmetric information, Moral Hazard, Health Insurance, Porpensity Score Matching, IV.

JEL codes: D82, G22, I11.

*This paper has benefited from comments by Magnus Johannesson, Michael Lechner, Douglas Lundin, Mark Pauly, Alexander Spearmann, and seminar participants at Stockholm School of Economics, Stockholm University, Uppsala University, 2nd Conference on Evaluation Research in Mannheim 2004, 26th Nordic Health Economics Study Group Meeting, and EEA congress 2006. Per Johansson, Jeffrey Smith and Niklas Zethraeus have made suggestions on the estimations. I am grateful to the National Dental Service in Värmland, Sweden, especially to Jörgen Paulander, for providing data and odontological expertise, and to Ann-Britt Emilsson for generously answering questions on registration practices. All remaining cavities in the analysis are mine. Financial support from Finanspolitiska Forskningsinstitutet is gratefully acknowledged.

**Centre for Health Economics, Stockholm School of Economics, P.O. Box 6501, SE-113 83 Stockholm, Sweden, e-mail: erik.gronqvist@hhs.se

1 Introduction

Presence of asymmetric information is potentially a large problem in insurance markets, especially when insuring health risks. If agents have private information on their behavior or realized health state, insurance contracts may lead to moral hazard. When an agent takes out an insurance contract his incentives for precautionary activities are reduced, increasing the risk level accordingly. In health insurance, there is also an important ex-post moral hazard effect, as suggested by Pauly (1968). Insurance coverage reduces the marginal price of services, so an agent will demand more comprehensive care than a priori agreed upon, once he becomes ill. Whether moral hazard is a problem depends on the extent of asymmetric information and whether agents act on informational asymmetries. Ultimately, this is an empirical question.

Assessing the presence of moral hazard in an insurance setting can be characterized as a program evaluation problem, where treatment refers to coverage by the insurance scheme being evaluated. Now, treatment is not random, but the result of a self-selection process giving rise to a potential endogeneity problem. Adverse selection gives a theoretical understanding of the endogeneity, and contract theory suggests that both the treatment-effect of insurance coverage (i.e. moral hazard) and the selection bias come from the same source: asymmetric information. Agents taking out a health insurance policy may have higher health care expenditures because of moral hazard but, at the same time, agents with high risk are more likely to buy insurance coverage, due to adverse selection. The methodological problem in the empirical literature on asymmetric information is to determine causality, i.e. the selection problem needs to be resolved in order to identify unbiased estimates of the moral hazard effect. The fundamental problem is that the private information on which agents act, is generally unobservable to the researcher. A second problem is that micro data from insurance records only covers treated agents from the time they enter the program, thus making the construction of a counterfactual difficult.

In this paper observational data is used to assess the moral hazard effect of a voluntary full-coverage dental insurance introduced in Sweden 1999. Data is sufficiently rich to overcome the selection problem, and offers three different strategies to identify the causal effect of insurance coverage.

The introduction of the insurance can be characterized as a quasi-experiment since it was unexpected from the patients' perspective and since no voluntary insurance was available to this population prior to 1999, thereby making patients' dental health and prior dental consumption exogenous to the launch of the insurance. As a result, the set up of the insurance and the recurrent characteristic of dental problems makes it possible to observe a proxy for private information on dental risk. (i) With access to a proxy for private information, the source of selection can be controlled for directly in cross-sectional estimations. (ii) Data on dental consumption can be observed both within the program and during the exogenous pre-program period. This panel data feature enables selection to be netted out using a difference-in-differences approach, after having controlled for differing trends. The treatment effect is estimated with nearest neighbor propensity score, both in the cross-sectional and in the difference-in-differences setting.

Further, the decision to buy insurance is highly influenced by one's dentist. If dentists' attitudes to insurance are conditionally independent to their dental care practice styles, this influence can be used as an instrument. (iii) Having access to an instrument the treatment effect can be identified with IV, in the event of remaining endogeneity.

In spite of the methodological problems, there are unequivocal experimental evidence of a moral hazard effect in health insurance from the RAND health insurance study; see for example, Manning, Newhouse, Duan, Keeler, Leibowitz, and Marquis (1987), Manning, Benjamin, Bailit, and Newhouse (1985) and Newhouse (1993).

Advances in studying moral hazard with observational data include Cardon and Hendal (2001) who estimate a structural model of employer-provided health insurance and utilizing the quasi-experimental feature that the menu of contracts from which employees can choose differs across employers, and Abbring, Chiappori and Pinquet (2003) who estimate the dependence of claims occurrence in dynamic model of experience rated automobile insurance, where a negative contagion indicates moral hazard since past accidents increase the marginal cost of new accidents. A similar strategy with dynamic data is also used by Israel (2004). I add to this literature by relating the identification of moral hazard—separating it from adverse selection—to the program evaluation literature.

Further, in most empirical work on health insurance, providers have direct incentives to expand the amount of services, obscuring moral hazard with supply-induced-demand. In the setting at hand, however, providers have no private economic incentive to induce demand, making it possible to directly observe the moral hazard effect generated by patients. If moral hazard is present for dental care—services which most people feel discomfort consuming—it could be widely expected for other types of health care services.

The next section present the methodology used to assess the moral hazard effect, controlling for selection bias. Section three describes the empirical setting of the dental insurance quasi-experiment and the data. Results of the estimations are presented in section four, and is then followed by a concluding discussion.

2 Empirical Strategy

At any point in time, an agent can only be observed in one state, either with or without insurance. Thus, in the program period, $t+1$ (the period when agents are offered to purchase insurance) dental consumption of policyholders cannot be observed in the non–insurance state, i.e. $Y_{0t+1}|D=1$ in figure 1. A policyholder must therefore be compared to an individual without coverage, possibly himself in a different period. The problem in evaluating moral hazard using non-experimental data is that agents purchasing insurance are generally different from those choosing to remain without coverage, which makes it easy to confound moral hazard with adverse selection. In terms of figure 1, dental care consumption for patients without insurance, $Y_{0t+1}|D=0$, is not directly comparable to the utilization policyholders would have outside coverage, $Y_{0t+1}|D=1$, where Y_{0j} is the non-treatment outcome in period j and where $D=1(=0)$ denotes (non-) policyholders. An appropriate counterfactual is therefore needed to identify the causal effect of insurance coverage; i.e., the moral hazard effect.

The moral hazard effect will be estimated with propensity score matching using both cross-sectional and conditional difference-in-differences identification. In the event of remaining endogeneity the effect is identified with IV.

[Figure 1 about here]

2.1 Cross-Section Model

One way of finding a counterfactual is to use information from non-treated agents in the program period. The cross-sectional model is identified if the non-treatment outcome is the same for both participants and non-participants, and the analysis therefore needs to take into account that agents with private information of high risk may be more likely to purchase insurance; ie., that policyholders are not directly comparable to agents without coverage.

Chiappori and Salanié (2000) and Dionne, Gouriéroux and Vanesse (2001) suggest a general framework for testing the presence of asymmetric information in a cross sectional setting based on a positive correlation property (also discussed by Chiappori, Jullien, Salanié and Salanié, 2006). The outcome of the insurance, Y, e.g. dental care expenditures, can be modelled as a function of X; observable risk factors. The decision to purchase insurance, D, is also a function of observable risk factors X, since this is the information used by the insurer to price contracts. That is

$$Y = f(X) + \varepsilon, \tag{1}$$

$$D = 1[g(X) + \eta > 0]. \tag{2}$$

Hence, conditional on all information available to the insurer, X, presence asymmetric information implies that $cov\,(\varepsilon, \eta) > 0$. The causal direction in this relation is not known from this test, only the presence of asymmetric information. It is stressed that the X-vector must exhaust all information used by the insurer, so as to avoid a spurious relation in $cov\,(\varepsilon, \eta)$.[1]

In order to identify the moral hazard effect in this setting, however, selection on unobservables must be eliminated. Now, contract theory gives that the unobserved component determining selection is private information on risk, P, which is also an unobservable variable in the outcome equation. Therefore, by appending P to the X-vector in equations (1) and (2) the unobserved selection would be eliminated so that any remaining dependence, $cov(\varepsilon, \eta) > 0$, would be due to moral hazard. That is, if private information on risk, P, were available to the researcher, the moral hazard effect could be estimated in a cross-sectional setting.

The dental insurance in Värmland, Sweden, provides a setting where past dental consumption is a proxy for private information on risk (see section 3). With access to private, P, and public, X, information on dental risk, any influence of adverse selection can be eliminated. Conditional on P and X, the difference in outcome between the insurance and the non-insurance state is independent of treatment

$$(Y_{1t+1}, Y_{0t+1}) \amalg D \mid X, P, \tag{3}$$

and the cross-sectional model thereby identifies the average treatment effect of insurance coverage for those treated.

2.2 Difference-in-Differences Model

Another way of studying the treatment effect is to analyze how dental care consumption changes as an individual enters the insurance. By observing the same agent over time, and in two states, the pre-program utilization can be used as a counterfactual; i.e., $Y_{0t}|D = 1$ in figure 1. In order to identify moral hazard, however, the non-treatment outcome must be the same in both periods. The problem here, is that dental health deteriorates continuously with age, so an agent's expected dental care utilization is increasing over time, thereby implying a trend component; that is,

$E(Y_{0t}|D=1) < E(Y_{0t+1}|D=1)$. A before-after comparison will therefore overestimate the moral hazard effect.

The differences in utilization levels across periods can be corrected by using information from non-participants. Since non-participants are observed in the non-treated state in both periods, the trend can be identified, given that the mean difference in non-treatment outcome is the same for the two groups

$$E(Y_{0t+1} - Y_{0t}|D=1) = E(Y_{0t+1} - Y_{0t}|D=0). \tag{4}$$

With adequate panel data the treatment effect on the treated can be estimated with difference-in-differences by taking the mean difference, of the program and pre-program period outcomes, between participants and non-participants

$$E(Y_{1t+1} - Y_{0t+1}|D=1) = E(Y_{1t+1} - Y_{0t}|D=1) - E(Y_{0t+1} - Y_{0t}|D=0), \tag{5}$$

where the identifying assumption is that equation (4) holds.

However, dental health may not deteriorate at the same rate for policyholders and non-policyholders. Agents with a rapid decline in dental status, i.e. a steeper trend, may be more inclined to purchase dental insurance. Now, given that caries can be characterized as a deteriorating lifelong infectious disease and that the propensity for teeth loosening increases with age, the trend in dental care utilization will be related to current dental health. Differences in trends between treated and non-treated agents are therefore attributable to the current dental care risk level; and is captured by the public information on risk, X. That is, the trend is assumed to be independent of insurance status conditional on the covariate vector X[2]

$$(Y_{0t+1} - Y_{0t}) \amalg D \mid X. \tag{6}$$

As the dental insurance in Värmland provides access to both X and P (see section 3), the moral hazard effect on the treated can be estimated with conditional difference-in-differences; that is, any adverse selection in cost level is netted out by taking differences, while any selection on trends is assumed to be based on observables.

2.3 Matching Estimation

The cross-section and difference-in-differences models are implemented with a one-to-one nearest neighbor propensity score matching estimator.

The idea with propensity score matching is to gauge the dental care utilization of a policy holder against that of an otherwise comparable person without insurance.

Matching rests on the conditional independence assumptions (CIA)—equation (3) and (6)—stating that conditional on the variables jointly affecting selection and outcome, the potential outcomes are independent of treatment status. This means that all selection on observables is removed, making the sample random conditional on the confounding factors. Hence, for agents with the same covariates, the difference in outcome due to insurance status is solely attributable to moral hazard. The CIA implies that moral hazard effect is non-parametrically identified (See for example Heckman, Lalonde and Smith, 1999).

To reduce the dimensionality of the heterogeneity in covariates, agents are matched on basis of propensity scores, $s(I)$, i.e. the probability of treatment conditional on the covariates, where $I = (X, P)$ in the cross-sectional model and $I = X$ in the difference-in-differences model. Rosenbaum and Rubin (1983) showed that given that the CIA and the common support property[3] hold, conditioning on $s(I)$ is equivalent to conditioning on the full I-vector. Matching on propensity scores, $s(I)$, instead of I

facilitates the implementation of matching.

The matching protocol is implemented as follows. A probit estimate is first used to calculate propensity scores for both treated and non-treated agents. Each treated agent i is then matched to the agent j in the non-treated sub-sample with the closest propensity score. As agents i and j are comparable in all relevant aspects specified by the covariates, any remaining difference in outcome is attributable to treatment, i.e. moral hazard. Each control unit can be used to construct more than one match.[4]

2.4 Instrumental Variable Estimation

The moral hazard effect is identified with matching if all relevant variables determining selection are included in the analysis. Now, if P does not capture all private information in the cross-sectional setting, or if X does not capture variation in the trend, then the CIA conditions, equations (6) and (3), will not hold. With selection on unobserved variables not captured by the data, the matching estimator will produce biased results.

Selection on unobservables can be characterized as endogeneity, $E(\varepsilon|D) \neq 0$ in a regression framework $Y = f(I, D, \varepsilon)$, where ε is the error term and I the observable risk. In the cross-sectional model $I = (X, P)$ and $Y = Y_{.t+1}$, and in the difference-in-differences model, $I = X$ and $Y = (Y_{.t+1} - Y_{.t})$. The endogeneity problem can be solved by instrumenting for insurance status, balancing out the bias in the sample. To implement IV, a good and valid instrument, Z, is needed. Differences in attitudes towards the dental insurance in Värmland across dentists will be used as an instrument (see section 3).

3 Empirical Setting

The public dental insurance in Sweden, covering all individuals from the year they turn 20, has become less generous over time. A gradual decrease in the level of subsidy and the introduction of a linear coverage have made individuals more exposed to the risk of high dental care costs.

In order to reduce this risk exposure, the National Dental Service in the region of Värmland, Sweden, introduced a voluntary dental insurance in January 1999, supplementing the public insurance. All patients were offered to *subscribe* to dental care; that is, at a fixed annual fee, a subscription contract provides free dental services during a two-year contract period. The dental care subscription is, in effect, a full-coverage voluntary dental insurance provided by a public monopolist.

A key idea with this voluntary insurance scheme is to reduce dental costs by substituting reparatory treatment for preventive services, and services provided by less expensive staff categories. When an agent signs up for the insurance, he accepts to comply with a prevention program, provided by dental hygienists, to counteract ex-ante moral hazard. Access to dental care, within the insurance, is also restricted and provided on the basis of odontological needs, in order to reduce ex-post moral hazard. It is also important to note that dentists within the National Dental Service are employed with a fixed salary and have no private economic interest in providing services.

The price of a contract is set after an oral examination of the patient, which evaluates dental risk in four dimensions (general risk, technical risk, caries risk, and parodental risk) and for each dimension, there are 6 to 8 risk indicators, where each indicator is gauged on a four-graded scale. Based on the sum of these scores patients are clustered into one of 16 risk classes. Contracts are priced according to risk class, with annual prices in 1999 ranging from 295 SEK (32 EURO) for the lowest risk class to 11 000 SEK (1 200 EURO) for the highest. The dental service in Värmland only uses these risk indicators when assessing risk, and does not take explicit account of realized dental costs.

The risk classification is partly used by the dental service in Värmland to assess the dental status of the population. Information on risk classes is therefore available for those who choose to purchase insurance as well as for non-purchasers.

3.1 A Dental Insurance Quasi-Experiment

An interesting feature of this insurance is that it was offered to a population of patients who, prior to 1999, had not been able to buy supplementary dental coverage. Hence, any variation in dental care risk across agents is not generated by differences in prior insurance status. The potential for moral hazard in anticipation of the insurance was also negligible, as the introduction was unexpected from the patients' perspective. Officials with the dental service in Värmland state that the opportunity to purchase the insurance was mainly brought to the individuals' attention by their dentist during check-up visits after January 1st 1999. This makes patients' prior dental consumption and dental health unaffected by the choice of insurance in 1999, and the introduction of the insurance can be characterized as a quasi-experiment in so far as it provides an exogenous variation in dental risk within each risk class. A randomization of dental risk across agents—if it were possible—at the time when the insurance was introduced would generate similar circumstances; that is, differences in risk are generated by factors exogenous to the launch of the insurance.

The main dental problem, caries, can be characterized as a life-long infectious disease. Caries arises when bacteria on the teeth produce acid, gradually dissolving the enamel on the surface of the teeth. Eventually this may result in a cavity; that is, the bacteria will penetrate the tooth and inflame the pulp. Every time we eat something, the harmful acid is produced and will be active for around half an hour. Caries is a relatively slow process and with preventive activities it can be stopped or even reversed at an early stage, e.g. with tooth brushing, dental flossing and fluoride rinsing.

If a person has a history of prior caries, he has a larger probability of getting new problems. Bacteria will grow more easily if the enamel has already been coarsened by prior caries, or grow in the seam between a prior filling and the tooth. Moreover, a filled tooth will need future maintenance or replacement, and is also more fragile. Consequently, the consumption of dental care has been observed to be highly correlated over time; see, for example, Powell (1998). Past dental consumption is therefore a good indicator of dental risk, and a predictor of future costs.

If agents in Värmland have private information about their dental risk, this could consequently be proxied with their past dental consumption, since the dental service in Värmland does not explicitly use realized dental costs in its risk assessment. Table 1 shows that past consumption of dental care, measured in terms of costs, rises with higher risk classes. Now, even if patients are clustered into 16 risk classes, there is still heterogeneity in risk among patients within each risk class, and since prior use of dental care is a predictor of future usage it would capture intra-group variation in risk. Hence, the impact of asymmetric information in the decision to purchase insurance can be observed. More specifically, dental cost in the two years preceding the offer to buy insurance is used as a proxy for private information.

An important question is whether past dental consumption captures private information on risk, given that oral examinations are used to price contracts. As a validity test of the proxy for private information, dental costs for the two years 2000 to 2001 are regressed on dental costs during the two preceding years (1998 to 1999) and on dummy variables for each risk class; for details see Grönqvist (2006).[5] Past dental costs alone explain 9 percent of the variation in dental costs during the subsequent period, while risk classes alone explain 11 percent of the variation. When both past dental costs and risk classes are used as regressors, 14 percent of the variation in dental costs 1999 to 2001 are explained. Hence, not all of the information contained in past dental consumption is captured by the risk classification system. Past dental consumption captures an additional 27 percent of the variation in future dental costs—not captured by the risk classification system—and can therefore be used as a measure of private information on which to act.

3.2 Data

Data comes from an administrative database on dental care. The sample consists of those patients who, in 1999, were given an offer to subscribe to dental care, and where a date for the offer can be determined.[6] Patients need to visits their dentist for a check-up visit in 1999, and also need to be registered with the National Dental Service in Värmland during the period 1997 through 2001, so that their consumption of dental care can be observed for a two-year period both before and after the offer.

The sample consists of 19407 patients of whom 33 percent, or 6400 individuals, chose to purchase an insurance contract. Table 1 shows that policyholders are mainly clustered in the low and middle risk classes, with 90 percent of the policyholders belonging to the eight lowest risk classes. None of the policyholders in the sample belong to any of the three top risk classes, and these risk classes are therefore excluded from the analysis. In every risk class, policyholders have higher dental costs in the program period than in the pre-program period, giving a first indication of moral hazard using only a simple before-after comparison.

[Table 1 about here]

3.2.1 Variables

The Y-variable in the analysis, *Post Cost*, contains dental costs during the two-year period following the insurance offer.[7] In the difference-in-differences estimations the dependent variable, $Y_{\cdot t+1} - Y_{\cdot t}$, is the difference between *Post Cost* and *Previous Cost*, where *Previous Cost* is the dental costs during the two-year period immediately preceding the insurance offer.[8] Dental costs can be decomposed into eight groups of procedures, *Examination*, *Prevention*, *Surgical procedures*, *Endo-surgical procedures*, *Preservatory treatment*, *Fixed dentures*, *Removable dentures*, and *Acute visits.* Descriptive statistics are reported in table A.1 in the Appendix.

As a robustness test, the analysis is also carried out with dental care utilization measured as visits instead of costs. Since different types of visits are not comparable from a resource perspective, visits will be analyzed separately for each of the eight decomposed group of procedures (see table A.2 in the Appendix for descriptive statistics on visits).

The X-vector—capturing the public information on risk used by the insurer to price insurance contracts—contains a dummy variable for each risk class. In addition, the X-vector also contains a dummy variable for each of the 43 dental care clinics within the dental service in Värmland. Treatment styles may vary across dental clinics due to praxis variation, as do agents' inclination to purchase insurance. Age and gender are also included in the X-vector observable factors.

The P-vector of private information contains the variable *Previous Costs* (i.e. dental costs the two-years proceeding the insurance offer). Grönqvist (2004b) finds that the impact of past dental consumption on the decision to buy insurance varies across risk classes, indicating private information on preferences as discussed by Finkelstein and Poterba (2006). To capture this heterogeneity in response to private information—caused by private information on preferences—a separate variable is used for each risk class. To this end, an interaction variable, *Dummy risk class g * Previous costs*, is included for each risk class $g = 1, 2, ..., 13$.

The Z-variable—i.e. the instrument for insurance coverage—utilizes the fact that a patient's inclination to purchase insurance is influenced by the attitudes of his dentist. Each patient within the National Dental Service is registered with a specific dentist providing all basic dental care, and the process of being allocated to a dentist is essentially random. Few individuals change dentists within the National Dental Service unless they move. Still, the inclination to purchase insurance varies substantially across dentists. This variation must therefore reflect the attitudes of the attending dentist towards the insurance.[9] Attitudes to whether the National Dental Service should run a private insurance scheme rather than just act as a public provider of dental care (within the public insurance) are likely to vary. In Sweden, there is little, if any, experience of private health insurance schemes, and the notion of private insurances in the health care sector is controversial. Given a certain conformity in treatment styles within each clinic, the attitudes of dentists will not have any direct effect on dental care utilization.

To capture the variation in attitudes, each *Dentist's share* of patients with insurance is used as an instrument.[10] The validity of the instrument relies on the assumption that a dentist's share of policyholders does not have any direct influence on dental care utilization, after controlling for X and P; that is, after controlling for dental risk and praxis variation at the clinic level. This assumption implicitly rules out that dentists with a taste for costly services would be more prone to market the insurance to their patients.

The dentist's share of policyholders is a good instrument. A probit of the decision to purchase insurance, $D = 1[g(X, P, Z) + \eta > 0]$, shows that the instrument has a strong influence on the decision, after controlling for X and P. The hypothesis that Z does not have any additional influence on D is strongly rejected in a LR-test: the $\chi^2(1)$-statistic is 467 and 468, respectively, for the difference-in-differences and cross-sectional model, giving p-values<0.001.

3.2.2 Validity of the Instrument

The identifying assumption in IV is that the instrument is valid, which here implies that each dentists' share of policyholders must be independent of dental costs, conditional on X, D (and P in the cross-sectional case). The validity of an instrument is essentially an untestable assumption, but in this setting, however, it is possible to perform a test based on the quasi-experimental feature of the dental insurance. That is, it is possible to test whether dentists with a large share of policyholders carried out more expensive dental care before the insurance was launched. Dental care utilization before the insurance was introduced is exogenous to the later insurance status, as the introduction constitutes a quasi experiment. This validity test depends on treatment styles being constant over time.

The validity of the instrument is tested by regressing *Previous Cost* on *Dentist's Share*, controlling for risk classification, age, gender and praxis variation at the clinic level. Table 2 report that dentists' share of policyholders is not related to dental costs prior to the insurance, indicating that it is a valid instrument.

[Table 2 about here]

4 Results

The results show evidence of moral hazard in the private dental insurance in Värmland. The estimated treatment effects of insurance coverage on the treated are reported in table 3.

[Table 3 about here]

The overall effect of insurance coverage is large: for the two-year contract period, dental costs increase with between 661 and 898 SEK depending on the estimator. The estimated effect implies dental costs to be 51 to 84 percent higher for participants than they would have been without coverage. The estimated effect for total cost is slightly larger in the cross-sectional estimates than in the difference-in-differences estimates. The Hausman-Wu test does not suggest that there is a remaining selection problem, thereby indicating that IV gives inefficient estimates.

Now, the increase in total costs may in part be driven by the way the insurance scheme is run. There is an explicit strategy to find potential dental problems in an early stage of progression, and contracted patients are therefore called on for checkups with shorter average intervals. When patients sign up for the insurance, they also agree to follow a prevention program to avoid future reparatory services. To isolate the moral hazard effect examinations and prevention must be separated from total costs. Total costs are therefore decomposed into three groups of procedures to be analyzed separately; examinations, prevention, and *other* procedures.

4.1 Examination and Prevention Costs

Insurance coverage increases the examination costs by 164 to 169 SEK (33 to 34 percent) when considering propensity score matching estimates. The increase in prevention costs is even larger, between 297 and 326 SEK (287 and 437 percent). The Wu-Hausman test indicates that there may be an endogeneity problem estimating the effect on prevention with the cross-sectional model. The IV estimate is, however, larger still and suggests that almost all prevention may be due to insurance coverage.

For examinations and prevention procedures it is difficult to separate the effect driven by the mechanics of the program from the higher demand for services following from insurance coverage, as the two effects go in the same direction.

4.2 Moral Hazard Components

Total cost minus costs for examinations and prevention—here labeled *other costs*—is a better measure of the changed incentives, since the cost components partly driven by the insurance scheme have been eliminated. Propensity score matching indicates a moral hazard effect for *other costs* in the range of 306 to 406 SEK (51 to 82 percent). The IV estimates give lower, and insignificant, estimates while the Hausman-Wu test only indicates endogeneity in the cross-sectional specification. The validity of the IV estimates depends crucially on the assumption that dentists with a preference for costly services are not more inclined to persuade their patients to purchase the insurance. When tested, the dentist's share of policyholders was a valid instrument, but the IV results must be interpreted with some care.

While the difference-in-differences specification point at a statistically, as well as economically, significant moral hazard effect, the cross-sectional setting does not provide evidence of moral hazard for *other costs*; i.e., total cost minus costs for examinations and prevention.

The treatment effect of insurance coverage can be further qualified by decomposing *other costs* into (1) reparatory services, (2) dentures and (3) acute treatment.

4.2.1 Reparatory services

The matching estimates indicate a moral hazard effect on reparatory services of between 225 and 240 SEK (46 and 51 percent). The IV estimates, however, are not significant and the Wu-Hausman statistic indicates endogeneity.

Reparatory services, in turn, consist of three different types of treatment; surgical, endo-surgical and preservatory procedures. Results for each type are reported in table A.3 in the Appendix. For surgical procedures, e.g. tooth-extraction, and endo-surgical procedures, e.g. root filling, there is an unambiguous moral hazard effect. The costs for surgical procedures increase with 11 to 23 SEK (27 and 79 percent) according to the matching estimates. The effect on endo-surgical procedures is 41 to 49 SEK (149 to 253 percent). For preservatory treatment, e.g. ordinary fillings, the moral hazard effect is less clear-cut. Costs increase with 176 to 212 SEK (42 to 55 percent) as measured by matching, while IV gives insignificant estimates, and the Wu-Hausman statistic indicates problems with remaining unobserved heterogeneity.

[Table 4 about here]

[Table 5 about here]

The treatment effect on reparatory services also differs between low risk classes—risk class 1 to 6—and high risk classes—risk class 7 to 13—see tables 4 and 5. For low risk classes, propensity score matching indicates an effect in the range of 115 to 123 SEK (40 to 44 percent), and there is no indication of problems with remaining selection. For high risk classes, the effect is 444 to 524 SEK (82 to 114 percent), but these numbers may, however, be inflated by selection: the IV estimates are not a significant and the Wu-Hausman test indicates that unobserved heterogeneity may be a problem.

The moral hazard effect for reparatory services is most distinct in low risk classes, and for surgical and endo-surgical procedures. There is less evidence of moral hazard in high risk classes and for preservatoty treatments.

4.2.2 Dentures

The next component of other costs is dentures. The difference-in-differences specification does not give any clear evidence of moral hazard, with the matching estimate at 39 SEK (76 percent) not reaching significance. The cross-sectional model does, however, indicate a significant effect of 74 (464 percent). This number is close to policyholders' average cost of dentures (94 SEK) and the pattern is similar for both fixed and removable dentures (see table A.3 in the Appendix). The Wu-Hausman test does not indicate any remaining unobserved heterogeneity, neither for the difference-in-differences nor for the cross-sectional models

The moral hazard effect on dentures differs across risk classes (see tables 4 and 5). In high risk classes, the treatment effect is considerably larger and significant, whereas in low risk classes it is only significant for the cross sectional estimate. Hence, the general pattern for dentures indicates that there may be a moral hazard effect of insurance coverage, particularly in high risk classes.

4.2.3 Acute Care

The cost of acute treatment—the last component of other costs—increases with 15 to 22 SEK (20 to 35 percent) due to insurance coverage, according to the matching estimates. The Wu-Hausman does not indicate problems with selection. Acute visits are exclusively initiated by agents, while the use of other types of dental care may be influenced by the dentist or the insurance scheme itself. The costs for acute treatment therefore give the most direct measure of the changed incentives.

4.3 Tests of Robustness and Validity

4.3.1 Pre-program Trends

In order for the conditional difference-in-differences model to capture the moral hazard effect, cost trends in the non-treated state must be the same for policyholders and non-policyholders when controlling for covariates; that is, the CIA in equation (6) must hold. The CIA is essentially an untestable assumption, but it is possible to test if trends were different before the insurance was launched. In the difference-in-differences analysis dental costs during the two-year period immediately preceding the insurance offer were used. For a subsample, however, dental costs can be traced for a longer period: for 14001 patients dental costs can be traced during three years (to 1996) and for 6974 patients during four years (to 1995). The proportion of patients purchasing insurance remain stable in the reduced samples (32 and 28 percent, respectively).

To test for differences in pre-program trends, the sum of percentage yearly cost changes—$\sum_{t=k}^{98}(Cost_{i,k}-Cost_{i,k-1})/Cost_{i,k-1}$, where k is 1997 and 1996, respectively—is regressed on the subsequent insurance status and the X-vector. The hypothesis that policyholders and non-policyholders have the same trend in the non-treated state cannot be rejected from this test, see table A.4 in the Appendix. Even if four years is a short period to establish a trend, the test still provides support for the validity of the difference-in-differences model; i.e., there is no evidence of selection on cost trends after controlling for public information on risk.

4.3.2 Common Support and Balancing Scores

The idea with propensity score matching is to remove the imbalance in covariates, where the propensity score reduces the heterogeneity of the observed covariates into one dimension. Figure A.1 shows the unmatched propensity scores to have a common support over the whole distribution of propensity scores for both difference-in-differences and cross-sectional estimates.

The balanced propensity scores reduce the imbalance of the underlying covariates. Tables A.5 and A.6 in the Appendix report diagnostics suggested by Rosenbaum and Rubin (1985); percentage bias reduction and two different t-tests for imbalance.[11] For most covariates, the imbalance is substantially reduced and statistically insignificant.

4.3.3 Estimating Selection on Observables with OLS

Propensity score matching identifies the moral hazard effect by controlling for selection on observables. An alternative would be to use OLS, which also requires confounding factors to be controlled for; i.e., the CIA in equations (3) and (6) to hold. In addition OLS (i) adds a linear model assumption and (ii) compares weighted units across the whole sample and not from the common support.

As a robustness test the treatment effect on the treated is therefore estimated with OLS; see, table A.7 in the Appendix. The OLS estimates are similar to the matching estimates both in size and level of significance, indicating that the results are robust to the choice of estimator when identifying the treatment effect on observables.[12]

4.3.4 Dentist Visits of Outcome

As a second robustness test, the treatment effect is also estimated using dental visits as the outcome variable.[13] Table A.8 in the Appendix reports the effect on all decomposed procedures, and confirms the presence of moral hazard. The effects are somewhat smaller across most procedures compared with using dental costs as outcome, but the significance levels, in turn, are higher and endogeneity seems to be less of a problem when using visits. The smaller effects may indicate that the calculated cost variable overstates the effect, but it can also be an indication of moral hazard operating both in the intensive and the extensive margin. The latter is consistent with findings in the RAND dental insurance experiment (Manning et al, 1985).

5 Discussion

Empirical studies of asymmetric information in insurance markets easily confound moral hazard with selection, since agents select on private information that is generally unobservable to the researcher. Experimental data would be ideal for studying moral hazard, but in many situations, social experiments are not feasible. Absent experimental data, the economist must resort to methods based on observational data. If the private information on which the agent acts were available, the sample would, in fact, be random conditional on observables; that is, each policyholder could be matched to an otherwise comparable agent without insurance. Differences in outcome due to selection can also be eliminated with a conditional difference-in-differences methodology, where selection in cost levels is netted out. If there is selection on unobservables, however, one can instead use an instrumental variable approach.

This paper exploits all these approaches to assess the scope of moral hazard in dental insurance. Taken together, the three identification strategies give empirical evidence of moral hazard. There is an clear increase in examinations and prevention costs as a result of coverage, but this may in part be driven by the way the insurance is run, as there is an explicit strategy for avoiding costly treatment by detecting dental problems at an early stage of progression and investing in prevention. Total costs less examination and prevention are therefore used as a measure of the changed incentives; i.e., the moral hazard effect. These costs increase with 47 to 51 percent for the difference-in-differences estimates. The estimated effect is even higher in the cross-sectional setting, but these higher estimates may be inflated by remaining selection.

For dentures there is clear evidence of moral hazard, primarily among high risk classes. It is not surprising to find moral hazard for dentures since these are very expensive services and if agents' budget constraint is binding, it is most likely to be for these services. With the marginal cost of dentures being zero, the utilization is likely to increase. It is more surprising to find moral hazard for services like tooth-extraction and root filling. For reparatory services the moral hazard effect is most pronounced among low risk classes. Dental care within the insurance is provided on the basis of odontological needs, so the utilization decision is made in interaction between dentist and agent. Note that dentists are employed by the National Dental Service and have no direct private stake in the insurance scheme. If access to services had been free, instead of being provided on basis of odontological needs, the moral hazard effect may thus have been even higher. The most direct evidence of moral hazard is the increased propensity to seek acute dental care, since the decision to initiate a contact is exclusively made by the agent. The estimated moral hazard effect for acute services is between 15 and 22 percent.

Irrespective of whether the increased utilization of dental care is driven by changed incentives or the mechanics of the program, it may induce inefficiency, given that agents are not liquidity constrained. As argued by Pauly (1968), agents receive services that they would not have voluntarily purchased outside the insurance, and this over-consumption implies an inefficiency. If this loss is larger than the gain from being insured, agents may not find the insurance worth its price. In fact, Grönqvist (2006) finds that about 10 percent of the policyholders opt out after their first contract period. However, the observed increase in the consumption of dental care need not necessarily imply inefficiencies. If agents under-invest in prevention, e.g. due to asymmetric information or bounded rationality, it could be feasible for the National Dental Service in Värmland to subsidize contracts and boost prevention, in order to capitalize on better dental health in the future.

The estimated moral hazard effects are generally larger in cross-sectional estimates than in the difference-in-differences, and the question is how to interpret this difference. Remaining unobserved heterogeneity is more likely to be a problem in the cross-sectional estimates. Selection on cost levels is accounted for in the difference-in-differences estimates by taking differences. The potential for selection bias instead lies in treated and untreated agents having different cost trends. Selection on cost trends is here assumed to only be made on observable risk factors such as age, gender and risk class. For cross-sectional estimates, on the other hand, the potential selection is in cost levels. To identify the moral hazard effect, the proxy variable for private information needs to capture the differences in unobserved dental risk, which seems to be a stronger requirement for identification. If there is private information affecting the insurance status not captured by the proxy variable, the estimated moral hazard effect will be inflated by selection. However, if agents select into the insurance on basis of high temporary dental costs in the pre-program period, the extent of moral hazard will be underestimated in the difference-in-differences estimates. Irrespectively, the results give strong evidence of a significant moral hazard effect; its exact size needs to be interpreted with some care, however.

The conservative interpretation of the results adopted here is to rely on the matching estimates, unless the Wu-Hausman test indicates endogeneity. The IV is then considered to give unbiased and consistent estimates of the treatment effects, and any discrepancy between IV and the matching estimates is interpreted as adverse selection. The Wu-Hausman test is essentially testing if the IV estimate—based on selection on unobservables—is different from the effect estimated on observables. The IV estimates are generally smaller than the matching estimates across all types of dental procedures. Now, the matching estimates capture the average treatment effect on the treated (ATT); i.e., the effect on those agents who voluntarily choose the purchase insurance. IV instead captures the local average treatment effect (LATE). In this setting this refers to the effect on those patients who would not have bought the insurance had they not been persuaded by their dentist, and this effect is likely to be smaller than the ATT. The difference between the IV and matching estimates may therefore reflect the estimators capturing different effects, and the Wu-Hausman test would thus not necessarily indicate remaining endogeneity but rather differences in effects. A more generous interpretation would thus give even stronger support for moral hazard in dental insurance.

References

Abbring, Jaap H., Chiappori, Pierre-André and Pinquet, Jean "Moral Hazard and Dynamic Insurance Data", *Journal of the European Economic Association*, vol. 1(4), 767-820.

Angrist, Joshua D., and Krueger, Alan B., "Empirical Strategies in Labor Economics" ch. 23 in Ashenfelter, A. and Card, C. (Eds), *Handbook of Labor Economics, Volume 3*, Elsevier Science, X (1999).

Cardon, James, and Hendel, Igal, "Asymmetric Information in Health Insurance: Evidence From the National Health Insurance Survey," *RAND Journal of Economics*, vol 32 (2001), 408-427.

Chiappori, Pierre-André, Jullien, Bruno, Salanié, Bernard and Salanié, Francois, "Asymmetric Information in Insurance: General Testable Implications", *The RAND Journal of Economics*, (forthcoming 2006)

Chiappori, Pierre-André, and Salanié, Bernard, "Testing for Asymmetric Information in Insurance Markets", *Journal of Political Economy*, vol 108 (2000), 56-78.

Dionne, George, Gouriéroux, Christian, and Vanesse, Charles, "Testing for Evidence of Adverse Selection in the Automobile Insurance Market: A Comment", *Journal of Political Economy*, vol 109 (2001), 444-453.

Efron, Bradley, and Tibshirani, Robert J., *An Introduction to the Bootstrap*, Monographs on Statistics and Applies Probability, No. 57, Chapman & Hall, New York (1994).

Finkelstein, Amy, and Poterba, James, "Testing for Adverse Selection with 'Unused Observables'", *NBER Working Paper* No. 12112 (2006)

Leuven, Edwin, and Sianesi, Barbara,"PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing", http://ideas.repec.org/c/boc/bocode/s432001.html. version 3.0.0. (2003).

Grönqvist Erik, "Information Updating and Insurance Dropout" in Chiappori PA and Gollier C (Eds), *Competitive Failures in Insurance Markets: Theory and Policy Implications*, MIT Press, 2006.

Grönqvist Erik, "Does Adverse Selection Matter? Evidence from a Natural Experiment" SSE/EFI Working Paper Series in Economics and Finance No. 575 (2004).

Heckman, James J., Lalond, Robert J., and Smith, Jeffrey A., "The Economics and Econometrics of Active Labor Market Programs" ch. 31 in Ashenfelter, A. and Card, C. (Eds), *Handbook of Labor Economics, Volume 3*, Elsevier Science, X (1999).

Israel, Mark, "Do We Drive More Safely When Accidents Are More Expensive? Identifying Moral Hazard from Experience Rating Schemes", Mimeo, Kellogg School of Management, 2006.

Lechner, Michael, "Identification and Estimation of Causal Effects of Multiple Treatments under the Conditional Independence Assumption", in: Lechner, M. and Pfeiffer, F. (Eds), *Econometric Evaluation of Labour Market Policies*, Heidelber ca/Springer, (2001).

Manning, Willard G, Benjamin, Bernedette, Bailit, Howard L., Newhouse Joseph P., "The Demand for Dental Care: Evidence from a Randomized Trial in Health Insurance", *Journal of American Dental Association*, vol 110 (1985), 895-902.

Manning, Willard G, Newhouse, Joseph P., Duan, Naihua, Keeler, Emmett B., Leibowitz, Arleen, and Marquis, Susan, "Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment", *American Economic Review*, vol 77 (1987), 251-277.

Newhouse, Joseph, *Free For All?: Lessons from the RAND Health Insurance Experiment*, Harvard University Press, Cambridge (1993)

Pauly, Mark, "The Economics of Moral Hazard: Comment" *American Economic Review*, vol 58 (1968), 531-527.

Powell, L.V., "Caries Prediction: A Review of the Literature", *Community Dental Oral Epidemiology*, vol 26 (1998), 361-71.

Rosenbaum, Paul R., and Rubin, Donald B., "The Central Role of the Propensity Score in Observational Studies for Causal Effects", *Biometrika*, vol. 70 (1983), 41-50.

Rosenbaum, Paul R., and Rubin, Donald B., "Constructing a Control Group Using Multivariate Matched Sampling Methods That Incorporate the Propensity Score", *The American Statistician*, vol. 39 (1985), 33-38.

Windmeijer, Frank, "ExpEnd, a Gauss Programme for Non-Linear GMM Estimation of Exponential Models with Endogenous Regressors for Cross Section and Panel Data" *CEMMAP Working Paper* 14/02 (2003).

Wooldridge, Jeffrey M., *Econometric Analysis of Cross Section and Panel Data*, MIT Press, Cambridge (2002).

Notes

[1]Dionne, Gouriéroux and Vanesse (2001) show that restrictive functional forms in equations (1) and (2) can result in a spurious covariance, thereby falsely indicating the presence of asymmetric information. Chiappori and Salanie (2000) suggest that the covariance should be estimated non-parametrically.

[2]The proxy for private information—past dental consumption (see section 3)—is not used as a covariate in equation (6) since past dental consumption enters the dependent variable when forming the difference $Y_{.t+1} - Y_{.t}$. Section 4.3 provides a validity test of equation (6) by testing for differences in the pre-program trend.

[3]The common support property states that there are both treated and untreated agents in each neighborhood satisfying the CIA, formally $0 < \Pr(D = 1) < 1$.

[4]The matching protocol is implemented in STATA 9 with the PSMATCH2 routine by Leuven and Sinanesi (2003).

[5]The validity test is performed on the group of patients who did not purchase insurance during the period 1999 through 2001 (n=36 241). The reason for restricting the test to this group is that dental care within the insurance may be influenced by moral hazard or guided by clinical guidelines related to risk classes, thus generating a spurious relation.

[6]To compare dental costs for participants and non-participants during the contract period, a shadow contract period needs to be constructed for non-participants. A date for the insurance offer must therefore be identified. For participants, it is easy to see when the offer was made by considering when the contract starts. For non-participants, the date of the insurance offer is set to the date of the (first) check-up visit during 1999. Offers to buy contracts are usually made by the dentist during a check-up visit. Agents without check-up visits in 1999 are excluded from the sample, since a definite date for the offer in 1999 cannot be determined.

[7]For patients without insurance, each specific dental procedure is registered in the

program period, just as in the pre-program period. For patients with insurance, however, each specific procedure is not registered. Instead, dental care is registered within broader groups of procedures, and as time usage by staff category (dentist, hygienist, nurse). Hence, the registration of dental care is not directly comparable for patients with and without a contract. However, for about a third of all services, the specific procedure and time usage are registered for policyholders, roughly 24 000 registrations. Expected time usage can thus be calculated for each specific procedure, thereby creating a key. Under the assumption that the treatment style for any specific procedure does not differ between policyholders and non-policyholders, this key can be used to calculate time usage for the services consumed by non-policy holders. As dental consumption is expressed in time for both groups, the groups become comparable. Post Cost is then calculated by summing the time usage using the National Dental Service's time tariff. From 1999, there is no explicit time tariff for hygienist and nurse services. These are calculated as fractions of the time tariff for dentists using the price list 1995-1998 as a key.

[8]It is calculated applying gross prices to each specific procedure chargeable to patients, i.e. the amount charged by the National Dental Service.

[9]Dentists within FT are employed with a fixed salary, and have no private stake in the insurance.

[10]The instrument, *Dentist's share*, is defined as the number of policyholders among all risk classified patients registered with the National Dental Service in Värmland (n=54 669).

[11]The percent reduction in bias for a covariate is $100(1 - b_M/b_I)$, where b_I and b_M are the initial and matched differences in covariate means, respectively. The two-sample t-statistic compare if the distributions of the covariates in the treated and matched control groups have the same mean, while the paired t-statistics looks for systematic matched pair differences.

[12]The similarity between the matching and the OLS results is likely to be due to treated and non-treated units having a common support spanning the whole range of

propensity scores (see figure A.1 in the Appendix), and that a large number of the explanatory variables are binary. See Angrist and Krueger (1999) for a discussion on OLS versus matching

[13]When dental visits are analyzed in the cross-sectional setting, IV must be implemented in count data. The problem here is that the unobserved heterogeneity, Θ, in count models enters linearly in the data-generating process, $exp\left(I'\beta + D'\gamma + \Theta\right) + \varepsilon$. The standard nonlinear IV estimators, e.g. GMM, are not consistent in this case, as they assume the unobserved heterogeneity to be additively separable from the parametric structural model (Mullahy, 1997). To this end, Mullahy (1997) suggests consistent GMM estimation based on a transformation of the residual function, while Wooldridge (2002 pp. 663-666) suggests an alternative two-stage Poisson QML approach. Both these estimators are implemented. The transformed GMM estimation suggested by Mullahy (1997) is implemented with the ExpEnd Gauss routine provided by Windmeijer (2002). Wooldrige also suggests an endogeneity test based on the two-stage Poisson QML estimates.

Tables

Table 1: Insurance Status and Dental Cost by Risk Class

Risk class	N		Pre. Cost		Post Cost	
	Insur.	No insur.	Insur.	No insur.	Insur.	No insur.
All	6400	13007	1360	1659	2297	1803
1	123	152	359	488	1000	816
2	428	327	454	558	1024	810
3	492	466	579	604	1242	895
4	925	1113	682	739	1434	1025
5	1000	1364	976	914	1823	1227
6	1016	1555	1254	1122	2160	1395
7	966	1814	1593	1568	2735	1707
8	808	2215	2250	1804	3213	1919
9	363	1581	2603	2226	4178	2332
10	189	1129	3394	2628	4543	2513
11	53	773	3684	2800	5787	2783
12	30	369	4474	3064	6962	3040
13	7	149	3864	4053	7931	3517

Table 2: Validity Test of the Instrument

	Coef.	p-value
Pre. Cost		
Constant	-781	0.009
Dent. Share	67.1	0.782
Age	53.4	<0.001
Age^2	-0.476	<0.001
Gender	-108	<0.001
FE risk class	Yes	
FE clinics	Yes	
N	19407	
Adj. R2	0.146	

Table 3: Moral Hazard Effect in Dental Cost

	Tot. Cost	**Exam.**	**Prev.**	**Other**	*Rep.*	*All Dent.*	*Acute*
	1	2	3	4=1-2-3	*4a*	*4b*	*4c*
Cross-Sectional estimates							
PSM	898	169	326	403	*285*	*73.6*	*21.8*
	84%	34%	437%	81%	*66%*	*464%*	*35%*
	(<0.001)	(<0.001)	(<0.001)	(<0.001)	*(<0.001)*	*(<0.001)*	*(<0.001)*
IV	714	207	453	54.0	*-45.5*	*25.3*	*23.9*
	57%	45%	-	6%	*-6%*	*39%*	*40%*
	(<0.001)	(<0.001)	(<0.001)	(0.660)	*(0.636)*	*(0.601)*	*(0.305)*
	[0.192]	[0.058]	[0.007]	[0.003]	*[0.001]*	*[0.253]*	*[0.947]*
Difference-in-Differences estimates							
PSM	768	164	297	306	*240*	*38.6*	*14.5*
	64%	33%	287%	51%	*51%*	*76%*	*21%*
	(<0.001)	(<0.001)	(<0.001)	(<0.001)	*(<0.001)*	*(0.264)*	*(0.031)*
IV	661	208	381	72.4	*-5.3*	*84.0*	*-11.3*
	51%	46%	1866%	9%	*-1%*	*1527%*	*-12%*
	(0.006)	(<0.001)	(<0.001)	(0.752)	*(0.964)*	*(0.611)*	*(0.705)*
	[0.671]	[0.151]	[0.105]	[0.338]	*[0.046]*	*[0.760]*	*[0.396]*

Note: P-values for treatment effects are within parenthesis. For PSM the P-values are based on approximate standard errors suggested by Lechner (2001). The numbers within brackets are P-values of the Wu-Hausman endogeneity test for the IV estimates. Percentage increase is calculated as $TE/Counterfactual\ Cost$ (CC), where $CC = (Post\ Cost - TE)$. "-" indicates that TE is larger than $Post\ Cost$.

Table 4: Moral Hazard Effect in Dental Cost, Risk Class 1 to 6

	Tot. Cost	**Exam.**	**Prev.**	**Other**	*Rep.*	*All Dent.*	*Acute*
	1	2	3	4=1-2-3	*4a*	*4b*	*4c*
Cross-Sectional estimates							
PSM	569	155	253	162	*115*	*15.8*	*15.8*
	71%	32%	-	51%	*40%*	*167%*	*55%*
	(<0.001)	(<0.001)	(<0.001)	(<0.001)	*(<0.001)*	*(0.012)*	*(0.002)*
IV	459	209	283	-33.0	*8.85*	*-24.1*	*-37.2*
	50%	49%	-	-6%	*2%*	*-49%*	*-45%*
	(<0.001)	(<0.001)	(<0.001)	(0.765)	*(0.924)*	*(0.428)*	*(0.114)*
	[0.677]	[0.089]	[0.312]	[0.138]	*[0.363]*	*[0.268]*	*[0.031]*
Difference-in-Differences estimates							
PSM	497	160	226	111	*123*	*-33*	*13*
	57%	33%	942%	30%	*44%*	*-56%*	*41%*
	(<0.001)	(<0.001)	(<0.001)	(0.004)	*(<0.001)*	*(0.113)*	*(0.058)*
IV	484	219	249	16.0	*90.9*	*-36.9*	*-1.7*
	54%	52%	25831%	3%	*29%*	*-59%*	*-4%*
	(0.014)	(<0.001)	(<0.001)	(0.930)	*(0.432)*	*(0.748)*	*(0.956)*
	[0.962]	[0.158]	[0.626]	[0.682]	*[0.959]*	*[0.859]*	*[0.628]*

Note: P-values for treatment effects are within parenthesis. For PSM the P-values are based on approximate standard errors suggested by Lechner (2001). The numbers within brackets are P-values of the Wu-Hausman endogeneity test for the IV estimates. Percentage increase is calculated as $TE/Counterfactual\ Cost$ (CC), where $CC = (Post\ Cost - TE)$. "-" indicates that TE is larger than $Post\ Cost$.

Table 5: Moral Hazard Effect in Dental Cost, Risk Class 7 to 13

	Tot. Cost	**Exam.**	**Prev.**	**Other**	*Rep.*	*All Dent.*	*Acute*
	1	2	3	4=1-2-3	*4a*	*4b*	*4c*
Cross-Sectional estimates							
PSM	1420	185	455	780	*515*	*196*	*38.9*
	135%	37%	605%	162%	*110%*	-	*50%*
	(<0.001)	(<0.001)	(<0.001)	(<0.001)	*(<0.001)*	*(<0.001)*	*(<0.001)*
IV	1042	228	600	215	*-67.1*	*109*	*89.1*
	73%	50%	-	21%	*-6%*	*305%*	*324%*
	(<0.001)	(<0.001)	(<0.001)	(0.324)	*(0.689)*	*(0.232)*	*(0.028)*
	[0.149]	[0.150]	[0.042]	[0.009]	*[<0.001]*	*[0.471]*	*[0.204]*
Difference-in-Differences estimates							
PSM	1314	173	402	739	*444*	*246*	*20.0*
	113%	34%	314%	141%	*82%*	-	*21%*
	(<0.001)	(<0.001)	(<0.001)	(<0.001)	*(<0.001)*	*(0.005)*	*(0.105)*
IV	839	213.32	505	121	*-91.1*	*199*	*-31.3*
	51%	45%	1975%	11%	*-8%*	-	*-21%*
	(0.054)	(<0.001)	(<0.001)	(0.773)	*(0.657)*	*(0.520)*	*(0.548)*
	[0.509]	[0.330]	[0.229]	[0.315]	*[0.012]*	*[0.717]*	*[0.365]*

Note: P-values for treatment effects are within parenthesis. For PSM the P-values are based on approximate standard errors suggested by Lechner (2001). The numbers within brackets are P-values of the Wu-Hausman endogeneity test for the IV estimates. Percentage increase is calculated as $TE/Counterfactual\ Cost$ (CC), where $CC = (Post\ Cost - TE)$. "-" indicates that TE is larger than $Post\ Cost$.

Figures

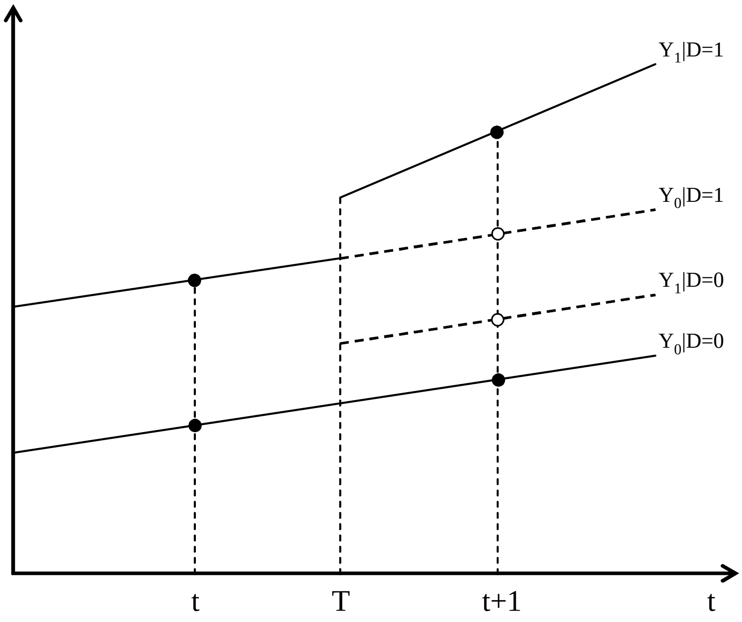


Figure 1: Characterization of the identification problem

Appendix

Table A.1: Descriptive Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Insurance	19407	0.330	0.470	0	1
Gender	19407	0.506	0.500	0	1
Age	19407	45.0	15.8	22.0	95.0
Dent. Share	19407	0.171	0.138	0	0.621
Pre. Cost	19407	1560	2132	0	54557
Examination	19407	268	183	0	1233
Prevention	19407	352	416	0	3796
Surgical	19407	36.2	167	0	3391
Endo Surgical	19407	86.6	392	0	6660
Preservatory	19407	425	663	0	12550
Denture. fixed	19407	252	1403	0	46276
Denture. removable	19407	34.5	341	0	9678
Acute	19407	81.6	206	0	3492
Post Cost	19407	1966	1493	0	22211
Examination	19407	663	246	0	2392
Prevention	19407	401	503	0	6752
Surgical	19407	52.5	195	0	4447
Endo Surgical	19407	68.5	317	0	9405
Preservatory	19407	595	771	0	8846
Denture. fixed	19407	78.5	430	0	15034
Denture. removable	19407	11.0	90.0	0	2523
Acute	19407	83.4	216	0	3244
Diff. Cost	19407	406	2195	-53586	21129
Examination	19407	396	294	-850	2184
Prevention	19407	48.7	450	-3191	5529
Surgical	19407	16.2	251	-3391	4447
Endo Surgical	19407	-18.1	492	-6660	9405
Preservatory	19407	169.5	811	-9424	8846
Dental. fixed	19407	-173.5	1437	-46276	15034
Dental. removable	19407	-23.5	342	-9678	1791
Acute	19407	1.78	270	-3241	2992

Table A.2: Descriptive Statistics on Visits

Variable		Obs	Mean	Std. Dev.	Min	Max
Pre. Visits						
	Examination	19407	0.798	0.563	0	5
	Prevention	19407	1.18	1.36	0	14
	Surgical	19407	0.081	0.317	0	5
	Endo Surgical	19407	0.060	0.260	0	4
	Preservatory	19407	0.811	1.14	0	12
	Denture. fixed	19407	0.084	0.376	0	8
	Denture. removable	19407	0.017	0.142	0	4
	Acute	19407	0.221	0.550	0	7
Post Visits						
	Examination	19407	1.94	0.552	0	5
	Prevention	19407	1.36	1.51	0	15
	Surgical	19407	0.111	0.369	0	5
	Endo Surgical	19407	0.075	0.338	0	8
	Preservatory	19407	1.20	1.39	0	16
	Denture. fixed	19407	0.076	0.353	0	6
	Denture. removable	19407	0.024	0.194	0	7
	Acute	19407	0.251	0.613	0	8
Diff. Visits						
	Examination	19407	1.14	0.75	-3	5
	Prevention	19407	0.181	1.40	-12	12
	Surgical	19407	0.030	0.470	-5	5
	Endo Surgical	19407	0.015	0.416	-4	7
	Preservatory	19407	0.385	1.41	-9	11
	Denture. fixed	19407	-0.007	0.479	-8	6
	Denture. removable	19407	0.007	0.224	-3	6
	Acute	19407	0.030	0.734	-6	7

Table A.3: Moral Hazard Effect in Dental Cost for Types of Treatment

	Exam.	Prev.	Surg.	Endo sur.	Pres.	Dent. f.	Dent. r.	Acute
Cross-Sectional estimates								
PSM	169	326	23.4	49.1	212	67.9	5.8	21.8
	34%	437%	81%	253%	55%	640%	110%	35%
	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(0.004)	(<0.001)
IV	207	453	19.6	-1.8	-63.3	26.0	-0.8	23.9
	45%	-	60%	-3%	-10%	50%	-7%	40%
	(<0.001)	(<0.001)	(0.360)	(0.959)	(0.423)	(0.580)	(0.939)	(0.305)
	[0.058]	[0.007]	[0.775]	[0.075]	[0.001]	[0.306]	[0.483]	[0.947]
Difference-in-Differences estimates								
PSM	164	297	23.2	41.0	176	33.5	5.08	14.5
	33%	287%	79%	149%	42%	75%	86%	21%
	(<0.001)	(<0.001)	(<0.001)	(0.001)	(<0.001)	(0.311)	(0.596)	(0.031)
IV	208	381	-10.4	31.2	-26.1	136	-52.2	-11.3
	46%	1866%	-17%	83%	-4%	-236%	-83%	-12%
	(<0.001)	(<0.001)	(0.710)	(0.570)	(0.771)	(0.392)	(0.170)	(0.705)
	[0.151]	[0.105]	[0.212]	[0.854]	[0.035]	[0.475]	[0.097]	[0.396]

Note: P-values for treatment effects are within parenthesis. For PSM the P-values are based on approximate standard errors suggested by Lechner (2001). The numbers within brackets are P-values of the Wu-Hausman endogeneity test for the IV estimates. Percentage increase is calculated as $TE/\textit{Counterfactual Cost}\ (CC)$, where $CC = (\textit{Post Cost} - TE)$. "-" indicates that TE is larger than *Post Cost*.

Table A.4: Pre-program Cost Trends

	1996-1998		1995-1998	
	Coef.	p-value	Coef.	p-value
Constant	2.40	(0.012)	2.96	(0.032)
Contract	0.035	(0.712)	0.124	(0.376)
Age	0.009	(0.595)	0.013	(0.569)
Age2	1.20E-06	(0.994)	-8.0E-05	(0.711)
Gender	-0.049	(0.525)	-0.177	(0.105)
FE risk class	Yes		Yes	
FE clinics	Yes		Yes	
N	14001		6974	
Adj. R2	0.015		0.021	

Note: The dependent variable is the trend in dental costs, which is culcualted as the sum of percentage yearly cost changes, $\sum(cost_{i,t1} - cost_{i,t-1})/cost_{i,t-1}$.

Table A.5: Covariate Balance for the Difference-in-Differences Model

Variable	Insurance	Non-insurance		Bias reduction	2 sample t-test	Paired t-test
		Unmatched	Matched	Percent	P-value	P-value
Age	39.7	47.6	39.4	96.0	0.144	0.051
Gender	0.501	0.509	0.499	76.5	0.818	0.750
D gr1	1.92E-02	1.17E-02	1.44E-02	35.7	0.033	0.016
D gr2	6.69E-02	2.51E-02	6.81E-02	97.0	0.778	0.709
D gr3	7.69E-02	3.58E-02	7.92E-02	94.3	0.621	0.527
D gr4	1.45E-01	8.56E-02	1.37E-01	87.0	0.213	0.068
D gr5	1.56E-01	1.05E-01	1.58E-01	97.6	0.846	0.787
D gr6	1.59E-01	1.20E-01	1.73E-01	64.9	0.036	0.004
D gr7	1.51E-01	1.39E-01	1.53E-01	85.0	0.786	0.707
D gr8	1.26E-01	1.70E-01	1.20E-01	86.9	0.320	0.165
D gr9	5.67E-02	1.22E-01	5.77E-02	98.6	0.819	0.757
D gr10	2.95E-02	8.68E-02	3.03E-02	98.6	0.795	0.734
D gr11	8.28E-03	5.94E-02	6.25E-03	96.0	0.176	0.118
D gr12	4.69E-03	2.84E-02	3.28E-03	94.1	0.207	0.160
D gr13	1.09E-03	1.15E-02	7.81E-04	97.0	0.564	0.564
D clin1	1.88E-03	3.92E-03	6.25E-04	38.9	0.045	0.033
D clin2	1.34E-02	3.18E-02	1.34E-02	100	1.000	1.000
D clin3	2.81E-03	2.00E-03	2.81E-03	100	1.000	1.000
D clin4	1.56E-04	1.38E-03	0.00E+00	87.3	0.317	0.317
D clin5	1.44E-02	2.01E-02	1.45E-02	97.3	0.941	0.929
D clin6	3.75E-03	2.08E-02	2.34E-03	91.8	0.149	0.095
D clin7	3.03E-02	2.64E-02	3.14E-02	71.7	0.721	0.665
D clin8	4.45E-02	8.69E-03	4.67E-02	93.9	0.553	0.435
D clin9	1.41E-02	5.18E-02	1.28E-02	96.7	0.539	0.310
D clin10	1.45E-02	4.54E-03	9.84E-03	53.1	0.016	0.011
D clin11	2.55E-02	3.58E-02	2.58E-02	97.0	0.911	0.877
D clin12	7.34E-03	2.20E-02	4.22E-03	78.7	0.020	0.006
D clin15	1.24E-01	1.02E-01	1.49E-01	-13.0	0.000	0.000
D clin16	9.22E-03	6.85E-02	1.11E-02	96.8	0.290	0.040
D clin17	4.66E-02	6.10E-02	4.59E-02	95.7	0.866	0.782
D clin18	1.14E-01	2.50E-02	1.20E-01	93.7	0.322	0.142
D clin19	4.44E-02	5.45E-02	5.22E-02	22.9	0.039	0.000
D clin20	2.92E-02	3.70E-02	3.05E-02	83.9	0.678	0.537
D clin21	3.13E-04	1.54E-04	6.25E-04	-96.9	0.414	0.414
D clin22	7.81E-04	1.46E-03	4.69E-04	54.0	0.479	0.480
D clin23	3.91E-03	6.88E-02	5.31E-03	97.8	0.240	0.007
D clin24	4.14E-02	1.87E-02	3.83E-02	86.2	0.366	0.309
D clin25	1.06E-02	9.99E-03	8.28E-03	-272	0.171	0.112
D clin26	4.45E-02	2.30E-02	4.06E-02	81.9	0.274	0.191
D clin27	1.00E-02	8.61E-03	8.13E-03	-35.0	0.263	0.226

Table continues

Table A.5: continued

Variable	Insurance	Non-insurance		Bias reduction	2 sample t-test	Paired t-test
		Unmatched	Matched	Percent	P-value	P-value
D clin 28	6.06E-02	3.80E-02	6.09E-02	98.6	0.941	0.920
D clin 29	2.66E-03	1.23E-02	3.13E-03	95.1	0.621	0.602
D clin31	2.97E-03	2.17E-02	2.81E-03	99.2	0.869	0.827
D clin32	1.48E-02	1.23E-02	1.36E-02	50.8	0.550	0.483
D clin33	2.34E-03	1.14E-02	2.34E-03	100	1.000	1.000
D clin34	6.02E-02	5.84E-02	6.11E-02	45.7	0.824	0.711
D clin35	1.42E-02	2.18E-02	1.20E-02	71.0	0.277	0.149
D clin36	3.23E-02	4.07E-02	2.75E-02	41.8	0.108	0.018
D clin37	3.13E-02	1.89E-02	2.86E-02	78.5	0.378	0.274
D clin38	4.66E-02	2.21E-02	4.11E-02	77.7	0.131	0.059
D clin39	4.22E-03	7.69E-04	9.38E-04	4.9	0.000	0.000
D clin40	3.75E-03	3.54E-03	2.50E-03	-486	0.205	0.170
D clin41	6.31E-02	6.15E-03	6.16E-02	97.3	0.715	0.541
D clin42	9.38E-03	1.95E-02	7.03E-03	76.9	0.142	0.108

Table A.6: Covariate Balance the Cross-Sectional Model

Variable	Insurance	Non-insurance		Bias reduction	2 sample t-test	Paired t-test
		Unmatched	Matched	Percent	P-value	P-value
Age	39.7	47.6	39.2	93.8	0.027	0.015
Gender	0.501	0.509	0.498	67.4	0.750	0.740
Cost gr1	6.9	5.70	5.04	-54.9	0.087	0.080
Cost gr2	30.4	14.0	26.6	76.5	0.132	0.120
Cost gr3	44.5	21.6	38.2	72.2	0.090	0.080
Cost gr4	98.5	63.2	93.6	86.1	0.473	0.468
Cost gr5	153	95.8	147	89.8	0.537	0.526
Cost gr6	199	134	194	92.0	0.722	0.720
Cost gr7	240	219	250	57.4	0.622	0.620
Cost gr8	284	307	271	42.1	0.544	0.544
Cost gr9	148	271	138	92.3	0.561	0.559
Cost gr10	100	228	149	61.8	0.063	0.063
Cost gr11	30.5	166	36.4	95.7	0.516	0.511
Cost gr12	21.0	86.9	18.7	96.6	0.781	0.782
Cost gr13	4.23	46.4	3.64	98.6	0.834	0.834
D gr1	1.92E-02	1.17E-02	1.63E-02	60.6	0.203	0.185
D gr2	6.69E-02	2.51E-02	6.64E-02	98.9	0.915	0.905
D gr3	7.69E-02	3.58E-02	7.31E-02	90.9	0.421	0.367
D gr4	1.45E-01	8.56E-02	1.52E-01	87.8	0.252	0.206
D gr5	1.56E-01	1.05E-01	1.63E-01	86.3	0.278	0.248
D gr6	1.59E-01	1.20E-01	1.64E-01	85.7	0.387	0.364
D gr7	1.51E-01	1.39E-01	1.53E-01	86.4	0.805	0.795
D gr8	1.26E-01	1.70E-01	1.19E-01	83.3	0.205	0.187
D gr9	5.67E-02	1.22E-01	5.06E-02	90.6	0.126	0.117
D gr10	2.95E-02	8.68E-02	2.98E-02	99.5	0.917	0.916
D gr11	8.28E-03	5.94E-02	9.22E-03	98.2	0.569	0.556
D gr12	4.69E-03	2.84E-02	2.97E-03	92.7	0.115	0.116
D gr13	1.09E-03	1.15E-02	7.81E-04	97.0	0.564	0.564
D clin1	1.88E-03	3.92E-03	1.88E-03	100	1.000	1.000
D clin2	1.34E-02	3.18E-02	1.28E-02	96.6	0.756	0.742
D clin3	2.81E-03	2.00E-03	3.13E-03	61.6	0.745	0.739
D clin4	1.56E-04	1.38E-03	3.13E-04	87.3	0.564	0.564
D clin5	1.44E-02	2.01E-02	1.17E-02	53.3	0.186	0.172
D clin6	3.75E-03	2.08E-02	3.59E-03	99.1	0.884	0.876
D clin7	3.03E-02	2.64E-02	3.34E-02	19.2	0.314	0.306
D clin8	4.45E-02	8.69E-03	5.08E-02	82.6	0.097	0.060
D clin9	1.41E-02	5.18E-02	1.34E-02	98.3	0.762	0.752
D clin10	1.45E-02	4.54E-03	7.81E-03	32.8	0.000	0.000
D clin11	2.55E-02	3.58E-02	2.67E-02	87.9	0.657	0.648
D clin12	7.34E-03	2.20E-02	8.59E-03	91.5	0.427	0.404
D clin15	1.24E-01	1.02E-01	1.30E-01	73.2	0.313	0.245

Table continues

Table A.6: continued

Variable	Insurance	Non-insurance		Bias reduction	2 sample t-test	Paired t-test
		Unmatched	Matched	Percent	P-value	P-value
D clin16	9.22E-03	6.85E-02	9.84E-03	98.9	0.716	0.646
D clin17	4.66E-02	6.10E-02	4.61E-02	96.8	0.900	0.887
D clin18	1.14E-01	2.50E-02	1.22E-01	90.9	0.155	0.071
D clin19	4.44E-02	5.45E-02	4.75E-02	69.2	0.399	0.357
D clin20	2.92E-02	3.70E-02	3.17E-02	67.8	0.411	0.390
D clin21	3.13E-04	1.54E-04	1.56E-04	1.60	0.564	0.564
D clin22	7.81E-04	1.46E-03	4.69E-04	54.0	0.479	0.480
D clin23	3.91E-03	6.88E-02	4.69E-03	98.8	0.499	0.398
D clin24	4.14E-02	1.87E-02	3.81E-02	85.6	0.342	0.329
D clin25	1.06E-02	9.99E-03	1.08E-02	75.2	0.932	0.931
D clin26	4.45E-02	2.30E-02	3.97E-02	77.5	0.173	0.161
D clin27	1.00E-02	8.61E-03	9.06E-03	32.5	0.585	0.581
D clin 28	6.06E-02	3.80E-02	5.89E-02	92.4	0.682	0.671
D clin 29	2.66E-03	1.23E-02	3.13E-03	95.1	0.621	0.622
D clin31	2.97E-03	2.17E-02	3.13E-03	99.2	0.873	0.853
D clin32	1.48E-02	1.23E-02	1.64E-02	38.5	0.476	0.471
D clin33	2.34E-03	1.14E-02	1.25E-03	87.9	0.144	0.090
D clin34	6.02E-02	5.84E-02	6.28E-02	-53.9	0.532	0.492
D clin35	1.42E-02	2.18E-02	1.55E-02	83.4	0.559	0.540
D clin36	3.23E-02	4.07E-02	3.20E-02	96.2	0.920	0.919
D clin37	3.13E-02	1.89E-02	2.95E-02	86.1	0.571	0.549
D clin38	4.66E-02	2.21E-02	3.92E-02	70.0	0.040	0.033
D clin39	4.22E-03	7.69E-04	2.19E-03	41.1	0.042	0.042
D clin40	3.75E-03	3.54E-03	5.31E-03	-632	0.188	0.189
D clin41	6.31E-02	6.15E-03	5.84E-02	91.8	0.267	0.089
D clin42	9.38E-03	1.95E-02	6.72E-03	73.8	0.093	0.088

Table A.7: OLS estimates of the Moral Hazard Effect in Dental Cost

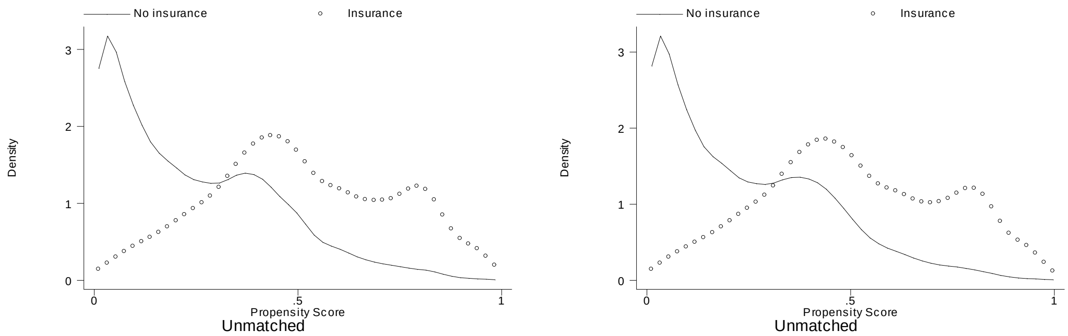
	Tot. Cost	**Exam.**	**Prev.**	**Other**	*Rep.*	*All Dent.*	*Acute*
	1	2	3	4=1-2-3	*4a*	*4b*	*4c*
Cross-Sectional estimates							
OLS	892	161	326	406	*279*	*79.7*	*25.4*
All risk classes	83%	32%	432%	82%	*64%*	*816%*	*44%*
	(<0.001)	(<0.001)	(<0.001)	(<0.001)	*(<0.001)*	*(<0.001)*	*(<0.001)*
OLS	513	151	233	128	*92.4*	*9.15*	*12.5*
Risk class 1-6	60%	31%	1339%	36%	*30%*	*57%*	*39%*
	(<0.001)	(<0.001)	(<0.001)	(<0.001)	*(<0.001)*	*(0.051)*	*(<0.001)*
OLS	1383	176	440	767	*524*	*173*	*38.4*
Risk class 7-13	127%	35%	491%	155%	*114%*	-	*49%*
	(<0.001)	(<0.001)	(<0.001)	(<0.001)	*(<0.001)*	*(<0.001)*	*(<0.001)*
Difference-in-Differences estimates							
OLS	761	165	307	289	*225*	*34.1*	*13.8*
All risk classes	63%	33%	329%	47%	*46%*	*62%*	*20%*
	(<0.001)	(<0.001)	(<0.001)	(<0.001)	*(<0.001)*	*(0.193)*	*(0.004)*
OLS	475	161	225	89.7	*85.1*	*-16.8*	*12.7*
Risk class 1-6	53%	33%	879%	23%	*27%*	*-40%*	*40%*
	(<0.001)	(<0.001)	(<0.001)	(0.001)	*(<0.001)*	*(0.342)*	*(0.006)*
OLS	1121	171	415	535	*409*	*88.3*	*15.2*
Risk class 7-13	83%	33%	361%	74%	*71%*	*157%*	*15%*
	(<0.001)	(<0.001)	(<0.001)	(<0.001)	*(<0.001)*	*(0.079)*	*(0.073)*

Note: P-values for treatment effects are within parenthesis. Percentage increase is calculated as $TE/Counterfactual\ Cost\ (CC)$, where $CC = (Post\ Cost - TE)$. "-" indicates that TE is larger than $Post\ Cost$.

Table A.8: Moral Hazard Effect in Visits for Different Types of Treatment

	Exam.	Prev.	Surg.	Endo sur.	Pres.	Dent. f.	Dent. r.	Acute
Cross-Sectional estimates								
PSM	0.245	1.05	0.039	0.054	0.342	0.072	0.013	0.053
	14%	341%	54%	258%	40%	1682%	111%	27%
	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(0.005)	(<0.001)
IV	0.224	0.988	0.041	0.045	0.302	0.035	0.011	0.048
2SQML	13%	265%	59%	147%	34%	83%	89%	24%
	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)
	[0.567]	[0.014]	[0.404]	[0.373]	[0.122]	[0.555]	[0.044]	[0.048]
IV-T	0.396	1.61	NA	NA	-0.117	NA	NA	0.045
GMM	26%	-			-9%			22%
	(<0.001)	(<0.001)			(0.319)			(0.530)
Difference-in-Differences estimates								
PSM	0.225	0.942	0.032	0.050	0.267	0.046	0.014	0.033
	13%	225%	40%	195%	29%	151%	140%	15%
	(<0.001)	(<0.001)	(0.008)	(<0.001)	(<0.001)	(<0.001)	(0.013)	(0.075)
IV	0.317	1.08	0.049	-0.006	-0.174	0.054	-0.031	0.043
	20%	393%	81%	-7%	-13%	237%	-57%	21%
	(<0.001)	(<0.001)	(0.344)	(0.897)	(0.268)	(0.314)	(0.215)	(0.596)
	[0.261]	[0.322]	[0.880]	[0.196]	[0.012]	[0.957]	[0.065]	[0.862]

Note: P-values for treatment effects are within parenthesis. For PSM the P-values are based on approximate standard errors suggested by Lechner (2001). For IV 2SQML P-values are not corrected for the inclusion of a predicted regressor. The numbers within brackets are, for IV, P-values of the Wu-Hausman endogeneity test, and for IV 2SQML, P-values of Wooldridge's endogeneity test (Wooldridge 2002 p. 665). NA indicates that IV-T GMM estimates are not available because the Hessian is not negative definite. Percentage increase is calculated as $TE/\mathit{Counterfactual\ Cost}$ (CC), where $CC = (\mathit{Post\ Cost} - TE)$. "-" indicates that TE is larger than $\mathit{Post\ Cost}$.



Difference in Differences **Cross Section**

Figure A.1: Gaussian Kernal Density Distribution of Propensity Scores Before Matching.