Policy Research Working Paper 4821

# Social Health Insurance vs. Tax-Financed Health Systems—Evidence from the OECD

*Adam Wagstaff*

The World Bank
Development Research Group
Human Development and Public Services Team
January 2009

Policy Research Working Paper 4821


## Abstract

This paper exploits the transitions between tax-financed health care and social health insurance in the OECD countries over the period 1960–2006 to assess the effects of adopting social health insurance over tax finance on per capita health spending, amenable mortality, and labor market outcomes. The paper uses regression-based generalizations of difference-in-differences and instrumental variables to address the possible endogeneity of a country's health system. It finds that adopting social health insurance in preference to tax financing increases per capita health spending by 3–4 percent, reduces the formal sector share of employment by 8–10 percent, and reduces total employment by as much as 6 percent. For the most part, social health insurance adoption has no significant impact on amenable mortality, but for one cause—breast cancer among women—social health insurance systems perform significantly worse, with 5-6 percent more potential years of life lost.


This paper—a product of the Human Development and Public Services Team, Development Research Group—is part of a larger effort in the department to analyze health financing and delivery issues. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The author may be contacted at awagstaff@worldbank.org.

# Social Health Insurance vs. Tax-Financed Health Systems—Evidence from the OECD

by

Adam Wagstaff[a]

[a] Development Research Group, The World Bank, Washington DC, USA

**Contact author:** Adam Wagstaff, World Bank, 1818 H Street NW, Washington, D.C. 20433, USA. Tel. (202) 473-0566. Fax (202)-522 1153. Email: awagstaff@worldbank.org.

**Keywords:** Social health insurance; labor markets; health finance; health sector reform.

**Acknowledgements:** I am heavily indebted to Rodrigo Moreno-Serra: this paper employs methods and utilizes computer code that he and I jointly developed for two papers on the impacts of social health insurance in central and eastern Europe and central Asia. He also provided me with some helpful comments on the paper; due to timing, not all of them have been incorporated in this version. The findings, interpretations and conclusions expressed in this paper are entirely mine, and do not necessarily represent the views of the World Bank, its Executive Directors, or the countries they represent.

## 1. Introduction

The developed and developing world is currently the midst of a lively debate about the relative merits of social health insurance (SHI) and tax-financed health systems.[1] Skeptics of SHI argue that SHI discourages firms from hiring workers, and hence reduces employment and encourages informal labor markets. They also point to lack of coverage among certain groups during the often long period between establishing SHI and achieving universal coverage, and systematic variations in benefit packages and quality of care across subpopulations. By contrast, SHI advocates argue that it can provide an important additional source of revenue for the health system, and that by separating the purchasing of health care from its provision and encouraging selective contracting between providers (including private sector ones), SHI systems are able to achieve better quality health care at a lower cost than tax-financed health systems.

Like many intriguing and important debates, this one is being conducted on a flimsy evidence base.[2] The most basic question—still only partially answered—is how, on average, SHI and tax-financed systems perform in terms of key health sector and labor market outcomes. Of course, no two SHI systems are exactly alike, and no two tax-financed systems are alike; indeed, the health systems in one camp are increasingly borrowing ideas from the other. Nonetheless, each system still has certain hallmarks, and trying to establish the system-wide (i.e. aggregate)

[1] In Latin America, arguably the two biggest concerns have been closing gaps in coverage and the impact of SHI on the labor market (cf. e.g. Baeza and Packard 2006). These concerns have been evident too in the German reform debate (cf. e.g. Busse and Riesberg 2004; Schmidt 2006). In 1999, France eliminated the linkage between coverage and contributions, and made a commitment to universal coverage for all residents; at the same time, it widened its SHI tax base from earnings to include nonwage income. In Africa and Asia, several developing countries are in the process of trying to extend and deepen SHI coverage, or are thinking of introducing SHI. Arguably the biggest concern in these countries has been and continues to be extending coverage beyond the formal sector (cf. e.g. Hsiao and Shaw 2007).

[2] See Wagstaff (2007) for a review of issues and the evidence to date.

impacts of choosing one system over another is a useful exercise. Such impacts are unlikely, however, to be accurately estimated simply by comparing countries that finance their health care through SHI with countries that do so through general revenues, the reason being that whether a country has a SHI system or a tax-financed system in place is likely to be endogenous—unobservable factors correlated with the outcome of interest are also likely to be correlated with the choice of system.[3] Two recent papers by Wagstaff and Moreno-Serra (2008a; 2008b) (hereafter WMS) attempt to get round this problem by focusing on *changes* in the way countries finance their health care. If the problematic unobservable variables are time-invariant, or if the unobservables of the SHI and tax-financed countries follow a parallel trend, a simple-differences-in-differences regression will reveal the effect of having one system in place (say, SHI) rather than the other. WMS test the parallel trend assumption (PTA) against two alternative specifications: one where the unobservables follow a country-specific linear trend, the other where they follow a system-specific trend that is not necessarily linear. They also test whether these three specifications properly account for the potential reverse causality, and estimate the impacts of SHI via instrumental variables (IV) where they do not.

This paper employs the same methods as WMS but on a (mostly) different group of countries, namely those in the OECD, and over a longer time period, namely 1960-2006.[4] Between 1967 and 1986, no fewer than 10 OECD countries abandoned SHI in favor of the tax-financed NHS-type model. By contrast, in the 1990s, all four of the OECD's new central European countries abandoned the tax-financed Semashko model in favor of SHI. The OECD countries provide a potentially important learning experience for other countries: it was in

[3] Van der Zee and Kroneman (2007), for example, compare mean outcomes between SHI and tax-financed OECD countries at different points in time. Implausibly, the authors' conclusions are made on the basis of simple comparisons of means. No attempt is made to control for confounders, whether observed or unobserved.

[4] WMS focused on the countries of central and east Europe, and central Asia, and covered the period 1990-2004.

Germany and neighboring countries that SHI began, and where arguably it has become most sophisticated; and it was in the UK and Scandinavia where the tax-financed Beveridge NHS model emerged and has developed most. By contrast, the countries examined by WMS—namely those in Central and Eastern Europe, and Central Asia—contain SHI countries that are still making adjustments to their new systems as well as tax-financed countries that have often struggled with the transition to a market economy following the breakup of the former Soviet Union. One might argue that the higher health spending but similar health outcomes associated with SHI unearthed by WMS could be due to a combination of immaturity on the part of the region's new SHI systems and fiscal stress among the region's tax-financed systems.

The present paper differs from the two aforementioned studies in another respect, namely that in its analysis of the impacts of SHI on health outcomes the present paper restricts its attention to causes of death (and age groups) identified by Nolte and McKee (2008) as amenable to medical care, i.e. causes where timely and effective medical care can result in a death being avoided. WMS looked at many more causes of death than are included in the present paper, of which five are among the 33 considered to be amenable to medical care; however, for three of the five causes of death, they used the standardized death rate, which reflects deaths among people aged 75 and over among whom deaths are less easily avoided. So, in addition to focusing on more mature SHI and tax-financed health systems, this paper focuses on causes of death that ought to be amenable both to medical care and to changes in health financing and delivery arrangements. If SHI systems do leave some groups uncovered or covered less well than others, this ought to get reflected in higher death rates from causes amenable to medical care. If SHI systems are better funded, or if they are more efficient in the purchasing and provision of health care, this too ought to get reflected in amenable mortality data.

The organization of the paper is as follows. Section 2 provides a brief history of health system reforms in the OECD countries, and discusses the ways that choice of one system over another might affect health spending, health status, and labor market outcomes. Section 3 presents a brief summary of the methods used, section 4 the data, and section 5 the empirical results. Section 6 presents the paper's conclusions.

## 2. OECD health systems, and health sector and labor market outcomes

The OECD countries—with the exception of the United States, which relies largely on private insurance and out-of-pocket payments, and is excluded from the analysis in this paper—rely largely on general revenues or SHI to finance health care. This section provides a brief review of the transitions from and to SHI, and speculates on the possible impacts on the outcomes analyzed in the empirical analysis. These include: per capita national health spending; amenable mortality; employment; and the formal-sector share of total employment. Too few years' worth of data on measures of health system throughput—inpatient admissions, length of stay, etc.—are available to allow these to be analyzed, or variables building on these (such as cost per inpatient admission).

### *2.1 Transitions to and from SHI*

In terms of health financing, the (now) 29 OECD countries (excluding the United States) fall into four groups (cf. Figure 1).[5] Group (i) comprises countries that have maintained a SHI system since the early 1960s: this group includes Austria, Belgium, France, Germany, Japan, Korea, Luxembourg, Mexico, the Netherlands, Switzerland and Turkey. Except in Mexico and

[5] This section relies heavily on the Health Systems in Transition (HiT) series, downloadable from http://www.euro.who.int/observatory/Hits/TopPage, and Saltman and Dubois (2004).

Turkey, SHI contributions account for the vast majority of health revenues in this group of countries, and the vast majority of the population is covered by SHI. In Mexico and Turkey, out-of-pocket payments and taxes finance a sizeable share of health spending. Group (ii) includes those who have maintained predominantly tax-financed systems since the 1960s: these countries include Australia, Canada, Finland, Ireland, New Zealand and the United Kingdom. The commitment to tax finance in some countries in this group—notably Australia, Canada, Finland, and Sweden—increased in the 1960s and early 1970s, but none of these countries was a SHI country at any stage in the period studied. Group (iii) includes those who relied largely on SHI contributions but then opted for a universal tax-financed health system at some stage since 1960: this group includes Denmark, Greece, Iceland, Italy, Norway, Portugal, Spain and Sweden. Group (iv) comprises the central European countries that had a SHI system prior to the post-war Communist takeover in the late 1940s, operated a tax-financed Semashko system under Communism, and then reverted to SHI system in the 1990s: this group includes the Czech Republic, Hungary, Poland, and the Slovak Republic.

The fundamental difference between SHI and tax-financed systems is that SHI systems raise revenues largely from earnings-related contributions levied largely on formal sector workers while tax-financed systems draw their revenues from taxes and nontax government revenues. The two systems also differ, however, in their delivery arrangements: traditionally SHI systems have been more likely to contract with providers (public and private) rather than operate directly-managed facilities (the model many tax-financed systems operate), and as a result the provider payment mechanisms in SHI systems tend to be more high-powered than those in tax-financed systems; furthermore, most tax-financed systems operate a GP gatekeeper system, whereas most SHI systems do not. These differences are likely to have consequences for health spending, health outcomes, and employment.

### *2.2 SHI vs. tax-finance vis-à-vis health spending*

The collection of earnings-related contributions in SHI systems is typically done within the SHI system, while a tax-financed system's revenues are collected via the tax system. Other things equal, this is likely to make a SHI system more costly. These extra costs are often seen as a price worth paying by SHI advocates who see earnings-related contributions as more stable and more reliable than tax revenues: people may be willing to contribute more if they know their contributions are going to provide health insurance coverage for them and other contributors than if the monies end up in a large pool of tax revenues whose allocation between the health sectors and other sectors is left to the whims of policymakers. In practice, evasion and avoidance of SHI contributions is a major issue in many SHI countries, and some ministries of finance reduce the tax revenues they assign to the health sector as SHI revenues grow, so SHI revenues are not necessarily additional.

Differences in health spending between the two systems are also likely to reflect differences in delivery arrangements. As mentioned, traditionally SHI systems have been more likely to contract with providers (public and private) rather than operate directly-managed facilities. Again, there are exceptions—some SHI countries (Mexico is an example) rely in part or wholly on their own network of directly managed facilities, and there are some tax-financed systems that have used or are starting to use a contracting approach—but historically the separation of purchasing and provision has been more pronounced in the SHI countries. This separation is likely to entail additional costs to the system, especially where there is selective contracting: the costs include the selection of providers, the drawing-up and enforcement of contracts, fraud, and so on.

On the other hand, the purchaser-provider split seen in SHI systems may put downward pressure on health spending—a reason why many tax-financed systems have experimented with it. Whether it does so in practice is likely to depend in part on the payment methods used. Many SHI countries paid providers fee-for-service (FFS), which encourages the provision of services on which providers can earn high margins; it is well known to be a payment system that is associated with high per capita health spending levels. Only fairly recently have SHI systems moved away from FFS and many have yet to do so (Table 1). By contrast, tax-financed systems have traditionally (again, there are exceptions) tended to use lower-powered incentives that are less likely to lead to high levels of per capita expenditure. Where FFS has been used, it has tended to be in the context of paying primary care providers fees to provide specific preventive measures, such as immunizations. Only fairly recently have tax-financed systems begun to use more high-powered incentives for hospitals, and even here the shift has tended to be toward diagnosis-related groups or similar methods rather than toward FFS.

Another factor affecting whether the purchaser-provider split commonly seen in SHI systems puts downward pressure on total health spending is the way the purchaser interacts with providers. Surprisingly, perhaps, it was not until fairly recently that SHI systems or tax-financed systems began to contract selectively with competing providers. In Belgium, Hungary, and Switzerland, sickness funds contract with all health care providers, and in the Czech Republic, Germany, the Netherlands and the Slovak Republic selective contracting has begun only recently. If US evidence is a guide, this might be expected to put downward pressure on costs.

However, some tax-financed purchasers have also started to contract selectively, so this may not be a factor in understanding any expenditure differences between the two types of system. By contrast, SHI systems appear to have more of a tradition of making use of both public and private providers, whereas in tax-financed systems it has tended to be at the primary care level that private providers have been used. Whether this will work to give SHI systems lower spending is unclear.

A final factor likely to be important in driving spending differences between SHI and tax-financed systems is the absence of a GP-gatekeeper in a typical SHI system. The Netherlands is unusual among SHI systems in having a GP-gatekeeper (some other SHI countries have tried but largely failed to introduce it), while Finland and Iceland are unusual among tax-financed systems in not having one (Table 1). Countries with GP gatekeepers are known empirically to have lower spending per capita (Gerdtham and Jönsson 2000).

### *2.3 SHI vs. tax-finance vis-à-vis health outcomes*

On balance, then, it seems plausible a prior that SHI systems will spend more per capita on health than tax-financed systems. Insofar as this is true, one might be tempted to conclude that SHI systems ought to do better than tax-financed systems in terms of health outcomes. This conclusion is not necessarily warranted, however, since the very differences that make SHI systems more likely to spend more are also likely to affect the health outcomes they achieve for a given level of spending.

One factor working against SHI systems is that coverage is not necessarily universal in a SHI system but is typically so in a tax-financed system. The southern European OECD countries achieved universal coverage only after switching from SHI to tax finance, and many SHI systems did not achieve universal coverage until some time after the start of the period under study (Korea, for example, achieved universal coverage only in 1989). Some SHI systems in the OECD (notably Turkey and Mexico) have yet to achieve universal coverage. It is typically the case that the groups not covered by the SHI program—who are often reliant on programs and facilities run by the health ministry—will be disadvantaged in terms of resources: in Mexico, for example, the health ministry spends less than 50% per capita than the social security institute that operates the SHI program (OECD 2005). Moreover, several SHI systems operate (or used to operate) multiple SHI schemes alongside one another. In most countries, gaps in benefit packages have been reduced if not eliminated, but in some countries they remain: in Austria, for example, the highest-spending sickness fund spends 63% more per person than the lowest-spending sickness fund (Koettl 2008). In countries where SHI schemes compete with one another for enrollees usually under a risk-adjusted capitation system (as is currently the case in the Czech Republic, Germany, the Netherlands, the Slovak republic, and Switzerland), there is the added worry of risk-selection, so that depending on the risk-adjustment formula some groups will prove less profitable and may be avoided by insurers and as a result underserved (van de Ven *et al.* 2003; van de Ven *et al.* 2007).

These gaps and inequalities in coverage in SHI systems are likely to translate into inequalities in per capita health spending, which in turn will work to reduce average health below what it would have been in the absence of spending inequalities if health spending is subject to diminishing returns in the production of health.[6] This ignores, however, the possibility that tax-

[6] The argument is analogous to Rodgers's (1979) argument about income inequalities lowering mean health status if income is subject to diminishing returns in the production of health.

financed systems may also be prone to spending inequalities across the population. While many OECD countries now have schemes to steer resources between different areas to reflect population numbers (if not population health care needs), these efforts typically date back to the late 1970s and early 1980s (the resource-allocation working party (RAWP) formula was introduced in England and Wales only in 1977, for example). Today, in fact, it is not clear that inequities in health care utilization between the better off and the less well off are any higher in the SHI countries in the OECD than in the tax-financed countries (Masseria *et al.* 2004). This may not always have been the case, however.

Differences between SHI and tax-financed systems in the organization and payment of providers are also likely to get reflected in health outcomes for a given level of health spending. The aforementioned greater use of the GP gatekeeper model in tax-financed systems seems likely to push tax-financed systems toward better health outcomes, since the primary care system offers scope for greater use of preventive care, better management of chronic problems, a better chance of early detection and treatment of illness, a more coordinated approach to care across different providers, and so on. It is also sometimes argued that the more integrated approach to provision in a tax-financed system offers a better chance of a close relationship between the public health system and the provision of personal health care—a theme that is returned to in the conclusions.

The greater use of contracting providers in SHI systems is also likely to get reflected in health outcomes for a given level of spending. The traditional payment methods of a tax-financed system and its traditionally heavy emphasis on a directly-managed provider network may work well to contain overall spending levels, but do little to encourage higher volumes of care (Ellis and Miller 2008); SHI's traditional emphasis on FFS is likely to do better in this regard. Having insurers contract among competing providers might also be expected to encourage lower costs if the US experience is anything to go by (cf. e.g. Bodenheimer 2005) but as mentioned above, this has only recently started in the OECD's SHI countries, and has also started in some tax-financed systems too. In any case, extra volumes of care may not necessarily translate automatically into better health outcomes if the measures that produce the extra volume of care—the payment method, or the selective contracting—adversely affects the quality of care. Under FFS providers may deliver additional services whether or not they are medically necessary, but under the budget payment system seen frequently in tax-financed systems providers are likely to end up under-serving patients (Ellis and Miller 2008). The US evidence on the impact of competition on quality is mixed; however, the evidence from the UK which started encouraging public hospitals to compete with one another for contracts with tax-financed public payers, suggests that quality may have been adversely affected by competition (Propper *et al.* 2004; Propper *et al.* 2008). In any case, this is a moot point, given that neither SHI nor tax-financed systems have historically relied upon competition between providers.

### *2.4 SHI vs. tax-finance vis-à-vis labor market outcomes*

Overall, then, the likely impacts on health outcomes of the choice between SHI and tax-financed systems is unclear a priori. By contrast, the impact of the choice on labor market outcomes is somewhat clearer. SHI contributions are in effect a payroll tax and hence, according to the textbook treatment of payroll taxes, are likely to reduce employment levels and encourage informal employment. This argument begs, however, the question of what the labor market effects would be of whatever taxes were used instead of a payroll tax. As far as employment is concerned, it has been argued that, at least in the OECD countries, payroll, income and consumption taxes all have broadly similar effects on employment (cf. Nickell 2004; Nickell *et al.* 2005).

On the other hand, SHI can provide a disincentive for people to join and stay in formal employment, while the same is not true of general revenue-financing (cf. e.g. Belev 2003; Baeza and Packard 2006; Datta 2006; Levy 2007). The incentive is likely to be especially large in the case where among informal sector workers SHI enrollment is voluntary (as in Germany) or is only weakly enforced (more likely to be the case in countries like Mexico and Turkey). In such a setting, people can rely on private insurance (in some countries they are required by law to have it if they do not join the SHI scheme), or they may opt for informal employment arrangements and resort to using the health ministry's public health care system, albeit facing higher out-of-pocket payments than they would have done under the SHI system and probably ending up with worse quality care. These incentives become even more of an issue if being in formal employment means being drawn into the tax system and having to contribute to a pension scheme that because of limited life expectancy may only pay out for a few years if any.

## 3. Methods

The methods used are the same as those used by WMS (2008a; 2008b) to which the reader is referred for more detail. Let $y_{it}$ be the outcome of interest in country $i$ at time $t$, $SHI_{it}$ be a dummy variable taking on a value of 1 if country $i$ has a SHI health financing system at time $t$, and $X_{it}$ be a vector of covariates thought to potentially influence both outcomes and the SHI adoption decision. The basic model takes the form:

$$(1) \qquad y_{it} = X_{it}\gamma + \delta SHI_{it} + e_{it},$$

where the $e_{it}$ capture unobservable variables and statistical noise. The interest is in the coefficient $\delta$ which gives the impact of SHI on $y_{it}$. If $SHI_{it}$ and $e_{it}$ are correlated (i.e. SHI status is endogenous), estimation of eqn (1) by pooled OLS would result in a biased estimate of $\delta$. It could be that countries with unobserved characteristics that led to higher-than-expected levels of, say, self-employment may deliberately choose not to adopt SHI because of the difficulty of having a contribution-based financing system with large numbers of self-employed. Or it might be that certain changes or events occurred broadly around the same time that SHI was introduced; if we do not capture these in our model but instead implicitly include them in $e_{it}$, and if they affect the outcomes of interest, our estimate of $\delta$ will be biased.

### *3.1 Difference-in-differences models*

The simplest way to allow for such a correlation is to let:

$$(2) \qquad e_{it} = \alpha_i + \theta_t + \varepsilon_{it} \,,$$

where $\theta_t$ is a period-specific intercept, $\alpha_i$ is a country-specific effect which captures time-invariant unobservables that are potentially correlated with SHI status, and $\varepsilon_{it}$ is an idiosyncratic error term (iid over $i$ and $t$). Substituting eqn (2) in eqn (1), and taking first-differences, gives:

$$(3) \qquad \Delta y_{it} = \Delta X_{it}\gamma + \delta \Delta SHI_{it} + \xi_t + \Delta\varepsilon_{it} \,.$$

In the special case where the $X_{it}$ are omitted, eqn (3) is the standard difference-in-differences (DID) estimator (cf. e.g. Wooldridge 2002 p.284).[7]

Eqn (3) assumes a parallel or common trend: the $\theta_t$ do not depend on the value of $SHI_{it}$, and therefore the health systems that switch from tax-finance to SHI or vice versa and those that

[7] Standard errors need to be adjusted for clustering at the country level to allow for serial correlation (cf. Bertrand *et al*. 2004; Cameron and Trivedi 2005 p.705).

do not switch exhibit the same trend. A model that allows this parallel trend assumption (PTA) to be relaxed is the 'random trend' (RT) model (cf. e.g. Wooldridge 2002 p.316). Eqn (2) is replaced by the assumption

(4) $$e_{it} = \alpha_i + \theta_t + k_i t + \varepsilon_{it}.$$

This allows for the possibility that different countries have different trends, as reflected in different values of $k_i$. Substituting eqn (4) in eqn (1), and taking first differences, gives:

(5) $$\Delta y_{it} = \Delta X_{it}\gamma + \delta \Delta SHI_{it} + \xi_t + k_i + \Delta\varepsilon_{it},$$

which can be estimated using a fixed effects estimator. If the $k_i$ are jointly insignificant, eqn (5) collapses to eqn (3), which would provide some evidence in support of the PTA. Yet even if the $k_i$ were jointly significant, the PTA would still be a reasonable assumption if the $k_i$ are uncorrelated with $SHI_{it}$; this can be tested via a generalized Hausman test (cf. Wooldridge 2002 p.290).[8]

The RT model, while less restrictive than the DID estimator, assumes the trend in unobservables is linear, and that the trend is specific to the country and is not modified by a switch from tax-finance to SHI or vice versa. Another model that allows the PTA to be relaxed is the 'differential trend' (DT) model of Bell et al. (1999). They assume:

(6) $$e_{it} = \begin{cases} \alpha_i + k_s m_t + \varepsilon_{it} & if \quad SHI_{it} = 1 \\ \alpha_i + k_n m_t + \varepsilon_{it} & if \quad SHI_{it} = 0 \end{cases},$$

---

[8] This can be implemented by estimating an augmented version of eqn (5) using a random effects estimator—adding the within-country panel means of the original covariates which vary over *i* and *t* as regressors, and testing the null hypothesis of insignificance of the additional SHI term (with cluster-robust standard errors). Non-rejection of this hypothesis would suggest that the $k_i$ are uncorrelated with $SHI_{it}$.

where $m_t$ is an unobserved trend, the influence of which on $y_{it}$ is allowed to differ between SHI and non-SHI systems. Incorporating this assumption into eqn (1), and taking first differences, gives:

(7) $$\Delta y_{it} = \Delta X_{it}\gamma + \delta\Delta SHI_{it} + k_n \Delta m_t + (k_s - k_n)\Delta(m_t SHI_{it}) + \Delta\varepsilon_{it}\,.$$

To estimate eqn (7), the $\Delta m_t$ are replaced by first differences of year dummies and the $\Delta(m_t SHI_{it})$ are replaced by first differences of interactions between year dummies and the SHI status dummy. The estimating equation is thus:

(8) $$\Delta y_{it} = \Delta X_{it}\gamma + \delta\Delta SHI_{it} + \sum_{\tau=2}^{T}\beta_\tau \Delta YEAR_\tau + \sum_{\tau=2}^{T}\varphi_\tau \Delta(YEAR_\tau SHI_{it}) + \Delta\varepsilon_{it}\,,$$

which can be estimated by pooled OLS. In this model the impact of SHI varies over time, but one can estimate the average impact of SHI over time:

(9) $$MEAN\ SHI\ IMPACT = \hat{\delta} + \sum_{\tau=2}^{T}\hat{\varphi}_\tau / T - 1\,.$$

The PTA assumption in this model implies $k_s$=$k_n$. This can be tested indirectly by testing the nonlinear restriction:

(10) $$\frac{\sum_t m_t (k_s - k_n)}{\sum_t m_t k_n} = \frac{(k_s - k_n)\sum_t m_t}{k_n \sum_t m_t} = \frac{\sum_{\tau=2}^{T}\varphi_\tau}{\sum_{\tau=2}^{T}\beta_\tau} = 0\,.$$

## 3.2 *Testing for and dealing with reverse causality*

Although the DID, RT and DT models all allow for some correlation between SHI and the original error term $e_{it}$, they entail specific assumptions that may not adequately capture the endogeneity of SHI. Like WMS, the analysis below uses the test of reverse causality proposed by Gruber and Hanratty (1995). In each of the three models, a lead dummy variable is included indicating whether SHI will be adopted the following year. If causality goes from SHI to the outcome variable, the coefficient on the lead dummy will be zero. A nonzero coefficient would point towards causality running the other way or some other type of endogeneity that cannot be captured by the model in question.

For outcomes where none of the three models above is able to address the endogeneity of $SHI_{it}$, estimation is undertaken using instrumental variables (IV), using lags of the potentially endogenous SHI variable as instruments, as well as "traditional" instruments. If these instruments are valid (i.e. exogenous and strong in the sense of being highly correlated with the instrumented SHI dummy—assumptions which can be tested), they should control for any kind of endogeneity in eqn (1), including that arising from country-specific effects and trends. Eqn (1) can thus be consistently estimated by two-stage least squares (IV-2SLS) or using the more efficient two-step generalized method-of-moments (IV-GMM) estimator (cf. e.g. Cameron and Trivedi 2005).

The first lag of the SHI dummy ($SHI_{i,t-1}$) is used as an instrument[9], along with two other time-invariant variables: a dummy indicating whether the country was a World War II Axis member (Germany, Italy and Japan) or occupied by one or more of the Axis power[10]; and a dummy indicating whether the country is an English-speaking country. Many SHI countries had a fledgling SHI system in place prior to World War II, but Germany was quick to mould the health systems of the countries it occupied to conform to the Bismarkian SHI model, and many

[9] Although under weak exogeneity one could theoretically use more lags of the SHI variable as instruments, only the first lag is included due to the lack of variation over time in our SHI dummy (transitions between tax-funded health and SHI systems occur only once, if at all, in our sample), leading to the redundancy of additional lags and loss of degrees of freedom for overidentifying restrictions tests.

[10] Countries with a value of 1 for the Axis dummy are: Austria, Belgium, Czech Republic, Slovak Republic, Denmark, Finland, France, Germany, Greece, Hungary, Italy, Japan, Korea, Luxembourg, Netherlands, Norway, Poland.

countries left the system largely intact after 1945. Japan occupied Korea between 1910 and 1945, and while it was not until the early 1960s that SHI began to take off in Korea, the influence of Japan's system is much in evidence (Anderson 1989; Bong-Min 1996). By contrast, political leaders in the English-speaking countries pushed for tax-finance early after World War II, though universal tax-financed health care did not immediately become a reality except in the UK where it was adopted in 1948. The relevance of the three instruments is assessed through LM and Wald versions of under-identification and weak identification tests based on the *rk* statistic recently proposed by Kleibergen and Paap (2006) which account for non-iid structures of the error terms in the estimated equations. Cluster-robust versions of Hansen's *J*-statistic tests are used to check the exogeneity of the instruments in the estimated models.

## 4. Data

The empirical analysis uses a SHI indicator constructed from various sources, and annual data from OECD Health Data for the period 1960-2003.[11]

### *4.1 Social health insurance status*

$SHI_{it}$ is defined as taking a value of one if in country *i* at time *t* if the cell in Figure 1 is shaded gray. The SHI status dummy is equal to 1 in about half (716 observations) of the 1,363 country-year combinations for which there are non-missing values of the indicator.

For the most part, the classifications in Figure 1 are innocuous. However, there are some ambiguities worth mentioning. In group (i), there are several ambiguities. Germany was two countries from 1945-89, and operated two different health systems; no allowance is made for this

[11] The GDP deflator was taken from the World Bank's World Development Indicators database.

fact. Strictly, Korea began SHI only in 1963, but it was a largely out-of-pocket system before then, and coding it a tax-financed system before 1963 would likely be a bigger error. Mexico and Turkey began setting up a SHI system in the 1940s, but universal coverage has yet to be achieved in either country; nonetheless, both countries manifestly aspire to be fully-fledged SHI systems. In group (ii), there are no coding ambiguities. Nor are there any in group (iii)—all are countries that had a fledgling SHI system in place but abandoned it in favor of universal tax-financed health care. In group (iv), the only ambiguity is Poland, which reverted to SHI, having had a fledgling SHI system before World War II, but levies a dedicated income tax on everyone to finance it rather than an earnings-related payroll tax on formal sector workers. The results presented below are based on Poland being coded a SHI country from 1998 onwards; the results are similar if Poland is coded a tax-financed country throughout the period.

### *4.2 Outcome variables*

The dependent variables include total health expenditure (THE) per capita, potential years of life lost (PYLL) for nine 'amenable' causes of death, employment as a share of the population of working age, and the share of people employed who are paid a salary or a wage.

THE per capita is converted at purchasing power parities (PPP) and expressed in real terms using the country's GDP deflator. PYLL indicate the number of years of life (up to the age of 70) that are lost as the result of a death. They are calculated by summing up deaths occurring at each age and multiplying this by the number of remaining years to live up to the selected age limit of 70. To improve comparability over time and across countries, the OECD Health Data PYLL calculations are age-standardized to a reference population. The nine 'amenable' causes selected for the analysis below are the nine causes that appear both in the OECD Health Data and

Nolte and McKee's (2008) list of amenable causes. These are: malignant neoplasms of the colon and rectum (ICD10 C18-21); malignant neoplasm of the breast (females only) (C50); malignant neoplasm of the cervix uteri (females only) (C53); diabetes (E10-4); ischemic heart disease (ICD10-I20-5); cerebrovascular disease (I60-9); influenza and pneumonia (J10-28); maternal death (O00-99); and perinatal deaths other than stillbirths (P00-96).[12]

Employment includes formal-sector workers (employees paid a wage or salary) as well as the self-employed. The variable is defined as employment as a percentage of the population aged 15-65. The formal-sector employment variable is defined as the number of people with a salaried or waged job, expressed as a fraction of the number of people in employment.

Descriptive statistics are shown in Table 2. Health spending per capita is higher in SHI country-year combinations than in tax-financed country-year combinations. SHI country-year combinations have higher PYLL for six of the nine amenable causes of death, and have lower rates of formal sector employment and employment. These are, of course, simple bivariate comparisons, and tell us nothing about the causal effect of SHI on these outcomes.

### *4.3 Covariates in the estimating equation*

As in WMS (2008b), the only covariate included is GDP per capita. There is some evidence that in central and eastern Europe and central Asia SHI schemes emerged first in countries with higher initial (i.e. pre-transition) per capita income levels, whilst tax-based funding prevailed in countries with lower initial per capita income (Preker *et al.* 2002).

---

[12] Nolte and McKee (2008) argue that only half of ischemic heart disease deaths below the age of 74 are avoidable, and use an upper limit of 49 for diabetes. The OECD Health Data PYLL use an age limit of 70 for all ICD codes, so there is nothing that can be done about the latter. In both cases, we might expect that the choice of health system will matter less for these two causes of death than for other causes.

The exclusion of other possible covariates means that the results will capture the overall effects of SHI, holding constant only per capita income. One could include, in addition to per capita income, variables capturing the fraction of the population covered by insurance, the provider payment mechanism in force, whether or not there is a GP gatekeeper, and so on. But these are all features of the health system that are associated with whether the system is SHI or tax-financed. If these were included, the coefficient on SHI would be interpreted as the effect of SHI *holding these other institutional features constant*. The specification used in the present study, by contrast, sheds light on the overall effects of SHI, without shedding light on the channels by which the choice between SHI and tax finance makes a difference, if indeed it does.

Of course, whether or not the coefficient on SHI can be so interpreted depends on the specification not excluding other factors that are associated with the adoption of SHI or tax finance and the outcomes of interest. Were eqn (1) to be estimated by ordinary least squares (OLS), the case for interpreting $\delta$ as the impact of SHI would clearly be weak. But the use of the various generalizations of differences-in-differences and IV makes the case much stronger, especially as tests are reported that shed light on the validity of the assumptions underlying the various models. The results ought, therefore, to be consistent with a correlation between $e_{it}$ and $SHI_{it}$ in eqn (1) caused by omitted relevant variables being lumped into the error term.

## 5. Results

This section presents first the results of the specification tests, and then the estimates of the impacts of SHI.

### *5.1 Specification tests*

Table 3 reports the results of the reverse causality tests. For the per capita health spending variable, and most PYLL variables, there is no evidence that the DID, random trend and differential trend models fail to account adequately for any correlation between $e_{it}$ and $SHI_{it}$ in eqn (1). The exceptions are the ischemic heart disease PYLL variable, where the coefficient on the lead value of $SHI_{it}$ is significant at the 5% level in all three models, and cerebrovascular disease, where the coefficient on $SHI_{it+1}$ is significant at the 10% level in the random trend model and almost so in the basic DID model. In the case of the formal sector share labor market outcome, there is evidence that the three generalized differences-in-differences models fails adequately to take into account the correlation between $SHI_{it}$ and the error term in eqn (1). This is not an unexpected result, given that SHI hinges on formal-sector contributions; it is less easy to rationalize why there appears to be reverse causality in the cases of ischemic and cerebrovascular disease but not in the cases of other diseases.

Table 4 reports the results of the parallel-trend assumption (PTA) tests for the random-trend and differential-trend models, i.e. eqns (5) and (8). The only two relevant outcomes where the PTA is rejected are the influenza and pneumonia PYLL variable, and the pregnancy PYLL variable. In the former, the PTA is rejected (at the 10% level only) in favor of the random trend model but not in favor of the differential trend model, while in the case of the latter the PTA is rejected (at the 1% level) in favor of the differential trend model but not in favor of the random trend model.

Table 6 shows the relevant IV diagnostic test statistics for all the outcomes, though the interest lies with ischemic heart disease, cerebrovascular disease, and the formal-sector share of employment variables, for which IV is the preferred estimator. The three instruments are highly relevant for all outcomes, with high partial *F* statistics (not shown in Table 6). Kleibergen-Paap LM and Wald tests also strongly reject the null hypotheses of model under-identification and weak instruments respectively (Stock and Yogo 2002). The results of the Hansen tests for over-identifying restrictions, by contrast, are mixed. In the case of five outcomes, the null hypothesis of over-identification is rejected at the 10% level or lower, casting doubt on the validity of the instruments for these outcomes. In only one of these cases, however, is IV the indicated estimation strategy on the basis of the reverse causality tests, namely ischemic heart disease; for this variable, therefore, there are no reliable estimates of the impact of SHI. In the case of cerebrovascular disease, the over-identification restrictions are not rejected, but the rejection is not as decisive as one might like. By contrast, the Hansen test results suggest the over-identifying restrictions are valid in the case of both labor market outcomes, one of which failed the endogeneity test with all three DID models.

### *5.2 Estimates of SHI impacts*

Table 5 reports the coefficient estimates (and associated p-value for the null hypothesis of a zero SHI impact) for the DID models and Table 6 the IV estimates.[13]

In the case of health spending per capita, the simplest model consistent with the data is the basic DID model, i.e. eqn (3). This suggests that SHI raises total health spending per capita by 3.5%; the estimated impact is significant at the 1% level. The random trend model gives a similar result, as does the differential trend model, though the effect is not significant in the latter.

[13] The models are estimated using the natural logarithm of the dependent variable.

Turning to the estimates of SHI on potential years of life lost from amenable mortality, the preferred model for colon cancer is also the basic DID model; there is no sign in this model (or any other) of SHI having a significant effect. The same is not true, however, of breast cancer, where the preferred model is again the DID model: this model and the random trend model suggest that SHI raises mortality by 5-6%, though in both models the effect is significant at the 10% level but not at the 5% level. No effect is found for either cervical cancer or diabetes; the DID model is the preferred model in both cases. In the case of ischemic heart disease, there is, as already mentioned, no set of estimates that passes the relevant diagnostic tests. In the case of cerebrovascular disease, the preferred estimates are the IV estimates. However, the hypothesis of invalid over-identification restrictions is not decisively rejected, and in any case the lead SHI dummy test of exogeneity is not decisively rejected. The fact that the DID and IV estimates in this case give rather different results is therefore a cause for concern: neither suggest a significant impact at the 5% level, but the estimates are of opposite sign, and the IV results point to a positive impact of SHI that is significant at the 10% level. This result should be probably interpreted with caution. In the case of influenza and pneumonia, there is some ambiguity about the validity of the PTA assumption; in the event, none of the estimates is anywhere near being significant. The differential trend model is the preferred model in the case of pregnancy, but the estimated impact of SHI is not significant. Finally, in the case of the perinatal causes PYLL variable, the basic DID model is the preferred model; once again, there is no evidence of any impact of SHI.

Turning to the labor market outcomes, the preferred model in the case of the variable capturing the formal sector share of employment is the IV model, which, in this case, passes all the diagnostic tests decisively. The estimates suggest that SHI reduces the formal sector share of employment by 8-10%; the estimated effects are significant at the 1% level. The results on the employment level are less clear-cut. The simplest model consistent with the data is the basic DID model which suggests a small and insignificant negative effect. However, the IV estimates pass the various diagnostic tests decisively, and therefore can be considered legitimate estimates: they suggest a significant negative effect of SHI on employment of the order of 5-6%, but only the GMM estimate is significant and even then only at the 10% level.

## 6. Summary and discussion

The transitions between SHI and tax finance that have occurred in the OECD countries since 1960 provide an opportunity to assess the systemwide impacts of adopting SHI in preference to a tax-financed health system. While reforms during the 1990s and 2000s have left the distinction between the two models increasingly blurred, the two systems have—and for the most part continue to have—certain distinctive hallmarks. SHI systems provide coverage in exchange for earnings-related contributions levied largely on formal-sector workers; tax-financed systems, by contrast, rely on general government tax (and nontax) revenues. Gaps in coverage are commonplace in SHI systems, and traditionally differences have existed in benefit packages and expenditures per enrollee. SHI systems rarely have a GP acting as a gatekeeper to second-level providers, while tax-financed systems almost always do; the latter traditionally have, as a result, put greater emphasis on primary care. OECD SHI systems mostly do not directly manage their own provider network, but instead contract with providers—often public and private sector ones; fee-for-service (FFS) has been the commonest payment method, though this is changing.

The questions asked by this paper are whether, on balance, having a SHI system rather than a tax-financed system results in higher or lower health spending per capita, and whether having one system or the other (*not* holding health spending constant) results in better or worse outcomes with regard to amenable mortality, higher or lower levels of employment, and a higher or lower share of formal-sector workers in total employment. Econometric models are estimated on panel data from 29 OECD countries over the period 1960-2006. Particular attention is paid to the likelihood that whether a country has a SHI or tax-financed system in place at a particular moment in time may be correlated with excluded and potentially unobservable variables that also affect the outcomes of interest, i.e. SHI status may be endogenous. The models estimated include a generalized differences-in-differences model, a random trend model, and a differential trend model: the first assumes unobservables remain constant over time; the second allows them to change over time but linearly, albeit at a different rate in each country; the third allows them to grow nonlinearly, but at the same rate in all countries within a given health system group. The parallel trends assumption implicit in the first model is tested in the second two models. Where these assumptions fail adequately to capture the endogeneity of SHI status, as reflected in a significant coefficient on SHI status one year in the future, instrumental variables (IV) are used, and diagnostic tests are undertaken to assess the validity of the identifying restrictions.

The results are clear-cut on two questions: SHI raises per capita total health spending by 3-4%, and reduces the formal-sector share of employment by 8-10%. With regard to health outcomes, there is absolutely no evidence in the results that SHI systems achieve lower rates of amenable mortality. In fact, the evidence suggests that with regard to premature mortality from breast cancer among women, SHI systems perform worse, with 5-6% higher potential years of life lost (PYLL). The lack of positive health impacts in favor of SHI systems is despite their 3-

4% higher health spending per capita. With regard to employment levels overall, the IV results are consistent with SHI reducing total employment as a share of the working-age population by 5-6%, though the differences-in-differences estimates which are also valid point to a smaller (and insignificant) impact.

The results are similar (though somewhat smaller in magnitude) to those of Wagstaff and Moreno-Serra (WMS) (2008a; 2008b) who employ similar methods to look at the impacts of SHI transitions among a group of 28 (mostly different) countries in central and eastern Europe and central Asia over the period 1990-2003. They too found that SHI raised health spending per capita (by 11% in their comparable model), increased self-employment (by 17%), and reduced total employment as a share of the working-age population (by 10%). WMS, however, found no significant impacts in either direction on health status indicators; it is worth noting, though, that many of their cause-specific mortality indicators are not ones that are considered amenable to medical care, and their indicators reflected deaths among all age groups not just those among age groups where mortality is considered amenable to medical care.

The findings in this paper (and in WMS) with regard to SHI increasing health spending are not new, but rather provide reassurance that the often-noted higher spending levels in SHI countries in a typical cross-section reflects causality and is not a spurious association. The finding in this paper and in WMS with regard to the size (in employment terms) of the formal sector is novel, but is consistent with speculation by previous authors. The results in the two papers suggest that policymakers should take seriously the suggestion (cf. e.g. Baeza and Packard 2006) that SHI may encourage informality in labor markets.

The results in this paper concerning the impacts of SHI on health outcomes suggest that the WMS conclusions (SHI countries spend more but do not achieve better health outcomes) need to be nuanced slightly. The estimates reported in this paper suggest that SHI systems may actually achieve *worse* outcomes with respect to at least one cause of premature death that is amenable to medical care, namely breast cancer. This finding is not implausible. Coleman et al. (2008) have noted that reducing mortality from breast cancer (and other cancers, including cancer of the colon) requires the efficient implementation of organized, population-based screening programs which can provide detection and diagnosis at an early stage of disease (p.4), while Allin et al. (2004) concluded that “countries with SHI models of funding have less comprehensive national public health activities than those with tax-funded systems” (p.18). This, they note, reflects a different historical focus, with SHI systems focusing on individual members of insurance schemes and tax-financed systems focusing more broadly on the entire population. As a result, public health programs—including screening—tend in tax-financed systems to be part of a population-wide integrated program, while in SHI countries they are often a set of disjointed activities undertaken by several often poorly coordinated health system “actors”.[14] The finding that SHI systems perform worse on rates of amenable mortality where strong population-based public health programs are required is consistent with the Arbelaez et al.’s (2004) finding that the introduction of managed competition in Colombia’s SHI system led to fragmentation of responsibilities and discontinuity in the prevention and treatment of tuberculosis.

[14] In some SHI countries, many public health programs are undertaken by the public health authorities, and hence beyond the relationship between sickness funds and providers. In others, at least some public health activities are undertaken within this relationship, funded by sickness funds, delivered by private physicians, and overseen by public health authorities. Germany adopts this model for screening, but not altogether successfully according to Allin et al. In still other countries, public health activities are provided by other organizational structures that bring together sickness funds, providers, public health authorities and others. This is how the Netherlands organizes its cancer screening program. Whichever of the three approaches is used, the degree of coordination and integration seems bound to be less than it would be in an efficiently organized population-oriented tax-financed system.

The results of this paper suggest, then, that SHI systems, on balance, have certain characteristics that make them more expensive than tax-financed systems, do no better in terms of most health outcomes that are amenable to medical care despite the extra spending, may do worse in respect of outcomes that require strong population-level public health programs, and do worse in terms of encouraging informal labor markets and discouraging employment. It is, of course, possible that not all of the attributes of SHI systems pull in the same direction. SHI advocates might argue, for example, that its typically favored purchaser-provider split model where purchaser organizations contract with providers rather than manage them directly can lead to lower expenditures and better health outcomes. This could be true, and the results above are consistent with it being true. But they are also consistent with it not being true. And even if it were true, the question arises: How far policymakers can choose the good elements from one health system model without getting the bad ones? A SHI system can easily increase its reliance on tax revenues, as several have recently done. But can a tax-financed system with a directly managed provider network shift to a purchaser-provider-contracting model without risking coherence and integration in its public health programs? Questions such as these will doubtless occupy much of the future debate on health system design, and the SHI and tax-financed systems of the rich OECD countries will probably continue to borrow ideas from one another, making broad-brush studies such as this will less useful than studies that allow the pros and cons of specific aspects of each type of system to be determined. For the developing world, however, and for countries considering a wholesale switch between SHI and tax finance, broad-brush studies like this will continue to be helpful in deciding whether, on balance, SHI systems or tax-financed systems spend more, achieve better health, and lead to better labor market outcomes.

Figure 1: SHI systems in the OECD, 1960-2006


Sources: WHO Health in Transition series (HiTS) http://www.euro.who.int/observatory/Hits/TopPage, Saltman and Dubois (2004), Carrin and James (2005).

Table 1: Health financing, provider payment methods, and existence of gatekeeper—late 2000s

| | SHI/tax-finance | Primary care providers | Hospitals | Gatekeeper |
|---|---|---|---|---|
| Austria | SHI | Capitation + FFS | DRGs + FFS (outpatient) | No |
| Belgium | SHI | FFS | Budgets + FFS | No |
| Czech Republic | SHI | Capitation + FFS (previously FFS) | Budgets (previously FFS) | No |
| France | SHI | Capitation + FFS | Global budgets | No |
| Germany | SHI | Capitation | DRGs | No |
| Greece | SHI | FFS | Per diems | No |
| Hungary | SHI | Capitation | DRGs | No |
| Japan | SHI | FFS | FFS | No |
| Korea | SHI | FFS | FFS | No |
| Luxembourg | SHI | FFS | Global budgets | No |
| Mexico | SHI | Budgets | Budgets | No |
| Netherlands | SHI | Capitation | DRGs | Yes |
| Poland | SHI | Capitation | DRGs | No |
| Slovak Republic | SHI | Capitation + FFS (previously FFS) | Budgets (previously FFS) | Partial |
| Switzerland | SHI | FFS | Global budgets + per diems | No |
| Turkey | SHI | Salary | Global budgets + variations | No |
| Australia | Tax | FFS | Global budgets | Yes |
| Canada | Tax | FFS | Global budgets | Yes |
| Denmark | Tax | Capitation + FFS | Global budgets | Yes |
| Finland | Tax | Salary, capitation + FFS | FFS | No |
| Iceland | Tax | Salary | Global budgets | No |
| Ireland | Tax | Capitation | Global budgets | Yes |
| Italy | Tax | Capitation + FFS | DRGs | Yes |
| New Zealand | Tax | Capitation + FFS | Global budgets | Yes |
| Norway | Tax | Salary, capitation + FFS | DRGs | Yes |
| Portugal | Tax | Budgets | Global budgets | Yes |
| Spain | Tax | Salary | DRGs | Yes |
| Sweden | Tax | Salary | Global budgets + variations | Yes |
| United Kingdom | Tax | Capitation + allowances + FFS | DRGs (previously budgets) | Yes |

Sources: WHO Health in Transition series (HiTS) http://www.euro.who.int/observatory/Hits/TopPage and others.

## Table 2: Descriptive statistics

| | Full sample | | | SHI = 1 | | | SHI = 0 | | |
|---|---|---|---|---|---|---|---|---|---|
| | Mean | SD | Obs | Mean | SD | Obs | Mean | SD | Obs |
| Tot. health exp. per capita | 174.30 | 166.46 | 872 | 235.87 | 210.91 | 409 | 119.92 | 81.45 | 463 |
| PYLL cancer of colon | 93.85 | 32.70 | 1158 | 87.77 | 29.32 | 594 | 100.24 | 34.82 | 564 |
| PYLL breast cancer | 224.61 | 67.76 | 1158 | 207.79 | 73.50 | 594 | 242.33 | 55.97 | 564 |
| PYLL cervical cancer | 60.20 | 42.27 | 1158 | 63.50 | 49.86 | 594 | 56.72 | 32.09 | 564 |
| PYLL diabetes | 56.60 | 46.59 | 1162 | 64.99 | 61.25 | 598 | 47.70 | 18.51 | 564 |
| PYLL ischemic heart disease | 527.32 | 291.84 | 1162 | 423.92 | 211.29 | 598 | 636.95 | 323.89 | 564 |
| PYLL cerebrovascular disease | 243.25 | 132.84 | 1162 | 257.21 | 141.66 | 598 | 228.46 | 121.17 | 564 |
| PYLL influenza pneumonia | 229.75 | 366.43 | 1162 | 302.81 | 472.14 | 598 | 152.28 | 169.80 | 564 |
| PYLL pregnancy | 29.33 | 54.45 | 1154 | 41.40 | 70.74 | 597 | 16.39 | 21.46 | 557 |
| PYLL perinatal causes | 746.39 | 623.71 | 1162 | 862.32 | 687.16 | 598 | 623.48 | 521.64 | 564 |
| Formal sector employment | 0.77 | 0.13 | 1112 | 0.74 | 0.14 | 575 | 0.81 | 0.09 | 537 |
| Employment | 0.66 | 0.08 | 1163 | 0.65 | 0.08 | 620 | 0.67 | 0.08 | 543 |

Note: Health expenditure per capita is at 1960 prices and converted at PPP. PYLL are expressed per 100 000 population. Formal sector employment is as a share of total employment. Employment is as a share of the population aged 15-65.

## Table 3: Tests of reverse causality

| | DID model<br>Lead SHI dummy test on eqn (3) | | Random trend model<br>Lead SHI dummy test on eqn (5) | | Differential trend model<br>Lead SHI dummy test on eqn (8) | |
|---|---|---|---|---|---|---|
| Dependent variable | Coef | p-value | Coef | p-value | Coef | p-value |
| Tot. health exp. per capita | -0.007 | 0.641 | -0.001 | 0.953 | -0.012 | 0.572 |
| PYLL cancer of colon | -0.076 | 0.198 | -0.073 | 0.198 | -0.075 | 0.279 |
| PYLL breast cancer | -0.084 | 0.189 | -0.081 | 0.206 | -0.094 | 0.167 |
| PYLL cervical cancer | -0.032 | 0.782 | -0.026 | 0.822 | -0.039 | 0.736 |
| PYLL diabetes | 0.052 | 0.253 | 0.058 | 0.316 | 0.046 | 0.381 |
| PYLL ischemic heart disease | -0.031 | 0.003 | -0.027 | 0.014 | -0.031 | 0.003 |
| PYLL cerebrovascular disease | 0.079 | 0.102 | 0.081 | 0.085 | 0.075 | 0.137 |
| PYLL influenza pneumonia | -0.060 | 0.212 | -0.068 | 0.128 | -0.069 | 0.205 |
| PYLL pregnancy | 0.037 | 0.774 | 0.024 | 0.853 | 0.047 | 0.740 |
| PYLL perinatal causes | 0.007 | 0.909 | 0.005 | 0.944 | -0.013 | 0.869 |
| Formal sector employment | -0.011 | 0.004 | -0.009 | 0.014 | -0.011 | 0.001 |
| Employment | -0.002 | 0.780 | -0.002 | 0.757 | -0.001 | 0.821 |

Table 4: Tests of the parallel trend assumption

| Dependent variable | Random trend model Generalized Hausman test on eqn (5) chi-square | p-value | Differential trend model Non-linear restriction test on eqn (10) *F* | p-value |
|---|---|---|---|---|
| Tot. health exp. per capita | 2.450 | 0.118 | 0.290 | 0.595 |
| PYLL cancer of colon | 0.720 | 0.395 | 1.280 | 0.270 |
| PYLL breast cancer | 0.790 | 0.375 | 1.350 | 0.257 |
| PYLL cervical cancer | 0.570 | 0.450 | 0.060 | 0.809 |
| PYLL diabetes | 0.000 | 0.966 | 0.090 | 0.771 |
| PYLL ischemic heart disease | *1.090* | *0.296* | *0.650* | *0.428* |
| PYLL cerebrovascular disease | 0.190 | 0.664 | 0.470 | 0.501 |
| PYLL influenza pneumonia | 3.720 | 0.054 | 0.170 | 0.683 |
| PYLL pregnancy | 2.360 | 0.125 | 8.220 | 0.009 |
| PYLL perinatal causes | 2.170 | 0.141 | 0.120 | 0.728 |
| Formal sector employment | *3.170* | *0.075* | *5.550* | *0.028* |
| Employment | 1.130 | 0.289 | 0.100 | 0.756 |

Note: Italicized numbers indicate estimates that are to be ignored because diagnostic tests have shown them to be irrelevant or unreliable.

Table 5: DID estimates of SHI effects on health and labor outcomes

| Dependent variable | DID model Coef | p-value | Random trend model Coef | p-value | Differential trend model Coef | p-value |
|---|---|---|---|---|---|---|
| Tot. health exp. per capita | 0.035 | 0.001 | 0.040 | 0.001 | 0.037 | 0.261 |
| PYLL cancer of colon | -0.034 | 0.412 | -0.029 | 0.468 | -0.037 | 0.525 |
| PYLL breast cancer | 0.055 | 0.095 | 0.059 | 0.082 | 0.065 | 0.245 |
| PYLL cervical cancer | 0.020 | 0.805 | 0.027 | 0.738 | 0.005 | 0.949 |
| PYLL diabetes | 0.000 | 0.996 | 0.007 | 0.927 | -0.118 | 0.376 |
| PYLL ischemic heart disease | *0.009* | *0.760* | *0.013* | *0.635* | *0.131* | *0.139* |
| PYLL cerebrovascular disease | -0.125 | 0.175 | -0.124 | 0.187 | -0.065 | 0.125 |
| PYLL influenza pneumonia | 0.075 | 0.339 | 0.068 | 0.403 | 0.050 | 0.619 |
| PYLL pregnancy | -0.031 | 0.767 | -0.043 | 0.690 | -0.135 | 0.361 |
| PYLL perinatal causes | 0.012 | 0.599 | 0.010 | 0.690 | 0.098 | 0.396 |
| Formal sector employment | *0.000* | *0.944* | *0.002* | *0.699* | *0.006* | *0.788* |
| Employment | -0.003 | 0.407 | -0.003 | 0.370 | 0.025 | 0.153 |

Notes: In all models, the natural logarithm of the dependent variable is used. Italicized numbers indicate estimates that are to be ignored because diagnostic tests have shown them to be irrelevant or unreliable.

Table 6: IV estimates of SHI effects on health and labor outcomes

| | 2SLS<br>Coef. | p-val | GMM<br>Coef. | p-val | KP LM | KP Wald | Hansen J |
|---|---|---|---|---|---|---|---|
| Tot. health exp. per capita | *-0.154* | *0.05* | *-0.167* | *0.02* | 0.00 | 12363 | 0.044 |
| PYLL cancer of colon | *-0.179* | *0.08* | *-0.111* | *0.24* | 0.00 | 11004 | 0.034 |
| PYLL breast cancer | -0.263 | 0.06 | -0.217 | 0.06 | 0.00 | 11004 | 0.188 |
| PYLL cervical cancer | -0.059 | 0.72 | -0.067 | 0.65 | 0.00 | 10938 | 0.179 |
| PYLL diabetes | 0.235 | 0.15 | 0.226 | 0.15 | 0.00 | 10999 | 0.904 |
| PYLL ischemic heart disease | *-0.489* | *0.01* | *-0.551* | *0.00* | 0.00 | 11119 | 0.076 |
| PYLL cerebrovascular disease | 0.245 | 0.06 | 0.216 | 0.07 | 0.00 | 11119 | 0.220 |
| PYLL influenza pneumonia | *0.504* | *0.08* | *0.511* | *0.06* | 0.00 | 11119 | 0.021 |
| PYLL pregnancy | *0.953* | *0.00* | *0.944* | *0.00* | 0.00 | 11797 | 0.010 |
| PYLL perinatal causes | *0.285* | *0.21* | *0.286* | *0.16* | 0.00 | 11119 | 0.059 |
| Formal sector employment | -0.097 | 0.01 | -0.084 | 0.01 | 0.00 | 8353 | 0.492 |
| Employment | -0.053 | 0.14 | -0.060 | 0.08 | 0.00 | 8298 | 0.413 |

*Notes:* Italicized numbers indicate estimates that are to be ignored because diagnostic tests have shown them to be irrelevant or unreliable. The reported IV point estimates are from 2SLS estimation, where the excluded instruments are the first lag of the SHI dummy, whether the country was a World War II Axis member or occupied by one or more of the Axis power, and a dummy indicating whether the country is an English-speaking country. P-values are reported for the Kleibergen-Paap LM and Hansen tests, whereas the values reported for the Kleibergen-Paap Wald tests of weak identification correspond to the estimated *rk* Wald *F* statistics; in our case, the critical value for the weak identification tests (tabulated by Stock and Yogo 2002, 10% maximal IV size) is 22.30. In the last column, the joint null hypothesis of the over-identification tests is that the instruments are uncorrelated with the error term and that the excluded instruments are correctly excluded from the main equation.

**References**


Allin, S., E. Mossialos, M. McKee and W. Holland (2004). Making decisions on public health: a review of eight countries. Copenhagen, World Health Organization.

Anderson, G. F. (1989). "Universal health care coverage in Korea." Health Aff (Millwood) **8**(2): 24-34.

Baeza, C. and T. Packard (2006). Beyond Survival: Protecting Households from Health Shocks in Latin America. Washington, DC, World Bank.

Belev, B., Ed. (2003). The Informal Economy in the EU Accession Countries: Size, Scope, Trends and Challenges o the Process of EU Enlargement. Sofia, Center for the Study of Democracy.

Bell, B., R. Blundell and J. Van Reenen (1999). "Getting the Unemployed Back to Work: The Role of Targeted Wage Subsidies." International Tax and Public Finance **6**(3): 339-60.

Bertrand, M., E. Duflo and S. Mullainathan (2004). "How Much Should We Trust Differences-in-Differences Estimates?" Quarterly Journal of Economics **119 1**: 249-75.

Bodenheimer, T. (2005). "High and Rising Health Care Costs. Part 1: Seeking an Explanation." Ann Intern Med **142**(10): 847-854.

Bong-Min, Y. (1996). "The role of health insurance in the growth of the private health sector in Korea." International Journal of Health Planning and Management **11**(3): 231-252.

Busse, R. and A. Riesberg (2004). Health Care Systems in Transition: Germany. Copenhagen, WHO Regional Office for Europe on behalf of the European Observatory on Health Systems and Policies.

Cameron, A. and R. Trivedi (2005). Microeconometrics: Methods and Applications. Cambridge, Cambridge University Press.

Carrin, G. and C. James (2005). "Social Health Insurance: Key Factors Affecting the Transition Towards Universal Coverage." International Social Security Review **58**(1): 45-64.

Coleman, M. P., D. M. Alexe, T. Albreht and M. McKee (2008). Responding to the challenge of cancer in Europe. Responding to the challenge of cancer in Europe. M. P. Coleman, D. M. Alexe, T. Albreht and M. McKee. Ljubljana, Institute of Public Health of the Republic of Slovenia.

Datta, S. (2006). "Is Germany gearing up for health care reform?" Retrieved 10/31/06, from http://www.frost.com/prod/servlet/cif-econ-insight.pag?docid=77082735.

Ellis, R. P. and M. M. Miller (2008). Provider Payment Methods and Incentives. International Encyclopedia of Public Health. H. Kris. Oxford, Academic Press**:** 395-402.

Gerdtham, U. G. and B. Jönsson (2000). International Comparisons of Health Expenditure: Theory, Data and Econometric Analysis. Handbook of Health Economics. A. Culyer and J. Newhouse. Amsterdam, North Holland. **1:** 11–53.

Gruber, J. and M. Hanratty (1995). "The Labor-Market Effects of Introducing National Health Insurance: Evidence from Canada." Journal of Business and Economic Statistics **13**(2): 163-73.

Hsiao, W. C. and R. P. Shaw, Eds. (2007). Social Health Insurance for Developing Nations. Washington, D.C., World Bank.

Kleibergen, F. and R. Paap (2006). "Generalized reduced rank tests using the singular value decomposition." Journal of Econometrics **133**(1): 97-126.

Koettl, J. (2008). Yardstick competition among multiple non-competing health insurance funds: The case of Austria. Washington DC, The World Bank. Mimeo.

Levy, S. (2007). "Pueden los programas sociales disminuir la productividad y el crecimiento economico? Una hipotesis para Mexico. (With English summary.)." El Trimestre Economico **74 3**: 491-540.

Maria Patricia Arbelaez, M. B. G., Alvaro Franco, Roman Restrepo, Doracelly Hincapié, Erik Blas, (2004). "Tuberculosis control and managed competition in Colombia." The International Journal of Health Planning and Management **19**(S1): S25-S43.

Masseria, C., E. van Doorslaer and OECD Health Equity Research Group Members (2004). Income-Related Inequality in the Use of Medical Care in 21 OECD Countries. Paris, OECD. OECD Health Working Papers, No. 14.

Nickell, S. (2004). Employment and Taxes, Centre for Economic Performance, LSE, CEP Discussion Papers.

Nickell, S., L. Nunziata and W. Ochel (2005). "Unemployment in the OECD Since the 1960s: What Do We Know?" Economic Journal **115**(500): 1-27.

Nolte, E. and C. M. McKee (2008). "Measuring the health of nations: updating an earlier analysis." Health Aff (Millwood) **27**(1): 58-71.

OECD (2005). OECD Reviews of Health Systems: Mexico. Paris, OECD.

Preker, A., M. Jakab and M. Schneider (2002). Health financing reforms in central and eastern Europe and the former Soviet Union. Funding health care: options for Europe. A. D. E. Mossialos, J. Figueras and J. Kutzin,. Buckingham, Open University Press.

Propper, C., S. Burgess and D. Gossage (2008). "Competition and Quality: Evidence from the NHS Internal Market 1991-9." Economic Journal **118 525**: 138-70.

Propper, C., S. Burgess and K. Green (2004). "Does Competition between Hospitals Improve the Quality of Care? Hospital Death Rates and the NHS Internal Market." Journal of Public Economics **88 7-8**: 1247-72.

Rodgers, G. B. (1979). "Income and inequality as determinants of mortality: an international cross-section analysis." Population Studies **33**: 343-51.

Saltman, R. B. and H. F. W. Dubois (2004). The historical and social base of social health insurance systems. Social health insurance systems in western Europe. R. B. Saltman, R. Busse and J. Figueras. Maidenhead, United Kingdom, Open University Press.

Schmidt, U. (2006). Health policy and health economics in Germany. Washington DC, Friedrich Ebert Foundation. Speech.

Stock, J. H. and M. Yogo (2002). Testing for weak instruments in linear IV regression. Cambridge, Mass., National Bureau of Economic Research. NBER, Technical Working Paper # T 0284.

van de Ven, W. P. M. M., K. Beck, F. Buchner, D. Chernichovsky, L. Gardiol, A. Holly, L. M. Lamers, E. Schokkaert, A. Shmueli, S. Spycher, C. Van de Voorde, R. C. J. A. van Vliet, J. Wasem and I. Zmora (2003). "Risk adjustment and risk selection on the sickness fund insurance market in five European countries." Health Policy **65**(1): 98.

van de Ven, W. P. M. M., K. Beck, C. Van de Voorde, J. Wasem and I. Zmora (2007). "Risk adjustment and risk selection in Europe: 6 years later." Health Policy **83**(2-3): 162-179.

van der Zee, J. and M. Kroneman (2007). "Bismarck or Beveridge: a beauty contest between dinosaurs." Bmc Health Services Research **7**(1): 94.

Wagstaff, A. (2007). Social health insurance reexamined. Washington DC, World Bank. Policy Research Working Paper # WPS 4111.

Wagstaff, A. and R. Moreno-Serra (2008a). "Europe and Central Asia's great post-communist social health insurance experiment : impacts on health sector outcomes." J Health Econ (in press).

Wagstaff, A. and R. Moreno-Serra (2008b). Social health insurance and labor market outcomes: Evidence from central and eastern Europe, and central Asia. Advances in Health Economics and Health Services Research, volume 19: Health Care Financing in Low- and Middle-Income Countries (in press). D. Chernichovsky and K. Hanson. Bingley, United Kingdom, Emerald Group Publishing

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