THE UNIVERSITY OF NEW SOUTH WALES

Faculty of Business

School of Economics
UNSW, Sydney 2052
Australia

http://www.economics.unsw.edu.au

The Impact of the Indonesian Health Card Program: A Matching Estimator Approach

Meliyanni Johar

School of Economics Discussion Paper: 2007/30

ISSN 1323-8949

ISBN 978 0 7334 2564 6

The Impact of the Indonesian Health Card Program: A Matching Estimator Approach

Meliyanni Johar

University of New South Wales, Australia

Abstract

This study evaluates the effectiveness of a pro-poor nation-wide health card program in Indonesia which provides free basic health care at public health facilities. To quantify the effect of the program, it departs from the traditional regression-based approach in the literature to employ propensity score matching to reduce the selection bias due to non-random health card distribution. The setting of the program and the richness of the data set support this strategy in providing accurate estimates of the program's effect on its recipients. The result finds that in general the health card program only has limited impact on the consumption of primary health care by its recipients. This finding suggests the presence of other factors that are counteracting the generous demand incentive.

Keywords: Impact evaluation, health sector reform, Indonesia

JEL classification code: I1

I. INTRODUCTION

Inequalities in access to health care have become a prominent policy agenda in many countries worldwide. Interventions of various forms were introduced by governments and non-government institutions as an attempt to minimise these asymmetrical situation with targeted programs for the poor being the common theme. Some countries directly provide health goods (e.g., the United Kingdom) while others combine public provision with subsidised health insurance for the poor (e.g., Australia, Malaysia, Singapore). The common justification for subsidisation is that health care services are particularly costly for the poor yet they are more likely to face adverse health shocks (Wagstaff, 2005; Whitehead et al., 2001; Xu et al, 2003; Case et al., 2002). Further, it is believed that there are positive externalities from a healthy population.

In Indonesia, health care payments are largely out-of-pocket and cash at the time of purchase or service provision. The health insurance market is underdeveloped with less than 20 percent of the population covered by at least one form if health insurance as at 2000 (IFLS). As a result, sick individuals with subsistence or low income and without insurance may be unable to obtain the necessary medical treatment; 56 percent of the population lived on less than $2 per day. Health care utilisation rates are low and unchanged, despite marked increases in the incidences of both communicable and non-communicable diseases. Restricted access to adequate health care has also been linked to critical health statistics, such as under-five mortality figure of 38 per 1,000 live births due to preventable factors (MoH) and the highest maternal mortality rate among Southeast Asian nations (WHO).

The health card program of 1994 is one of the government's major efforts to improve the nation's health conditions by promoting equality in access to primary health care. It is a nation-wide project involving all public health facilities, including hospitals. The program targets poor households and provides full price subsidy to medical expenses at public health facilities by all members of the household. Health cards are distributed at the household level by the village heads as the program administrators, based on a list of criteria that reflects welfare. Substantial regional heterogeneity makes it implausible to have a common eligible rule, and so eligibility determination is essentially decentralised and varies across communities. In principle, there is no limit on the frequency of health card use, but recipient households can lose the health card if income increases with the re-assessment typically is performed annually.

The aim of this study is to measure the effectiveness of the health card program. To my knowledge, this is the program's first formal evaluation at the microeconomic level. So far, assessments are based on aggregate statistics on the number of health card disseminated in one area and the recorded number of patients with health card received by public health facilities (Widianto, H. at MoH, 2007, pers. comm., 12 January). Results form aggregate analyses however are likely to be contaminated with other factors. Accurate evaluation of the program is necessary, particularly in the face of constrained resources, as policymakers must ensure that ongoing programs are not wasteful. The outcome of this study therefore will provide valuable feedback to policymakers, and suggest appropriate directions for future policy.

This thesis also contributes to the general health and development literature in at least three significant ways. First, it may extend our knowledge about the interaction between demand incentives, which health cards are a form of it, and health care choice in developing countries which we know very little about. Most of the existing evidences in the literature are based on developed countries, which have distinct environmental settings to developing countries. For example, in developing countries, transport and time costs associated with the subsidised care are often non-negligible, and consumption of traditional medicines is prevalent. In addition, public health facilities are often inadequate and rated as having low quality by potential patients (see Filmer et al., 2000 for cross countries review). In the presence of these other costs and readily accessible alternatives, it is not clear whether a price subsidy as a demand inducement will result in recipient households increasing their consumption of formal health care. In the literature, there have been numerous studies focusing on targeted programs in relatively poor countries, especially African countries, which find mixed conclusions about households' response (for example Newman et al., 2002; Castro-Leal et al., 2000; Pradhan et al., 1998).

However, evidence from middle-income developing countries like Indonesia is relatively thin.

Second, as the evaluation method, this study employs the Propensity Score Matching and Difference-in-Differences (PSM-DID) estimator, which has gained popularity in labour economics to produce causal inference (Smith and Todd, 2000, 2005; Heckman et al., 1997), but has not been widely used in the health literature. In non-experimental setting, it is well recognised that inferences drawn from comparing the outcomes of the program recipients and non-recipients are likely to be biased due to selection problems. The issue in evaluation studies therefore has centred on selection bias and the construction of the comparison group. Traditionally, health studies have relied on instrumental variable (IV) regressions to correct for sample selection bias, hypothesising that units are different in unobservable ways. However, it is known that valid instruments are rarely available, and IV regression's results rely in strong assumptions for causal inference (Imbens and Angrist, 1994). PSM oppositely supposes that sample units are different by their observed characteristics, and does not require IV(s). In the case of the health card program, selection on observables may be more appropriate, as health card eligibility does not require households' initiative to apply for the program, but is determined based on their state of economy that are observed by the program administrators. Further, the combination PSM and DID allows selection on unobservables, as long as they are time-invariant. The reliability of PSM-DID to provide causal estimates, under reasonable conditions, has been found by many studies (see Smith and Todd, 2005; Heckman et al., 1997, 1998; Augurzky and Schmidt, 2001).

Third, this is a demand-side study that incorporates supply-side variations into the analysis. Commonly, supply variables are excluded from demand studies due to lack of data. Yet, it is known in the literature that supply influences demand either directly through additional arrangements or indirectly through increasing health awareness among potential patients. In fact, many data have supported this hypothesis by featuring positive correlation between health care utilisation rates and physician density (McGuire, 2000;

Dow et al., 2000). Analyses without supply side factors therefore may be invalid due to omitted variable bias.

The data used in this study comes from the Indonesian Family Life Survey (IFLS) in 1993 (IFLS1), 1997 (IFLS2) and 2000 (IFLS3). The IFLS may be the only large sample longitudinal study on Indonesian households. It is a very rich data source collected at the household and individual levels that supports the empirical strategy. Most studies in various literature on Indonesia use the national annual household survey (*Susenas*) by the Central Bureau of Statistics, which is an ongoing large sample study but only repeated cross-sections. Alternatively, the IFLS is also nationally representative as it uses the same sampling frame as those constructed for *Susenas* in selecting its enumeration areas. Furthermore, the household survey is accompanied by community and facility surveys, which provide information about health care supply conditions in communities where households lived. The reliability of the data set is formally documented in Thomas et al. (2001).

The results show that the health card program has only limited impact on its recipients' health care consumption. The presence of a health card increases the number of visitations to public facilities by the younger household members, especially to receive curative treatments, but has no significant positive effect on other household members. However, the initial levels of health care consumption by children are very low. The other effect of health card availability is increasing contraceptive enrolments among eligible females in the recipient households, but in this case, the demand incentive was paralleled by supply expansion of contraceptive services in public health facilities in the late 1990s. Considering the health card program is fairly generous in its design, this finding may suggest that either a subsidy (or price-based incentive in general) to general health care is ill-suited as a form of intervention to increase health care utilisation, or there are factors that are counteracting the demand inducement, or both. As policy implications, program redesign or redirection of resources therefore may yield a larger impact.

II. BACKGROUND AND LITERATURE

The health card program was designed to allow poor households to obtain at least basic health care services. Health cards are distributed at the household-level, and cover fully the costs of medical treatments of all household members at all public health facilities, including hospitals and village midwife. Covered services include basic outpatient and inpatient care, diagnostic testing, contraceptive treatments, and children and maternal care. In principle, there is no limit on the frequency of health card use, and neither the Ministry of Health nor the public facilities have a formal procedure in place to monitor the use of health card.

Eligibility is determined using a list of criteria that reflects welfare, such as lack of permanent income source, failure to provide two basic meals per day and inability to purchase basic health care services. The village head in each community act as the program administrator and decides on the weights attached to each criterion. This procedure is used instead of the common fixed-income threshold rule because regional heterogeneity affects price and income levels, and accordingly poverty lines. Once household is identified as poor, health card is issued. No application by households is needed. Public health facilities are informed on the number of health card disseminated in the area, and are reimbursed based on user costs. Provision for this outlay is set according to historic records on the number of poor households in the area.[1]

In the past, there are only a few large scale health projects in the country that had been formally assessed. Gertler and Molyneaux (2000) investigate the effects of public contraceptive subsidies organised by the National Family Planning and Coordinating Board on contraceptive use and fertility rates. They use IV regressions to correct for non-random program placement, and argue that district-level variables such as eligible population, per capita subsidy level and their lagged values are valid IVs. Several data sources are combined to provide 36 quarterly data for the decade 1985-1994. The

[1] Since 2000, the health card program is reformed to be contained within a social system (JPKM) managed by a state-own insurance company. Under the reformed system, fund-holding entities were set up at district level to replace the central government in managing resource allocation and service provision in their areas. Nonetheless, the design of the health card program is unchanged.

individual-level information is obtained from the contraception calendar of women in the 1991 and 1994 annual Demographic and Health Survey while the remaining data sources are extracted from various unpublished Indonesian data sets. The study finds that the project has no impact on fertility rates (as measured by the proportion of women aged 15 – 49 years old who gave birth in the district in that quarter), and only a small but significant impact on contraceptive prevalence in a district. For the latter, the study also finds that the effect of contraceptive subsidies is smaller than the effect of increased number of clinics.

Pradhan et al. (2007) analyse the immediate impact of a World Bank project in 1998 as a response to increased poverty level post the 1997 financial crisis. The Social Safety Net (SSN) program transfers lump-sum payments to targeted districts to be allocated to different sectors. The study is based on *Susenas* data set, appended with administrative district data. To tackle selection problems, it relies on cross-sectional PSM by island (Indonesia is an archipelagic state) to match households in 1999 to households in 1998. Least Square regression is then applied on the matched sample to estimate the program effect. By exploiting variations in funds across districts, the study finds that spending in the health sector significantly increases (decreases) households' outpatient contact rate at public (private) facilities by 0.1% point in a month time.

There are only a handful of health evaluation studies that involve non-parametric approach, in particular PSM-DID. Three recent studies from developing countries that use matching are Trujillo et al. (2005) in analysing the impact of a targeted subsidised health insurance on health care use in Columbia, Galiani et al. (2005) in estimating the effect of privatisation of water system on children's mortality in Argentina, and Wagstaff and Yu (2007) in examining the performance of a World Bank's investment in China. All studies reported some positive effects of the studied program. The China project (which is a combination of demand and supply expansion) for instance reduces households' out-of-pocket health expenses, particularly on drugs, thereby reducing the likelihood of catastrophic health expenditure to poor households. The project however has no or even negative impact on the number of doctor visits. The problem in the study is lack of common support. Many observations have to be dropped to make the counterfactual and the program beneficiary samples comparable. Resultantly, the effect of the program is estimated tolerating observations that are just off-support. Hence, hitherto, there are very few convincing studies in the literature based on matching technique. One of the contributions of this thesis is to fill this gap by exploiting the availability of large rich data sets.

III. METHODOLOGY

The objective of this study is to accurately estimate the effect of health card availability on its recipients' health care utilisation. Specifically, the interest is to compare outcomes of the recipients to the counterfactual, that is, their outcomes when health card is not available at the same point in time. The counterfactual however is never observed. If health cards are randomly distributed, its effect can be measured by comparing the outcomes of recipient and non-recipients. Let Y_{it} denotes individual i's health care utilisation at time t, where $t = 0$ indicates pre-treatment period and $t = 1$ indicates treatment period. D is the treatment indicator which takes a value of 1 if household is treated, and 0 otherwise. The average treatment effect of the treated (ATT) is given by:

$$ATT = E(Y^d - Y^c \mid d = 1) \quad . \tag{1}$$

(1) is an unbiased estimator of the ATT as random allocation implies orthogonality of the outcome variable and the treatment status; $Y_{i1} \perp D_i$.

However, in non-experimental setting, there are at least two reasons to believe that treatment is not random. First, recipient and non-recipient households are different. For instance, it is suggested that poor households have low taste for formal medical care (Akin et al., 1998). If this is the case, then the ATT will be downwardly biased. Second, eligibility determination is decentralised and varies across communities. Selection by the program administrators may also relate to the outcome variables. In the absence of a controlled randomised experiment, the primary task therefore is to use estimation to create the counterfactual or control units under most reasonable conditions. The following discusses the matching technique in constructing these counterfactuals.

In the presence of longitudinal data, matching technique can be combined with DID estimator, which has been extensively used in policy evaluation studies. Let t_0 indicates pre-treatment period and t_1 indicates treatment period. The first underlying assumption behind matching in a longitudinal data context is:

$$Y_{t_1}^c - Y_{t_0}^c \perp d \mid X , \qquad (2)$$

where X is a vector of strictly exogenous variables that are unaffected by the treatment or anticipation of the treatment (Heckman et al., 1998). Equation (2) states the conditional independence or 'ignorability' assumption which requires that conditional on observables X, the outcome of the control units progresses in the same way as that of the treated units had they not been treated. Health card distribution may not be random, and this is not ignorable with respect to the outcome variables. However, health card allocation becomes ignorable if all factors that are influencing the allocation and are related to health care utilisations are in X.

The second fundamental assumption is that it is possible to find the counterfactual of each treated unit in the control group. This is the common support condition; $0 < \Pr(D = 1 \mid X) < 1$. Unlike in regression-based methods, matching explicitly tests for the fulfilment of this condition, thereby avoiding "off-support" inferences. By off-support we mean, attempt to establish the treatment effect when the treated and non-treated units in the data are incomparable, leading to biased ATT estimates (Heckman et al., 1997; Cobb-Clark and Crossley, 2003).

When there are many covariates to match however, matching can suffer from dimensionality problem. One solution suggested by Rosenbaum and Rubin (1983) is to match on a propensity score $P(X)$ that summarises information given by X A useful result is that under the above two assumptions, the conditional independence assumption also applies for $P(X)$:

$$Y_{t_1}^c - Y_{t_0}^c \perp d \mid P(X) \quad . \qquad (3)$$

(3) is a weak requirement in the sense that it conditions on fixed-effects and does not rule out selection on the basis of time-invariant variables. The first fundamental assumption

requires there is no "hidden bias" due to unobserved factors affecting both treatment status and health care choice. In observational data, this endogeneity of treatment status is common, and cross-sectional PSM estimator is generally inconsistent in this case. However, one may think these unobservables to be taste or tradition, which is household-specific and changes very slowly, or not at all, over time. If so, the combination DID and PSM generates a powerful estimator of ATT by first eliminating the effects of time-invariant unobservables, which is difficult to deal with in traditional cross-section matching, and then uses matching technique to construct counterfactual for each treated unit (Smith and Todd, 2000).

To calculate the ATT though, we only need mean conditional independence. The ATT now can be written as:

$$ATT = E(Y_{t_1}^d - Y_{t_0}^d \mid P(X_i), d = 1) - E(Y_{t_1}^c - Y_{t_0}^c \mid P(X_i), d = 0), \qquad (4)$$

which essentially is the conditional difference-in-differences (DID) estimator. First-differencing will eliminate effects of inherent or time-invariant households' and regional characteristics. Meanwhile, double-differencing will eliminate macroeconomic effects, if they are homogenous across units, as well as potential bias arising from differences in survey questionnaires. Later, I will explore the validity of the assumption of universal macroeconomic effects assumption. Given the two fundamental assumptions, equation (4) writes the consistent estimator of the ATT.

To find a matching-pair for each treated unit, the PSM method searches for a unit in the control group with the closest propensity score to that treated unit. Treated units that are unable to find a match lie outside the region of common support, thus they are excluded in calculating the ATT. Notice that unlike regression approach, matching makes minimal assumption about the functional form of the relationship between health card availability and health care utilisation, which is implicitly assumed to be linear in the standard regression method. Given the common support restriction, the sample-counterpart for the ATT can be written as:

$$ATT = \sum_{i \in D} \left[(Y_{it1} - Y_{it0}) - \sum_{j \in C} W_{ij} (Y_{jt1} - Y_{jt0}) \right] w_i \qquad (5)$$

where W_{ij} is the weight placed on a control unit j for a treated unit i. Equation (5) states that the ATT is found by comparing the change in outcomes of the treated unit with that of the counterfactual, which is given by the weighted average of the change in outcomes of the untreated units, where the weights reflect the propensity score. The weighting function may vary according to the selected definition of a close neighbourhood and this depends on the researchers rather than being implicit in the estimator, as in the case of linear regression. This flexibility in aggregating the treatment effect makes matching more amenable to heterogenous treatment effect context. Several possible definitions for W_{ij} include the nearest neighbour which gives zero weights to all matched pairs except the one closest to the treated unit, several matched pairs within a specified distance, or the entire sample of matched-pairs with a weighting function that accounts for relative closeness. In general, studies have favoured the latter definition, which maximises the sample size, with weights given by kernel weights (Dehejia and Wahba, 1998; Heckman et al., 1997; Galiani et al., 2005). Another advantage of this method from the standard regression is that health care variables contain a lot of zeros, which is likely to create bias in estimation of linear slopes.

In this study, the propensity scores are estimated using logit model – one may opt to use non-parametric technique, but it would just bring back the dimensionality problem –, and the Kernel weights is assumed to be Gaussian. Since analytical formula to compute the standard error of the ATT estimates is not available, I follow the common practice to compute the standard errors using bootstrap method. With kernel weight, because it does not run into the discontinuity that arises in other method, particularly in nearest neighbour matching, the reservations that have been forwarded about bootstrapping do not apply (see for example Abadie and Imbens, 2006). Further, bootstrapping has additional advantages as it takes into account both sampling errors in the propensity score estimates and errors due to multiple matches for a single treated unit. This is because each bootstrap repetition selects a sample (with replacement) from the sample of households, within the common support, and, in each replication, re-estimates the propensity score. The standard errors are found by bootstrapping with 200 replications.

IV. DATA AND VARIABLES

The data used in this study are derived from the Indonesian Family Life Survey (IFLS) in 1993 (IFLS1), 1997 (IFLS2) and 2000 (IFLS3) by RAND in collaboration with several local universities. The IFLS is a nationally representative data set covering 13 out of 27 Indonesian provinces where 83% of the population resides. It is a very rich data source collected at the individual, household and community levels, and its reliability had been formally documented (Thomas et al., 2001).

In this study, I focus on treated households in IFLS3 and use their characteristics in IFLS1 (Sample 1) and IFLS2 (Sample 2) as initial conditions. This is because only a small number of households were treated by IFLS2; most of the treated households in IFLS3 obtained the concession after 1997. I drop a minority of households (2%) that are treated in both IFLS2 and IFLS3 to make IFLS2 a valid control sample. Also, only non-split households are included in the samples as conditions at the origin households may be a poor representation of conditions after the separation.

The units of study are both household and individual. For the individual-level analysis, I focus on adults (age ≥15 years old), as young children's health and health care need vary greatly over time as they grow up (e.g., intensity of regular check-up naturally falls as the child gets older). Matching is performed at household-level because members of one household have the same treatment status, and individual characteristics, except those of the household heads, do not influence health card's availability. The ATT is then calculated at the individual-level, with all members of one household having identical propensity score of being treated. To incorporate individual-specific variables, individuals are categorised according to their relation to the household head in the treatment year: the household heads, the spouse of the household heads, children of the household head and a pool of other household members. It is important to keep in mind that 'children' does not imply young age. Indeed, the average age of children is close to 22 years in the treatment year. Meanwhile, other household members include parents, in-laws, step children, grandparents, nieces/spouses, uncles/aunts and servants. This division is chosen because household hierarchy may convey extra information about intrahousehold information sharing and resource allocation; there is only 1 health card per household to cover all household members. Further, it may be a nifty way to group individuals based on common characteristics. For example, spouses tend to be 30–50 years old females who are not working in the formal (taxed) labour market. The final samples consist of 5,262 households in Sample 1 and 4,580 households in Sample 2.

Dependent variables

Information about health card availability is obtained from the household heads. About 20% of households in each sample are treated households. To measure health care utilisations, I consider the number of inpatient and outpatient care received by individuals at public and private facilities. Public facilities include health centers and its subsidiaries, hospitals and village midwives. Meanwhile, private facilities are private hospitals and physicians. Traditional practitioners (e.g., religious healers) are excluded. For inpatient care, the reference period is 12 months to the survey, and for outpatient care, this period is 4 weeks. Outpatient care are further categorised according to their purpose, namely curative-type and preventative-type services. From the list of purposes in the survey, visitation for “treatment” and “medication” only are classified as curative-type service where as for other purposes such as medical check-ups, vaccination and an assortment of treatments are classified as preventative-type care. As suggested by equation (4), the dependent variables are the first-difference of each of these health care consumption measures.

In the samples, the proportions of non-user of formal health care are very large. For outpatient care, averaging the two samples, about 80 percent of members in treated households are non-users of public health care, and 93 percent are non-users of private health care. In comparison between health care utilisation by household members of treated and non-treated households, members of non-treated households consume slightly less public services, which is not unexpected as services at public facilities are typically cheaper than those at private clinics, and more private services. This picture of low health care utilisation confirms observations from *Susenas* data set (Lanjouw et al., 2000). Meanwhile, inpatient cases are very rare for both groups in both samples (1–3 percent).

Beside the common agreement that hospitalisation is an unpleasant experience, this rarity may be explained by high opportunity cost of inpatient days, as compensation payments for sick days are not available for most workers.

Table 1A and 1B present health care consumption pattern in the treatment year by household hierarchy for Sample 1 and 2 respectively. These are data a typical cross-sectional evaluation study would analyse. The absolute *t*-statistics for differences in means for two samples are reported in parentheses. In general, as anticipated, on average members of treated households (treated individuals) pay more visits to public facilities than members of non-treated households (non-treated individuals) to receive outpatient care. Because the initial utilisation levels are considerably low, the magnitudes of the differences according to treatment status are quite marked. Based on this snap-shot thus it appears that the health card program has some success in achieving its objective.

Treated household heads and spouses tend to use more public facilities and less of private facilities for outpatient care than non-treated heads and spouses. For spouses, this is mainly due to preventative-type services. Meanwhile, treated children visit public facilities more frequently than non-treated children to treat illness or obtain medication and this difference is significant in Sample 2. It is worth noting that the children sub-samples are quite different for Sample 1 and 2. The children sample is much smaller, and children are older in Sample 1 because many young adults in 2000 (i.e., the 15–21 year olds) are not in the 1993 (adult) survey – they answered children booklets in the IFLS1. It is tempting to track down these young adults' health care consumption to the children booklet. However, this is not pursued because the assumption of constant evolution of time-invariant individual heterogeneity made by the empirical strategy may be too strong (see Section III). In particular, in seven years time, 8–14 years-olds are likely to engage in different activities than schooling, and their immune system are likely to strengthen as they grow up to adulthood. For other household members in different households, there is no significant difference in health care consumption, although in general the directions of differences are consistent with those experienced by the core household members. Like the children samples, other members in Sample 1 and Sample 2 are quite different, and by comparing the age figures, it appears that non-core household members who remain in the house throughout the surveys tend to be elderly.

Table 1A: *Post-treatment health care consumption (Sample 1)*

	Head (D=0)	Head (D=1)	Spouse (D=0)	Spouse (D=1)	Child (D=0)	Child (D=1)	Other (D=0)	Other (D=1)
Outpatient care								
Public	0.141	0.172	0.201	0.225	0.069	0.130	0.166	0.210
		(1.647)*		(0.872)		(1.128)		(0.746)
Private	0.097	0.070	0.096	0.065	0.085	0.196	0.076	0.056
		(1.812)*		(1.831)*		(1.369)		(0.567)
Curative - public	0.055	0.079	0.091	0.110	0.032	0.065	0.094	0.161
		(2.032)**		(1.139)		(0.868)		(1.439)
Curative - private	0.037	0.027	0.039	0.033	0.068	0.0434	0.034	0.008
		(1.127)		(0.509)		(0.086)		(1.438)
Preventative – public	0.086	0.092	0.110	0.114	0.037	0.065	0.072	0.048
		(0.384)		(0.201)		(0.704)		(0.635)
Preventative – private	0.060	0.043	0.057	0.031	0.037	0.152	0.043	0.048
		(1.424)		(1.967)**		(1.862)		(0.185)
Inpatient care								
Public	0.013	0.012	0.014	0.017	0.011	0.000	0.009	0.008
		(0.253)		(0.384)		(0.700)		(0.097)
Private	0.007	0.013	0.009	0.003	0.016	0.043	0.007	0.008
		(1.696)*		(1.536)		(1.155)		(0.127)
Average age (in 2000)	50.61	50.74	43.56	43.06	27.31	28.13	58.98	60.90
N	3526	929	2769	726	188	46	445	124

Table 1B: *Post-treatment health care consumption (Sample 2)*

	Head (D=0)	Head (D=1)	Spouse (D=0)	Spouse (D=1)	Child (D=0)	Child (D=1)	Other (D=0)	Other (D=1)
Outpatient care								
Public	0.135	0.148	0.189	0.232	0.071	0.122	0.124	0.155
		(0.751)		(1.761)*		(3.542)***		(1.114)
Private	0.084	0.061	0.092	0.054	0.057	0.039	0.069	0.044
		(1.770)*		(2.373)***		(1.524)		(1.291)
Curative - public	0.055	0.061	0.087	0.113	0.031	0.065	0.069	0.087
		(0.576)		(1.653)*		(3.349)***		(0.877)
Curative - private	0.035	0.028	0.035	0.029	0.025	0.018	0.034	0.014
		(0.922)		(0.573)		(0.923)		(1.521)
Preventative - public	0.080	0.087	0.102	0.119	0.040	0.057	0.055	0.068
		(0.480)		(0.919)		(1.625)		(0.690)
Preventative - private	0.049	0.034	0.057	0.025	0.032	0.022	0.035	0.030
		(1.567)		(2.539)***		(1.232)		(0.346)
Inpatient care								
Public	0.012	0.012	0.012	0.020	0.010	0.011	0.015	0.005
		(0.030)		(1.770)*		(0.320)		(1.340)
Private	0.006	0.006	0.010	0.005	0.003	0.011	0.007	0.003
		(0.200)		(1.177)		(0.356)		(0.914)
Average age (in 2000)	48.50	48.89	41.69	41.51	22.01	21.47	38.46	40.05
N	3642	933	3026	759	2079	464	1029	269

Note: absolute *t*-statistics are in parentheses. *, ** and *** denotes statistical significance at 10, 5 and 1% level respectively. D=0 refers to non-treated individuals and D=1 refers to treated individuals. Figures are means in year 2000 in each sample. 'Child' refers to son/ daughter of the household head who answered adult booklets (15 years or older) in both sample years.

Lastly, with regard to inpatient care, the two samples provide different pictures. This may be explained by the very few hospitalisation cases in each year. Hence, although important, extra caution should be taken when dealing with inpatient care variables.

Nonetheless, it is still to be determined whether these differences in health care utilisation pattern are due to the health card program. Before doing the formal evaluation, Table 2A and B report the simple DID estimates for Sample 1 and 2 respectively. It can be seen that double-differencing eliminates almost all the previously observed differences in health care consumption between treated and non-treated individuals. This suggests that treated individuals have already used more public facilities, which generally provides cheaper services, than private facilities even without health card holding. Analyses that miss pre-treatment information therefore may provide misleading conclusion.

In Sample 1, it can be seen that there is almost no significant difference in health care consumption between treated and non-treated individuals. A possible explanation for this result is the relatively long time elapse between the control and the treatment samples (7 years), which allow individuals' health conditions to change perhaps substantially. But another reason could well be that the program has no effect. The picture is more promising in Sample 2, which involves much shorter time gap between samples, with treated children having significantly higher consumption of public health care to treat illness than non-treated children. Most of the increase in health care consumption occurred after the treatment as there were no significant differences between children's health care consumption pre-intervention. The table also shows that in general treated individuals rely less on private providers especially to perform preventative-type treatments.

Table 2A: *DID of Health Care Consumption by Treatment Status (Sample 1)*

	Head		Spouse		Child		Other	
	DID	|t-stat|	DID	|t-stat|	DID	|t-stat|	DID	|t-stat|
Outpatient care								
Public	0.030	1.196	0.030	0.822	0.006	0.054	-0.065	0.897
Private	-0.008	0.411	-0.003	0.126	0.044	0.447	-0.028	0.6511
Curative - public	0.021	1.475	0.024	1.112	0.006	0.126	0.022	0.424
Curative - private	0.010	0.905	0.002	0.141	-0.086	1.517	-0.027	1.358
Preventative - public	0.005	0.236	0.004	0.130	-0.012	0.122	-0.098	1.652*
Preventative - private	0.002	0.132	-0.006	0.322	-0.049	0.944	-0.001	0.032
Inpatient care								
Public	0.002	0.388	0.016	1.908*	0.016	0.560	-0.009	0.612
Private	0.006	1.182	-0.004	0.570	0.044	1.279	0.000	0.000
N	4455		3495		234		569	

Table 2B: *DID of Health Care Consumption by Treatment Status (Sample 2)*

	Head		Spouse		Child		Other	
	DID	|t-stat|	DID	|t-stat|	DID	|t-stat|	DID	|t-stat|
Outpatient care								
Public	-0.012	0.507	0.007	0.236	0.067	2.893***	0.072	1.612
Private	-0.021	1.228	-0.024	1.169	-0.054	2.253**	-0.006	0.187
Curative - public	-0.017	1.204	0.031	1.461	0.053	2.931***	0.051	1.504
Curative - private	0.000	0.006	0.011	0.891	-0.006	0.380	-0.015	0.675
Preventative - public	0.004	0.202	-0.015	0.551	0.023	1.126	0.020	0.591
Preventative - private	-0.021	1.714*	-0.037	2.217**	-0.048	2.623***	0.008	0.302
Inpatient care								
Public	0.000	0.005	0.010	1.486	0.003	0.439	-0.007	0.553
Private	0.000	0.060	-0.001	0.170	0.003	0.391	-0.005	0.722
N	4575		3787		2543		1298	

Note: *, ** and *** denotes 10, 5 and 1% significance level respectively.

Explanatory variables

Conditions that explain both eligibility and health care demand are relevant covariates. Eligibility is determined by household's welfare condition while health care demand is influenced by demographics, socioeconomic status, health care supply and environmental conditions. Fortunately, the IFLS data set is sufficiently rich to capture almost all of these variables.

The first set of variables consists of measures of household's welfare that determine both eligibility and health conditions. These variables include household compositions, value of asset, weekly expenditure, the proportion of household members who regard their general health as unhealthy, residential status (renting or not), source of drinking water, flooring materials and ownership of at least one form of health insurance. In addition, I include household's head characteristics such as age, gender, dialect, education and employment. Head's characteristics are observed by the program administrators, and so they influenced health card eligibility. For example, unemployment of the household head is listed as one of the eligibility criteria.

The second set of variables measure the extent of household's knowledge on health care facilities. Knowledge is important because it reflects the accessibility of a certain type of health facility and consideration to receive treatment at this facility. Furthermore, it has been suggested that the main way social relationships influence health care demand is through their effect on health knowledge (Andersen, 1995). Social networks disseminate references and updates on new products, and help locate appropriate health care providers (Weerdt and Dercon, 2006). As such, households with strong social network tend to be knowledgeable. This link suggests that we may reduce omitted variable problem associated with social network – which is hard to quantify – by including health knowledge as covariates. In the IFLS, spouses were asked whether or not they know the whereabouts of public hospitals, private hospitals, health centers, private practices, nurses or midwives and traditional practitioners that the family could go to. They respond on behalf of the entire household as the knowledge questions were not forwarded to other household members.

The final set of control variables deal with variations in the quantity and quality of health care providers. The availability of community-level data and facility surveys in the IFLS makes this study one of the few demand-side analyses that can jointly account for variations in the supply-side factors. It is well-known in the literature that supply influences demand through induced-demand or by altering people's behaviour towards illness. Information about the quantity of health facilities are obtained from the community data, while quality variables are obtained from the facility survey. To measure quality, I consider the availability of full-time health workers (e.g., GPs, dentists and nurses), birth services, laboratories and check-up equipment in facilities in the community. Furthermore, both demand and supply of health care may be affected by exogenous health shocks to the community. To control for this, two indicator variables for minor and major health shocks in the last three years are created. A minor (major) health shock refers to any health-related epidemics such as outbreak of diseases or flood which affected less than (at least) 50% of the local population. In Sample 1, 31 out of 312 IFLS communities (10 percent) experienced at least one major shock in the past 3 years and 132 communities (49 percent) experienced at least one minor shock in the same period. In Sample 2, the corresponding numbers are 41 (13 percent) and 100 (32 percent) communities.

In addition to these covariates, a dummy variable for urban area and 12 provincial dummy variables are included to control for regional heterogeneity. This is essential because Indonesia's population is very unevenly distributed. About 70% of the population lives in Java Island which has a land area of only 7% of the country's total dry land. As a result, development stages vary substantially across regions, and health resources are unevenly distributed (Lanjouw et al., 2000). Regional differences also carry exogenous variations such as differences in staple food, soil fertility and rainfall activity, which all affect local health conditions. Finally, sampling weights are used that take into account attrition in the survey (IFLS).

Table 3 reports selected summary statistics by treatment status. The full summary statistics are provided in Appendix A. All variables are characteristics as at pre-treatment period to ensure satisfaction of the exogeneity assumption (Caliendo and Kopeinig, 2005). Comparison between the two base years reveals high similarity between the two base samples. This is expected because most the IFLS3 households have complete information in the earlier two waves. Both years provide similar household head's characteristics, household compositions, initial level of health status with about 1 in 10 household members reported themselves generally unhealthy, and similar levels of health knowledge. The latter observation however is somewhat disappointing as it suggests that Indonesian's health knowledge has not improved over time. On the other hand, housing conditions were better in 1997. Meanwhile, comparing between treated and non-treated households, the table reveals that treated households tend to be headed by lowly educated heads, consist of more elderly members and fewer working-age adult members, have lower value of asset and expenditure, are less likely to have health insurance and live in relatively sub-standard quality houses with non-tile flooring and limited access to direct piped water for drinking and cooking. Treated households also tend not to be renters which may indicate long-term residents. Hence, although accompanied with considerable variations, average figures hint that the program have successfully reached its target. Using 1998 *Susenas* data set and benefit-incidence analysis, Lanjouw et al. (2000) support this conjecture that the poor does benefit from subsidised primary health care.

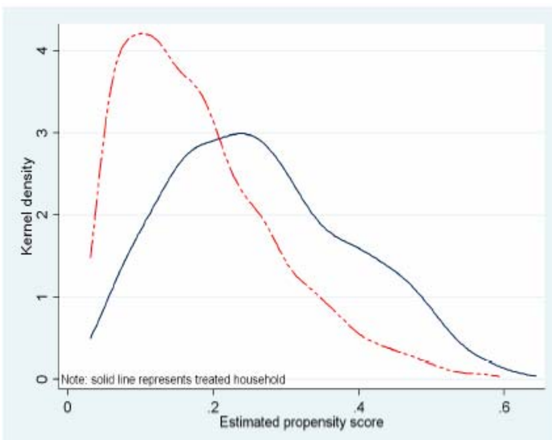
Table 3: *Selected descriptive statistics*

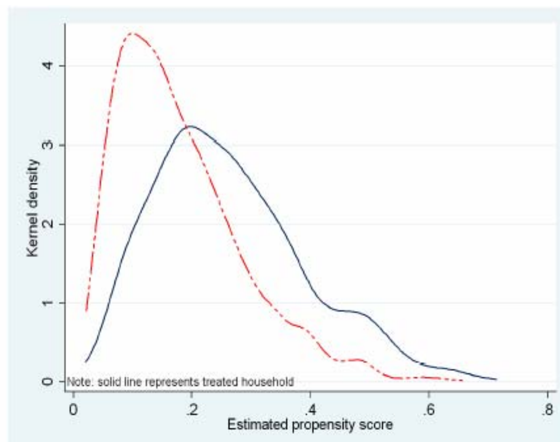
	Sample 1				Sample 2			
	Non-treated		Treated		Non-treated		Treated	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Household head's Characteristics								
Male	0.85	0.36	0.85	0.36	0.84	0.36	0.83	0.38
Primary school	0.50	0.50	0.58	0.49	0.55	0.50	0.61	0.50
Junior high school	0.13	0.33	0.10	0.30	0.11	0.31	0.10	0.29
Senior high school	0.15	0.35	0.10	0.29	0.12	0.33	0.11	0.26
College / higher	0.05	0.22	0.02	0.15	0.04	0.19	0.02	0.13
Household's Characteristics								
# under 6 years	0.59	0.77	0.61	0.77	0.49	0.72	0.52	0.70
# 6 – 14 years olds	1.05	1.12	1.00	1.10	1.05	1.08	1.04	1.05
# 15 - 49 years olds	2.46	1.44	2.22	1.29	2.94	1.73	2.70	1.57
# elderly (50+ years)	0.63	0.80	0.65	0.81	0.85	0.88	0.87	0.91
Not renting	0.80	0.40	0.84	0.37	0.85	0.36	0.88	0.32
Piped water	0.16	0.37	0.09	0.29	0.24	0.43	0.17	0.37
Ceramic floor	0.26	0.44	0.17	0.38	0.30	0.46	0.21	0.41
Cement floor	0.38	0.49	0.33	0.47	0.37	0.48	0.37	0.48
Bamboo floor	0.16	0.37	0.14	0.35	0.16	0.37	0.14	0.35
Insurance	0.19	0.39	0.14	0.35	0.20	0.40	0.13	0.34
% unhealthy	0.12	0.27	0.12	0.27	0.12	0.24	0.13	0.25
Log Asset	14.83	1.89	14.41	1.76	15.60	1.68	15.18	1.59
Log expenditure	10.73	1.21	10.43	1.14	10.94	0.96	10.76	0.85
Knowledge (spouse)								
Public hospital	0.68	0.47	0.62	0.49	0.59	0.49	0.57	0.50
Private hospital	0.37	0.48	0.29	0.46	0.34	0.47	0.29	0.45
Health centers	0.91	0.28	0.94	0.24	0.90	0.30	0.94	0.23
Private doctor	0.54	0.50	0.48	0.50	0.40	0.49	0.36	0.48
Midwife	0.66	0.47	0.66	0.48	0.77	0.42	0.78	0.42
Traditional healer	0.63	0.48	0.67	0.47	0.69	0.46	0.78	0.42
N	4,588		1,037		3,689		891	

V. RESULT

As discussed in Section III, to obtain unbiased estimates of the ATT, attention must be restricted to samples within the region of common support. In this case, imposing this condition only requires exclusion of a small number of households (less than 3 percent in both samples). This is encouraging as with large number of households within the common support, it is likely that the estimated ATTs are free of bias due to observables. Both samples readily pass the balancing tests as described by Becker and Ichino (2002) which ensures that the observed characteristics of matched-pairs are comparable.

Figure 1 and 2: *PSM in full sample*





Note: the left-side graph is for Sample 1 and the right-side is for Sample 2. The kernel density of the estimated propensity scores is calculated assuming Gaussian weight.

The quality of match is summarised in the above figures. It can be seen that both samples featured considerable overlapped regions, especially for Sample 2, suggesting that for many treated households there are many non-treated households with the same realisation of propensity score. They however are not perfect mapping. In fact, matching technique requires that every household constitutes a *possible* health card holder; the common support condition fails if households with certain characteristic are either always or never receive treatment (P(X_i) =0 or P(X_i) =1). So, the finding that no household in either sample has 0 or 1 realisation of its propensity score is actually encouraging in this sense, especially since the health card program is targeted to the poor.

Nonetheless, statistical significance of covariates in the matching equation means that households with health card have different characteristics than those without health card. Consequently, simple comparison between average outcomes of the treated and non-treated individuals (Tables 1A–2B) is unlikely to yield the true causal effects of health card availability. In general, it is found that health card holding is significantly more likely for households with inferior housing conditions, and have few assets and low income. The coefficients of household composition variables and household head's characteristics have the expected signs but are not significant. Rubin and Thomas (1996) however advise not to remove insignificant variables from the matching equation unless there is a strong reason to do so. With regard to the knowledge variables, the likelihood of health card holding increases with knowledge about the locations of public health centres and decreases with knowledge of private hospitals. This result may be explained by the tendency of private hospitals to be located in better-off neighbourhoods – with a few poor households – where there is demand for them.

Meanwhile, the relationship between household's health card status and supply variables is less clear. This makes sense because the inclusion of supply variables as covariates in the matching equation is justified on the basis that they influence health care utilisation, rather than because they influence one household's eligibility for a health card. It seems unreasonable to expect that changes in supply variables directly affect one household's welfare state, which in turn determines its eligibility for health card. With regards to the quantity of public health facilities in the community, the finding of lacking statistical insignificance is also consistent with the commitment of the primary health system in the country to provide at least one form of public provider within a defined distance or to arrange regular visitation by medical professionals if physical form of facility is not available. For the quality variables, in both samples, households with health card are more likely to reside in areas where birth services were initially limited, which is consistent with incidences of the family planning program in the late 1990s, but almost all of the health personnel variables are not statistically significant. Nevertheless, the direction of the coefficients of most personnel variables is negative. Unlike the supply-expansion part of the family planning program though, to my knowledge, there is no similar expansion in health personnel in the country. While, negative coefficients suggest that many treated households reside in areas with limited health personnel in the pre-intervention year. Thus, if excess capacity is unlikely, restricted supply conditions may inhibit the induced demand (due to health card availability) to be realised. This issue must be kept in mind by policymakers.

Compared to all samples, control households in the matched samples have relatively low income, more knowledge about public facilities and traditional healers and less knowledge about private facilities, indicating increased compatibility with treated households. Unlike in the case of matching by covariate however, the PSM method does not generally yield identical means of covariates for matched pairs because the propensity score is a summary measure that takes into account influences from all variables in the matching equation on the health card holding status. The PSM method instead places a larger emphasis on balancing covariates that are the key predictors of the treatment status as found in the logit regression.

Table 4A and 4B reports the ATT estimates for different household members in Sample 1 and 2, respectively. These are the main results. Differences between estimates in these tables with those in Table 2A–B reflect the confounding effects of observed covariates on health care consumption pattern. In general, the conditional estimates in Table 4A–B are smaller in absolute magnitude than the unconditional estimates in Table 2A–B suggesting that the overall effect of observables on the treatment status is positive. This direction is consistent with the fact that village heads disseminate health cards based on household's observed characteristics.

Table 4A: ATT *Estimates by Household Hierarchy (Sample 1)*

	Head			Spouse			Child			Other		
	Mean (C)	Mean (D)	ATT (s.e.)[B]	Mean (C)	Mean (D)	ATT (s.e.)[B]	Mean (C)	Mean (D)	ATT (s.e.)[B]	Mean (C)	Mean (D)	ATT (s.e.)[B]
Outpatient care												
Public	-0.010	0.029	0.039	-0.035	-0.011	0.025	-0.030	-0.023	0.007	-0.046	-0.107	-0.061
			(0.027)			(0.042)			(0.124)			(0.090)
Private	0.024	0.015	-0.009	0.019	0.003	-0.016	-0.004	0.070	0.073	0.014	0.000	-0.014
			(0.019)			(0.020)			(0.161)			(0.047)
Curative - public	0.018	0.037	0.019	0.036	0.045	0.009	0.011	0.023	0.012	0.067	0.054	-0.013
			(0.017)			(0.024)			(0.070)			(0.078)
Curative - private	0.018	0.005	-0.013	0.022	0.015	-0.006	0.036	-0.047	-0.082	0.025	0.000	-0.025
			(0.010)			(0.012)			(0.064)			(0.016)
Preventative - public	-0.026	-0.011	0.015	-0.067	-0.054	0.013	-0.041	-0.047	-0.005	-0.098	-0.161	-0.062
			(0.025)			(0.033)			(0.107)			(0.089)
Preventative - private	0.006	0.011	0.005	-0.002	-0.012	-0.010	-0.039	0.116	0.156	-0.012	0.000	0.012
			(0.017)			(0.017)			(0.142)			(0.052)
Inpatient care												
Public	0.003	0.004	0.000	-0.010	0.003	0.013	-0.006	0.000	0.006	-0.014	-0.018	-0.004
			(0.006)			(0.009)			(0.009)			(0.021)
Private	-0.001	0.007	0.008	-0.003	-0.009	-0.006	-0.008	0.047	0.054	0.009	0.000	-0.009
			(0.006)			(0.005)			(0.039)			(0.019)
N	3047	842		2425	661		149	43		360	112	

Table 4B: ATT *Estimates by Household Hierarchy (Sample 2)*

	Head			Spouse			Child			Other		
	Mean (C)	Mean (D)	ATT (s.e.)[B]	Mean (C)	Mean (D)	ATT (s.e.)[B]	Mean (C)	Mean (D)	ATT (s.e.)[B]	Mean (C)	Mean (D)	ATT (s.e.)[B]
Outpatient care												
Public	-0.008	-0.008	0.000	0.009	0.022	0.013	0.023	0.066	0.044	0.027	0.055	0.028
			(0.020)			(0.040)			(0.021)**			(0.045)
Private	0.014	0.006	-0.008	0.019	-0.020	-0.039	0.016	-0.010	-0.027	0.027	-0.004	-0.031
			(0.016)			(0.021)*			(0.017)			(0.025)
Curative - public	0.008	-0.001	-0.010	0.017	0.053	0.036	0.011	0.044	0.033	0.024	0.063	0.039
			(0.020)			(0.026)			(0.016)**			(0.038)
Curative - private	0.004	0.000	-0.004	0.008	0.011	0.003	0.008	0.006	-0.002	0.017	0.000	-0.017
			(0.014)			(0.013)			(0.008)			(0.015)
Preventative - public	-0.014	-0.006	0.008	-0.008	-0.022	-0.014	0.012	0.022	0.010	0.003	-0.008	-0.011
			(0.023)			(0.032)			(0.014)			(0.031)
Preventative - private	0.011	0.006	-0.005	0.012	-0.031	-0.043	0.009	-0.016	-0.025	0.010	-0.004	-0.014
			(0.014)			(0.017)***			(0.015)			(0.020)
Inpatient care												
Public	0.001	0.000	-0.001	-0.003	0.006	0.009	0.002	0.007	0.005	0.005	-0.004	-0.009
			(0.006)			(0.008)			(0.005)			(0.011)
Private	0.002	-0.001	-0.003	-0.001	-0.002	-0.001	0.006	0.006	0.000	-0.002	-0.004	-0.002
			(0.005)			(0.004)			(0.000)			(0.010)
N	3044	790		2442	638		2876	680		981	254	

Note: (s.e.)[B] denotes bootstrapped standard error with 200 replications. *, ** and *** denotes statistical significance at 10, 5 and 1% level respectively. C refers to control units and D is treated units. ATT are estimated for matched treated individuals in the region of common support.

For outpatient care, in Sample 1, none of the ATTs are statistically significant at any conventional significance levels. For household heads and spouses, the direction of the treatment effects for all types of care at public facilities though is positive. For children and other household members, some of the estimates have unexpected sign, but they may be unreliable due to small sample sizes. Zhao (2004), using Monte Carlo experiment, finds that the PSM method does not perform well in small sample, which is set at 500 in the study, compared to other method of matching (e.g., covariate matching) as the variance of the estimated treatment effect is too large. In this case, the sizes of both the children and other member samples in Sample 1 are less than 500 observations.

In Sample 2, the results persistently find that health card availability has no significant effect on treated heads' health care consumption. The ATT for every type of health care is very small. On the other hand, health card significantly lowers treated spouses' visitations to private providers to receive preventative-type treatments. Further investigation reveals that this result is largely driven by the spouses' demand for contraception. As have been mentioned previously, there had been a large expansion of contraceptive services in public facilities, which is part of the national family planning program in the late 1990s, that can be covered by health card. The consumption of preventative-type care at public facilities however is stable on average. A possible explanation for this is the much fewer occurrences of spouses' visits to private clinics compared to visits to public facilities (Table 1A–B). Given so, their visitations to private providers tend to be for a specific cause(s), such as to obtain contraception, while they visit public health providers for other services too. I will explore the effect of health card on contraceptive take-up next. On the other hand, treated spouses' visitations to public providers for curative-type treatments are higher after health card holding and this result is statistically significant for a positive alternative.

Health card availability especially benefited the son/daughters of the household head. In particular, it allows the children to pay more frequent visits to public health facilities to cure illness and receive medications. Because the initial consumption level is so low, the magnitude of the ATT is actually quite substantial when compared to the mean level pre-intervention; the availability of health card increases treated children's health care consumption by more than 80 percent. Their consumption of preventative-type treatments at public facilities also increases but is not statistically significant. Meanwhile, health card availability is associated with fewer visitations to private providers, but to a smaller magnitude than the overall increase in consumption of public health care, and it is not statistically significant for a causal relationship. For other household members, consistent with the results from Sample 1, health card has no meaningful effect on their health care consumption.

Meanwhile, Sample 1 and 2 provide consistent results with regard to inpatient care: in no case health card availability increases treated members' consumption of inpatient care. This is not unexpected given that inpatient treatment is a rarity for Indonesian households.

Contraceptive enrolment

The effect of health card on fertility decisions is also important as health card can cover both maternal care and contraceptive treatments; the former encourages a larger family size while the latter delays or prevents pregnancy. There are however reasons to believe that the effect of health card on encouraging larger family is small. These include uncertainty in continuous availability of health card and intensifying government's family planning campaign (which encourage 2 children per family). To investigate this matter, I consider contraceptive take-up during the study period and future plan to use contraceptive device for females who do not use contraception in the treatment year. In the IFLS, contraceptive booklets are forwarded to all females aged 15–54 year olds (eligible females). For the majority of households, there is only one eligible female in the household who is the spouse of the household head. There are about 3,799 eligible females in Sample 1 and 3,518 females in Sample 2. Table 5 reports the ATT estimates for eligible females in the two samples. 20 percent of them are members of treated households.

Table 5: *ATT estimates on contraceptive take-up*

	Sample 1			Sample 2		
	Mean (C)	Mean (D)	ATT (s.e.)[B]	Mean (C)	Mean (D)	ATT (s.e.)[B]
Enrollment at public facilities	0.141	0.208	0.067 (0.016)***	0.104	0.137	0.033 (0.015)**
Plan to use (if not using)	0.145	0.128	-0.017 (0.012)	0.141	0.142	0.002 (0.020)
N	2981	795		2788	725	

Figures under the (D) columns reveal the average proportions of new contraceptive enrolments by treated eligible females in matched households within 7 and 3 years time for Sample 1 and 2 respectively. Figures under the (C) columns are similarly defined for eligible females in the matched non-treated households. The mean for Sample 2 is lower due to enrolments prior to 1997.

The results are not surprising given the reinforcement to enroll is heightened by supply expansion in public facilities mentioned earlier. The ATT estimates suggest that treated eligible females take advantage of both health card coverage and the supply expansion to start using contraception. This result is consistent with Jensen (1996) which finds that Indonesian women's contraceptive behaviour is highly sensitive to the presence of subsidised facilities. Results from the RAND HIE have also suggested that preventative-type services are particularly price-sensitive due to reasons such as preventative-care as a luxury good as opposed to a normal good, and its high substitutability. Note though that enrollment does not guarantee continuity of use; females may fail to meet their next treatment if health card is not available then. Meanwhile, health card has no effect in altering attitude towards contraception among eligible females who, for various reasons (e.g., religious), choose not to use contraception.

Robustness Check

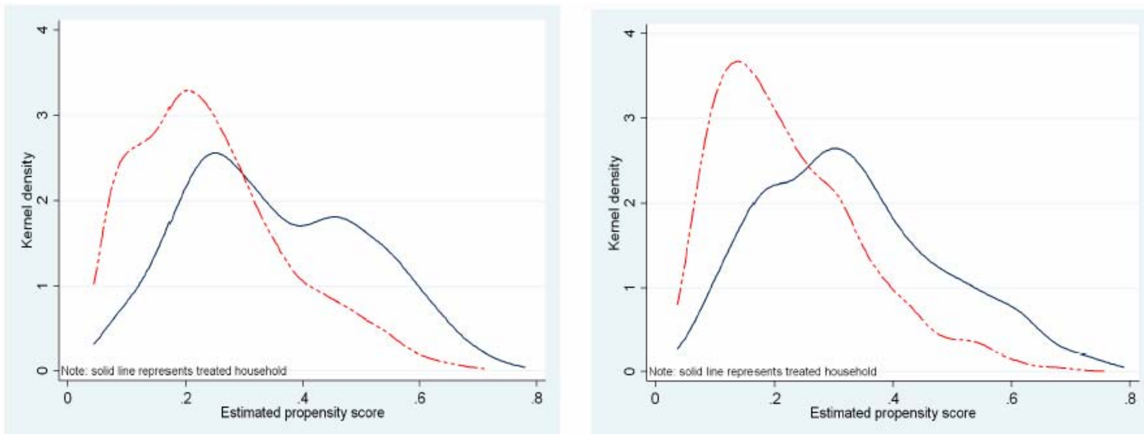
This section concerns with sensitivity of the results to assumptions made in producing them. The following explore several problems that may lead to biased results.

First, treatment effects may be heterogeneous and classification by household hierarchy does not sufficiently capture individual heterogeneity. For example, the age range for household heads is a wide 80 years. Other ways to slice the data hence is considered: by age group and by gender. The ATTs are re-calculated, and overall the results by age groups reflect the previous results with positive effects found among the younger cohorts – who are likely to be in the children sample in the earlier division by household hierarchy –, and negative effects on preventative care at private facilities for 30-49 years-olds, which is the age range of most spouses. While, the ATTs based on gender group are also calculated, and the results for males and females closely reflect the results for heads and spouses. One may also be tempted to slice the data by education levels, hypothesising that educated individuals have different health care consumption pattern than less educated individuals, but the sample sizes of the treated group for the higher education level sample get very small.

Second, the estimated ATTs may be confounded by the effect of relevant macroeconomic changes. So far, it is assumed that macroeconomic effects are homogenous across units – and so they are differenced away. The implication of this assumption is that all non-treated households are a potential comparison group. Nevertheless, it is often the case that macroeconomic movements affect different groups of the population differently. One way to investigate this possibility is to restrict the comparison group to households that share more similar characteristics with the treated households. Income level is commonly used to guide this division. However, in the face of substantial regional heterogeneity, this separation is complex (see Lanjouw et al., 2000; Booth, 1993 for discussion).

An alternative way may be to restrict the comparison group to households who reside in communities that issued a health card in 2000; no restriction is made on the treated households because health card can be used at any public health facilities in the country. In effect, this restriction eliminates exclusive neighbourhoods consisting of only rich households. Arguably, macroeconomic changes affect households in these communities more uniformly (compared to all sample). The resulting restricted sample contains of about 72 percent and 77 percent of all households in Sample 1 and Sample 2, respectively. The propensity scores are recalculated, and households are re-matched. The matching results are summarised in Figure 3 and 4.

Figure 3 and 4: *PSM in restricted sample*



Note: the left-side graph is for Sample 1 and the right-side is for Sample 2. The kernel density of the estimated propensity scores is calculated assuming Gaussian weight.

Similar to the previous matching results, there are considerable overlapped regions. But in these samples, the upper bounds of the estimated propensity score is closer to one compared to those obtained from the unrestricted samples. There are also fewer matched households with propensity score less than 0.2 in these samples than there are in the unrestricted samples. All of these results may reflect improved compatibility of matched pairs. In general, the estimated health card effects are larger in these samples as the means for the new control samples are smaller. This trend may suggest that macroeconomic crisis affects health care consumption pattern of the poor more than it affects the rich. This is consistent with the finding that income elasticity of health care is larger for poorer households than it is for poorer households (IRMS). However, the main thrust of the previous results is maintained that the health card program only have limited effects on health care consumption of its beneficiaries.

The last test but perhaps the most important of all is with regard to the first fundamental assumption in the PSM method that there is no selection on unobservables. Matching and the balancing tests adjust for biases due to non-overlapping support and differences in the distribution between treated and non-treated households. It however *assumes* that bias from the third source, that is, selection on unobservables, is zero. The above results therefore may be changed by factors that are not in the data. This possibility is rarely checked by researchers. Recently, Becker and Caliendo (2007) suggest an indirect check for this condition by asking the question how large the effect of the unobservables or "hidden bias" needs to be in order to reverse the results found by the PSM method. The test is based on non-parametric Mantel and Haenszel (MH) test statistic (1959) for binary outcomes with the null hypothesis that, given random sampling, the outcome variables are not affected by treatment status – otherwise, the outcomes may be positively or negatively affected by the treatment and the estimated treatment effect is said to be significant if the test statistic crosses a given critical value. As demonstrated in Aakvik (2001), the test involves comparing the number of treated individuals who are affected by the treatment and its expected number if the treatment has no effect. When the outcome variables are not binary in nature, Becker and Caliendo (2005) suggest transforming the variables so that an intended outcome is coded 1 and 0 otherwise.

To provide some underlying behind the test, consider a matched pair *i* and *j*, and let P_i and P_j be the probability that each individual receives treatment. The odds ratio that individuals are treated is:

$$\frac{P_i/(1-P_i)}{P_j/(1-P_j)} = \frac{P_i(1-P_j)}{P_j(1-P_i)} = \frac{\exp(\beta x_i + \gamma u_i)}{\exp(\beta x_j + \gamma u_j)}, \tag{6}$$

which becomes $\exp[\gamma(u_i - u_j)]$ if a matched pair has comparable observable covariates. In other words, theoretically, matched-pairs differ only by a factor γ and their unobservables *u*. The ignorability assumption requires that either $\gamma = 0$ or $(u_i - u_j) = 0$. An important result from Rosenbaum (2002) is that suppose the null hypothesis is false, the odds ratio that one of the matched pair receives treatment can be bounded by the following bounds:

$$\frac{1}{\exp(\gamma)} \leq \frac{P_i(1-P_j)}{P_j(1-P_i)} \leq \exp(\gamma). \tag{7}$$

The MH test statistic Q_{MH} follows the standard normal distribution in large sample and is given by:

$$Q_{MH} = \frac{|Y_1^D - \sum_{s=1}^{S} E(Y_{1s}^D)| - 0.5}{\sqrt{\sum_{s=1}^{S} \mathrm{var}(Y_{1s}^D)}}, \tag{8}$$

where Y_{1s}^D $(= N_{Ds}Y_{1s} / N_s)$ is the number of positive outcome for the treated sample (i.e., the number of treated individuals who use more health care post-treatment than they do pre-treatment) in stratum *s* of the sample, $N_{Ds} + N_{Cs} = N_s$, and $Y_{1s}^D + Y_{1s}^C = Y_{1s}$, where Y_{1s} is the total number of positive outcome in stratum *s*. Under the null hypothesis, there is no treatment effect and positive outcome is equally likely for treated and non-treated units. In this case, the strata are given by the estimated propensity score (Aavik, 2001).

Let $\Gamma = \exp(\gamma)$. Rosenbaum (2002) shows that for a given Γ and $u \in \{0,1\}$, the MH test statistic is bounded by two known distributions, which move apart from each other reflecting increased uncertainty about the test statistics in the presence of hidden bias. If $\Gamma = 1$, there is no hidden bias, and *i* and *j* have the same probability of being treated. The size of Γ hence reflects the extent matching results depart from the assumption of no hidden bias. The details about the MH test statistic and its bounds can be found in Aakvik (2001) and Becker and Caliendo (2005), which also offers a routine that calculates the upper and lower bounds and their respective probability values. For a given Γ, the upper bounds adjust the MH test statistics downwards when the ATTs are overestimated and the lower bounds adjust them upwards when the ATTs are underestimated.

In the case of the health card program, the direction of the hidden bias is not obvious. In the PSM-DID method, the hidden bias must come from time-varying unobserved heterogeneity. Note that selection on unobservables must not be confused with selection by observables, which we have explored so far and found to be in a positive direction. The former needs not follow the same pattern as the latter. Indeed, if there is a substantial selection by unobservables in the opposite direction, the true effect may be the reversed of the estimated effect. Here, given that most of the ATTs at public facilities have positive sign, which indicate benefit from health card availability, underestimation is somewhat less concerning than overestimation. In other words, if individuals with low value of unobservables are overrepresented in the treatment samples, the true effects will be larger and even more significant than what have been estimated – and this is not undesirable. We hence shall focus on the upper bounds which assume the estimated ATTs overestimate the true treatment effects. The test is performed for each sub-sample, specifying Γ at 0.05 interval from 1 to 2 (double the odds of being treated). $\Gamma = 2$ is considerably large given that the matching equation has included extensive background covariates. The outcome variables are transformed to binary variables (from count data), and the weighting functions are given radius matching with radius 0.01 (similar to the one used to produce results in Table 9); this test may be unsuitable for kernel matching which uses the entire sample (within the common support) as the matching pair (Becker and Caliendo, 2005).

The results from the tests suggest that already at small levels of Γ, health card availability no longer has significant effect on spouses and children in treated households (Sample 2). Under $\Gamma = 1$, the MH test statistic suggests significant treatment effects for spouses. But with a value of $\Gamma = 1.1$, the ATT for curative treatment at public facilities is no longer significant at 10 percent level (p-value for the upper bounds is 0.136). This means that the positive results might be reversed if spouses in treated and non-treated households are allowed to differ by just 10 percent in terms of unobserved characteristics. Meanwhile, the positive effect on treated children is more robust to influence from unobservables with $\Gamma = 1.3$ before the confidence interval of this effect includes 0. It is noteworthy though that $\Gamma = 1.1$ or $\Gamma = 1.3$ does not imply that unobserved heterogeneity exists and that there is no positive effect from health card availability. This test is not designed to directly justify or invalidate the ignorability assumption. Rather, it suggests extra caution must be taken when concluding positive effect of health card, because the confidence interval for this effect would include 0 if unobservables cause the odds ratio of treatment status to differ between the treatment and the control groups by small magnitudes.

In other cases where the treatment effects at $\Gamma = 1$ are insignificant, the bounds informs us the degree of hidden bias, positive and negative, the ATTs would become significant. Assuming 5 percent significance level, for household heads, the conclusion that health card has no effect on health care consumption in general is robust to a hidden bias that would increases Γ to 1.1 in Sample 1 and 1.4–1.55 in Sample 2. Meanwhile, for other members, the conclusion of no effect is robust to a hidden bias that would increase the odds of receiving treatment up to 1.8 (Sample 2). In short, these checks suggest that it is highly unlikely that the result of no positive effect is reversed.

Overall, the results are pointing to the same direction, independent of the chosen comparison group or estimation technique: there is lacking evidence of positive impact of the health card program.

Discussions

Reconciling these results with prior expectation and evidences from developed countries, the limited effect of health card is counterintuitive; by design, health card offers generous subsidy to those who were formerly restricted in access to formal health care. For instance, many public programs in developed countries, especially generous ones, are equipped with mechanism that discourages excessive usage or other rent-seeking behaviours (Manning et al., 1987; Riphan et al., 2003 provide contrasting evidence from developed countries). However, there are several explanations that may rationalise this result.

First, Poterba (1994) argues that government intervention through price subsidy is ill-suited when the price elasticity of demand for the subsidised good is low, or when there is large uncertainty and divergence in this elasticity across units. From Indonesia's own health experiment study (the IRMS), Gertler (1995 in Lanjouw et al., 2000) finds that the demand for health services in general at Indonesian public health facilities fit this condition; it is inelastic and varies greatly with income. Standard economic demand-supply theory predicts also a small change in quantity for a given price change when demand schedule is inelastic.

Second, a relatively stable health care consumption may be due to the fact that households are not selected exclusively on the basis of their health conditions; healthy individuals – by his/her own standard – need very little medical care in a given period of time. Meanwhile, health card availability does not mandate (higher) health care consumption; like other households, households with health card can choose whether or not to seek medical treatment, from whom they wish to receive medical treatment, and the level of treatments to be consumed. Of course, increased visitations to health facilities are not necessarily a good outcome. However, the serious problem of underutilisation of health care particularly by the poor is well-known. It is commonly argued that the poor have adapted to certain adverse health conditions that they do not consider themselves sick when others facing a similar condition do. Consequently, this so-called adaptation bias sets the tolerance level for sickness before concluding a need for medical care higher among poor households than other households in general.

A related issue is the variation in the extent of information given out by the village heads when distributing health cards. Because of the decentralised nature of the health card program (and the absence of official public information such as brochures explaining the program at least in the first few years of the program), households may receive different levels of information about the objective of the program and the functioning of health card, such as with regard to the scope of coverage. For instance, recipient households may understand that health card can be used in the event of illness but are uncertain about its use in other circumstances, such as preventative-type services and non-emergency procedures. There is in fact a body of literature arguing that recipient households are unlikely to misuse public sympathy (Miguel and Gugerty, 2005; Fehr and Gachter, 2000; Besley et al., 1993). These studies argue that the fear of social sanctions, such as isolation and shame, discourages recipient households from economising health card availability. The credibility of this theory has been highly praised in developing countries where there is great physical proximity between neighbours and households tend to be long-term and less mobile residents.

The third explanation relates to the adequacy of the public health system. It has been widely recognised that public health facilities are often inadequate in developing countries. A survey study by Filmer et al. (2000) for instance note the possibility of subsidised patients choosing full-price providers if they are of higher quality. Given that the IFLS data contains data on the supply side, future research should examine the supply conditions available to different households.

To sum up, limitedness of the effect of a public program is not a new result in the developing countries' context. In the case of the health card program, this result is somewhat expected as the demand schedule for general health care in public facilities is not price-sensitive to begin with. In addition, the demand incentive created by the program may also be counteracted by other costs associated with the subsidised care such as time costs and limited supply conditions. For the government, this finding has important implications, particularly with respect to resource allocation and designs for future policies. The above discussions suggest that to encourage health care use, at least in the short-run, the priority should be on expanding the public health system.

VI. CONCLUSION

It has been suggested that poor households have low health because they lack access to adequate health care. The intention of the 1994 health card program has been to protect the health status of Indonesian poor by allowing them access to adequate health care. By design, the program is fairly generous providing full coverage for a wide range of primary health care services at public health facilities and allowing for unlimited claims. However, the findings in this thesis show that in general recipient households do not exploit the presence of health card and increase health care utilisation. There is some indication that younger members in the household made more visits to treat illness, but this result is not robust. There is also evidence that health card availability encourages contraceptive take-up by eligible females in the recipient households, but in this case, the demand reinforcement was paralleled with the expansion of family planning services in the public health facilities in the late 1990s. Meanwhile, health care utilisation by household heads and other household members are unaffected by the presence of health card.

In coming to this conclusion, PSM-DID estimator is applied on two large longitudinal samples with different base year. This estimator has not been extensively used in the health literature, but has been shown to be powerful in providing causal estimates in labour applications. More specifically, important advantages of the matching technique over the standard regression methods include flexibility with regard to functional forms, involvement of an explicit test to ensure that the estimated program effects are supported by the data, and the combination between PSM and DID techniques eliminates the effects of time-specific heterogeneity and common macroeconomic effects. Further, matching technique is particularly suited for the current setting for the following two reasons. First, the richness of the data set allows matching technique to deal with selection biases due to observables considerably well. The availability of data on health care providers adds credit to this study, as it manages to jointly account for variations in demand and supply factors faced by different households that are often neglected in demand-sided studies. Second, the feature of the health card program that determines eligibility based on observed characteristics suggests that the assumption about selection on observables may be more plausible than assuming selection on unobservables made by the standard selection equation model.

This study has several important implications. First, the health card program, in spite of its well intention, is not well designed. Although more studies are needed to suggest a suitable program design, there are suggestions that households may be more responsive to incentives in specific health care services as opposed to a general coverage or to programs that are targeted to individuals as opposed to households. Second, limitedness of the effect of the health card program highlights the role of other factors in influencing the performance of the program. A major problem in developing countries is inadequate supply of health care. Policymakers should ensure that there is no mismatch between demand and supply in the sense that the program attempts to expand demand in areas where supply conditions could not deliver.

Finally, the results provide yet another reason that results from developed countries might not be generalised to programs applied in developing countries. For example, moral hazard behaviour associated with generous programs is a common concern in developed countries. In contrast, in developing countries, the cost of risky behaviour or reduced care is increasingly high in the absence of compensation payments and insurance. The results find no evidence of household members in recipient households reducing their consumption of preventative-type health care, as predicted by the moral hazard theory, as health card becomes available. In addition, in developing countries the role of informal health care, such as traditional healers and home-grown remedies is prevalent, and households consider them as acceptable substitutes for formal health care. In most cases, they are cheap and highly accessible. Future design for program aiming to encourage use of formal health care therefore may be accompanied with educational material to alter preferences. Implementing public programs may be more challenging for government in developing countries in the face of constrained resources, lack of knowledge and market imperfections.

References:

Aavik A. Bounding a matching estimator: the case of a Norwegian training program. Oxford Bulletin of Economics and Statistics 2001; 63(1); 115-142.

Abadie A, Imbens G W. On the failure of the bootstrap for matching estimators. NBER Technical Working Paper 2006 no.325.

Akin J S, Guilkey D K, Hutchinson P L, Mcintosh, M T. Price Elasticities of Demand for Curative Health Care with Control for Sample Selectivity on Endogenous Illness: An Analysis for Sri Lanka. Health Economics 1998; 7(6); 509-531.

Andersen R M. Revisiting the behavioral model and access to medical care: does it matter? Journal of Health and Social Behaviour 1995; 36; 1–10.

Angeles G, Guilkey D K, Mroz T A. The Effects of Education and Family Planning Programs on Fertility in Indonesia. Economic Development and Cultural Change 2005; 54(1); 165-201.

Augurzky B, Schmidt C. The Propensity Score: A Means to An End. IZA Discussion Paper 2001 no. 271.

Becker S O, Caliendo M. Sensitivity analysis for average treatment effect. The Stata Journal 2007; 7(1); 71-83.

Becker S O, Ichino A. Estimation of average treatment effects based on prepoensity scores. The Stata Journal 2002; 2; 358-377.

Besley T, Coate S, Loury G.The economics of rotating savings and credit associations. American Economic Review 1993; 83(4); 792–810.

Case A, Lubotsky D, Paxson C. Economic Status and Health in Childhood: the Origin of the Gradient. American Economic Review 2002; 92(5); 1308-1334.

Castro-Leal F, Dayton J, Demery L, Mehra K. Public Spending on Health Care in Africa: Do the Poor Benefit? Bulletin of the World Health Organization 2000; 78(1); 66-74.

Cobb-Clark D A, Crossley T. Econometrics for Evaluation: An Introduction to Recent Developments. Economic Record 2003; 79(247); 491-551.

Dehejia R H, Wahba S. Propensity Score Matching Methods for Nonexperimental Causal Studies. Review of Economics and Statistics 2002; 84(1); 151.161.

Dow W H, Gertler P, Schoeni R F, Strauss J, Thomas D. Health Care Prices, Health and Labor Outcomes: Experimental Evidence. RAND Working Paper 2000; California.

Filmer D, Hammer J S, Pritchett L H. Weak Links in the Chain: A Diagnosis of Health Policy in Poor Countries. World Bank Research Observer 2000; 15(2); 199-224.

Fehr E, Gachter S. Cooperation and punishment in public goods experiments. American Economic Review 2000; 90(4); 980–994.

Frankenberg E, Beegle K, Thomas D, Suriastini W. Health, Education and the Economic Crisis in Indonesia. RAND working paper 1999; California.

Frankenberg E. The Effects of Access to Health Care on Infant Mortality in Indonesia. Health Transition Review 1999; 5; 143-146.

Galiani S, Gertler P, Schargrodsky E. Water for Life: The Impact of the Privatization of Water Services on Child Mortality. Journal of Political Economy 2005; 113; 83–120.

Gertler P J, Molyneaux J W. The Asian Demographic Transition and Its Causes: The Impact of Targeted Family Planning Programs in Indonesia. Population and Development Review 2000; 26 (supplement); 61-85.

Heckman J, Ichimura H, Todd P E. Matching as an Econometric Evaluation Estimator: Evidence for Evaluating a Job Training Programme. Review of Economic Studies 1997; 64(4); 605-654.
------- Characterizing Selection Bias Using Experimental Data. Econometrica 1998; 66(5); 1017-1098.

Jensen E R. The Fertility Impact of Alternative Family Planning Distribution Channels in Indonesia. Demography 1996; 33(2); 153-165.

Lanjouw P, Pradhan M, Saadah F, Sayed H, Sparrow R. Poverty, Education and Health in Indonesia: Who Benefits from Public spending? World Bank 2000.

Lieberman S S, Marzoeki P. Health Strategy in a Post-Crisis, Decentralizing Indonesia. World Bank Report 2000; no. 21318-IND.

McGuire T 2000. Physician agency. In: Culyer and Newhouse (Eds), Handbook of Health Economics vol.1A. North Holland: Amsterdam; 2000.

Miguel E, Gugerty M K. Ethnic diversity, social sanctions, and public goods in Kenya. Journal of Public Economics 2005; 89(11); 2325-2368.

Newman J, Pradhan M, Rawlings L B, Ridder G, Coa R, Evia J L. An Impact Evaluation of Education, Health, and Water Supply Investments by the Bolivian Social Investment Fund. The World Bank Economic Review 2002; 16(2); 241-274.

Pitt M, Rosenzweig M, Gibbons D. The Determinants and Consequences of the Placement of Government Programs in Indonesia. World Bank Economic Review 1993; 73; 319-348.

Poterba J M. Government Intervention in the Markets for education and health care: how and why? NBER Working Paper 1994 no. 4916.

Pradhan M, Rawlings L, Ridder G. The Bolivian Social Investment Fund: An Analysis of Baseline Data for Impact Evaluation. World Bank Economics Review 1998; 12; 457–482.

Pradhan M, Saadah F, Sparrow R A. Did the Health Card Program Ensure Access to Medical Care for the Poor during Indonesia's Economic Crisis? World Bank Economic Review 2007; 21(1); 125-150.

Smith J, Todd P. A Critical Survey of Empirical Methods for Evaluating Active Labor Market Policies. Swiss Journal of Economics and Statistics 2000; 136(3); 1-22.
----- Does Matching Overcome Lalonde's Critique of Nonexperimental Estimators? Journal of Econometrics 2005; 125(1); 305-375.

Thomas D, Frankenberg E, Smith J P. Lost But Not Forgotten: Attrition and Follow-up in the Indonesia Family Life Survey. Journal of Human Resources 2001; 36(3); 556-592.

Trujillo A, Portillo J, Vernon J. The Impact of Subsidized Health Insurance for the Poor: Evaluating the Colombian Experience Using Propensity Score Matching. International Journal of Health Care Finance and Economics 2005; 5(3); 211-239.

Wagstaff A. Health Systems in East Asia: What Can Developing Countries Learn from Japan and the Asian Tigers? World Bank Working Paper 2005 no. 3790.

Wagstaff A, Yu S. Do health sector reforms have their intended impacts?: The World Bank's Health VIII project in Gansu province, China Journal of Health Economics 2007; 26(3); 505-535.

Weerdt J D, Dercon S. Risk-sharing networks and insurance against illness**.**
Journal of Development Economics 2006**;** 81(2); 337-356

Xu K, Evans D B, Kawabata K, Zeramdini R, Klavus J, Murray C J L. Household Catastrophic Health Expenditure: A Multicountry Analysis. Lancet 2003; 362(9378); 111-117.

van de Walle D. The Distribution of Subsidies through Public Health Services in Indonesia, 1978–87. World Bank Economic Review 1994; 8(2); 279-309.

Zhao Z. Using Matching to Estimate Treatment Effects: Data Requirements, Matching Metrics, and Monte Carlo Evidence. Review of Economic and Statistics 2004; 86(1); 91-107.

Appendix
Results for the matching equation

	Sample 1		Sample 2	
	Coeff.	|t-stat|	Coeff.	|t-stat|
Household head's Characteristics				
Age	-0.005	1.25	-0.007	1.53
Male	0.141	1.22	-0.045	0.38
Primary	0.126	1.18	0.197	1.71*
Junior school	-0.118	0.73	0.225	1.28
Senior school	-0.218	1.28	-0.148	0.76
College/ higher	-0.444	1.68	-0.224	0.70
Working	-0.085	0.76	-0.044	0.33
Muslim	0.043	0.25	0.346	2.01**
# acute problem	0.011	0.39	0.001	0.07
Household's Characteristics				
# < 6 years old	0.058	1.13	0.099	1.7*
# 6 -14 years old	0.036	1.00	0.065	1.60
# 15 - 49 years old	0.009	0.29	-0.002	0.07
# > 49 years old	0.052	0.76	0.113	1.80*
Speak Dialects	0.078	0.87	0.042	0.45
House self-owned	0.141	1.21	0.310	2.27**
Piped water	-0.347	2.7***	-0.247	2.19**
Ceramic floor	-0.549	4.19***	-0.571	4.19***
Cement floor	-0.407	3.86***	-0.296	2.51**
Bamboo floor	-0.180	1.20	-0.355	2.02**
Health cover	0.103	0.88	-0.179	1.45
Average unhealthy	0.099	0.66	0.114	0.64
Log asset	-0.073	2.99***	-0.131	4.29***
Log expenditure	-0.084	2.29**	0.010	0.16
Knowledge				
Know public hospital	-0.083	0.94	0.047	0.49
Know private hospital	-0.201	2.06**	-0.134	1.27
Know health centers	0.312	2.05**	0.562	3.44***
Know private doctor	0.174	1.97**	-0.045	0.46
Know midwife	0.101	1.22	0.083	0.82
Know traditional healer	-0.037	0.44	0.219	2.09**

Appendix (Continued)

	Sample 1		Sample 2	
	Coeff.	|t-stat|	Coeff.	|t-stat|
Supply				
Average full-time doctor	-0.214	1.77*	0.116	1.06
Average full-time dentist	0.258	2.14**	-0.131	1.23
Average full-time nurse	-0.031	0.96	-0.087	2.66***
Average full-time midwives	0.022	0.54	-0.027	0.55
Average full-time paramedics	-0.034	1.51	0.055	1.70*
Laboratories	0.173	1.29	0.195	1.18
# beds	-0.003	0.67	0.014	1.92*
Inpatient care	0.709	2.87***	0.461	1.84*
Check-up	0.502	2.94***	-0.426	0.68
Birth services	-0.666	3.41***	-0.492	2.61***
Private - inpatient care	0.195	0.63	0.804	2.62***
Private - minor surgery	-0.179	1.01	na	-
Private - check-up	0.040	0.40	-0.181	0.71
Private - birth services	-0.212	0.84	-0.825	3.60***
# private hospital^	0.066	0.93	0.038	1.03
# public hospital	0.081	1.45		
Village Characteristics				
Urban area	0.351	3.18***	0.426	3.65***
Minor shock	-0.006	0.08	0.081	0.86
Major shock	0.097	0.83	-0.256	1.92*
Constant	0.621	1.06	0.836	0.85
N	5262		4580	
Log Likelihood	-2390.6		-2065.9	
R squared	0.085		0.085	

Note: ^ in Sample 2, public and private hospitals are combined to make total number of hospitals. In addition to these variables, 12 regional dummy variables and sampling weight are included in the estimation. Estimates are logit coefficients, and *, **, *** denotes significance at 10, 5 and 1% level respectively.