IFAU – INSTITUTE FOR
LABOUR MARKET POLICY
EVALUATION

The information method
– theory and application

Per Engström
Patrik Hesselius

WORKING PAPER 2007:17

The Institute for Labour Market Policy Evaluation (IFAU) is a research institute under the Swedish Ministry of Employment, situated in Uppsala. IFAU's objective is to promote, support and carry out: evaluations of the effects of labour market policies, studies of the functioning of the labour market, evaluations of the labour market effects of measures within the educational system and evaluations of the effects of the social insurance on the labour market. Besides research, IFAU also works on: spreading knowledge about the activities of the institute through publications, seminars, courses, workshops and conferences; influencing the collection of data and making data easily available to researchers all over the country.

IFAU also provides funding for research projects within its areas of interest. The deadline for applications is October 1 each year. Since the researchers at IFAU are mainly economists, researchers from other disciplines are encouraged to apply for funding.

IFAU is run by a Director-General. The authority has a scientific council, consisting of a chairman, the Director-General and five other members. Among other things, the scientific council is responsible for giving proposals for decisions about what research grants should be approved. A reference group including representatives for employers and employees as well as the ministries and authorities concerned is also connected to the institute.

Postal address: P.O. Box 513, 751 20 Uppsala
Visiting address: Kyrkogårdsgatan 6, Uppsala
Phone: +46 18 471 70 70
Fax: +46 18 471 70 71
ifau@ifau.uu.se
www.ifau.se

Papers published in the Working Paper Series should, according to the IFAU policy, have been discussed at seminars held at IFAU and at least one other academic forum, and have been read by one external and one internal referee. They need not, however, have undergone the standard scrutiny for publication in a scientific journal. The purpose of the Working Paper Series is to provide a factual basis for public policy and the public policy discussion.

ISSN 1651-1166

The information method – theory and application*

Per Engström and Patrik Hesselius†

20th August 2007

Abstract

When estimating the extent of e.g. excess use of public benefits one traditionally uses direct monitoring. Such direct estimates are afflicted with an intrinsic negative bias since you only count what you find. This paper presents and assesses an alternative intuitive, yet relatively unexplored, approach that may reduce the bias by making use of the individual's own response to information of increased monitoring. Through an extensive randomized social experiment we apply the method to one particular Swedish public benefit: Parental Benefit for Temporary Childcare. In our view the application was successful: the results are interpretable and we are able to surface more hidden excess use through the information method. As a rough estimate we find that the information based estimate of excess use is 40 percent higher than the corresponding estimate based on ordinary random monitoring (22.5 percent compared to 16 percent). The method is potentially applicable to a large number of related fields, such as e.g. tax evasion and insurance fraud.

JEL: C51, C93, H55. Keywords: Monitoring, Social insurance, Randomized experiments

*The authors would like to thank David Card, Kenneth Carling, Matz Dahlberg, Richard Freeman, Peter Fredriksson, Anders Forslund, Per Johansson, Bertil Holmlund, Marten Lindeboom, Oskar Nordström Skans, Malin Persson, Knut Røed, Gerard van den Berg and seminar participants at IFAU and Tinbergen Institute for valuable comments and suggestions. Thanks also to Bengt-Olle Andersson and the people at Swedish Social Insurance Agency. Special thanks to Björn Blomqvist and Bettina Kashefi for supporting the project.

†Institute for Labour Market Policy Evaluation (IFAU) and Department of Economics, Uppsala University. E-mail addresses: per.engstrom@ifau.uu.se and patrik.hesselius@ifau.uu.se

Contents

1 Introduction

Any tax or transfer system may potentially be misused; it is virtually impossible to design an insurance, benefit, subsidy or tax that cannot be used in a way that is not in line with its set of rules and intentions. Consequently, the economic research in the fields of monitoring unemployment- and sickness insurance claimants, tax evasion, and private insurance fraud, is vast and ever growing.[1] This paper is concerned with the measurement of the extent of these illicit activities. Such measurements are intrinsically complex due to the fact that the subject does not want the activity to be observed. The literature has produced measures that are often based directly on the share of claims etc. that fail when monitored and closely inspected.[2] Such estimates are afflicted with an intrinsic weakness in that they are in virtually all cases downward biased; with limited resources and realistic legal restrictions you cannot expect to find all misuse in the selected subset. The purpose of this paper is to present an alternative method – the *information method* – to estimate the extent of misuse. The method is, to a varying degree, applicable in all of the areas mentioned above.

The key ingredient in the information method is a "modest" randomized experiment; "modest" in the sense that the treatment does not involve a different set of explicit benefit rules in relation to the control group, which is most often the case in randomized social experiments (see e.g. Moffitt, 2003, Heckman and Smith, 1995 and Burtless, 1995, for extensive surveys of the pros and cons of randomized social experiments). Instead, the only treatment is that a randomly selected group of people is made aware that their observance of the set of rules, regarding e.g. the benefit, will be monitored more intensely, for a limited time period, than what is normally the case.[3] Under rather general assumptions, such knowledge

[1] For surveys regarding the incentive problems in unemployment insurance see Fredriksson and Holmlund (2006) and Holmlund (1998); for a major randomised experiment on the effects of monitoring sickness insurance claimants see Hesselius et al (2005); for a seminal contribution on the theory of tax evasion see Allingham and Sandmo (1972) and for surveys see Sandmo (2005) and Schneider and Enste (2000 and 2002); and for a recent survey on private insurance fraud see Viaene and Dedene (2004).

[2] An important exception is the field of measuring tax evasion and the shadow economy, in which more innovative indirect estimation techniques are often used; such as consumption based estimates (Pissarides and Weber, 1989), Money demand methods, (Tanzi, 1983) and electricity based estimates (Lacko, 1998).

[3] When describing the logic behind the information method we will stick to the ex-

will naturally lower the excess use of the benefit in the selected group.[4] By comparing the average use of the benefit in the treatment group with the average use in the control group, we may produce an indirect estimate of the average excess use pertaining to the default monitoring intensity regime (which is here implicitly assumed to be our target variable); this is the basic intuition of the information method. In addition, the announced increase in monitoring of the treatment group is indeed carried out. And the average number of withdrawals that fail, despite the announced increase in monitoring, is an additional ingredient in the full estimate of excess use.

The paper derives the theoretical properties of the estimates generated by the information method. It is shown that the method produces estimates that are afflicted with potentially smaller biases than the method based on traditional random monitoring does. In other words, a larger part of the hidden excess use may be exposed; On the aggregate rather than the individual level, however. The intuition behind this attractive property is that the information based estimate identifies the full amount of excess use that is eliminated by the announcement of increased monitoring, while traditional random sampling only partially identifies this measure.

In principle, the information method can be applied to a vast number of circumstances that involve estimating the general level of observance of some regulations or laws. The basic conditions for when the information method may be applied are twofold. First, it must be possible to increase the level of monitoring for some randomly selected subset of individuals. And second, there must be a measurable individual variable that captures the individuals' "engagements" in the phenomenon that we wish to study; in the example of a benefit this variable simply consists of the individuals' use of the benefit, and in the case of tax evasion it may consist of the declared individual income or the level of tax deductions made. How-

ample of a public benefit, since this is the institution that will be subject to our application of the method.

[4] Unemployment insurance has often been examined using randomised experiments featuring monitoring regimes that differ in intensity and design. The disciplinary effect that monitoring has on e.g. job search intensity is well documented, see e.g. Van Den Berg and Van Der Klaauw (2006), Dolton and O'Neill (1996), Benus and Johnson (1997) and Hägglund (2006). However, for reasons discussed in this paper (see section 4.1 below), one may not interpret such disciplinary effects as estimates of pure excess use, when studying a very complex benefit such as unemployment insurance.

ever, given that these basic requirements are satisfied, there may still be a number of intrinsic characteristics, of the phenomenon that we wish to study, that disqualifies the information method as a fruitful estimation technique. The theoretical part of the paper aims at identifying, in very general terms, when the information method is likely to succeed and when the results are most easily interpretable.

The information method is not completely new to the literature (the name is, however). We have found at least one prior application. In 1995 an extensive experiment was carried out by the local tax authority in Minnesota to test different strategies to inspire the citizens' inclinations to pay taxes (see Coleman, 1996, for the details in the experiment and Slemrod et al, 2001, for a scientific study of tax evasion based on the experiment). Some treatment groups were informed that their tax returns would be subject to increased inspection. Slemrod et al (2001) interprets the induced increase in declared taxes as a potential measure of tax evasion. This is an intuitive interpretation and it is indeed the key identification strategy used in the information method. However, the interpretation may be more or less adequate and fruitful under different circumstances. This is why we choose to include a formal treatment of the method. To our knowledge, our theoretical presentation of the method, as a potential improvement compared to traditional random monitoring, is new to the literature. In addition to the Minnesota experiment there is a British study carried out in 2002 along approximately the same lines (Hasseldine et al, 2007).[5] However, the focus in this study is not explicitly on estimating tax evasion but rather on evaluating different tax compliance enforcement strategies.

In addition to the theoretical analysis we present an application of the method. During the spring of 2006 the Swedish Social Insurance Agency carried out an extensive randomized experiment on one of its social benefits: Parental Benefit for Temporary Childcare (Tillfällig Föräldrapenning för Vård av Barn), which from now on is denoted by its Swedish abbreviation, VAB. We designed and directed the experiment.[6]

[5] One important difference, though, is that the British study was directly aimed at entrepreneurs with a turnover less than £ 15 000. There is indeed support for the stylized fact that self-employed individuals are more likely to engage in tax evasion; see e.g. Slemrod et al (2001) and Joulfaian and Rider (1998); it is also the very identifying assumption in the expenditure based approach, see e.g. Pissarides and Weber (1989), Engström and Holmlund (2006), Apel (1994), Schuetze (2002), Lyssiotou et al (2004) and Tedds (2004).

[6] For a full report on the study see Engström et al (2007).

The VAB-Benefit has some intrinsic features that make it well adjusted for a pilot study of the information method. The most important feature is that it has a very simple set of rules which means that excessive use of the benefit is likely to be deliberate and well defined. In our view the application of the information method on the VAB-Benefit was a success. The method indeed seems to be able to expose otherwise hidden excess use; our point estimates roughly indicate that we found 40 percent more excess use with the information method compared to traditional random monitoring. Furthermore, we provide a robustness check that analyses whether the increased estimate stems from some sort of over-reaction to the notice of increased monitoring; and we find no evidence of such.

The outline of the paper is as follows. The first sections (2-4) present the theoretical properties of the information method. In Section 2 we set up the basic optimization problem facing an individual that may choose to misuse a public benefit (the analogous analysis for e.g. tax evasion would render the same properties). Based on this behavioral model, Section 3 defines our estimators and compares their properties. We proceed, in section 4, by extending the behavioral model along some different realistic margins in order to characterize some caveats of the information method. In Section 5 we turn to the application of the method. The properties of the VAB-Benefit are described, followed by a brief description of the experiment. Before presenting the results we also give a brief data description and some words about the statistical method that we employ. The section ends with a robustness analysis regarding whether the individuals over-react to the notice of increased monitoring. Section 6 summarizes and concludes the paper.

2 The basic behavioral model

There is a large population of size N. Assume further that each individual (i) has need (i.e. legitimate use) for some benefit $b_i(\Omega)$, where Ω is a vector of policy parameters and general macro variables. Each individual potentially misuses the benefit to an exogenously given monetary equivalent of θ.[7] The total observed withdrawal of the benefit (w_i) can be written $w_i = b_i + \theta I_i$, where I_i is an indicator variable that takes value 1 when

[7] We may introduce endogenous θ without changing the basic qualitative results. We will return to these issues in the extensions presented below.

the individual actually misuses the benefit and zero otherwise. There is a risk of getting caught when misusing the benefit given by $p(m_i)$ where m_i is the level of monitoring aimed at individual i.[8] It holds that $p'_m > 0$ so that increased monitoring increases the risk of getting caught with excess use.

There is a fixed heterogeneous moral cost of misuse κ_i, that captures all heterogeneity in the model. The cost of getting caught with misuse (c) is taken as exogenous. The expected utility for individual i is then written as,

$$U_i = b_i(\Omega) + I_i\left[\theta - (\kappa_i + p(m_i)c)\right]. \tag{1}$$

The simple behavioral rule consists of a breakpoint moral cost of misuse ($\bar{\kappa}_i$) over and at (by assumption) which the individual will choose not to engage in misuse. This breakpoint cost is given by equalizing the gain from misusing with the expected cost,

$$\theta = \bar{\kappa}_i + p(m_i)c \rightarrow \bar{\kappa}(m_i) = \theta - p(m_i)c. \tag{2}$$

We directly see that $\bar{\kappa}'(m_i) < 0$ holds, so that the risk of an individual misusing the benefit is decreasing in the level of monitoring aimed at him. This simply formalizes the very natural property that we may reduce the misuse among some group of individuals by letting them know that they are exposed to increased monitoring. The identifying mechanism of the information method is built around this simple logic.

3 Estimating the extent of excess use

A natural policy objective is to minimize the average amount of excess use in the economy among all individuals that are eligible for the benefit. Let all individuals face the same ex ante monitoring intensity denoted m. This variable (m) should be thought of as a joint measure of the probability of getting monitored as well as the degree or extent of monitoring when actually being exposed. It is not crucial to our analysis whether the individuals' perceived monitoring intensity actually is correct or not. However, in order to simplify the presentation we assume that the perceived baseline m is the true m. Conditional on m the average excess use,

[8] m_i may or may not be known to individual i. The behavioral model simply requires that the individual takes one m_i as given and acts on it; whether this anticipation is correct or not will be clear from the context.

denoted $\mu_\theta(m)$, can be expressed as:

$$\mu_\theta(m) \equiv \theta \frac{\sum_N I_i(m)}{N} = \theta \frac{N_\theta(m)}{N} = \theta E\left(I_i(m)\right) \tag{3}$$

where $N_\theta(m)$ is the number of individuals that misuse the benefit. The key property of $\mu_\theta(m)$ is that $\mu'_\theta(m) \leq 0$ holds since $\bar{\kappa}'(m) < 0$ gives $N'_\theta(m) \leq 0$.

We will now compare two different ways of estimating $\mu_\theta(m)$ (and then an additional third way). The first is based on traditional random monitoring that is not pre-announced to the subjects. The other is based on the same random monitoring, with the only difference that the monitoring is announced in advance to the individuals that will actually be monitored. We call the first approach *uninformed monitoring* and the second approach *pre-informed monitoring*; the difference lies in whether the individuals that are monitored know in advance about the monitoring or not. In both cases we assume that a monitored individual will be exposed to monitoring of degree m^*. When not informed you thus perceive monitoring intensity m, and when informed you perceive monitoring intensity m^*; obviously $m^* \geq m$ holds – with equality only when the general level of monitoring involves monitoring every withdrawal, which is a special case we may abstract from.[9] At this point we also abstract from the case when type two errors are possible, i.e. when there is a risk of getting caught when using the benefit legitimately; we will look closer at this important extension below. Furthermore, we assume that an individual who gets informed that he is subject to monitoring correctly estimates the true value $p(m^*)$, and acts accordingly. In reality there is of course the possibility of both over and under estimation of $p(m^*)$.[10] We also make the simplifying assumption that the non-informed individuals – i.e. both those that will end up being monitored in the ordinary uninformed monitoring scheme and those who will not be monitored at all – still estimate

[9] In this theoretical part of the paper we assume, for simplicity, that the treated individuals get informed that their potential withdrawals of the benefit indeed will be monitored (with 100 percent certainty). This assumption is made only for simplicity. The qualitative features of the estimates does not change if the treatment consists of declaring that You will be monitored with X percent certainty. Or, for that matter, merely declaring that you will be subject to intensified monitoring, without being explicit about the actual risk of monitoring. This latter approach will be the one applied in the actual experiment (see the application section below).

[10] The key is only that the informed individual perceives an increased monitoring intensity; he need not necessarily perceive the correct one.

the risk of getting caught at the lower default level $p(m)$. In other words, we assume that we may measure the variable $\mu_\theta(m)$ without affecting its true value in the non-informed population.[11]

3.1 Estimate based on uninformed monitoring

Let M_i be an indicator variable taking on value 1 if the individual is caught and zero otherwise. The expected value of M_i for a random monitored individual, who actually misuses (i.e. $I_i = 1$), will be,

$$E(M_i|I_i = 1) = p(m^*)$$

and zero otherwise. The uninformed estimate is defined as

$$\hat{\mu}_\theta(m) \equiv \theta \frac{\sum_n M_i I_i(m)}{n}, \tag{4}$$

where n is the number of people that were randomly selected for monitoring.

When taking the expected value of $\hat{\mu}_\theta(m)$ we get

$$E(\hat{\mu}_\theta(m)) = p(m^*)\theta \frac{\sum_n E\left(I_i(m)\right)}{n} = p(m^*)\mu_\theta(m) \leq \mu_\theta(m). \tag{5}$$

As long as monitoring is not perfect, $p(m^*) < 1$, it should come as no surprise that the traditional uninformed estimate has a downward bias. Its negative bias, $\hat{B}$, will be

$$\hat{B} = \left(1 - p(m^*)\right) \mu_\theta(m) \geq 0. \tag{6}$$

Without having some prior knowledge of the size of $p(m^*)$ there is not much more we can do based on uninformed random monitoring. We will always end up with a weakly negative bias; the lower the risk of getting caught when monitored, the higher the negative bias will be. Sometimes it is theoretically possible to have $p(m^*)$ very close to unity, but in practice it is often extremely costly and/or legally/ethically infeasible. As we will see below, pre-informed monitoring may then be an attractive alternative.

[11] For a discussion regarding this type of "contamination problem" see e.g. Moffitt (2004). In our case an involuntary partial treatment of the control group – say due to media attention in response to the experiment – would most likely generate a negative bias since it would tend to lower the general excess use of the benefit.

3.2 Estimate based on pre-informed monitoring

Now consider a possible estimate of $\mu_\theta(m)$ based on informing the randomly selected individuals in advance, i.e. the pre-informed monitoring. We may then base the estimate on not only the average excess use but also on the average reduction in total use of the benefit. The estimator is thus based on the latent assumption that all the information induced reduction in use of the benefit can be attributed to excess use.[12] Denote the estimate based on the information method by $\tilde{\mu}_\theta(m)$. We then define,

$$\begin{aligned}\tilde{\mu}_\theta(m) \quad \equiv \quad & \underbrace{\left[\frac{\sum_{N-n} w_i(m)}{N-n} - \frac{\sum_n w_i(m^*)}{n}\right]}_{\text{Reduction in use of benefit caused by information}} \\ & + \quad \underbrace{\theta \frac{\sum_n M_i I_i(m^*)}{n}}_{\text{Remaining excess use found}} \quad ,\end{aligned} \tag{7}$$

where n now represents the number of people informed that they will be monitored. The expected value of this estimate is

$$E(\tilde{\mu}_\theta(m)) = \mu_\theta(m) - (1 - p(m^*))\mu_\theta(m^*). \tag{8}$$

and its negative bias is thus

$$\tilde{B} = (1 - p(m^*))\mu_\theta(m^*) \tag{9}$$

Now, since $\mu_\theta(m)$ has negative slope, and $m^* > m$ holds by assumption, it holds that $\tilde{B} \leq \hat{B}$. This shows that we can get closer (in terms of lower bias) to the true variable $\mu_\theta(m)$ through the information method, than through traditional uninformed random monitoring.

We may also derive the difference in negative bias between the two estimators. From (6) and (9) we then get,

$$\hat{B} - \tilde{B} = (1 - p(m^*))\left(\mu_\theta(m) - \mu_\theta(m^*)\right). \tag{10}$$

From (10) some intuitive properties of the information based estimate emerge. Given a reduction in excess use rendered by information, i.e.

[12] There may be different reasons why this assumption does not hold in reality. Technically we may describe violations of this key assumption as $b(\Omega) \equiv b(\Omega', m^*)$, so that also the legitimate use of the benefit is affected by the knowledge of being selected for monitoring. We will return to this extension, and the trouble it causes, below.

$\mu_\theta(m) - \mu_\theta(m^*) > 0$, the information based method performs better, in relative terms, the harder it is to catch misusers (i.e. the lower $p(m^*)$ is). And given $p(m^*)$, the information method performs better the more elastic the excess use is to notice of increased monitoring; in the extreme case when all excess use is eliminated through information, i.e. $\mu_\theta(m^*) = 0$, it is clear from equation (9) that the information based estimate is unbiased. The intuitive key to the result lies in realizing that in the pre-informed estimate we identify all of the reduction in excess use rendered by knowledge of increased monitoring. Let μ_θ^A denote the average excess use that remains even when the individuals are informed of increased monitoring, and let μ_θ^B denote the average excess use that is eliminated through information. By definition we thus have $\mu_\theta(m) \equiv \mu_\theta^A + \mu_\theta^B$. The information approach identifies the whole of μ_θ^B while only part of μ_θ^A, i.e. $p(m^*)\mu_\theta^A$. While, on the other hand, traditional random monitoring only identifies parts of both μ_θ^B and μ_θ^A, that is $p(m^*)\left(\mu_\theta^A + \mu_\theta^B\right)$.

3.3 Consistent estimation of excess use

It turns out that the uninformed monitoring combined with informed monitoring allows us to go one step further. We may estimate the probability to get caught when monitored, $p(m^*)$, consistently. Once we have a consistent estimate of $p(m^*)$, we only need to scale up the average excess use found with uninformed monitoring by the inverse of the estimate of $p(m^*)$, in order to get a consistent estimate of $\mu_\theta(m^*)$.

3.3.1 Step one – consistent estimate of $\hat{p}(m^*)$

Define $\hat{p}(m^*)$ as,

$$\hat{p}(m^*) \equiv \frac{\theta \frac{\sum_n M_i I_i(m)}{n} - \theta \frac{\sum_n M_i I_i(m^*)}{n}}{\frac{\sum_{N-n} w_i(m)}{N-n} - \frac{\sum_n w_i(m^*)}{n}}. \tag{11}$$

Our estimate of the probability to get caught, conditional on misusing and being monitored, is given by the information induced change in detected excess use divided by the induced behavioral response. It is straightforward to confirm that the probability limit of this estimate is

$$\text{plim } \hat{p}(m^*) = p(m^*),$$

which proves consistency.

3.3.2 Step two – consistent estimate of $\mu_\theta(m)$

We simply proceed by defining our third estimate of $\mu_\theta(m^*)$, $\breve{\mu}_\theta(m)$, as

$$\breve{\mu}_\theta(m) \equiv \frac{\hat{\mu}_\theta(m)}{\hat{p}(m^*)}. \tag{12}$$

It is straightforward to confirm that this estimator is consistent, since

$$\text{plim } \breve{\mu}_\theta(m) = \frac{\text{plim } \hat{\mu}_\theta(m)}{\text{plim } \hat{p}(m^*)} = \frac{p(m^*)\mu_\theta(m)}{p(m^*)} = \mu_\theta(m).$$

It thus seems that we actually can do better and, at least in this basic setup, estimate the average misuse consistently. There are two problems with the estimator, however. One problem is that the estimate is a non-linear function of random variables. Therefore it is not straightforward to find its variance or probability distribution. One will have to resort to either bootstrapping or approximate methods as e.g. the Delta method. Another problem with the estimator is that it relies on two separate parallel studies, the informed estimate and the non-informed estimate, which may make the study rather costly to perform.

4 Extensions and robustness

In this section we will explore a number of different ways to extend the basic model. Our focus will be on how these extensions affect the bias and asymptotic bias of our three estimates of average excess use, $\hat{\mu}_\theta(m)$, $\tilde{\mu}_\theta(m)$ and $\breve{\mu}_\theta(m)$.

4.1 When the legitimate use is affected by the notice of intensified monitoring

In the basic setup above we made the simplifying assumption that the legitimate use of the benefit, $b_i(\Omega)$, is independent of the experiment in general and of the actual risk of getting monitored in particular. As we will discuss below, there are a number of reasons why this may not be the case in reality. We start with a short general treatment of the effects on our estimators when the legitimate use of the benefit is endogenous to the experiment.

4.1.1 The general effects of endogenous b_i

Now let each individual have two different potential levels of legitimate demand of the benefit: $\bar{b}_i$ and $b_i^{m^*}$, where $\bar{b}_i$ denotes the default level of legitimate use and $b_i^{m^*}$ is the level of legitimate use conditional on the individual being informed about monitoring. The traditional random sampling estimate is naturally indifferent to this extension. However, the information based estimate is not. The estimate is still given by eq. (7), but the expected value of this estimate will now be,

$$E(\tilde{\mu}_\theta(m)) = \mu_\theta(m) - (1 - p(m^*))\mu_\theta(m^*) + E(\bar{b}_i) - E(b_i^{m^*}), \qquad (13)$$

which shows that when knowledge of monitoring reduces (increases) the legitimate demand for the benefit, there will be an upward (downward) pressure on the estimate.

Regarding the third estimate, $\breve{\mu}_\theta(m)$, we get that the probability limit of eq. (11) now is,

$$\text{plim } \hat{p}(m^*) \equiv \frac{1}{1 + \frac{E(\bar{b}_i) - E(b_i^{m^*})}{\mu_\theta(m) - \mu_\theta(m^*)}} p(m^*),$$

so that $E(\bar{b}_i) > E(b_i^{m^*})$ will asymptotically underestimate the true risk of getting caught, and vice versa. In accordance with this we get, directly from eq. (12), that $E(\bar{b}_i) > E(b_i^{m^*})$ gives a positive asymptotic bias to the previously consistent estimate, and vice versa.

All in all, these properties are very intuitive. If we cause a(n) reduction (increase) in the legitimate use of the benefit by informing individuals that they are subject to monitoring, the information method will potentially over-estimate (additionally under-estimate) the true excess use. The analogous holds for the third, previously consistent, estimator, $\breve{\mu}_\theta(m)$.

4.1.2 Being monitored is costly in itself

Some benefits are such that the actual monitoring process involves the individual in a way that may be considered cumbersome and associated with various personal costs. Such activities may involve: visiting a doctor, filling out forms, attending interviews, social costs of being "inspected" etc. Since the individual may avoid monitoring by temporarily restraining from using the benefit, such a scenario could lead to information about monitoring reducing the legitimate use of the benefit. As we have shown above such behavior tends to produce an upward bias in the estimates based on the information method. It is therefore very important to, as much as possible, reduce the effort needed from the individual in the monitoring process. In cases when substantial individual effort is impossible to avoid, one should interpret the information based estimates with a great deal of caution. Unless it is safe to assume that the legitimate use of the benefit is very inelastic to the provision of individual effort, one should in these cases avoid using the information method altogether.

4.1.3 Type two error – getting caught without misusing

It may sometimes be a non-negligible risk that the monitoring authorities make mistakes and find some individuals erroneously guilty of excess use. Apart from the general problems this causes – the traditional random sample estimate is potentially afflicted with an obvious positive bias – there is an additional problem associated with using the information method under such circumstances. The risk of being falsely accused of misusing a benefit may indeed restrain individuals from using their full legitimate share; and when getting notice of being subject to monitoring the problem obviously increases. The information based estimates are thus, also in this case, afflicted with a potentially positive bias. And if the net bias is positive, the use of the information method will produce a larger bias than traditional random monitoring will.

The problem could also occur if the individual is not totally sure of the exact border between legitimate use and misuse. In such case, when informed about monitoring, it is possible that people err on the safe side and therefore not attain their full legitimate use of the benefit.

The first, rather self-evident, lesson to be learned from this is that good quality of the monitoring process is of great importance. Furthermore, the results obtained from information-based estimates, in cases when the benefit's set of rules are hazy, should be interpreted with caution. It is also suggestible to let the information about monitoring be accompanied by a pedagogically designed information pamphlet describing the rules of the benefit and that the reason for them being selected has nothing to do with prior mistrust.

4.1.4 The information makes individuals discover the benefit

It is impossible to inform individuals about them being subjects to monitoring without implicitly also informing about the benefit's existence. Everyone does not know about the existence of all benefits that they are eligible for. And some people may know about the existence of some benefits that they may or may not be eligible for, but they do not normally bother to check in detail. In such case the information about them being the subject to monitoring may make them actively aware of the benefit and the exact rules that apply. Such a scenario would tend to increase the use of the benefit for the group informed about monitoring, which would induce a negative bias in the information based estimates.

One remedy would be to only include those who have used the benefit in recent history in the study. But this changes the possible inference in a potentially undesired way. Another potential remedy is to include a third group of individuals who only get informed about the existence and rules of the benefit. If such a group increases their use of the benefit to a large extent, one has an indication that there may be a problem. Taking this strategy one step further, one could add the increase in average use, caused by the pure information of the benefit, to the information based estimate. But such a strategy makes the rather strong assumption that the pure information pamphlet effect is the same irrespective of whether the pamphlet is accompanied by a letter informing about monitoring, or not.

4.2 Heterogeneous ways of misusing

So far we have made the simplifying assumption that there was only one single way to misuse the benefit. In reality this may not be the case. There are often different ways to misuse and they are often heterogeneous in both monetary payoff and in risk of getting caught.

Let there now be two different ways of misusing the benefit. The difference between the two is that one is hard to detect (superscript h) and one is easy to detect (superscript e). The price you pay for the excess use that is hard to detect is that the monetary equivalent of misusing is lower. The individual has the opportunity to use either one of these two strategies of misusing, but not both at the same time. Given these extensions to the basic model above it is straightforward to show that a notice of increased monitoring may increase the average surfaced misuse that is hard to detect ($\mu_\theta^h(m)$) while the average easily detected misuse ($\mu_\theta^e(m)$) naturally still decreases.

Assume further that the variable of interest is as before $\mu_\theta(m)$, i.e. the average excess use in the economy. This measure may now be divided into the two parts: $\mu_\theta^h(m)$ and $\mu_\theta^e(m)$. The average excess use is thus given by,

$$\mu_\theta(m) = \mu_\theta^h(m) + \mu_\theta^e(m). \tag{14}$$

We will now compare the properties of our three different estimators, $\hat{\mu}_\theta(m)$, $\tilde{\mu}_\theta(m)$ and $\breve{\mu}_\theta(m)$.

4.2.1 Uninformed estimation

Let the two probabilities of getting caught, when monitored and misusing, be $p^e(m^*)$ and $p^h(m^*)$ respectively. The expected value of the traditional random sampling estimate is now given by,

$$E\left(\hat{\mu}_\theta(m)\right) = p^e(m^*)\mu_\theta^e(m) + p^h(m^*)\mu_\theta^h(m).$$

And the negative bias of this estimate is,

$$\hat{B} = \left(1 - p^e(m^*)\right)\mu_\theta^e(m) + \left(1 - p^h(m^*)\right)\mu_\theta^h(m). \tag{15}$$

4.2.2 Information based estimate

The corresponding expected value for the pre-informed monitoring is given by,

$$E(\tilde{\mu}_\theta(m)) = \mu_\theta^e(m) + \mu_\theta^h(m) - \left(1 - p^e(m^*)\right)\mu_\theta^e(m^*) - \left(1 - p^h(m^*)\right)\mu_\theta^h(m^*). \tag{16}$$

Its negative bias is thus,

$$\tilde{B} = \left(1 - p^e(m^*)\right)\mu_\theta^e(m^*) + \left(1 - p^h(m^*)\right)\mu_\theta^h(m^*) \tag{17}$$

The first thing to note is that we still have a lower bound estimate of the average excess use; the information method will not overestimate the excess use. The second important question concerns the relative bias: does the information method still outperform traditional random sampling in terms of bias? The relative bias is given by comparing (15) with (17); we get,

$$\hat{B}-\tilde{B} = (1-p^e(m^*))\left[\mu_\theta^e(m)-\mu_\theta^e(m^*)\right]+\Big(1-p^h(m^*)\Big)\Big[\mu_\theta^h(m)-\mu_\theta^h(m^*)\Big] \tag{18}$$

The first term on the right hand side of (18) is non negative since increased monitoring (weakly) decreases the easily detected excess use. However, the second term may be negative since increased monitoring may increase the excess use that is hard to detect. It is thus theoretically possible that a relatively large increase in low risk misuse, caused by knowledge of monitoring, may make the information approach inferior to traditional random sampling, in terms of bias. The extreme case makes the logic clear: let $p^e(m^*) = 1$ so that the easily detectable misuse is in fact detected in all monitored cases. The relative bias, (18), then simplifies to $\hat{B} - \tilde{B} = \left(1 - p^h(m^*)\right)\left[\mu_\theta^h(m) - \mu_\theta^h(m^*)\right]$. In this extreme case we will do worst with the information based approach as soon as the misuse that is hard to detect increases with knowledge of monitoring. Note that in this case we have no gain from using the information method when identifying $\mu_\theta^e(m)$ since monitoring works perfectly. In terms of estimating $\mu_\theta^h(m)$, we only gain in case knowledge of monitoring actually decreases this type of excess use.

4.2.3 Is consistent estimation still possible?

We start by proceeding as before and simply estimate the probability of getting caught (irrespective of type of misuse) conditional on misuse and actually being monitored. It is straightforward to show that, when defining $\hat{p}(m^*)$ in analogy with (11), the probability will now be,

$$\text{plim } \hat{p}(m^*) = \alpha p^e(m^*) + (1-\alpha)p^h(m^*),$$

where,

$$\alpha \equiv \frac{\mu_\theta^e(m)-\mu_\theta^e(m^*)}{\mu_\theta^e(m)+\mu_\theta^h(m)-\mu_\theta^e(m^*)-\mu_\theta^h(m^*)}.$$

The probability limit of $\breve{\mu}_\theta(m) \equiv \hat{\mu}_\theta(m)/\hat{p}(m^*)$ then simplifies to,

$$\text{plim } \breve{\mu}_\theta(m) = \frac{\beta p^e(m^*) + (1-\beta)\, p^h(m^*)}{\alpha p^e(m^*) + (1-\alpha)p^h(m^*)}\Big(\mu_\theta^e(m)+\mu_\theta^h(m)\Big), \tag{19}$$

where,

$$\beta \equiv \frac{\mu_\theta^e(m)}{\mu_\theta^e(m) + \mu_\theta^h(m)}.$$

A sufficient condition for the estimator to still be consistent is thus $\alpha = \beta$. Furthermore, since $p^e(m^*) > p^h(m^*)$, a sufficient condition for the estimator to have an asymptotically non positive bias is $\beta \leq \alpha$. Given the definitions of α and β this condition simplifies to,

$$\frac{\mu_\theta^h(m) - \mu_\theta^h(m^*)}{\mu_\theta^h(m)} \leq \frac{\mu_\theta^e(m) - \mu_\theta^e(m^*)}{\mu_\theta^e(m)}. \tag{20}$$

A sufficient condition for the combined estimate, $\breve{\mu}_\theta(m)$, to have a non positive asymptotic bias is thus that the easily detectable aggregate excess use should be more elastic with respect to information about increased monitoring than the excess use that is hard to detect. In other words, as long as information induces a relatively large decrease in the easily detectable excess use we do not asymptotically overestimate excess use. However, the estimate suffers from the same qualitative problems as the information based estimate, i.e. the (negative) bias may be large if the excess use that is hard to detect increases with information.

4.2.4 Discussion

The problems described in this section extends to at least two relatively closely related cases. The first case is when the individual can make some endogenous investment in "stealth technology", i.e. making it harder to detect misuse. If so, knowledge of monitoring may increase investments in stealth technology, which has the same kind of adverse effects on the information estimate as the case explored in this section.

The other case concerns the realistic possibility of engaging in excessive use to a varying degree. It is straightforward to extend the model by endogenizing θ, so that the choice of misusing is not just dichotomous. When informed about monitoring the workers may switch to a lower level of excess use which naturally lowers the risk of detection. If the probability to get caught is very elastic to changes in the level of excess use, we could also get an analogous effect as the one described here.

In potential applications of the information method one must ideally assess the risk of these problems occurring. The problems described here are of different gravity in different circumstances. It must be judged on a case by case basis whether there is a real risk that one may do worse with the information method. The problem is generally less severe in cases when quitting the misuse altogether, under the duration of the experiment, is likely to be associated with a relatively low individual cost. When the monitoring process has a high degree of intrinsic randomness, i.e. when pure chance decides to a large degree who gets caught when monitored while misusing, the information method performs the best. Note also that some degree of asymmetric information regarding the exact means and margins of monitoring would by this logic work in favor of using the information method. However, in the opposite scenario, when the individuals are aware of some very low risk way to misuse the benefit, that they may also rather costlessly switch to, the information method is likely to perform rather badly.

5 Application – the VAB-Benefit

In this section we will present an application of the information method on Swedish data. During the spring of 2006 we designed and directed an extensive experiment on the VAB-Benefit. The experiment aimed at determining the level of excess use of the benefit, and it was administered by the Swedish Social Insurance Agency who also administers the benefit itself.

5.1 About the VAB-Benefit and how it can be misused

Basically, the VAB-Benefit gives a parent right to economic compensation when he has to take time from work in order to take care of a sick or infectious child. VAB-benefits will not be paid when a parent receives sick-pay, sickness-insurance benefit, is on vacation, is doing military service or is in custody in the correctional system. As a first order approximation, the VAB-benefit can be drawn for 60 days per year for each child that has not reached the age of 12.[13] The cost of the VAB-Benefit is about 3 billion SEK yearly (7 SEK $\approx$ \$ 1). The compensation level for the VAB-Benefit is at present 80 percent of the wage income. However, the part of the income that is over a cap (approximately 300 000 SEK per year during the experiment period) is not compensated. The compensation level is

[13] See Engström et al (2007) for an extended presentation of the rules attached to the VAB-Benefit.

approximately the same as the sick-pay compensation level. However, the sick pay system involves an uncompensated qualification day, which the VAB-Benefit does not – when using the VAB-Benefit you thus get reimbursed from day one but when on sick-leave you do not.

A number of potential ways to misuse the VAB-Benefit emerges. Perhaps the most obvious one is to use the VAB-Benefit instead of using sick-pay, in order to escape the uncompensated qualification day. Another way to misuse the benefit is to simply claim the VAB-Benefit without staying home from work, and thus receive double payments. Other potential ways of misusing include black market work or receiving some other benefit at the same time (e.g. unemployment benefit). The monitoring of withdrawals includes both contacts with employer and schools/pre-schools as well as cross-checking in registers for multiple benefits.

5.2 The experimental setup and execution

The experiment can be described as follows. First, a randomized selection of the eligible individuals was selected to be included in a treatment group. These individuals were informed through letters notifying them that they were included in a randomized group whose use of the VAB-Benefit were to be exposed to intensified monitoring for a period of time.[14] The individuals in the treatment group were not informed whether they actually were to be monitored or how substantial the risk for this was. It turned out that about 30 percent of all individuals in the treatment group, who used the benefit during the experiment period, got their withdrawals monitored.

The treatment group consisted of three different sub-groups of sizes, 28 543 (A), 7 051 (B) and 6 655 (C): group A, received both the information letter about increased monitoring and a short pamphlet explaining the set of rules attached to the benefit; group B received only the pamphlet (with an explanatory text declaring that the pamphlet was part of a campaign); and group C received only the information letter about increased monitoring. This division of the treatment groups partially aimed at assessing the margin discussed in section 4.1.4, i.e. to what extent does this type of information make the individuals "discover" the benefit. We also wanted to keep an eye on whether the reaction differed to a large extent when there was no accompanying information pamphlet; an over-reaction to the letter is arguably more plausible when there is no accompanying pamphlet

[14] See Appendix for the exact reproduction of the information letter.

(see sections 4.1.2 and 4.1.3). In terms of the estimates of excess use, the relevant treatment groups are A and C, since they get informed about being subject to increased monitoring; we will often merge these two groups, in which case the joint group will be referred to as group AC.

Letters and information pamphlets, to the selected caretakers in the different treatment groups, were sent out on the 24th of March 2006. The majority of those reached its receivers four days later on the 28th. The experiment period started on the 29th of March and carried on until the 31st of May.[15] The Swedish Social Insurance Agency also established a temporary call-center and a special e-mailbox where those who received the letters (group A and C) could turn with questions; this was one of the strategies for reducing the risk of an over-reaction to the notice of increased monitoring (see sections 4.1.2 and 4.1.3).

From those who were selected to the three treatment groups, a certain share (about 30 percent) had their withdrawals from the VAB-Benefit examined. The chosen individuals' withdrawals were closely checked through contacts with employers and pre-schools/schools, and also crosschecked against the unemployment benefits etc. The absence was checked with the employer during the compensated time along with the grounds for absence (care for sick child, vacation, compensatory leave, work allocation, or else); while the child's absence status, and reason for potential absence, was checked with school/pre-school.

5.3 Data

The data were collected from the administrative registry of the Swedish Social Insurance Agency and cover all withdrawn VAB-benefits in Sweden during the periods of interest. Our population is defined by those who are eligible for the benefit – in principle every caretaker, with a few exceptions, with children in the age 1-11 years old.[16] This means that the population, approximately, consists of all Swedish caretakers of children born between 1st of June 1994 and 20th of March 2005. The selection into

[15] The exact timing of the experiment had to be determined by practical considerations. Much effort was put into finalizing the experiment before the summer (when the VAB-withdrawals are the lowest).

[16] The individuals who were unregistered, taken into custody, doing military service, caretakers with protected identity and persons missing from the tsunami-catastrophe were not included in the treatment group or remaining population.

treatment groups were then determined by randomly selected birthdates of the caretakers.

Table 1. Descriptive statistics for the treatment groups (A, B and C) and the remaining population (D).

Group	Number of individuals	Share of women	Average age
A	28 543	0.516	38.26
		[0.003]	[0.04]
B	7 051	0.512	38.24
		[0.006]	[0.08]
C	6 655	0.514	38.26
		[0.006]	[0.08]
D	1 272 993	0.515	38.29
		[0.0004]	[0.006]

Note: Standard errors in parentheses.

Table 2. Percentage share withdrawing VAB-Benefit an average day during the reference period (2005-10-01 until 2006-02-28).

Group	Woman	Men	All
A	1.37	0.88	1.14
	[0.022]	[0.020]	[0.015]
B	1.42	0.82	1.13
	[0.046]	[0.035]	[0.029]
C	1.50	0.95	1.23
	[0.050]	[0.043]	[0.033]
D	1.40	0.86	1.14
	[0.003]	[0.003]	[0.002]

Note: Standard errors in parentheses.

Table 1 shows the number of individuals, age and gender distribution in the different groups. In total 42 249 individuals were included in some treatment group; 28 543 were in group A, 7 051 in group B and 6 655 in group C. The remaining population (group D) contains 1 272 993 individuals. It should be noted that there are no statistical significant gender and age differences when comparing the different treatment groups (A, B and C) to the remaining population (D).[17] Neither are there any significant differences between these variables when comparing the treatment groups in pairs.

We have used a reference period against which the different groups' withdrawals may be compared (see the exact reason for this in the next subsection where we describe the statistical method). We decided to use the six months that preceded the experiment period, minus a "buffer month" in between the reference period and the actual experiment period.[18] Our reference period was thus made out of the 5 months between 05-10-01 and 06-02-28.

If the random selection process has been correct, the expected withdrawal patterns within the different groups are identical. Table 2, that presents the average fraction of men and woman that made withdrawals, as well as the average number of VAB-Benefit days during the reference period, shows that there were no significant differences between the groups A, B and the remaining population (D). But for the small group C, the withdrawal is significantly larger than it was in group D. One explanation for this could be pure chance: we may have got an unusual high average need of the benefit among the caretakers in group C.[19] If we instead look at group A and C jointly (the AC-group), the significant differences between all remaining groups (AC, B and D) disappear. Since the C-group's constitution, concerning a number of variables – such as age, gender, sector, income etc. – does not differ significantly from the other groups, we con-

[17] This also holds for a number of other characteristics, such as: sectoral belonging, income, education etc.

[18] The reason for using a "buffer month" is that some people linger with their applications for VAB-Benefit for a few weeks. When an individual, in the end of March, receives the letter about intensified monitoring, there is thus a possibility that he reduces earlier withdrawals retroactively. This could produce an underestimation of the letter's effect.

[19] It should also be noted that the difference does not disappear when we cut the top of the withdrawal distribution, which means that it is not due to a small number of extreme outliers, producing the higher number of withdrawals in group C.

clude that the random selection has been correctly carried out and that the C-group, simply by chance, has been assigned a larger fraction of individuals with a somewhat increased demand for VAB-Benefit.

5.4 Statistical method

Our estimate of the total excess use of the VAB-Benefit will have two components: it is partly made out of the potential decrease in withdrawals induced by the letters, and partly of the results of monitoring in itself. The statistical analyses behind the two measures are separate. We will therefore describe the underlying analysis of each component separately. Finally, we will describe how we aggregate the two separate measures into a total measure of excess use.

5.4.1 The letter effect

When estimating the effect of the information letter we are able to choose from two unbiased estimators: the traditional "Difference-Estimator" (D-Estimator) or the "Difference-in-Differences-Estimator" (DiD-Estimator). Since both estimators are unbiased, the choice of which one to use could be made based on the estimators' variances. Theoretically, the DiD-Estimator outperforms the D-Estimator when there is substantial, time-invariant, individual variation in the variable. The intuition is that the DiD-Estimator eliminates the individual specific component already when taking the (first) difference at the individual level over time, while with the D-estimator the individual specific component is simply averaged out. However, if there is a large time variation there may be a high cost, in terms of increased variance, of taking the first difference in order to normalize against earlier withdrawal pattern. In our case the sample variances of the two different estimators are approximately the same. However, we have reason to believe (see the data-section above) that we have, by chance, gotten a rather skewed sample in the small C-group; this group shows an exceptionally high withdrawal pattern during the reference period. If this is due to individual-specific components, rather than time-specific components, there are problems using the simpler D-Estimator. Given a skewed sample, with regards to the individual-specific component, the D-estimator is biased, since the individual-specific component is not eliminated at the individual level. However, the DiD-Estimator may still be

We confine our presentation to cover only groups based on gender, sector, educational level, region, age and income. The results are for the whole AC group and refers to withdrawn amounts (the results based on days of VAB-use are approximately the same). The estimates for each category are presented in Table 4 and can be summarized with two significant findings. Firstly, we find significant differences between the sexes. Excess use of the benefit by men is estimated to 28 percent of the amount withdrawn by men, while the same measure for women is about 19 percent. This should be interpreted in the sense that a given amount of benefit handed out to a man is likelier to stem from excess use, than if it was paid out to a woman. Secondly, we find that excess use decreases with education attainment. In the group with only nine-year compulsory school the excess use is estimated to 41.3 percent, while the corresponding number for the group with at most high-school education is 26.7 percent. However, the group with the lowest education level is relatively small, both in size as well as in share of withdrawals. When turning to the excess use of the group with the highest education, post high-school, we find a substantial decrease in excess use. The excess use is estimated to 11.2 percent in the group with the highest education. All three estimates are significantly different from each other.

5.7 Robustness – do people over-react to the letter?

As described in section 4.1.2 and 4.1.3, the potential risk for positively biased estimates of excess use is the risk of over-reactions to the information about increased monitoring. A number of measures have been taken in the experimental design to reduce the risk of such reactions (see Engström et al, 2007, for a more thorough discussion of these issues). However, it is naturally of great value to measure the extent or find indications of such reactions. Fortunately, this particular benefit provides us with a fruitful strategy to go about this.

The schools and pre-schools do not usually admit sick children, due to the risk of contamination as well as the increased workload. Since the sickness status of the child is an often binding constraint, a parent who overreacts to the letter must find a substitute caretaker. An obvious candidate is the partner, since the letter is very clear on that it is only "Your" individual use of the VAB-Benefit that will be subject to increased monitoring. Even if grandparents who are no longer working in some cases could step in, the partner must in many cases be the most obvious choice. Since we have data on the partners, we can examine whether an overreaction is at hand by studying the implicit reaction of the partner to the ones exposed to the treatment (in particular group A and C). Since the effect on the partner is likely to be larger if the two caretakers actually live together we try to limit the sample to the traditional nuclear family. However, we have not got the information on whether the caretakers of the child live together or not. In the cases when a child's caretakers have changed, we do not know who the current caretakers are. For this reason we have restricted our population to the observations where the children only have had two caretakers during their lifetime and when those and the child was registered in the same parish and when none of the caretakers have children in the age of 1 to 11 years together with another partner. We also exclude all "families" where both caretakers have had any kind of treatment (A, B or C). Given this selection of population we have a remaining 83.6 percent of the original population, which is close to official numbers from Statistics Sweden for the year 2004 that shows that 78 percent of the children in the age 1 to 11 years old lives with their original parents.[24]

We find that the letters' cross-effects on the partners in the main treatment group, AC, consist of an insignificant reduction in withdrawal of VAB-Benefit (2.6 percent with standard deviation 1.9 percent). If the effect had been positive, it had been consistent with the over-reaction hypothesis described above, and we would have an indication of a potentially serious positive bias in our estimate of excess use. Since such reaction did not occur we may be less concerned with the risk of over-reaction. However, if there are two counteracting effects, one substitution and one disciplinary (if some interpret the letter as the whole family being monitored even though the letter was very clear on that it was the receiver, and not a particular child, that was the focus for monitoring) we can get a net result that does not show any effect in the aggregate. It is principally possible that what we measure is the sum of a positive effect (when you let the partner take care of the sick child) and a negative contamination effect (the disciplinary effect on the partner). However, two counteracting effects tend to increase the variance of the variable in question. A comparison of the distribution of the differences in paid amount between the AC

[24] It is natural that we get a somewhat higher number than Statistics Sweden, since we include those parents who are separated but still live in the same parish.

Heckman, J and J Smith, (1995), Assessing the Case for Social Experiments, J*ournal of Economic Perspectives* 9:2, 85-110.

Hesselius, P, P Johansson and L Larsson (2005), Monitoring sickness insurance claimants: evidence from a social experiment, IFAU Working Paper 2005:15.

Holmlund B (1998), Unemployment insurance in theory and practice, *Scandinavian Journal of Economics*, 100, 113-41.

Hägglund P (2006), Are there pre-programme effects of Swedish active labour market policies? IFAU Working Paper 2006:2.

Johansson, E (2000), An Expenditure-Based Estimation of Self-Employment Income Underreporting in Finland, Working Paper 433, Swedish School of Economics and Business Administration, Helsinki.

Joulfaian, D and M Rider (1998), Differential Taxation and Tax Evasion by Small Business, *National Tax Journal* 51, 675-687.

Lacko, M (1998), The Hidden Economies of Visegrad Countries in International Comparison: A Household Electricity Approach, Hungary: Towards a market economy, 1998, 128-52, Cambridge; New York and Melbourne: Cambridge University Press.

Lyssiotou, P, P Pashardes and T Stengos (2004), Estimates of the Black Economy Based on Consumer Demand Approaches, *Economic Journal,* 114, 622-640.

Moffitt Robert A. (2004), The role of randomized field trials in social science research, American Behavioral Scientist, 47, 506-540.

OECD (2004), Compliance Risk Management: Managing and Improving Tax Compliance, Guidance Note, Centre for Tax Policy and Administration.

Pissarides, C and G Weber (1989), An Expenditure-Based Estimate of Britain's Black Economy, *Journal of Public Economics 39*, 17-32.

Sandmo, A (2005), The Theory of Tax Evasion: A Retrospective View, *National Tax Journal* 58, 643-663.

Working Papers

2007:1 de Luna Xavier & Per Johansson "Matching estimators for the effect of a treatment on survival times"

2007:2 Lundin Daniela, Eva Mörk & Björn Öckert "Do reduced child care prices make parents work more?"

2007:3 Bergemann Annette & Gerard van den Berg "Active labor market policy effects for women in Europe – a survey"

2007:4 Andersson Christian "Teacher density and student achievement in Swedish compulsory schools"

2007:5 Andersson Christian & Nina Waldenström "Teacher supply and the market for teachers"

2007:6 Andersson Christian & Nina Waldenström "Teacher certification and student achievement in Swedish compulsory schools"

2007:7 van den Berg Gerard, Maarten Lindeboom & Marta López "Inequality in individual mortality and economic conditions earlier in life"

2007:8 Larsson Laura & Caroline Runeson "Moral hazard among the sick and unemployed: evidence from a Swedish social insurance reform"

2007:9 Stenberg Anders "Does adult education at upper secondary level influence annual wage earnings?"

2007:10 van den Berg Gerard "An economic analysis of exclusion restrictions for instrumental variable estimation"

2007:11 Forslund Anders & Kerstin Johansson "Random and stock-flow models of labour market matching – Swedish evidence"

2007:12 Nordin Martin "Immigrants' return to schooling in Sweden"

2007:13 Johansson Mats & Katarina Katz "Wage differences between women and men in Sweden – the impact of skill mismatch"

2007:14 Gartell Marie, Ann-Christin Jans & Helena Persson "The importance of education for the reallocation of labor: evidence from Swedish linked employer-employee data 1986–2002"

2007:15 Åslund Olof & Hans Grönqvist "Family size and child outcomes: Is there really no trade-off?"

2007:16 Hesselius Patrik & Malin Persson "Incentive and spill-over effects of supplementary sickness compensation"

2007:17 Engström Per & Patrik Hesselius "The information method – theory and application"