

TI 2019-010/V
Tinbergen Institute Discussion Paper



Financial Protection Against Medical Expense

Owen (O.A.) O'Donnell¹

¹ Erasmus University Rotterdam, University of Lausanne

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and VU University Amsterdam.

Contact: discussionpapers@tinbergen.nl

More TI discussion papers can be downloaded at <http://www.tinbergen.nl>

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 598 4580

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900

FINANCIAL PROTECTION AGAINST MEDICAL EXPENSE

OWEN O'DONNELL

*Erasmus School of Economics, Erasmus University Rotterdam
Tinbergen Institute
Faculty of Business and Economics, University of Lausanne*

January 2019

Abstract

Financial protection is claimed to be an important objective of health policy. Yet there is a lack of clarity about what it is and no consensus on how to measure it. I address the ambiguity of meaning by considering three questions: Protection of what? Protection against what? Protection with what? The proposed answers lead to the suggestion that financial protection is about shielding nonmedical consumption from the cost of healthcare using formal health insurance and public finances, as well as informal and self insurance mechanisms that do not impair earnings potential. Given this definition, I evaluate four approaches to the measurement of financial protection: a) consumption smoothing over health shocks; b) the risk premium; c) catastrophic healthcare payments; and, d) impoverishing healthcare payments. The first of these does not restrict attention to medical expenses, which limits its relevance to health financing policy. The second rests on assumptions about risk preferences. No measure is entirely satisfactory in its treatment of medical expenses that are financed through informal and self insurance instruments. By ignoring these sources of imperfect insurance, the catastrophic payments measure overstates the impact of out-of-pocket medical expenses on living standards, while the impoverishment measure does not credibly identify poverty caused by them. It is better thought of as a correction to the measurement of poverty.

JEL Classification: D12, D31, D80, I13, I15

Keywords: health policy, financial protection, medical expenses, catastrophic payments, poverty

Acknowledgements: This paper is forthcoming in A.M. Jones (ed.) Oxford Encyclopedia of Health Economics. Oxford: Oxford University Press. doi: 10.1093/acrefore/9780190625979.013.27. It grew out of joint research with Gabriela Flores. I thank Gabriela for sharing her thoughts and helping me clarify my own. I am grateful to an anonymous reviewer for comments. The research is supported by the Swiss Agency for Development and Cooperation (SDC) and the Swiss National Science Foundation through the Swiss Programme for Research on Global Issues for Development (r4d programme) grant "Inclusive social protection for chronic health problems" (Grant number: 400640_160374 PI: Jürgen Maurer).

odonnell@ese.eur.nl

Introduction

Healthcare can be expensive. Paying the full cost of treatment would strain the finances of many and leave some bankrupt or destitute. Risk pooling can ensure that financial wellbeing is not threatened by medical needs. A multitude around the globe lacks this financial protection. To gauge their number and depth of predicament, measures of financial hardship caused by payments for healthcare and of economic vulnerability generated by exposure to uncertain medical expenses are needed. Such measures allow assessment of potential and realised gains from extending health insurance coverage. They are important tools to inform the setting of policy agendas and to evaluate reforms. Yet, despite “financial protection” often being cited as a major health policy objective, there is a lack of clarity over what it means. Consequently, there is no consensus on how to measure it. This impedes the design of efficient and equitable health financing instruments. While it is recognised that reasonable minds can differ regarding what financial protection metrics should aim to capture, the purpose of this paper is to bring greater clarity to the definition of the concept and so improve interpretation of measures of it.

Financial protection against medical expense is one of two primary motivations for universal health coverage (UHC), which the 2010 World Health Report (World Health Organization, 2010) thrust to the top of global health agenda. To monitor progress toward UHC, the World Health Organization (WHO) and the World Bank rely on two measures of financial protection. One purports to show the proportion of a population impoverished by health payments. The other is the proportion incurring so-called “catastrophic payments”, which are defined as spending on healthcare that is large as a share of the household budget or some measure of capacity to pay (World Health Organization and The World Bank, 2015, 2017). One variant of the latter has been adopted by the United Nations as a UHC target indicator (3.8.2) for the Sustainable Development Goals (United Nations, 2017). It was chosen only after lobbying persuaded the responsible UN committee that its proposed use of the nominal population health insurance coverage rate would provide an uninformative measure of effective financial protection (McIntyre et al., 2016). This episode is indicative of the lack of agreement on what financial protection is and how to measure it.

Judged by frequency of use, the impoverishment and catastrophic payments measures have been hugely successful.¹ This is probably due to intuitive appeal and ease of application. Everyone can understand an indicator that is claimed to reveal the proportion of individuals made poor because they had to pay for healthcare. The proportion who spent more than, say, a tenth of the household budget on healthcare is also easily understood. This comprehensibility makes these potent indicators for the purpose of bringing health financing to the attention of policy makers. Each measure can be calculated easily from data on only two variables – total household expenditure and spending on healthcare. However, these undoubted advantages are also potential sources of weakness. The indicators do not capture what they are frequently claimed to measure. The impoverishment metric identifies individuals made poor because of payments for healthcare only under a highly unrealistic assumption about the standard of living a household would have reached if it had not incurred such payments. A payment labelled “catastrophic” may have no such impact on wellbeing if it is financed by instruments, such as savings, credit and transfers from friends and relatives, that make it possible to cover healthcare costs while maintaining consumption of other goods.

Notwithstanding these limitations, the minimal data requirements of the impoverishment and catastrophic payments indicators make them potentially useful proxies for conceptually stronger measures of (lack of) financial protection against medical expense. However, this proxy status is not always recognised. Preoccupation with the indicators risks diverting attention from the efficiency and equity of health financing that are the primary concerns. Further, the dependence of the indicators on general economic conditions, most specifically the rate of poverty for the impoverishment measure, means that they are likely to emit a rather dull signal of progress toward UHC through sufficient, efficient and fair financing of healthcare.

At the country level, there is also appreciation of the need for measures of financial protection against medical expense without consensus emerging on what they should be. In the US, the sizeable uninsured population prior to the 2010 Patient Protection and Affordable Care Act (PPACA), the large shares of incomes spent on health insurance premiums, which grew immediately after the landmark reform (Kowalski, 2014), and the substantial fraction of medical

¹ A google scholar search for articles with titles including the words health or medical in combination with a) catastrophic payment, b) catastrophic expenditure, c) financial catastrophe or d) impoverishment turns up 305 articles (duplicates excluded). Searched 20.8.2018.

costs that are paid out-of-pocket through coinsurance have ensured that the strain health payments place on household finances continues to attract the attention of the media, politicians and academics. Two National Research Council panels have proposed a poverty measure that takes account of the burden of medical expenditures and called for another metric to capture exposure to economic risk arising from future medical care needs (National Research Council, 1995; National Research Council and Institute of Medicine, 2012). These suggestions bear similarities to the impoverishment and catastrophic payments measures used for UHC monitoring, although important details and interpretations differ. However, evaluations of the PPACA (Brevoort et al., 2018; Hu et al., 2018), as well its precursor in Massachusetts (Mazumder and Miller, 2016) and of expansions of social health insurance for the elderly (Medicare) (Finkelstein and McKnight, 2008; Engelhardt and Gruber, 2011) and the poor (Medicaid) (Gross and Notowidigdo, 2011; Finkelstein et al., 2012; Finkelstein et al., 2018), have not relied on these indicators to measure the extent to which coverage extensions have improved protection of financial wellbeing against the cost of medical care. Instead, they have estimated the welfare gain from facing reduced medical expenditure risk. This is much more strongly grounded in theory than the proxy indicators of financial protection. However, it rests on potentially restrictive assumptions about risk perceptions and preferences.

In the next section, I consider alternative definitions of financial protection as an objective of health policy. I propose that it is the protection of opportunities for nonmedical consumption against the burden and risk of medical expense using formal health insurance provided through markets, social insurance agencies and tax-financed public provision, as well as informal means of pooling risks that do not operate through market and government institutions, and self insurance instruments that allow households to spread the costs of medical care through time without sacrificing long-term potential to earn income. I then assess consistency with this definition of four measures of financial protection: a) the extent to which consumption is smoothed over health shocks; b) the risk premium – willingness to pay (WTP) in excess of a fair premium – to insure currently uninsured medical expenses; c) catastrophic healthcare payments; and, d) impoverishing healthcare payments. Each measure suffers from theoretical and/or practical weaknesses. The first does not restrict attention to medical expenses, which limits its relevance to health financing policy. The second rests either on difficult-to-verify assumptions about risk preferences or on hypothetical statements of WTP. None of them treats medical expenses financed through informal

and self insurance instruments in an entirely satisfactory way. By ignoring these sources of insurance, which can be far from optimal but do not correspond to no insurance, the catastrophic payments measure overstates the impact of medical expenses on living standards, while the impoverishment measure does not credibly identify poverty caused by them. It is better thought of as a correction to the measurement of poverty. In the final section, I discuss the health policy usefulness of the different metrics.

What is Financial Protection?

Financial protection is frequently mentioned in health policy, particularly global health policy, documents and discussions. Sometimes reference is made to “financial risk protection” or the even vaguer “financial protection in health” (Waters et al., 2004; Wagstaff, 2010; Moreno-Serra et al., 2011). Attention is usually focused on the lack of financial protection. Roughly, this is understood as the degree to which financial wellbeing is jeopardized by healthcare needs. There is no agreement, however, on the precise meaning. Some want to capture unaffordable and so forgone medical treatments. Others deem this beyond the scope of the financial impact of using healthcare. Some would extend the financial impact to include that running through the effect of illness or disability on earnings. Others judge this to lie outside the remit of health financing policy that is concerned with paying for healthcare. Some restrict attention to out-of-pocket (OOP) payments for healthcare and are troubled by reliance on them irrespective of how they are financed. Others view savings and remittances as efficient means of financing spending on healthcare that protect consumption of other goods, and are concerned about formal health insurance crowding out these forms of self and informal insurance.

With the aim of clarifying the meaning of financial protection and, consequently, improving its measurement and interpretation, I address three questions. Protection of what? Protection against what? Protection with what? Not everyone will agree with my answers. I hope that posing the questions is nonetheless useful in distinguishing the meanings that can be given to the concept.

Protection of What?

Health

It is often argued that the measures of financial protection used by the WHO and the World Bank for UHC monitoring are deficient because they do not capture consequences of forgone needed but unaffordable healthcare (Moreno-Serra et al., 2011; Moreno-Serra et al., 2013; McPake, 2018). While these consequences are undoubtedly of major concern, they manifest in the deterioration of health, not (directly) in the worsening of household finances. Health suffers as a result of the sick going without effective treatment. The prospect of not being able to afford needed medical care is a health risk, not a financial risk. It should be captured by a measure of access to healthcare, not by a measure intended inform of the extent to which payments for healthcare strain household finances (Wagstaff, 2010; National Research Council and Institute of Medicine, 2012).

Unaffordable healthcare will be reflected in the level and distribution of health payments. Measurement of access, including affordability, is therefore an essential complement to measurement of financial protection. Restricting attention to the latter would lead to highly misleading conclusions from comparison of health financing systems and evaluation of coverage extensions.

Consider a hypothetical example in which all treatments are lumpy. You get treatment, or you do not. There are no half measures. Assume the poor cannot afford any treatment at market prices. Their welfare is not threatened by medical expenses. Of course, they are highly vulnerable to sickness. Partially subsidised healthcare that makes at least some treatments affordable to the poor is introduced. The poor now make use of healthcare and pay for it. Clearly, the subsidy has not reduced the welfare of the poor. Quite the opposite. It has given them options they did not have previously. The improved access to affordable healthcare reduces the risk to their health. But it has also raised their exposure to the risk that medical expenses eat into the resources they have available to spend on other goods. They have gone from facing a risk of getting sick, not being able to afford treatment and so experiencing a deterioration in health, to facing a risk of getting sick and incurring medical costs that will leave them with less to spend on food, housing, clothing or whatever.

This example is obviously contrived.² While there are some real world examples of the extension of subsidised health insurance coverage apparently raising OOP payments and worsening measures of financial protection (Wagstaff and Lindelow, 2008; Bernal et al., 2017), increased coverage would be expected to reduce both the health risk of forgone unaffordable care and the financial risk of paying for care that is affordable, but only marginally so. Ideally, any evaluation of the gains from the extension of insurance coverage captures both the health benefits from improved access to otherwise unaffordable healthcare and the consumption benefits from reduced exposure to previously affordable, but nonetheless onerous, OOP medical expenses. These are two distinct welfare gains from health insurance. Without placing a monetary value on the health benefit from improved access, it is not possible to capture both effects in one measure. Each can be measured separately. Neither should be ignored.³

The dominant theory of the demand for health insurance ignores the access motive (Zeckhauser, 1970; Cutler and Zeckhauser, 2000). It views insurance solely as an instrument for reducing medical expenditure risk. Whether or not the person is insured, they will incur medical expenses should they fall sick. The purchase of insurance is motivated by the reduction or elimination of financial risk through partial or complete reimbursement of those expenses. The welfare gain from insurance comes from risk averse, expected utility maximizing individuals facing less variability in potential medical expenses. John Nyman (1999b, a, 2003) challenges this view that health insurance is demanded mainly to reduce medical expenditure risk. He contends that the primary motivation is to facilitate the transfer of resources to the state – sickness – in which they are most highly valued. Curative medical treatment has no utility when you are healthy. But it is highly valued when you are sick. It is critical to ensure access to sufficient resources when sick to allow spending a great deal, perhaps well in excess of current income, on medical care in

² Lumpy medical technology is not the only reason a subsidy could raise the health spending of those subsidised. Access to fully subsidised care could alert them to health conditions and unsubsidised effective treatments of which they were previously ignorant (Bernal et al., 2017). The supply-side response to subsidised care may also bring health facilities in closer proximity and reduce the indirect costs of access (Finkelstein, 2007).

³ There is evidence of positive effects of extended coverage on measures of financial protection in a variety of contexts (Finkelstein and McKnight, 2008; Engelhardt and Gruber, 2011; Limwattananon et al., 2015; Barnes et al., 2017). See also the US studies of the PPACA and Medicare/Medicaid expansions cited in section 1. Evidence of health gains from extended coverage is rarer (Gruber et al., 2014; Cesur et al., 2017) but is obtained from some studies that also find positive effects on financial protection (Finkelstein et al., 2012; Miller et al., 2013; Shigeioka, 2014). Many studies estimate positive effects on healthcare utilisation but are not able to establish whether this is due to access to previously unaffordable, effective treatment or is moral hazard. Coverage of the poor in Peru is estimated to have reduced their likelihood of being unable to afford treatment when sick by a quarter (Neelsen and O'Donnell, 2017).

that state. This is done by entering a contract in which the healthy agree to transfer resources to the sick. According to this perspective, the welfare gain from insurance arises from health benefits obtained through access to medical care that would not have been affordable without insurance. Using the Medical Expenditure Panel Survey (MEPS), Nyman and Helen Trenz (2015) estimate that one third of the medical spending on insured Americans would have been unaffordable, in the sense of being beyond their disposable incomes and assets, if they had not been insured.

If each of these theories is interpreted as representing *the* motivation for health insurance, then they are conflicting. Both cannot capture what solely motivates each person to demand insurance or to potentially benefit from it. But if each theory is interpreted as representing *a* motivation for health insurance, then they are not mutually exclusive. Insurance can be valued both to protect against the threat to nonmedical consumption arising from risky medical expenses and to ensure sufficient resources are available to afford medical care that is needed in some sickness states. The relative importance of the two motivations may vary from individual to individual. Each must be measured in order to quantify the total value of health insurance.

There is an important caveat to the argument that financial and health risks are distinct. Forgoing needed but unaffordable healthcare impacts on household finances through the deterioration of labour productivity. The financial risk includes an earnings risk in addition to the medical expenditure risk. However, as I will argue below, if a measure of financial protection is to be used to inform health financing policy, which is concerned with the financing of healthcare, then it should capture only the medical expenditure risk and not the earnings risk.

Wealth

Some argue that the use of savings and the divesting of assets to pay for healthcare is indicative of a lack of financial protection (Ruger, 2012; Moreno-Serra et al., 2013). Admittedly, the “financial protection” label gives the impression that finances are the focus of concern. But this argument overlooks the precautionary motive for holding savings and other assets (Leland, 1968; Kimball, 1990). Acquisition of assets during periods of health and their disposal in times of sickness facilitates at least partial smoothing of nonmedical consumption across these states. A household with assets has greater financial protection against medical expenditure risk than one that is without assets. Liquid assets provide most protection, while productive assets that are relied on to maintain long-term earnings potential provide the least. Ideally, a measure of financial protection

identifies a household that can cover medical expenses by divesting liquid assets as being better protected than one that is forced to sell productive assets, and still better protected than another that is without any assets and must immediately cut back on its current nonmedical consumption to pay for healthcare.

Wealth holdings will often be less efficient than formal health insurance in protecting economic wellbeing against the threat of medical expenses. The cost of some treatments will exceed the wealth that many households are able to accumulate and will be afforded, if at all, only by cutting back on current consumption. Many risks cannot be spread across periods by a single household as efficiently as they can be pooled across households through insurance. However, this is not always the case. If risks are not serially correlated over time, then it is more efficient to cover them, provided they are not too large, through wealth rather than insurance that is offered at a less than actuarially fair price (Gollier, 2003). I will return to this later when discussing instruments of financial protection.

Consumption

I contend that financial protection is not about the protection of health because that is a distinct domain of wellbeing that is the ultimate target of measures of access to healthcare. Neither should financial protection be about protection of wealth because that only has instrumental value. Rather, measures of financial protection ought to provide information on the extent to which opportunities for nonmedical consumption are protected from healthcare costs.

Suppose wellbeing is generated by health and consumption of nonmedical goods. Maintaining health through the utilisation of medical care inevitably involves someone, at some time, sacrificing opportunities to consume other goods. This is true whether medical care is paid for out-of-pocket or financed by private insurance, social insurance or taxation. The issue is how the health financing burden is distributed across individuals and over periods. Financial protection is about efficiency and equity in the distribution of this burden.

To achieve efficiency, the burden should not fall exclusively on users of healthcare at the time of sickness. Otherwise, a health shock will be accompanied by a consumption shock. The risk averse are willing to pay a price in excess of the fair price, i.e. their expected medical expenses, for insurance that removes the uncertainty over those expenses and so their nonmedical

consumption. This risk premium is one possible measure of the lack of financial protection and the potential gain from insurance.

To achieve equity, many would prefer to avoid the financing burden falling on those with limited ability to pay. Otherwise, paying for medical care could reduce consumption to the poverty threshold or below. If there is intolerance of poverty, then there will be concern about poverty caused or exacerbated by payments for healthcare. This is the perspective that underpins the impoverishment measure of financial protection that is intended to enumerate the number of individuals made poor or poorer through payments for healthcare (Wagstaff and Van Doorslaer, 2003). Another equity position, which has been proposed as an ethical foundation for the catastrophic payments measure, is that it is unfair for any household, whatever its standard of living, to exhaust a large fraction of its resources paying for healthcare (Wagstaff and Van Doorslaer, 2003). But it is also possible to find an efficiency motivation for this measure. Keeping health spending below a certain fraction of the household budget would ensure that it did not disrupt consumption of other goods to an excessive degree. This is the target of efficient smoothing of consumption over health shocks.

Protection Against What?

Illness impacts on consumption opportunities through earnings loss as well as medical expense. It is sometimes claimed that measures of financial protection are deficient because they do not capture the first of these two effects (Moreno-Serra et al., 2013). In fact, the catastrophic payments metric, which is a function of the ratio of OOP medical expenses to income (or total expenditure or another measure of capacity to pay), will rise in response to an illness-induced fall in income. However, this seems a deficiency of the measure, since changes in payments for healthcare are potentially confounded by changes in income.

Protection of earnings from ill-health is the objective of sickness and disability insurance, not health insurance or health financing more generally. If a measure is to inform a policy domain, then it should be sensitive to the instruments controlled within that domain (Wagstaff, 2010). Health financing policy levers operate on payments for medical care, not earnings. To ensure financial protection metrics are relevant to those charged with designing health financing policy, they should be restricted to summarising changes in the distribution of medical expenses. From this policy perspective, it is an advantage that a measure is unresponsive to the relationship

between health and earnings. Different metrics are required for the evaluation of sickness and disability insurance policies.

Protection can be sought against both the risk and the expected burden of medical expenses. The former provides the motivation to insure according to the standard theory of insurance (Zeckhauser, 1970; Cutler and Zeckhauser, 2000). The risk averse suffer anxiety from exposure to volatile future medical expenses. Insurance raises their wellbeing. Financial protection metrics are needed to quantify the potential gains from insurance. However, such measures may be little use in capturing the predicament of the chronically ill with predictable needs for care and medication. Subsidization of the high expected medical expenses of these individuals can be motivated by altruism, a sense of justice and by the containment of adverse selection in insurance markets. It can also be seen as insurance from a lifetime perspective, i.e. coverage of the risk of contracting chronic disease that would raise the premium in an unregulated insurance market.

Financial protection is about equitable and efficient shielding of nonmedical consumption opportunities from expected and risky medical expenses. How can this protection be delivered?

Protection with What?

A variety of instruments may be relied on to shelter consumption opportunities from the cost of medical care. A household may be insured and it may benefit from access to subsidised medicine. Insurance can be formal, informal and self-provided through saving. Formal insurance can be voluntarily purchased in a market or mandated by government. Public insurance, which can be defined to include public provision of subsidised care, can be financed through social insurance contributions or taxation. Tax financed care for the chronically ill and frail elderly is a way of insuring the population against contracting such illness and living to an old age at which nursing care is often required. Informal insurance pools risk across households without formal contracts or state enforced obligations. Transfers from relatives, friends and community support networks are relied on in times of medical need. Self insurance spreads risk over periods rather than pooling across households. Savings, possibly in the form of a Medical Savings Account (Pauly and Goodman, 1995), and other assets are accumulated in readiness for future medical expenses. Finally, a household may borrow to cover healthcare costs.

Which of these instruments provide financial protection? The answer partly determines how financial protection should be measured. If protection is considered to be delivered only

through formal insurance and public subsidies, then all OOP payments are of interest, irrespective of how households raise the money to make these payments. If informal and self insurance are seen as affording financial protection, then we must net out the OOP payments paid for using these coping strategies. If consumption that is sacrificed to pay a high insurance premium is considered to be indicative of a lack of financial protection, then the metric needs to go beyond OOPs to include premiums.

Informal and self insurance do indeed provide financial protection, at least to some extent. Consumption is better protected when there are support networks to call on and savings to draw on to help pay for healthcare. In a first best world of complete and perfect capital and insurance markets, insurance would be the preferred instrument to transfer resources between states of nature, e.g. healthy and sick, while saving and credit would be used only to transfer consumption opportunities between periods (Besley, 1995). In the real world, costs of administering insurance raise its price above the expected value of the loss covered. Saving can then be preferred to insurance at a less than actuarially fair premium to cover serially uncorrelated risks that are not too large as share of wealth. The standard single-period expected utility model of insurance predicts that a risk averse individual will not fully insure at an actuarially unfair premium (Mossin, 1968). Christian Gollier (2003) demonstrates, by means of simulation, that in a multi-period model, the option to spread risk across periods by using wealth to absorb the impact on consumption of a loss – time diversification – reduces the demand for actuarially unfair insurance further. A liquidity constraint – a limit on the amount that can be borrowed against lifetime expected income – gives an incentive to hold a buffer stock of wealth as a precaution to protect future consumption against risk (Deaton, 1991). If the household can accumulate a sufficient buffer stock, it will prefer to rely on this rather than purchase insurance at a cost in excess of the expected payout. This substitution of self insurance for formal insurance is constrained by existence of the liquidity constraint. At low wealth, there is a greater likelihood of being liquidity constrained should illness strike and so being forced to cut consumption to cover medical expenses. At high wealth, this is less of a risk. Hence, the amount of formal insurance demanded is predicted to fall as wealth rises. At sufficiently low wealth, there is complete reliance on formal insurance. At sufficiently high wealth, dependence switches entirely to self insurance.

This analysis implies that, in a second-best world, financing healthcare from savings, or the sale of other assets, is not necessarily indicative of a lack of financial protection. It may be the

preferred way of covering medical expenditure risks for some households when, as is usually the case, formal insurance cannot be offered at an actuarially fair price. The model suggests that households that are not wealth poor would prefer to rely on savings to cover expenses that are not too large a proportion of their wealth and to formally insure only risks of needing treatments that are sufficiently expensive to threaten exhausting a large fraction of household wealth. This is the logic that underpins proposals for Medical Savings Accounts combined with catastrophic insurance (Pauly and Goodman, 1995). Obviously, this does not offer viable protection to poor households that have little or no potential to accumulate wealth. It is also not a feasible plan for covering the risk of managing chronic health problems. The model assumes serially uncorrelated risks. Once this is relaxed, wealth becomes a less effective instrument of insurance. It would be rapidly depleted.

Private health insurance also protects against serially uncorrelated risks better than it does against correlated risks. Single-period insurance leaves individuals exposed to the risk of contracting a chronic illness that will result in a hike in the premium or denial of insurance cover (Hendren, 2013). While the market could possibly be regulated to ensure cover is offered to those with pre-existing conditions at better than fair premiums or it may deliver products that cover the premium risk (Cochrane, 1995; Pauly et al., 1995), public insurance may well be more efficient in providing this cover. Coverage of the costs of treating a condition or a sequence of conditions over years could generate a large part of the gains from social insurance. Using 17 years of longitudinal data on all hospital visits in New York, Amanda Kowalski (2015) estimates that 80 percent of the value of the public insurance safety net to young, initially privately insured individuals is missed if the gain from transiting from private to public insurance after the onset of a persistent health condition is ignored. There are two implications for the measurement of financial protection. First, longitudinal data are highly desirable. They offer the opportunity to estimate exposure to serially correlated medical expenses, which is a common characteristic of healthcare costs. Unfortunately, such data are rare, although increasing access to administrative data on hospital admissions and insurance claims is making them less so. Second, in private insurance systems, there is little rationale for confining attention to OOP payments, even from a risk perspective. A large part of the lack of financial protection in such systems is likely to take the form of premium risk. Public insurance removes this risk by linking contributions to earnings or income, rather than expected medical expenses.

There are, however, also limits to the financial protection provided by social health insurance and publicly financed healthcare. Effective coverage can fall well short of nominal coverage. Long waiting times, low quality care or simply unstaffed facilities can induce the nominally insured to pay out-of-pocket for treatment. This is the reason many lobbied the UN to rethink its initial proposal to measure financial protection by simply counting the proportion of a population covered by some form of prepayment for healthcare (McIntyre et al., 2016). An informative measure of financial protection tells us about the extent to which the consumption of a household is shielded from the cost of the healthcare it uses and may use.

I have argued that financial protection involves the use of formal, informal and self insurance, as well as public subsidies, to shield nonmedical consumption from the cost of medical care. Of course, consumption opportunities must be sacrificed to provide medical care. The purpose of instruments of financial protection is to distribute the costs efficiently and equitably across households and over periods. Evaluation of success in achieving this objective requires measurement. In the next section, I assess the information content of the main measures of financial protection.

Measurement of Financial Protection

Consumption Smoothing Over Health Shocks

If there were complete financial protection against medical expense, then household consumption of nonmedical goods and services would be maintained during periods of sickness that result in utilisation of healthcare. Perfectly smoothed consumption over health shocks appears to leave no potential for welfare to be raised through the extension of formal insurance. Doing so would crowd out informal coping strategies that protect consumption from the threat of illness. On the other hand, if consumption drops when illness strikes while preferences remain stable, then economic risks arising from ill-health are not fully pooled, leaving scope for welfare gain from the extension of formal (social) insurance.⁴ Much rests on how sensitive consumption is to ill-health.

Full smoothing of consumption over health shocks is a prediction of the theory of perfect risk sharing (Wilson, 1968). A collection of risk averse households will achieve a Pareto optimum

⁴ The assumption of stable preferences is important. If ill-health were to reduce the marginal utility of consumption, as some evidence suggests (Finkelstein et al., 2013), then even with full insurance, consumption would fall during periods of illness.

by pooling its resources each period and sharing them according to predefined rules that ensure the relative consumption enjoyed by each household is consistent, for example, with its wealth endowment, but is independent of stochastic income and losses, such as medical costs. This is the optimum that would be achieved by a benevolent social planner who maximizes the weighted average of utilities subject to the resource constraint of the economy. The first order conditions for this problem imply that the growth in the marginal utility of a household's consumption depends only on the growth in aggregate consumption. The consumption of each household will be insensitive to any idiosyncratic shock it experiences and will respond only to aggregate shocks that impact on all households simultaneously. Household-specific risks are fully insured. An Arrow-Debreu world of complete and perfect markets trading state-contingent commodities would be sufficient to achieve this optimum. More realistically, it might be obtained through a combination of private, social and informal insurance. Risks may be pooled through market transaction, government mandate, and norms, obligations and altruism arising from kinship, friendship and community values.

John Cochrane (1991) derives the most simple, and yet general, test of the full insurance hypothesis. Under the null, a cross-section regression of consumption growth on a variable representing an idiosyncratic shock, such as illness, will give a zero coefficient. Provided the shock variable is exogenous to any shift in preferences (possibly a strong assumption in the case of illness) and is uncorrelated with measurement error in consumption, complete pooling of risks over households through any combination of markets, institutions and social networks ensures that the consumption growth of a household is unaffected by occurrence of a negative outcome. Using data on food consumption from the US Panel Study of Income Dynamics (PSID) in the early 1980s, Cochrane does not reject the hypothesis for short duration illnesses (<100 days), but does for longer illnesses. This implies that all forms of risk pooling were insufficient to fully insulate even food consumption, which would be expected to be the least sensitive category of consumption to idiosyncratic shocks, from the more persistent, and presumably more serious, illnesses in the US at the time. This has been confirmed by subsequent analyses of the PSID. Using data from 1968 to 1992, Melvin Stephens (2001) finds that in the year of onset of a disability there is a non-significant 1% fall in food consumption but after six or more years this has increased to a significant 5.7% decline. Extending the length of the post-shock period further, Bruce Meyer and Wallace Mok (2018) estimate even larger effects. Food consumption falls by 2.5% in the year of

disability onset and by 7.6% after ten years. For chronic and severe disabilities, the effects are even larger: 8.9% in the year of onset and 18.2% after ten years.⁵ The increase in the effect with time since the onset of illness or disability suggests that initially households are able to partially protect their consumption of food but they are less insured over a longer horizon. In fact, the theory of perfect risk sharing does not extend to serially correlated health shocks or those that permanently reduce income net of medical expenses. Using a life-cycle model with incomplete insurance markets, Mariacristina Di Nardi, Svetlana Pashchenko and Ponpoje Porapakkarm (2017) demonstrate that the strong dynamics of health, characterised by persistence, can generate the steep income-health and wealth-health gradients observed in US data. Without insurance that extends over many periods, consumption cannot be held constant when health deteriorates irreparably.

Consumption smoothing over health shocks is more often tested in low- and middle-income countries (LMIC). In this context, the purpose is usually to assess whether informal risk sharing networks are sufficient to provide full insurance in the absence of formal contracts. If not, there is a potential welfare gain through the provision of formal social insurance. If informal arrangements are sufficient to protect consumption from shocks, then formal insurance would crowd them out and may not raise welfare. The evidence is mixed (Alam and Mahal, 2014).⁶ Using longitudinal data from three villages in southern India, Robert Townsend (1994) does not reject full insurance of consumption with respect to reported days of sickness. Abay Asfaw and Joachim Von Braun (2004) find full smoothing of total food consumption (home-produced and purchased) over illness of the head of household in rural Ethiopia but drops in expenditure on food and in non-food consumption. Using the same dataset and the body mass index (BMI) as a measure of nutritional consumption at the individual level, Stefan Dercon and Pramila Krishnan (2000) reject full risk sharing even *within* some (but not most) households: the BMI of poorer, southern Ethiopian women (but not men) falls as a result of sickness.

Illness is highly heterogeneous. Ability to hold consumption constant in the face of minor illness does not imply full insurance against major illness. Some studies reject complete smoothing

⁵ Meyer and Mok (2018) estimate that disability has even larger effects on food plus housing consumption.

⁶ Some studies rely on cross-sectional measures of reported health and so can say little about the consumption response to health changes (Skoufias and Quisumbing, 2005; Islam and Maitra, 2012; Sparrow et al., 2014; Wagstaff and Lindelow, 2014). Others obtain a noisy measure of health changes by taking the difference between reported illness in the last four weeks and the respective indicator one, two or even four years previously (Dercon and Krishnan, 2000; Asfaw and von Braun, 2004; Liu, 2016).

of consumption over more serious health events (Gertler and Gruber, 2002; Wagstaff, 2007; Gertler et al., 2009). A limitation of these studies is that households experiencing a decline in consumption may be more prone to illness. Although there may be an offsetting bias through (classical) measurement error in the health indicator. Two studies with designs capable of dealing with such endogeneity do not reject full consumption smoothing (Genoni, 2012; Mohanan, 2013). Of particular note is Manoj Mohanan's (2013) study of the victims of bus accidents in Karnataka (India), which finds that households are able to maintain expenditures on food and housing while incurring large medical expenses resulting from this plausibly exogenous health shock. However, this is achieved not by risk sharing across households but through the accumulation of debt, leaving the possibility of a negative long-term impact on consumption. This points to a general limitation on the extent to which the optimality of financial protection can be assessed from a test of consumption smoothing over health (and other idiosyncratic) shocks. Consumption may be held relatively stable through efficient pooling of risks across households. But it could also result from reliance on less efficient coping strategies.

In the absence of formal insurance, the monetary equivalent of the welfare gain from insuring a loss is proportional to the relative difference in the marginal utilities of consumption between the two states of the world (e.g. healthy and sick), which is approximately equal to the relative drop in consumption multiplied by the coefficient of relative risk aversion (RRA) (Chetty, 2008).⁷ Equivalently, the welfare loss from being underinsured is proportionate to the variability in consumption scaled by the misery caused by that volatility, which is reflected in the curvature of the utility function – the sole determinant of risk aversion in this expected utility model. Raj Chetty and Adam Looney (2006) point out that even if there is relatively little variation in consumption, there can still be a large welfare loss under current arrangements and the potential for substantial gain from the provision of formal insurance if risk aversion is high, which is a

⁷ Assume utility is a function of nonmedical consumption alone and there is no state dependence, $u(c)$ with $u'(c) \geq 0$, $u''(c) \leq 0$. Let c_h be consumption in the healthy state and $c_s = c_h - m$ be consumption when sick, where $m \geq 0$ is uninsured (by any means) medical expenses. Then the potential welfare gain from insuring those expenses relative to the gain from increased consumption in the healthy state is $W \propto \frac{u'(c_s) - u'(c_h)}{u'(c_h)}$. Taking a Taylor series expansion, this is approximated by $W \approx RRA \frac{\Delta c}{c}$, where $\frac{\Delta c}{c} = \frac{c_h - c_s}{c_h}$ and $RRA = -\frac{u''(c)}{u'(c)} c_h$ is the Arrow-Pratt coefficient of relative risk aversion (Chetty and Looney 2006).

characteristic of the poor under standard assumptions. The importance of this observation is magnified by the further realisation that the change in consumption and the degree of risk aversion are inversely related. The more risk averse is an individual, the more they are willing to resort to inefficient coping strategies, such as holding liquid but unproductive assets, that help protect consumption from bad outcomes, such as illness. The effort the risk averse expend on inefficient ways of protecting their current consumption at the cost of their long-run earnings potential is part of the welfare loss from the lack of formal insurance (Chetty and Looney, 2006).

This does not take us back to equating informal insurance and self insurance with no insurance. Transfers and remittances from relatives and friends can be effective in pooling many medical expenditure risks, if not the most extreme ones. As argued in the previous section, saving instruments and borrowing on reasonable terms can also be preferred to actuarially unfair insurance by middle- and high-income households with opportunities to accumulate wealth and to borrow against future earnings. Expansion of social health insurance is likely to crowd-out these forms of insurance, in addition to any private health insurance (Cutler and Gruber, 1996), which will lower the net increase in financial protection and the welfare gain generated. This cost should not be ignored given the likely high opportunity cost of the funds potentially available to provide social health insurance in LMIC. But if a household is divesting productive assets, possibly including the human capital of children withdrawn from school, to make it possible to purchase needed healthcare while maintaining nonmedical consumption at close to subsistence levels, then there will be a large potential welfare gain from insurance that will be undetected by a test of consumption smoothing. Recognising this, the literature is beginning to shift focus from testing consumption smoothing per se to examining how consumption is kept smooth in the face of health shocks.

There is evidence that Vietnamese households exposed to ill-health and hospitalization of a family member are able to hold their nonmedical expenditure constant, but they spend less on education (Mitra et al., 2016). Kai Liu (2016) finds a similar result in Chinese households without health insurance. Those with insurance are able to smooth their consumption over health shocks without sacrificing their children's education. Improved opportunities to self insure can also reduce the need to resort to inefficient coping strategies. There is evidence that microfinance programs that allow households to borrow on better terms have increased the ability of Indonesian

households to maintain their consumption during periods of ill-health (Gertler et al., 2009) and have reduced the sale of livestock, a common strategy to shield consumption when illness strikes in rural Bangladesh (Islam and Maitra, 2012).

In addition to ignoring the (in)efficiency of the financing instruments deployed to smooth consumption, a further limitation of a test of full insurance over health shocks is that it is a joint test of ability to informally insure against both medical expenses and earnings losses arising from illness. Consumption volatility arising from sickness-related disruptions to earnings may be as large as that due to sacrifices made to pay for healthcare. But only the latter is relevant to the potential welfare gains from the extension of health insurance. This medical expenses source of risk cannot be isolated using an empirical strategy that involves regressing the change in consumption on the change in health. While separate regressions can be run with earnings and medical expenditures as dependent variables, and this has sometimes revealed that loss of earnings is the greatest health-related economic risk (Di Nardi et al., 2017) even in environments with little or no formal health insurance (Gertler and Gruber, 2002), this does not tell us anything specific about informal insurance of medical expense specifically.⁸ A metric that informs of exposure to medical expenditure risk exclusively is required.

Welfare Gain From Health Insurance

The consumption smoothing literature aims to test the hypothesis of full insurance of health shocks. It does not provide a measure of financial protection, or the lack of it. The standard measure used in economics to capture the welfare loss from exposure to risk, and hence the potential gain from insurance that eliminates that risk, is the risk premium. This is the amount an individual is willing to pay (WTP) in excess of an actuarially fair insurance premium to remove the risk. Assume utility is derived from nonmedical consumption, which is equal to income (y), assumed fixed, net of stochastic medical expenses (M). Under expected utility (EU), the risk premium (π)

⁸ Sven Neelsen and colleagues (Neelsen et al., 2018) find that Thai households with no earnings insurance but with universal health insurance coverage, are able to use a combination of informal insurance, savings and credit to hold nonmedical expenditure constant when hit by a health shock that reduces earnings by a third and raises OOP payments by two-thirds due to gaps in coverage.

is defined implicitly by $u(y - E[M] - \pi) = EU(y - M)$, where E is the expectations operator.⁹ It is increasing with the variance of medical expenses and the degree of risk aversion.¹⁰ It captures the individual's exposure to risk and their dislike of risk, where risk is defined as volatility.

To estimate risk premiums, one needs to simulate the distribution of medical expenditure faced by each household, find the corresponding distribution of consumption by estimating or assuming the fraction of medical expenditure at each point in its distribution that is afforded by sacrificing spending on other goods as opposed to being financed through informal and self insurance strategies, and make assumptions about the nature of the utility function and the magnitudes of its parameters. Uncertainty about the validity of the estimates and assumptions used at each of these steps is likely to be considerable. Estimates of risk premiums should therefore be interpreted cautiously. Nonetheless, by providing a monetary value of the welfare gain from contemplated or implemented insurance that can be compared with its cost, the risk premium is a powerful tool for policy analysis.

Cost-benefit analyses of extensions to social health insurance in the US (Finkelstein and McKnight, 2008; Engelhardt and Gruber, 2011; Finkelstein et al., 2012; Mazumder and Miller, 2016; Finkelstein et al., 2018; Hu et al., 2018) and elsewhere (Shigeioka, 2014; Limwattananon et al., 2015) tend to find welfare gains through reduced exposure to medical expenditure risk that are large relative to efficiency costs arising from paying for the insurance. For example, Amy Finkelstein and Robin McKnight (2008) estimate that the introduction of Medicare – social health insurance for the elderly in the US – generated financial protection benefits that alone reached two fifths of the deadweight efficiency cost of the programme. This omits the value of any health benefits arising from improved healthcare access and affordability. Using the Oregon Health Insurance Experiment, Finkelstein, Nathaniel Hendren and Erzo Luttmer (2018) estimate the insurance value of Medicaid – social health insurance for US low income households – at 52-77%

⁹ $EU(y - M) = \int_0^y u(y - m) dF_M$, where F_M is the cumulative distribution of medical expenses. Upper case indicates a random variable and lower case a realization of it. The assumption that spending on medical care is bounded above by income could be relaxed in a multi-period model that allows for financing from accumulated assets and credit.

¹⁰ Taking Taylor series expansions of the left- and right-hand-side of the expressions that define the risk premium, gives

$$\pi \approx -\frac{1}{2} \frac{u''}{u'} \text{Var}(M) = \frac{1}{2} \text{ARA} \times \text{Var}(M)$$

where $\text{Var}(M)$ is the variance of medical expenses and ARA is the Arrow-

Pratt measure of absolute risk aversion. Note the similarity to the expression for the welfare loss of being unable to smooth consumption over a binary loss given in footnote 6.

of the societal cost of the programme.¹¹ However, one method and dataset gives a much lower estimate of 9%; an indication of how sensitive these estimates can be. The universal coverage scheme that extended health insurance to the whole population of Thailand in 2002 is estimated to have generated a welfare gain from reduced exposure to medical expenditure risk equal to 84-200% of the deadweight loss of the taxes levied to finance the landmark reform (Limwattananon et al., 2015).

In addition to its usefulness for evaluating implemented policies, the risk premium can be used to estimate potential welfare gains. For example, it has been used to argue that there is sufficient potential demand to make the provision of private health insurance viable in many LMIC (Pauly et al., 2009). An estimate of the value of full insurance that entirely eliminates medical expenditure risk is of limited relevance where a large informal economy constrains the extent to which healthcare can be financed through pre-payment or moral hazard renders full insurance inefficient. In those contexts, the risk premium for major risk insurance is more informative. This is the difference between the risk premium in the absence of any insurance and the excess over an actuarially fair premium the individual would be willing to pay to remove the residual risk remaining after covering major risks.¹² It gives the welfare gain from eliminating exposure to major risks. Martin Feldstein and Jon Gruber (1995) demonstrate potential welfare gains in the US from introducing major risk health insurance comprising 50% coinsurance combined with a ceiling on OOP payments at 10% of income.

Studies that use the risk premium to estimate the (potential) welfare gains from formal insurance typically ignore its crowding out of informal and self insurance. The impact of formal insurance on the distribution of OOP medical expenses is estimated. To get to the effect on the distribution of nonmedical consumption, it is usually assumed that \$1 more spent OOP on

¹¹ These estimates, which are derived from the results given in Table 3 of the paper, are obtained using the cost net of programme-financed transfers to those who would have incurred the cost of the unfunded care Medicaid beneficiaries would have received if the programme did not exist. These transfers are estimated to be 60% of total Medicaid expenditures. Hence, the insurance value is a much lower fraction of the gross cost of the programme to government.

¹² Consider major risk insurance that covers all expenses in excess of a deductible ζ , leaving the individual exposed to $M^o = \min(M, \zeta)$. The risk premium, π^o , indicating the welfare gain from removing the residual risk is defined implicitly by $u(y - E[M^o]) - \pi^o = EU(y - M^o)$. The risk premium for major risk insurance is the welfare gain from eliminating all risk less then welfare gain from removing the residual risk not covered by the partial insurance, $\hat{\pi} = \pi - \pi^o$.

healthcare corresponds to \$1 less consumption of other goods. This is clearly inconsistent with the evidence of at least partial consumption smoothing over health shocks.

Besides financing health payments through informal transfers, savings and borrowing, consumption can be made insensitive to medical bills by not paying them. Bankruptcy provides financial protection by placing a ceiling on liability for medical debt. It is an imperfect substitute for health insurance (Mahoney, 2015). Self-reported data on the causes of bankruptcy suggest medical expenses play a critical role in 17-54% of bankruptcies in the US (Himmelstein et al., 2005; Dranove and Millenson, 2006). Estimates based on more credible identification strategies put this fraction at 5-26% (Gross and Notowidigdo, 2011; Dobkin et al., 2018). Using state expansions of Medicaid in the 1990s, Tal Gross and Matthew Notowidigdo (2011) estimate that a 10 percentage point increase in Medicaid eligibility would reduce bankruptcies by 8%, and calculate the programme's benefit level that achieves the optimal trade-off between insuring uncertain medical expenses and containing moral hazard would be 14-24% higher if there were no bankruptcy laws to limit debt liability.

One side of the coin is that more health insurance reduces bankruptcies. The other is that the option of declaring bankruptcy reduces the demand for insurance (Mahoney, 2015). Even without taking the drastic step of bankruptcy, the uninsured may have some latitude to receive treatment while leaving their medical bills unpaid. This "implicit insurance" would appear to reduce the value of insurance to the uninsured (Finkelstein et al., 2017) and implies that part of the benefit from extending coverage goes to the patients who are indirectly subsidising them (Mahoney, 2015; Dobkin et al., 2018; Finkelstein et al., 2018; Garthwaite et al., 2018). However, leaving medical bills unpaid has long-term negative consequences for financial wellbeing through reduced access to credit. Carlos Dobkin et al (2018) find that a hospital stay reduces the credit limit of 25-64 year-old Americans by 5% four years after the admission. Kenneth Brevoort and colleagues (Brevoort et al., 2018) extend the standard model in which insurance is valued only because it reduces exposure to OOP payments to allow for an indirect benefit through improved creditworthiness resulting from avoidance of penalties for unpaid medical bills. They use the expansion of Medicaid resulting from the PPACA to quantify these effects, estimating that the reform reduced newly-acquired medical debt by 35% and that the financial benefits of Medicaid double when the indirect benefit is taken into account.

In the wake of the PPACA, US researchers have clearly become interested in estimating the value of formal health insurance taking account of its interaction with other strategies for coping with medical expenses. One might expect this interest to spread to other countries in which extensions of health insurance coverage are being considered.

Catastrophic Payments

Consumption smoothing can be tested only with longitudinal data on health and nonmedical consumption. A risk premium can be estimated with cross-section data on OOP medical expenditures but only under the strong assumption that those expenditures are independent and identically distributed both across households and through time. Only then will the cross-section distribution conditional on observable covariates be indicative of the distribution faced by any one household with those characteristics. This assumption is inconsistent with the dynamics of health and medical expenses, which are characterized by persistence and strong serial correlation arising from the chronic nature of many diseases (Feenberg and Skinner, 1994; French and Jones, 2004; Di Nardi et al., 2017). Calculation of a risk premium also requires making assumptions about preferences. Estimates of the degree of risk aversion vary wildly and the risk premium is highly sensitive to this parameter.

Faced with these limitations on the use of financial protection measures that are conceptually grounded in economic theory, many analysts have opted for a proxy measure that has minimal data requirements – a cross-section of household-level data on total and OOP medical expenditure – and does not appear to require making any assumption about preference parameters. The prevalence of catastrophic health payments, or catastrophic medical expenditures, is the proportion of households with healthcare payments that exceed some fraction of household resources, which have been measured by income, total expenditure or expenditure net of spending on food and possibly other items considered to be necessities (Berki, 1986; Feenberg and Skinner, 1994; Wagstaff and Van Doorslaer, 2003; Xu et al., 2003; Cylus et al., 2018). As mentioned previously, this measure – specifically, the proportion of the population with OOP health payments that exceeds 10 (or 25) percent of their total consumption – is being used to monitor progress with the UHC and SDG agendas (World Health Organization and The World Bank, 2015; United Nations, 2017; World Health Organization and The World Bank, 2017). In addition to being easy to calculate, the measure is intuitively appealing. Both characteristics have contributed to its

widespread application (Limwattananon et al., 2007; Van Doorslaer et al., 2007; Xu et al., 2007; Dmytraczenko and Almeida, 2015; Cylus et al., 2018; Wang et al., 2018; Yerramilli et al., 2018). It has been effective in drawing attention to the strain that utilisation of healthcare can place on household finances when formal insurance is lacking. And yet, an inevitable result of the measure's lack of conceptual basis is that the information it provides on financial protection is ambiguous. Healthcare may account for a large fraction of household expenditure without necessarily causing any drop in living standards, never mind a catastrophic one.

It is not always clear whether the aim of those calculating and using the catastrophic payments measure is to identify households experiencing excessive financial burdens of healthcare costs or to capture exposure to medical expenditure risk. Most applications, including those done in the context of UHC monitoring, restrict attention to OOP payments, which suggests that the intention is to focus on risk exposure. But in this respect the measure has some important deficiencies (Flores and O'Donnell, 2016).

The proportion of households in a cross-section with OOP medical spending in excess of some fraction of resources is not necessarily indicative of the typical exposure to risk and it tells us nothing about the distribution of risk. A given prevalence of catastrophic payments is consistent with one situation in which the same households always spend a large share of their budgets on healthcare and with another in which all households face an equal chance of spending excessively in each period but only a fraction of them actually incur the loss. In the first case there is essentially no risk, while in the second there is a lot of risk that is evenly spread.

Even if the chance of incurring OOP payments above some threshold fraction of resources were estimated for each household, this would still be a very partial measure of risk. It contributes to the expectation of particularly burdensome expenses but does not capture their variability (Gruber and Levy, 2009; National Research Council and Institute of Medicine, 2012). Attempts to move from prevalence to the magnitude of catastrophic payments have simply used the average extent to which OOP spending exceeds the threshold (Wagstaff and Van Doorslaer, 2003). However, for expected utility maximising, risk averse individuals, utility is a nonlinear function of consumption and so the average (excess) health payment does not correspond to the average loss from risk exposure. Perversely, only under risk neutrality does the average payment inform of

the average impact on welfare. Unlike the risk premium, the catastrophic payments metric does not capture the value attached to peace of mind from insuring medical expenses.

A further limitation of the measure is that it makes no allowance for the exercise of informal and self insurance, such as intra community and extended family risk sharing, dissaving and use of credit, that may cushion the impact of OOP payments that are large in relation to current income or total expenditure (Flores et al., 2008). While much of the consumption smoothing literature lies at one extreme in treating these coping strategies as optimal forms of insurance, most of the catastrophic payments literature lies at the other. The implicit assumption in applications that interpret the catastrophic payments indicator as a measure of risk is that none of the OOP spending is financed from informal and self insurance. Only then does increased OOP spending on healthcare (possibly in excess of some fraction of the household budget) imply a squeeze on nonmedical consumption and a drop in living standards. Some argue that by ignoring recourse to coping strategies the measure understates the economic impact of OOP payments since it misses the long-term consequences of depleted savings and accumulated debt (Ruger, 2012). While these effects are indeed overlooked, ignoring financing through informal and self insurance will overestimate, possibly grossly, the short-term disruption to consumption.

Gabriela Flores and colleagues (2008) use data from the Indian National Sample Survey on reported means of financing healthcare payments to obtain a measure of catastrophic payments that takes account of informally and self insured spending. Households are asked the amount of OOP payments financed by each of current income, saving, borrowing, sale of assets and all other sources. The amount paid out of current income is assumed to be at the expense of nonmedical consumption, while the remaining sources protect that consumption. Catastrophic payments are then defined by OOP spending that pushes out nonmedical consumption relative to income, which, consistent with the Haig-Simons definition (Atkinson and Stiglitz, 1980), is defined as total consumption (x) net of medical spending that is financed from informal and self insurance. That is, catastrophic payments are incurred when $\frac{m - \gamma m}{x - \gamma m} > \tau$, where m is OOP spending, $\gamma \in [0,1]$ is the proportion of this that is not financed from current income and τ is some threshold fraction. This ratio is less than $\frac{m}{x}$, which is usually used to identify catastrophic payments, provided there is

some financing from coping strategies ($\gamma > 0$).¹³ Flores et al (2008) find that this adjustment dramatically reduces the estimated prevalence of catastrophic payments for hospital inpatient treatment in India. Using the standard method that ignores the sources of financing OOP payments, around 2.3% of households are estimated to incur catastrophic payments defined as spending more than 10% of the household budget on inpatient treatment. Adjusting for the exercise of coping strategies, the estimate is almost ten times smaller. This is because, on average, households report paying for only a quarter (rural) to a third (urban) of inpatient costs from current income, with most of the remainder coming from saving and borrowing. No doubt there is less reliance on informal and self insurance to pay for outpatient care and medicines. And one should be cautious about the reliability of self-reported means of financing spending on any good.¹⁴ Nevertheless, the economic misery caused by medical expenses will depend on how they are financed and estimates of catastrophic payments are likely to be highly sensitive to allowing for informal and self insurance.

Of course, while recourse to saving and borrowing can avoid the immediate sacrifice of consumption to pay for healthcare, an opportunity cost will be incurred at some time. This will be either when savings are built up or when debts are repaid. But this is true for all sources of health financing. Insurance premiums, social insurance contributions and tax also push out nonmedical consumption. Like saving and borrowing, they avoid doing so in the period when healthcare is used. Saving and borrowing are inefficient means of health financing in comparison with the first best of actuarially fair insurance. But, as argued in the previous section, they are not necessarily less efficient than insurance that is offered at an unfair premium. Equity is a separate issue. A health financing system in which households draw down savings or accumulate debt to pay for healthcare may well be considered inequitable by many. Those who see injustice in recourse to these sources of finance will object to their removal from the measure of catastrophic payments. This is a perfectly legitimate ethical position. But defending it requires that the proposed measure of catastrophic payments be consistent with an explicit definition of equity in health financing

¹³ Since spending on healthcare is included in total consumption ($x = c + m$), $m < x$.

¹⁴ The India NSS is one of the very few surveys that collects data on the amount of healthcare spending financed from different sources. Even without this type of data, one can still adjust measures of catastrophic payments for recourse to coping strategies provided there is at least binary information on whether there is any financing from those sources (Flores and O'Donnell, 2016).

(Wagstaff and van Doorslaer 2003). The rationale for restricting attention to OOP payments then becomes less clear. The distribution of health insurance premiums, and indeed other prepayment sources of finance, also affects fairness in health financing. Why should it be inequitable to spend more than ten percent of the household budget on OOP payments for healthcare and yet it not be considered unfair to spend the same fraction on a health insurance premium that covers the expected cost of managing a chronic condition? Restricting attention to OOP makes sense only if the aim is to capture risk exposure, and even then it ignores premium risk.

An obvious difficulty faced by anyone applying a catastrophic payments measure is how to set the threshold above which OOP payments are labelled “catastrophic”. One approach is to set the threshold as a constant fraction of income or total consumption, with the latter being more feasible in economies with a large number of non-salaried workers in the informal sector. The rationale for this approach could come from an efficiency motivation for measuring catastrophic payments. In the presence of administrative costs or moral hazard, full insurance is not optimal. The second-best policy is a deductible with full coverage above it (Arrow, 1963; Drèze and Schokkaert, 2013). If risk aversion decreases with income, then welfare is maximised by setting a lower deductible for poorer households (Drèze, 1981; Feldstein and Gruber, 1995; Gollier, 2003). Viewed from this perspective, the prevalence of catastrophic payments defined as OOP spending in excess of a constant fraction of income could be seen as indicating the number of households with spending in excess of what would be optimal if there were an income-contingent deductible. Of course, this would be a convincing interpretation only if the threshold fraction of income were derived from a welfare maximisation problem. To date, this has not been done in the literature.

Since the poor spend a larger proportion of their resources on necessities, it has been argued that the threshold should increase not only in absolute terms with income but also in relative terms (Ataguba, 2012). Alternatively, the denominator of the OOP spending ratio used to identify catastrophic payments can be changed from total income/consumption to either variable net of spending on food (Wagstaff and Van Doorslaer, 2003; Xu et al., 2003) or a wider bundle of goods considered to be necessities (Cylus et al., 2018). Changing the denominator and allowing the (relative) threshold to vary with it would seem adjust twice for the income elasticity of health spending.

Arguments for different denominators tend to be motivated by distributional concerns and seem to imply that the primary purpose of the catastrophic payments metric is to assess equity (World Health Organization and The World Bank, 2017; Cylus et al., 2018). Indeed, Adam Wagstaff and Eddy van Doorslaer (2003) motivate the measure by an ethical principle: “no one ought to spend more than a given fraction ... of their income on health care.” (p.923). In their application, they use total consumption to proxy income. They observe that unless spending on food is excluded from the consumption aggregate, then the prevalence of catastrophic payments tends appear higher among the “rich” than the “poor”. Removing spending on all necessities (not only food) from the denominator can skew the distribution even more towards the “poor” (Cylus et al 2018). Such analyses of the distributional incidence of catastrophic payments are likely to be strongly confounded by the means of financing OOP spending when total consumption is used to rank households from poor to rich. A household that draws down its savings, sells assets, borrows heavily or resorts to some other coping strategy to pay a large medical bill OOP will be observed to have a high level of total expenditure and a large OOP share of this expenditure. This does not make it rich. Nor does it necessarily mean that its living standard has been catastrophically impacted. It could well be a middle-income, or even a poor, household with total spending that is temporarily inflated by medical expenses. Expenditure gross of payments for healthcare is an inappropriate metric to judge how well off a household is (Deaton and Zaidi, 2002). It should not be used to assess the distributional incidence of catastrophic payments irrespective of which measure of household resources is used as the denominator of the ratio that defines those payments.

Removing spending on necessities from the denominator, as is done in the assessments of financial protection in the WHO European region (Cylus et al 2018), brings the measure closer to one of the affordability of medical payments (Gruber and Perry, 2011), which I discuss toward the end of the next sub-section. The purpose of such a metric is to establish whether a household has scope within its budget to purchase nonmedical necessities after meeting healthcare costs. The catastrophic payments indicator with non-necessities in the denominator establishes whether health payments absorb more than some fraction of the household budget that is available after meeting nonmedical needs (n): $\frac{m}{x-n} > \tau$. The more items of expenditure that are defined as necessities, presumably the less concern there will be about medical expenses squeezing the remaining items, and so the higher τ should be set.

To deal with some of the limitations of the catastrophic payments indicator, Gabriela Flores and I (Flores and O'Donnell, 2016) propose a measure that is intended to capture household level exposure to high, possibly even predictably high, healthcare costs. We use estimates from cross-section quantile regressions of OOP payments to simulate the distribution of uninsured (by any means) medical expenses faced by each household. The regressors include binary indicators of whether OOP payments are financed from a number of coping strategies, which allows us to predict the household-specific distribution of OOP spending that is not self or informally insured. In order to focus on exposure to high medical expenses, we censor draws from this distribution below some threshold value (ζ), which could be defined in relation to expected expenditure or household income, and confine attention to the distribution of censored expenses, $m^* = \max(\zeta, m)$, and the corresponding distribution of censored nonmedical consumption, $c^* = y - m^* = \min(c, \xi)$, where $\xi = y - \zeta$. Our measure (R) is the maximum amount by which the household would be prepared to reduce consumption below the threshold value rather than face the censored distribution, $u(\xi - R) = EU(C^*)$.¹⁵

Unlike the risk premium for full insurance, and even the risk premium for major risk insurance, this measure is sensitive only to medical expenses in excess of the defined threshold. It can be used to identify households exposed to particularly burdensome medical expenses without comparisons being confounded by differences in exposure to more moderate expenses. A second attraction of the measure is that it decomposes into the probability of incurring medical expenses in excess of the threshold, the loss due to predictably low consumption if they do and the further loss arising from the volatility of medical expenses above the threshold. This can be useful to distinguish households afflicted by a chronic illness, for which the first two components of the decomposition may be large, from others exposed to the risk of acute sickness, which raises the value of the third component. A third advantage is that the measure can be used to identify households facing predictably high medical expenses, which again may arise because of the presence of a chronic illness. The standard risk premium is non-zero only if medical expenses are volatile. And it responds to deviations both below and above the expected value. Our measure can be made unresponsive to better than expected outcomes by setting the threshold at the (household-

¹⁵ $EU(C^*) = \int_0^y u(y - m^*) dF_{M^*} = \int_0^y u(c^*) dF_{C^*} = \int_0^\xi u(c) dF_C + (1 - F_C(\xi))u(\xi)$.

specific) expectation of medical expense, or some multiple of it, and so focusing on downside risk. Households facing expenses that exhaust a predictably large fraction of their resources can be identified by setting the threshold in relation to income or wealth. Such households would not benefit from the offer of actuarially fair single-period health insurance, which would be expensive given pre-existing conditions. But they would benefit from regulated and subsidised insurance. Consumption smoothing tests and risk premium estimates focus entirely on the loss from volatile medical expenses. Yet healthcare payments are not made only, or even mainly, for treatments of unexpected episodes of acute illness. Non-communicable diseases (NCD) are now the leading causes of death, as well as the main burden of disease, in middle- and even low-income countries. The consumption smoothing literature addresses ability to insure temporary increases in medical expenses. It ignores households that are consistently forced to maintain consumption at low levels because chronic illness permanently raises payments for treatment and medication.

Unlike the risk premium for major risk insurance, our measure does not give the value of insurance that protects against exposure to particularly large medical expenses. It is not the appropriate measure to use in a cost-benefit analysis of such insurance. But it can be used to identify households that would benefit from subsidised coverage of medical expenses that predictably exhaust a large fraction of the household budget.

Relative to the simple catastrophic payments measure, the disadvantage of our measure is that, like the risk premium, it is more demanding of data and computation. But both are becoming less binding constraints. The minimum additional data requirement is the availability of predictors of OOP payments. Household budget surveys provide the required demographic and socioeconomic characteristics. Measures of acute sickness are not required. Preferably (but not essentially), indicators of chronic illness are available. Many household (if not budget) surveys provide these. If allowance is to be made for informal and self insurance (again this is not essential), then some information is needed on how OOP payments are financed. Such data are rarely available. It would cost little to add a question to household surveys asking whether saving, borrowing, remittances, etc. had been used to finance the OOP payments reported. Hopefully, the value of such information will become increasingly recognised and the question will be added to more surveys that collect data on health payments.

Health Payments, Affordability and Poverty

National governments and international organizations strive to reduce poverty. Protection against medical expenses that threaten or exacerbate poverty strengthens the social safety net. Quantification of these expenses can potentially make important contributions to anti-poverty and health financing policy. The WHO and World Bank (World Health Organization and The World Bank, 2015, 2017) claim to estimate impoverishment caused by healthcare payments by taking the difference between poverty (P) assessed on the basis of total household expenditure (or income) net of those payments, ($P_2 = P(z, x - m)$, where z is the poverty line) and poverty assessed on expenditure gross of them ($P_0 = P(z, x)$): $\Delta P = P_2 - P_0 > 0$ (Wagstaff and Van Doorslaer, 2003).¹⁶ The simplicity and low data requirements of this measure make it extremely attractive and explain its widespread use. Unfortunately, it will identify poverty caused by healthcare payments only under extremely strong assumptions.

The outcome of interest is poverty, captured either by a prevalence measure, such as the poverty headcount, or one that is also sensitive to intensity, such as the poverty gap. With the exception of US applications, attention is usually confined to OOP payments for healthcare, although the rationale for excluding private insurance premiums is not clear. The simple impoverishment measure, ΔP , appears to compare poverty given current OOP payments for healthcare with a counterfactual in which there are no such payments. This raises the question of how the healthcare currently paid for OOP is to be financed. If the aim is to estimate by how much poverty could be reduced if OOP finance were replaced with pre-payment financing, then the counterfactual distributional incidence of the latter would need to be specified. Implicitly, the counterfactual contemplated is one in which OOP payments are eliminated and replaced with some means of health finance that ensures no one who is poor (assessed on expenditure gross of OOP) contributes and no one who is not poor is required to make a contribution that would make them poor.

¹⁶ $P(z, x)$ represents any measure of poverty decreasing in a measure of household resources (x) and the poverty threshold (z), which could be defined to vary with household characteristics. For example, it includes the Foster, Greer and Thorbecke (1984) class of decomposable poverty measures.

Even without considering how healthcare would be financed in the absence of OOP payments, construction of the counterfactual requires an assumption about how households finance those payments and, consequently, what each household's consumption would be if it did not incur them. The simple measure relies on the assumption that expenditure gross of OOP payments indicates what nonmedical consumption would be in their absence. That is, it assumes that OOP payments are paid for entirely by sacrificing consumption of other goods. There is no reliance on informal and self insurance strategies. This is inconsistent with the evidence on consumption smoothing reviewed previously, which suggests that there is extensive use of coping strategies that prevent consumption falling by the full extent of medical expenses and sickness-related earnings losses. For households on the margins of poverty, one would expect all available means of coping to be exercised to prevent consumption falling below a subsistence level. Given that the poor and near-poor are likely to be highly risk averse and may resort to sub-optimal self insurance strategies to pay for needed healthcare while maintaining consumption near subsistence, stability of consumption at this level does not imply that the potential welfare gain from formal insurance is low (Chetty and Looney, 2006). But it does mean that expenditure gross of OOP payments is likely to greatly overstate what the consumption of such households would be if they had not incurred those payments.¹⁷

The left-hand panel of Figure 1 uses data from the Philippines to show a popular illustration of the claimed impoverishing effect of OOP payments using the simple approach (Wagstaff and Van Doorslaer, 2003). The blue curve traces Jan Pen's parade of the income distribution (Pen, 1971). Individuals are ranked from lowest per capita household expenditure *gross* of OOP payments on the extreme left to the highest of this quantity on the extreme right. The height of the curve indicates the gross expenditure of each individual relative to a \$2.15-a-day poverty line. Those below the horizontal line at this threshold are poor by that measure of household resources. The red "paint drops" indicate OOP payments. According to the measure, individuals who fall from a position in the parade above the poverty line to below it after subtracting OOP payments are impoverished. Those who start below the threshold and fall further below it after netting out OOP payments are claimed to have been pushed deeper into poverty.

¹⁷ For this reason, I refer to expenditure/consumption *gross* and *net* of OOP payments rather than *pre* and *post*, or *before* and *after*, OOP payments, which is the terminology often used in studies that claim to estimate their impoverishing effect.

The right-hand chart presents the same data from a different perspective.¹⁸ Now individuals are ranked by expenditure *net* of OOP payments and those payments are represented by the blue spikes. Looked at in this way, spending on healthcare can make poor households appear not to be poor if poverty is assessed, as it often is, on the basis of household expenditure (or consumption) inclusive of medical expenses. But this is a deficiency of the measure of poverty (Deaton and Zaidi 2002), not identification of impoverishment as a result of those expenses.¹⁹ From this perspective, the increase in measured poverty when it is assessed on household resources net of OOP payments and compared with an assessment based on resources gross of those payments is more appropriately labelled “hidden poverty” (Flores et al., 2008) than “impoverishment” or “iatrogenic poverty” (Meesen et al., 2003).

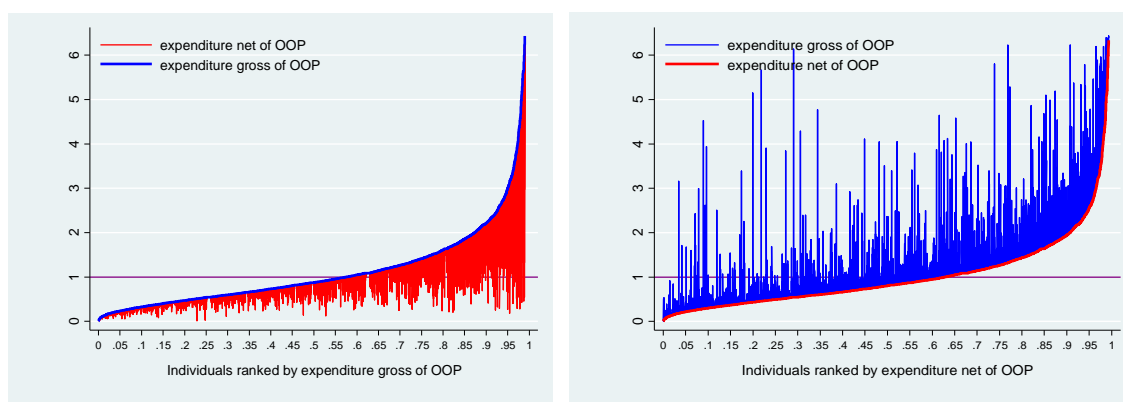


Figure 1: Pen’s parade based on expenditure gross (left) and net (right) of OOP payments

Notes: Data are from the Philippines World Health Survey 2003 and were analysed by Gabriela Flores and the author. Red in the left chart indicates OOP payments subtracted from gross expenditure. Blue in the right chart indicates OOP payments added to net expenditure.

The reality is that some, but not all, OOP spending on healthcare is financed from informal and self insurance coping strategies that partially shield nonmedical consumption. The counterfactual level of consumption if OOP payments were not incurred is given by neither consumption net of those payments nor consumption gross of those payments. For most, it will be somewhere in between. For those living at close to subsistence, if they do not forgo healthcare because it is unaffordable, then they will likely resort to extreme measures to avoid destitution and

¹⁸ Gabriela Flores came up with the idea of this chart.

¹⁹ Note that expenditures on durable goods are not included in the measure of living standards on which poverty status is assessed. Rather, an attempt is made to estimate and include the flow of services obtained from durables (Deaton and Zaidi 2002). Since medical expenditures can be viewed as payments for investment in health capital, their inclusion in the living standards is inconsistent with the treatment of expenditures on other durables.

the counterfactual will lie closer to consumption net of OOP payments. Then, the simple method will substantially overestimate the impoverishing effect of OOP payments. Those who normally would enjoy a standard of living comfortably in excess of the poverty line have greater latitude to pay for medical expenses from current income. Their counterfactual would be expected to lie closer to consumption gross of OOP payments, provided those payments are modest. If they were large, then a household that ordinarily is reasonably well-off would be unlikely to pay for them by sacrificing consumption of other goods to such a degree that it would be reduced to poverty. More likely, such a household would draw down savings and exhaust opportunities to borrow to keep itself out of poverty. Its informal and self insurance options may not allow perfect smoothing of consumption but still be sufficient to ward off poverty. If they are not, then there will indeed be medical impoverishment. But from the observation only of consumption gross and net of OOP payments in relation to the poverty line, it is not possible to distinguish this scenario from one in which a usually poor household runs down its savings, borrows or resorts to some other coping strategy to pay for healthcare while maintaining its subsistence level of consumption.

The simple method more plausibly identifies poor households that are mislabelled as non-poor by a deficient measure of poverty that counts spending on healthcare as part of consumption (Van Doorslaer et al., 2006; Flores et al., 2008; O'Donnell et al., 2008; Keane and Thakur, 2018). Van Doorslaer et al (2006), using data from eleven Asian countries in the late 1990s and early 2000s, estimate that taking OOP payments into account raises the estimate of poverty at the \$1-a-day threshold by 2.7 percentage points (pp), equivalent to 78 million people who were not counted as poor due to their spending on healthcare. Michael Keane and Ramna Thakur (2018) find that the adjustment raises the estimate of the poverty rate in India by 4.1 pp, or 50 million people, in 2011/12.

If healthcare needs were allowed for in construction of the poverty line, then it would be appropriate to include medical spending in the consumption aggregate. However, the stochastic nature of these needs makes it impossible to take them into account in an estimate of a constant level of expenditure or income considered consistent with an absolute or relative conception of poverty (O'Donnell et al., 2008). Consequently, it seems preferable to assess poverty on the basis of household resources net of spending on healthcare.

If data are available on the reported means of financing OOP payments, then those payments financed from informal and self insurance (γm) can be subtracted from total household expenditure to obtain a more plausible estimate of the nonmedical consumption the household can sustain. Let poverty assessed on the basis of this level of consumption be represented by $P_1 = P(z, x - \gamma m)$. Comparing this with the measure of poverty based on observed nonmedical consumption provides a more convincing estimate of impoverishment due to OOP payments: $P_2 - P_1 = P(z, x - m) - P(z, x - \gamma m) \geq 0$. Flores et al (2008) refer to this as ‘transient poverty’ due to spending on healthcare. Comparing poverty assessed on the basis of the sustainable level of nonmedical consumption with the conventional measure based on the aggregate of medical and nonmedical consumption provides a better estimate of what I referred to earlier as ‘hidden poverty’: $P_1 - P_0 = P(z, x - \gamma m) - P(z, x) \geq 0$ (Flores et al., 2008). This is poverty that is overlooked by the conventional measure because spending on healthcare inflates total household expenditure. The sum of these measures of transient and hidden poverty is the simple indicator that can be used to assess the consequences of OOP payments for the measurement of poverty in the absence of information on how they are financed, $P_2 - P_0$. Rather than identifying the impoverishing effect of health payments, this metric is more appropriately interpreted as revealing poverty that is overlooked because the conventional measure of poverty ignores the facts that spending on healthcare can both inflate total household expenditure when recourse is made to coping strategies and squeeze out nonmedical consumption.

Flores et al (2008) estimate that taking account of all OOP payments for inpatient care raised the poverty headcount by just over half a percentage point ($P_2 - P_0 = 0.55$) in India in 1995-96. Most (85%) of this correction is due to the increase in total household expenditure resulting from the use of coping strategies to finance payments for inpatient care ($P_1 - P_0$). The impoverishing effect of OOP payments pushing nonmedical consumption below the poverty line ($P_2 - P_1$) is less than one tenth of a percentage point. These results are unlikely to be indicative of the relative magnitudes of the hidden poverty and impoverishing effects generally. Spending on inpatient care tends not to account for the greater part of OOP payments and it is much more likely to be financed from coping strategies than smaller payments for items such as medicines that account for most OOP spending. Nonetheless, the general point holds that while ignoring OOP payments likely

results in the underestimation of poverty, ignoring that a substantial part of those payments are financed through informal and self insurance will greatly exaggerate their impoverishing effect.

In high-income countries, poverty tends to be assessed on the basis of income, which, unlike expenditure and consumption, is not mechanistically affected by OOP payments for healthcare, irrespective of how they are financed. Consequently, there is no potential for the poverty estimate to be downwardly biased by OOP payments financed by coping strategies. However, health payments still create a problem for the measurement of poverty. A constant poverty line, defined in either absolute or relative terms, cannot allow for resources required to satisfy medical needs that fluctuate. Spending that is made to cover healthcare expenses should therefore be subtracted from income prior to comparing the resulting measure of household resources available to meet nonmedical needs with the poverty line. This is what the US Census Bureau does to calculate its Supplemental Poverty Measure (Fox, 2017). Appropriately, spending on health insurance premiums, in addition to OOP payments, are removed from the measure of household resources.

The Census Bureau does not claim that comparison of the resulting poverty estimate with the official estimate that takes no account of medical expenses identifies medical impoverishment. Rather, both estimates are offered as alternative, imperfect measures of poverty. If one did attempt to estimate impoverishment in this way using the headcount measure of poverty, then there would be a bias in the downward direction. Some poverty created by payments for medical care will be missed since incomes may have been brought below the poverty threshold by illness-induced losses of earnings. This would be problematic in cross-country comparisons since the degree of income protection from illness and disability varies across countries. The impoverishing effect of health payments would be downwardly biased most in countries that provide the least protection against sickness-related earnings losses. This could be avoided by not using the headcount measure of poverty. Then, the impact of medical expenses on the intensity of poverty could be examined, although their indirect effect on income through labour supply behaviour should be considered.

In the US, the Patient Protection and Affordable Care Act has motivated interest in the affordability, rather than the impoverishing effect, of payments for healthcare (Gruber and Perry, 2011; National Research Council and Institute of Medicine, 2012). Payments that are literally unaffordable are not incurred and have no impact on consumption of other goods. Hence, unaffordable is not defined as that which is infeasible given the budget constraint. Rather,

healthcare payments may be deemed “unaffordable” if they would exceed the resources a household has available to spend on non-necessities. That is, they could only be afforded by cutting back on necessities. Implementation then depends on the definitions of healthcare payments, household resources and necessities. In the US context, there is little doubt that insurance premiums are as relevant as OOP payments. Indeed, the main motivation is to establish whether individuals could afford the plans they were mandated to purchase under the PPACA. If household resources are defined as income, then again no allowance is made for financing OOP payments through informal and self insurance. To allow partially for the latter, proposals have been made to add an annuitized value (or a simple proportion) of liquid assets to income (National Research Council and Institute of Medicine, 2012) or to replace income with consumption net of uncollateralised debt if the latter is greater (Gruber and Perry, 2011). Resources required to cover necessities could be defined by the poverty line or some multiple of it (National Research Council and Institute of Medicine, 2012) or an aggregate of spending on food, child care, housing, transport, taxes and an amount to cover all other necessities (Gruber and Perry, 2011). The latter is obviously highly subjective and potentially includes much spending that is plainly not on necessities. Using the poverty line has the advantage of being a recognised standard. This gives a measure that identifies payments for healthcare, inclusive of insurance premiums (i), as unaffordable if they exceed the gap between households resources, defined as income plus a proportion ($0 < \alpha < 1$) of assets (A), and a multiple ($k \geq 1$) of the poverty line: $m + i \geq y + \alpha A - kz$. This is obviously equivalent to resources net of health payments falling below a multiple of the poverty line. It is also close to the catastrophic payments measure that subtracts spending required to cover necessities from total expenditure in the denominator of the ratio used to identify such payments (Cylus et al., 2018).

Conclusion

Financial protection is about shielding nonmedical consumption from the cost of medical care that is utilized. Of course, someone must pay for healthcare. Health financing policy aims to distribute the costs of medical care efficiently and equitably across households and through time. Measures of financial protection should help monitor progress toward these goals and evaluate the effectiveness of policy reforms realizing them. None of the measures considered here achieves this perfectly.

Tests of consumption smoothing over health shocks inform of the extent to which all of the financing instruments households can call on are sufficient to protect their living standards when illness strikes. But the welfare interpretation of the test depends on which instruments are used. Taking kids out of school to cover medical expenses may allow consumption to be maintained close to subsistence in the short term, but it will reduce the consumption potential in the long term. Further, the test does not distinguish insurance of medical expenses from protection against earnings losses.

The risk premium for (major) health insurance deals with the latter limitation, but most applications go to the other extreme regarding the treatment of informal and self insurance. That is, they assume that coping strategies, such as transfers from relatives, saving and borrowing, provide no protection of nonmedical consumption from healthcare costs whatsoever. Consequently, the welfare gain from the extension of formal health insurance will be overestimated. This limitation can be corrected if there are data on the (reported) means of financing OOP payments, or if it is possible to estimate the impact of a health insurance reform on the distribution of nonmedical consumption directly.

By ignoring the fact that households can and do insure medical expenses informally and through saving/borrowing, the catastrophic payments measure likely overstates the extent to which spending on healthcare provokes economic catastrophe. Restricting attention to OOP payments gives the impression that those using this measure aim to capture risk exposure. But it is an extremely partial measure of risk. While the metric avoids the sensitivity of the risk premium to the parameterization of preferences, it is itself intrinsically sensitive to the choice of an ad hoc threshold budget share at which payments are labelled 'catastrophic'. A potential advantage over the risk premium is that it can be used to identify households facing predictably high healthcare costs, and not only those exposed to highly volatile costs. But then private health insurance premiums should be included in the measure of payments since these are a function of expected medical expenses, unlike social insurance contributions that are linked to income.

Re-calculation of poverty on the basis of household resources net of healthcare costs more plausibly identifies the extent to which those costs result in underestimation of poverty by the conventional measure rather than identifying medical impoverishment. As such, this approach is perhaps more relevant to the measurement of poverty than it is to health financing policy. More informative for the latter is the measurement of spending on healthcare that is 'unaffordable', or

affordable only under duress, in the sense that it is incurred by a household in poverty or on the margins of it. This may involve identifying households with income (total expenditure) net of OOP payments and insurance premiums below or within some multiple of the poverty line (or another estimate of expenditure required to cover nonmedical necessities) and counting the healthcare payments (OOP and premiums) incurred by those households. This would not measure the extent to which there is underinsurance of healthcare costs. Besides the inclusion of insurance premiums in the measure, there could be extensive reliance on coping strategies by households on the margins of poverty to pay for healthcare. But if equity in health financing includes shielding poor households from the costs of the care they use or expect to use, then the measure could be used to evaluate performance against that objective. The proportion of OOP and premium payments that are made by households in or close to poverty could be measured.

For the specific purpose of monitoring progress with UHC, the two measures currently employed suffer from an additional limitation. They are likely to be rather insensitive to actions governments can take to extend coverage and could well be more responsive to macroeconomic conditions that have little to do with health policy. The number of individuals identified as lying below the poverty line once health payments are subtracted from total expenditure will depend on the level of poverty and the density of the distribution of household expenditure just above the poverty line. Ministries of health cannot pull policy levers that shift these parameters. Admittedly, my suggestion to measure health payments incurred by those on the margins of poverty is open to the same criticism. However, by counting health payments incurred by the poor and near poor, rather than people claimed to be impoverished, health policymakers can be better motivated to take actions that impact on the distribution of those payments.

Through the income elasticity of health spending, the catastrophic payments measure will respond to economic growth. Economic expansion may result in more households being labelled as incurring catastrophic payments. Aggregate statistics, such as total health expenditure as a share of GDP, public health spending as a share to total public expenditure, OOP payments as a share of total health expenditure and the share of OOP payments incurred by the poor and near poor, do not measure financial protection but can be valuable complements to determine whether health financing policy is shifting in a direction consistent with the realisation of UHC.

REFERENCES

- Alam, K., Mahal, A., 2014. Economic Impacts of Health Shocks on Households in Low and Middle Income Countries: A Review of the Literature. *Globalization and Health* 10.
- Arrow, K.J., 1963. Uncertainty and the welfare economics of medical care. *American Economic Review* 53, 941-973.
- Asfaw, A., von Braun, J.B., 2004. Is Consumption Insured against Illness? Evidence on Vulnerability of Households to Health Shocks in Rural Ethiopia. *Economic Development and Cultural Change* 53, 115-129.
- Ataguba, J.E.-O., 2012. Reassessing catastrophic health-care payments with a Nigerian case study. *Health Economics, Policy and Law* 7, 309-326.
- Atkinson, A.B., Stiglitz, J.E., 1980. *Lectures on Public Economics*. McGraw Hill, New York.
- Barnes, K., Mukherji, A., Mullen, P., Sood, N., 2017. Financial risk protection from social health insurance. *Journal of Health Economics* 55, 14-29.
- Berki, S.E., 1986. A look at catastrophic medical expenses and the poor. *Health Affairs* 5, 138-145.
- Bernal, N., Carpio, M.A., Klein, T.J., 2017. The effects of access to health insurance: Evidence from a regression discontinuity design in Peru. *Journal of Public Economics* 154, 122-136.
- Besley, T., 1995. Savings, credit and insurance, in: Behrman, J., Strinivasan, T.N. (Eds.), *Handbook of Development Economics*. Elsevier, Amsterdam, pp. 2123-2207.
- Brevoort, K., Grodzicki, D., Hackmann, M.B., Koulayev, S., 2018. The credit consequences of unpaid medical bills, Department of Economics, UCLA, Los Angeles.
- Cesur, R., Güneş, P.M., Tekin, E., Ulker, A., 2017. The value of socialized medicine: The impact of universal primary healthcare provision on mortality rates in Turkey. *Journal of Public Economics* 150, 75-93.
- Chetty, R., 2008. Moral Hazard versus Liquidity and Optimal Unemployment Insurance. *Journal of Political Economy* 116, 173-234.
- Chetty, R., Looney, A., 2006. Consumption smoothing and the welfare consequences of social insurance in developing countries. *Journal of Public Economics* 90, 2351-2356.
- Cochrane, J.H., 1991. A Simple Test of Consumption Insurance. *Journal of Political Economy* 99, 957-976.
- Cochrane, J.H., 1995. Time-Consistent Health Insurance. *Journal of Political Economy* 103, 445-473.
- Cutler, D., Gruber, J., 1996. Does public insurance crowd out private insurance? *The Quarterly Journal of Economics* 112, 391-430.
- Cutler, D.M., Zeckhauser, R.J., 2000. The Anatomy of Health Insurance, in: Culyer, A.J., Newhouse, J.P. (Eds.), *Handbook of Health Economics*. Elsevier, Amsterdam, pp. 563-643.
- Cylus, J., Thomson, S., Evevovits, T., 2018. Catastrophic health spending in Europe: equity and policy implications of different calculation methods. *Bulletin of the World Health Organization* forthcoming.
- Deaton, A., 1991. Saving and liquidity constraints. *Econometrica* 59, 1221-1248.
- Deaton, A., Zaidi, S., 2002. Guidelines for constructing consumption aggregates for welfare analysis, LSMS Working Paper 135. World Bank, Washington, D.C.

- Dercon, S., Krishnan, P., 2000. In *Sickness and in Health: Risk Sharing within Households in Rural Ethiopia*. *Journal of Political Economy* 108, 688-727.
- Di Nardi, M., Pashchenko, S., Porapakkarm, P., 2017. The lifetime costs of bad health. NBER Working Paper No. 23963.
- Dmytraczenko, T., Almeida, G., 2015. *Toward Universal Health Coverage and Equity in Latin America and the Caribbean: Evidence from Selected Countries, Directions in Development*. World Bank, Washington D.C.
- Dobkin, C., Finkelstein, A., Kluender, R., Notowidigdo, M.J., 2018. The economic consequences of hospital admissions. *American Economic Review* 108, 308-352.
- Dranove, D., Millenson, M.L., 2006. Medical bankruptcy: myth versus fact. *Health Affairs* 25, W74-W83.
- Drèze, J., Schokkaert, E., 2013. Arrow's theorem of the deductible: Moral hazard and stop-loss in health insurance. *Journal of Risk and Uncertainty* 47, 147-163.
- Drèze, J.H., 1981. Inferring Risk Tolerance from Deductibles in Insurance Contracts. *The Geneva Papers on Risk and Insurance* 6, 48-52.
- Engelhardt, G.V., Gruber, J., 2011. Medicare Part D and the financial protection of the elderly. *American Economic Journal: Economic Policy* 3, 77-102.
- Feenberg, D., Skinner, J., 1994. The risk and duration of catastrophic health care expenditures. *The Review of Economics and Statistics* 76, 633-647.
- Feldstein, M., Gruber, J., 1995. A Major Risk Approach to Health Insurance Reform, in: Poterba, J.M. (Ed.), *Tax Policy and the Economy*, Volume 9. MIT Press, Cambridge, MA.
- Finkelstein, A., 2007. The Aggregate Effects of Health Insurance: Evidence from the Introduction of Medicare*. *The Quarterly Journal of Economics* 122, 1-37.
- Finkelstein, A., Hendren, N., Luttmer, E.F.P., 2018. The Value of Medicaid: Interpreting Results from the Oregon Health Insurance Experiment. *Journal of Political Economy* forthcoming.
- Finkelstein, A., Hendren, N., Shephard, M., 2017. Subsidizing health insurance for low-income adults: evidence from Massachusetts, NBER Working Paper No. 23668, Cambridge, MA.
- Finkelstein, A., Luttmer, E.F.P., Notowidigdo, M.J., 2013. What Good Is Wealth Without Health? The Effect of Health on the Marginal Utility of Consumption. *Journal of the European Economic Association* 11, 221-258.
- Finkelstein, A., McKnight, R., 2008. What did Medicare do? The initial impact of Medicare on mortality and out of pocket medical spending. *Journal of Public Economics* 92, 1644-1668.
- Finkelstein, A., Taubman, S., Wright, B., Bernstein, M., Gruber, J., Newhouse, J.P., Allen, H., Baicker, K., the Oregon Health Study Group, 2012. The Oregon Health Insurance Experiment: Evidence from the First Year. *Quarterly Journal of Economics* 127, 1057-1106.
- Flores, G., Krishnakumar, J., O'Donnell, O., Van Doorslaer, E., 2008. Coping with health care costs: implications for the measurement of catastrophic expenditures and poverty. *Health Economics* 17, 1393-1412.
- Flores, G., O'Donnell, O., 2016. Catastrophic Medical Expenditure Risk. *Journal of Health Economics* 46, 1-15.
- Foster, J., Greer, J., Thorbecke, E., 1984. A Class of Decomposable Poverty Measures. *Econometrica* 52, 761-766.

- Fox, L., 2017. The Supplemental Poverty Measure: 2016, Current Populations Reports. US Census Bureau, Washington, D.C.
- French, E., Jones, J.B., 2004. On the distribution and dynamics of health care costs. *Journal of Applied Econometrics* 19, 705-721.
- Garthwaite, C., Gross, T., Notowidigdo, M.J., 2018. Hospitals as Insurers of Last Resort. *American Economic Journal: Applied Economics* 10, 1-39.
- Genoni, M.E., 2012. Health Shocks and Consumption Smoothing: Evidence from Indonesia. *Economic Development and Cultural Change* 60, 475-506.
- Gertler, P., Gruber, J., 2002. Insuring consumption against illness. *American Economic Review* 92, 51-70.
- Gertler, P., Levine, D.I., Moretti, E., 2009. Do Microfinance Programs Help Families Insure Consumption Against Illness? . *Health Economics* 18, 257-273.
- Gollier, C., 2003. To Insure or Not to Insure?: An Insurance Puzzle. *The Geneva Papers on Risk and Insurance Theory* 28, 5-24.
- Gross, T., Notowidigdo, M.J., 2011. Health insurance and the consumer bankruptcy decision: Evidence from expansions of Medicaid. *Journal of Public Economics* 95, 767-778.
- Gruber, J., Hendren, N., Townsend, R., 2014. The great equalizer: health care access and infant mortality in Thailand. *American Economic Journal: Applied Economics* 6, 91-107.
- Gruber, J., Levy, H., 2009. The Evolution of Medical Spending Risk. *Journal of Economic Perspectives* 23, 25-48.
- Gruber, J., Perry, I., 2011. Realizing Health Reform's Potential: Will the Affordable Care Act Make Health Insurance Affordable?, The Commonwealth Fund publication 1493. The Commonwealth Fund, New York City.
- Hendren, N., 2013. Private Information and Insurance Rejections. *Econometrica* 81, 1713-1762.
- Himmelstein, D.U., Thorne, D., Warren, E., Woolhandler, S., 2005. Illness and injury as contributors to bankruptcy. *Health Affairs* 24, W65-W73.
- Hu, L., Kaestner, R., Mazumder, B., Miller, S., Wong, A., 2018. The effect of the Patient Protection and Affordable Care Act Medicaid expansions on financial well-being, NBER Working Paper 22170, Cambridge, MA.
- Islam, A., Maitra, P., 2012. Health shocks and consumption smoothing in rural households: Does microcredit have a role to play? *Journal of Development Economics* 97, 232-243.
- Keane, M.P., Thakur, R., 2018. Health care spending and hidden poverty in India, Business School Research Paper No. 2018 ECON 02, University of New South Wales Sydney.
- Kimball, M.S., 1990. Precautionary Saving in the Small and in the Large. *Econometrica* 58, 53-73.
- Kowalski, A., 2014. The Early Impact of the Affordable Care Act State-by-State. *Brookings Papers on Economic Activity*, 277-333.
- Kowalski, A., 2015. What Do Longitudinal Data on Millions of Hospital Visits Tell Us About Public Health Insurance as a Safety Net for the Young and Privately Insured? . NBER Working Paper 20887.
- Leland, H.E., 1968. Saving and Uncertainty: The Precautionary Demand for Saving. *The Quarterly Journal of Economics* 82, 465-473.

- Limwattananon, S., Neelsen, S., O'Donnell, O., Prakongsai, P., Tangcharoensathien, V., Van Doorslaer, E., Vongmogkol, V., 2015. Universal coverage with supply-side reforms: The impact on medical expenditure risk and utilization in Thailand. *Journal of Public Economics* 121, 79-94.
- Limwattananon, S., Tangcharoensathien, V., Prakongsai, P., 2007. Catastrophic and poverty impacts of health payments: results from national household surveys in Thailand. *Bulletin of the World Health Organization* 85, 600-606.
- Liu, K., 2016. Insuring Against Health Shocks: Health Insurance and Household Choices. *Journal of Health Economics* 46, 16-32.
- Mahoney, N., 2015. Bankruptcy as Implicit Health Insurance. *American Economic Review* 105, 710-746.
- Mazumder, B., Miller, S., 2016. The effects of the Massachusetts health reform on household financial distress. *American Economic Journal: Economics Policy* 8, 284-313.
- McIntyre, D., McKee, M., Balabanova, D., Atim, C., Reddy, K.S., Patcharanarumol, W., 2016. Open letter on the SDGs: a robust measure for universal health coverage is essential. *The Lancet* 388, 2871-2872.
- McPake, B.I., 2018. Crunching health expenditure numbers: important but treacherous terrain. *The Lancet Global Health* 6, e124-e125.
- Meesen, B., Zang, Z., van Damme, W., Devadasan, N., Criel, B., Bloom, G., 2003. Iatrogenic Poverty. *Tropical Medicine and International Health* 8, 581-584.
- Melvin Stephens, J., 2001. The Long-Run Consumption Effects of Earnings Shocks. *The Review of Economics and Statistics* 83, 28-36.
- Meyer, B.D., Mok, W.K.C., 2018. Disability, Earnings, Income and Consumption. *Journal of Public Economics*.
- Miller, G., Pinto, D.M., Vera-Hernández, M., 2013. Risk protection, service use and health outcomes under Colombia's Health Insurance Program for the Poor. *American Economic Journal: Applied Economics* 5, 61-91.
- Mitra, S., Palmer, M., Mont, D., Groce, N., 2016. Can Households Cope with Health Shocks in Vietnam? *Health Economics* 25, 888-907.
- Mohanan, M., 2013. Causal Effects of Health Shocks on Consumption and Debt: Quasi-Experimental Evidence from Bus Accident Injuries. *Review of Economics and Statistics* 95, 673-681.
- Moreno-Serra, R., Millet, C., Smith, P.C., 2011. Towards improved measurement of financial protection in health. *PLoS Medicine* 8, 1-6.
- Moreno-Serra, R., Thomson, S., Xu, K., 2013. Measuring and comparing financing protection, in: Papanicolas, I., Smith, P.C. (Eds.), *Health System Performance Comparison: An agenda for policy, information and research*. Open University Press, Maidenhead, pp. 223-254.
- Mossin, J., 1968. Aspects of Rational Insurance Purchasing. *Journal of Political Economy* 76, 553-568.
- National Research Council, 1995. *Measuring Poverty: A New Approach.*, in: Citro, C.F., Michael, R. (Eds.), *Panel on Poverty and Family Assistance: Concepts, Information Needs, and Measurement Methods*. Committee on National Statistics, Commission on Behavioral and Social Sciences and Education. , Washington, DC: .
- National Research Council and Institute of Medicine, 2012. *Medical Care Economic Risk: Measuring Financial Vulnerability from Spending on Medical Care*, in: O'Grady, M.J., Wunderlich, G.S. (Eds.), *Panel on Measuring Medical Care Risk in Conjunction with the New Supplemental*

- Income Poverty Measure. Committee on National Statistics, Division of Behavioral and Social Sciences and Education, and Board on Health Care Services, Institute of Medicine, Washington, DC.
- Neelsen, S., Limwattananon, S., O'Donnell, O., van Doorslaer, E., 2018. Universal health coverage: a (social insurance) job half done?, Erasmus School of Economics mimeo, Rotterdam.
- Neelsen, S., O'Donnell, O., 2017. Progressive universalism? The impact of targeted coverage on health care access and expenditures in Peru. *Health Economics* 26, e179-e203.
- Nyman, J.A., 1999a. The economics of moral hazard revisited. *Journal of Health Economics* 18, 811-824.
- Nyman, J.A., 1999b. The value of health insurance: the access motive. *Journal of Health Economics* 18, 141-152.
- Nyman, J.A., 2003. *The Theory of Demand for Health Insurance*. Stanford University Press, Stanford.
- Nyman, J.A., Trenz, H.M., 2015. Affordability of the Health Expenditures of Insured Americans Before the Affordable Care Act. *American Journal of Public Health* 106, 264-266.
- O'Donnell, O., van Doorslaer, E., Wagstaff, A., Lindelow, M., 2008. *Analyzing Health Equity Using Household Survey Data: A Guide to Techniques and their Implementation*. World Bank, Washington DC.
- Pauly, M., Kunreuther, H., Hirth, R., 1995. Guaranteed renewability in insurance. *Journal of Risk and Uncertainty* 10, 143-156.
- Pauly, M.V., Blavin, F.E., Meghan, S., 2009. How private, voluntary health insurance can work in developing countries. *Health Affairs* 28, 1778-1787.
- Pauly, M.V., Goodman, J.C., 1995. Tax Credits For Health Insurance And Medical Savings Accounts. *Health Affairs* 14, 126-139.
- Pen, J., 1971. *Income distribution: facts, theories, policies*. Praeger, London.
- Ruger, J.P., 2012. An Alternative Framework for Analyzing Financial Protection in Health. *PLoS Med* 9, e1001294.
- Shigeioka, H., 2014. The effect of patient cost sharing on utilization, health, and risk protection. *American Economic Review* 104, 2152-2184.
- Skoufias, E., Quisumbing, A.R., 2005. Consumption Insurance and Vulnerability to Poverty: A Synthesis of the Evidence from Bangladesh, Ethiopia, Mali, Mexico and Russia. *The European Journal of Development Research* 17, 24-58.
- Sparrow, R., Van de Poel, E., Hadiwidjaja, G., Yumna, A., Warda, N., Suryahadi, A., 2014. Coping with the Economic Consequences of Ill Health in Indonesia. *Health Economics* 23, 719-728.
- Townsend, R.M., 1994. Risk and Insurance in Village India. *Econometrica* 62, 539-591.
- United Nations, 2017. Resolution adopted by the General Assembly on Work of the Statistical Commission pertaining to the 2030 Agenda for Sustainable Development. Inter-Agency and Expert Group on SDGs A/RES/71/313, New York: United Nations.
- Van Doorslaer, E., O'Donnell, O., Rannan-Eliya, R.P., Somanathan, A., Adhikari, S.R., Garg, C.C., Harbianto, D., Herrin, A.N., Huq, M.N., Ibragimova, S., Karan, A., Lee, T.-J., Leung, G.M., Lu, J.-F.R., Ng, C.W., Pande, B.R., Racelis, R., Tao, S., Tin, K., Tisayaticom, K., Trisnantoro, L., Vasavid, C., Zhao, Y., 2007. Catastrophic payments for health care in Asia. *Health Economics* 16, 1159-1184.

- Van Doorslaer, E., O'Donnell, O., Rannan-Eliya, R.P., Somanathan, A., Adhikari, S.R., Harbianto, D., Garg, C.G., Herrin, A.N., Huq, M.N., Ibragimova, S., Karan, A., Ng, C.-w., Pande, B.R., Racelis, R., Tao, S., Tin, K., Trisnantoro, L., Vasasvid, C., Zhao, Y., 2006. Effect of health payments on poverty estimates in 11 countries in Asia: an analysis of household survey data. *The Lancet* 368, 1357-1364.
- Wagstaff, A., 2007. The economic consequences of health shocks: Evidence from Vietnam. *Journal of Health Economics* 26, 82-100.
- Wagstaff, A., 2010. Measuring Financial Protection in Health, in: Smith, P.C., Mossialos, E., Papanicolas, I., Leatherman, S. (Eds.), *Performance Measurement in Health System Improvement: Experiences, Challenges and Prospects*. Cambridge University Press, Cambridge, pp. 114-137.
- Wagstaff, A., Lindelow, M., 2008. Can insurance increase financial risk?: The curious case of health insurance in China. *Journal Health Economics* 17, 990-1005.
- Wagstaff, A., Lindelow, M., 2014. Are Health Shocks Different? Evidence from a Multishock Survey in Laos. *Health Economics* 23, 706-718.
- Wagstaff, A., Van Doorslaer, E., 2003. Catastrophe and impoverishment in paying for health care: with applications to Vietnam 1993-98. *Health Economics* 12, 921-934.
- Wang, H., Vinals Torres, L., Trivisa, P., 2018. Catastrophic health expenditure and financial protection in eight countries in the WHO South-East Asia Region. *Bulletin of the World Health Organization* forthcoming.
- Waters, H.R., Anderson, G.F., Mays, J., 2004. Measuring financial protection in health in the United States. *Health Policy* 69, 339-349.
- Wilson, R., 1968. The Theory of Syndicates. *Econometrica* 36, 119-132.
- World Health Organization, 2010. *The World Health Report 2010. Health systems financing: the path to universal coverage*. World Health Organization, Geneva.
- World Health Organization, The World Bank, 2015. *Tracking Universal Health Coverage: First Global Monitoring Report*, Geneva.
- World Health Organization, The World Bank, 2017. *Tracking universal health coverage: 2017 global monitoring report*, Geneva.
- Xu, K., Evans, D.B., Carrin, G., Aguilar-Rivera, A.M., Musgrove, P., Evans, T., 2007. Protecting Households from Catastrophic Health Spending. *Health Affairs* 26, 972-983.
- Xu, K., Evans, D.E., Kawabate, K., Zeramini, R., Klavus, J., Murray, C.J.L., 2003. Household catastrophic health expenditure: a multicountry analysis. *Lancet* 362, 111-117.
- Yerramilli, P., Fernández, Ó., Thomson, S., 2018. Financial protection in Europe: a systematic review of the literature and mapping of data availability. *Health Policy* 122, 493-508.
- Zeckhauser, R., 1970. Medical insurance: A case study of the tradeoff between risk spreading and appropriate incentives. *Journal of Economic Theory* 2, 10-26.