Chapter 9

Conclusions and discussion

In competitive individual health insurance markets, risk-rated premiums are observed to differ across subgroups of insured people, which are defined by rating factors such as: age, gender, family size, geographic area, occupation, length of contract period, the level of deductible, health status at time of enrollment, health habits (smoking, drinking, exercising) and — via differentiated bonuses for multi-year no-claim — to prior costs (Van de Ven et al. 2000). Financial transfers are needed in order to prohibit any problems of financial access to coverage for those at high risk. The first and best solution to increase financial access to coverage for those at high risk is to find a sponsor who organizes a regulatory system of risk-adjusted premium subsidies (Van de Ven et al. 2000). The financial transfers are then channeled via a so-called Risk Equalization Fund (REF). In that case, price competition is not distorted and therefore incentives for efficiency are not reduced. In all countries that apply risk-adjusted premium subsidies in their health insurance market, the sponsor organizes it in the form of risk equalization among insurers.

Although premiums are rated across many subgroups of insured people, a sponsor may not want to subsidize all premium rate variation observed in practice. The total set of risk factors that insurers use to rate their premiums can be divided in two subsets: the subset of risk factors that cause premium rate variation which the sponsor decides to subsidize, the S(ubsidy)-type risk factors; and the subset that causes premium rate variations which the sponsor does not want to subsidize, the N(on-subsidy)-type risk factors (Van de Ven and Ellis 2000, p. 768-769). In most countries, gender, health status and age (to a certain extent) will probably be considered S-type risk factors. Examples of potential N-type risk factors are: a high propensity for medical consumption, living in a region with high prices and/or overcapacity resulting in supply-induced demand, or using providers with an inefficient practice-style (Van de Ven et al. 2000). The sponsor determines the specific categorization of S-type and N-type risk factors. In case the government takes up the role of the sponsor, this categorization is ultimately determined by value judgments in society.

Under current law, Dutch government has decided to cross-subsidize cost variation between subgroups defined by the so-called S-type risk factors age, gender and health status only (MoHWS 2005, p. 23). Measures of age and gender are available in the administrations of all Dutch insurers and therefore can be included into the Dutch REF model relatively easily. However, the empirical possibilities to construct a health-based REF model at the individual level are rather limited. The 2004 Dutch REF equation (2.1) already contains an impressive range of health-based administrative adjusters, including Pharmacy-based Cost Groups (PCGs) and Diagnostic Cost Groups (DCGs). Currently, it is the most elaborate individual

level risk equalization model in the world. However, the question as to what extent even this extensive set of REF adjusters generates risk-adjusted premium subsidies (or cross-subsidies) that satisfy the policy goals of the Dutch government still remains unanswered. The central question of this study is therefore:

"To what extent does the 2004 Dutch risk equalization model induce risk-adjusted premium subsidies that meet the stated policy goals of the Dutch government and (how) can these subsidies be improved?"

In order to find an answer to this central question, three research questions must be formulated:

1. Given the definition of the basic benefits package, how can we calculate the cross-subsidies as intended by the Dutch government? (Chapters Three, Four and Five)
2. To what extent can the intended cross-subsidies be achieved by the health status measures included in the 2004 Dutch REF equation? (Chapter six)
3. To what extent can the intended cross-subsidies be achieved by alternative specifications of the 2004 Dutch REF model or by premium rate regulations? (Chapters Seven and Eight)

The main contribution of this study is the development and empirical application of a theoretical framework to determine the extent to which REF models induce effective cross-subsidies. In this study it is assumed that there is a periodic open enrollment requirement for a specified benefit package and a system of risk-adjusted premium subsidies. However, premium rates in the competitive individual health insurance markets are assumed not to be regulated.

In this study, a procedure is developed to test whether a given set of REF adjusters adequately compensates for cost variation caused by S-type risk factors (Chapter One). An overview is given of the relevant literature on (mainly) administrative measures of health status currently in use or under study. The methodology applied in subsequent chapters is described in more detail as well as in mathematical terms. Furthermore, guidelines are given on the interpretation of results and on the comparisons that are most relevant to find an answer to the central question of this study (Chapter Two).

For the application of the proposed test procedure, a tailor-made health survey was conducted amongst more than 50,000 sickness fund enrolees, such that the health status profile can be described much more broadly than with the REF adjusters alone (Chapter Three). The health status measures in this study are extensively tested for their completeness, reliability and validity (Chapter Four).

Throughout this study it is assumed that the Dutch government desires cross-subsidies for observed cost variation caused by the S-type risk factors age, gender and health status alone. For a limited sample of insured people, an alternative risk equalization model is then developed at the individual level that captures this observed cost variation as accurately as possible by including all measures of the S-type risk factors available from the administrative data sources and the health survey (Chapter Five). The so-called normative costs that follow from this alternative risk equalization model are then compared to REF predicted costs given the set of REF adjusters included in the 2004 Dutch REF equation: age, gender, insurance eligibility, region, PCGs and DCGs (Chapter Six). Following the same test procedure, cross-subsidies from alternative specifications of the REF model are tested for their effectiveness (Chapter Seven). Lastly, it is shown that an improvement of the REF model should be the preferred strategy to increase financial access to coverage rather than implicit cross-subsidies enforced by premium rate restrictions (Chapter Eight).

9.1 Conclusions

An answer to the first research question

Given the definition of the basic benefits package, the first research question determines how to calculate the cross-subsidies as intended by the Dutch government. In Chapter Five the normative costs are derived for a limited sample of insured people (N=18,617) under the assumption that society desires cross-subsidies for cost differences caused by the S-type risk factors age, gender and health status. The normative costs follow from a linear regression of observed costs on a broad array of health status variables from the health survey and administrative sources under the assumption that these form an adequate reflection of the S-type risk factors.

From the literature on risk adjusters as described in Chapter Two, it appears that the most promising candidates as measures of health status are self-reported measures of perceived health status, functional health status and chronic conditions. In order to guide the specific selection of health status measures, the conceptual model of Ruwaard and Kramers (1997) is applied. The selected health status measures are the eight SF-36 scales physical functioning (PF), role-physical (RP), bodily pain (BP), general health (GH), vitality (VT), social functioning (SF), role-emotional (RE) and mental health (MH). Three categories are based on the number of OECD (auditive, visual and mobility) limitations and three categories are based on the number of specific self-reported chronic conditions. The SF-36 is

a 36-item instrument for measuring health status and outcomes from the patient's point of view; it was designed for use in clinical practice and research, health policy evaluations and general population surveys (Ware and Hays 1988, Aaronson et al. 1998). PCGs and DCGs are finally added to the normative equation, as costs of medical care are not necessarily larger for lower scores on the above mentioned health status indicators (Newhouse 1989).

In Chapter Three the data that is used in this study is described. A broad array of health status measures is obtained by means of a tailor-made health survey which was conducted in 2001 amongst 50,022 Agis members. Gross response to the so-called "Agis Health Survey 2001" was 23,163 (46.3%). For the purpose of this study, 18,617 eligible records are included in the study sample, because valid SF-36 scale scores could be derived and the administrative records from the years 2001 and 2002 appeared to be both valid and available for these respondents. Reliability and validity of the eight SF-36 scales are tested positively in Chapter Four. Furthermore, a panel dataset is derived from the Agis social health insurance administration 1997-2002, as well as additional 2001 data which is kindly made available by Dutch research institutes (APE Public Economics and Prismant, located in The Hague and Utrecht, respectively).

Normative costs are then derived by ordinary least-squares, following equation (2.4) as the average observed costs 2002 for the subgroups defined by the so-called S-type adjusters age and gender, the non-rankordered versions of the PCGs and DCGs, the eight SF-36 scales, the number of OECD limitations, and the number of self-reported chronic conditions. Note that in contrast to the rank-ordered versions of the (rank-ordered) PCGs and DCGs which are included in the 2004 REF equation, in the normative equation, enrollees may be associated with multiple PCGs and DCGs. An answer to the first research question is that cross-subsidies which are in line with the policy goals of the sponsor can be calculated for a limited sample of survey respondents by following the approach developed in this study.

An answer to the second research question

In order to answer the second research question, the extent must be determined to which the health status measures included in the 2004 Dutch REF equation induce the cross-subsidies as intended by the Dutch government. Ideally, cross-subsidies are based on normative costs in order to be exactly in accordance with the policy goals of the sponsor. In that hypothetical case, for the subgroups defined by the S-type adjusters, REF predicted costs are equal to average observed costs. Therefore, the cross-subsidies which are implemented in practice can be tested by comparing REF predicted costs to normative costs for the subgroups defined by the S-type adjusters.

In Chapter Six, a test of the adjusters included in the 2004 Dutch REF equation is performed, given the normative costs as derived in Chapter Five. Following equation (2.6), the weighted average of the deviations of REF predicted costs from normative costs for the subgroups defined by the S-type adjusters equals 198. This number would be 687 in the absence of any risk adjusters in equation (2.1); that is, if REF predicted costs are equal to average observed costs for all insured people. Therefore, 1 – (198/687) x 100% = 71.2% of the cross-subsidies that the Dutch government desires can be achieved by the REF adjusters included in the 2004 Dutch REF equation. This finding answers the second research question.

REF predicted costs and normative costs can also be compared for subgroups defined by the REF adjusters (instead of the S-type adjusters). Deviations of REF predicted costs from normative costs may be attributed to N-type risk factors. REF predicted costs for disabled enrollees appear to be overcompensated for by 420 euros (15.1%), while enrollees on social welfare and self-employed enrollees are under compensated for by 327 euros (16.2%) and 162 euros (16.2%) relative to normative costs, respectively. Insured people living in the first regional cluster of zip codes are overcompensated for by 245 euros (13.6%) relative to normative costs, whereas enrollees living in the regional clusters 6, 7 and 8 are under compensated for by 190 euros (10.7%), 94 euros (5.4%) and 171 euros (11.4%) on average. Therefore, the assumption in the 2004 Dutch REF model that these REF adjusters are "unbiased" measures of health status differences must be rejected; i.e., in practice these REF adjusters lead to compensation for N-type cost variation. In this context it is important to notice that the study sample contains insurance members of only one Dutch health insurer, therefore the research results cannot be treated as representative of Dutch provinces.

Furthermore, it appears that REF predicted costs for self-employed enrollees are lower than those for employed enrollees, although normative costs are not for these subgroups. Apparently, the REF adjuster insurance eligibility fails to adequately capture S-type cost variation for self-employed enrollees. It should be noted that this problem can not be tackled by combining the subgroups of employed and self-employed enrollees into one subgroup, a solution which was heavily debated in the context of the 2004 specification of the Dutch REF equation.[151] In the end, the Dutch government decided to treat employed and self-employed insured people as separate subgroups in the 2004 REF equation, under the assumption that the difference could be attributed mainly to the S-type risk

151. Indeed, this solution would come at the expense of reduced risk-adjusted premium subsidies for employed enrollees.

factor health status. However, this decision has lead to lower premium subsidies for self-employed enrollees which from the results presented in this study cannot be justified given the policy goals as stated by the Dutch government.

Ideally, the gap between REF predicted costs and normative costs is removed such that the REF adjusters no longer induce cross-subsidies for cost variation caused by N-type risk factors. This may be achieved by an adjustment of the REF weights, for example by application of the so-called omitted variables approach or the so-called normative adjustment approach developed in this study. The omitted variables approach to remove N-type bias from the REF weights is proposed by Schokkaert, Dhaene and Van de Voorde (1998) and Schokkaert and Van de Voorde (2004). However, it turns out that following this approach, the reduction of the gap between REF predicted costs and normative costs is rather limited, at least given the study sample and the specific implementation of N-type adjusters in this study. An alternative procedure to adjust the REF weights is to replace them by so-called normatively adjusted REF weights that follow from equation (2.8) such that the bias is completely removed from the unadjusted REF weights.

The adjustment of the REF weights following the omitted variables approach leads to risk-adjusted premium subsidies to be in line with the policy goals of the Dutch government up to an extent of (1-201/687) x 100% = 70.7%. Alternatively, in case of a normative adjustment of the REF weights this performance outcome equals (1-209/687) x 100% = 69.6%. Apparently, a removal of the N-type bias from the REF weights at the same time slightly reduces the amount of S-type cost variation that is captured by the REF adjusters. Therefore, it is recommended to use these adjusted REF weights in a REF equation rather than the original REF weights in risk equalization models if this tradeoff is not too severe.

An answer to the third research question

In order to answer the third research question, the extent must be determined to which effective cross-subsidies can be achieved by alternative specifications of the 2004 Dutch REF model or by premium rate regulation. In Chapter Six, given the set of REF adjusters included in the 2004 Dutch REF equation, an adjustment of the REF weights is applied in order to remove the gap between REF predicted costs and normative costs *for the subgroups defined by the REF adjusters*. However, this procedure does not change the gap *for the subgroups defined by the S-type adjusters* and therefore does not improve the cross-subsidies to better compensate for cost variation caused by S-type risk factors. In Chapter Seven alternative specifications of the 2004 Dutch REF model are tested in order to reduce this gap for the subgroups defined by the S-type adjusters. This procedure may lead to cross-subsidies that are better in line with the policy goals of the sponsor.

As a first example, so-called paramedic cost groups (PMCGs), medical device cost groups (MDCGs) and mental pharmacy-based cost groups (MPCGs) are constructed from administrative data. Paramedic indicators of chronic diseases are used as indicators of physical limitations, medical devices are used as indicators of functional problems and pharmaceutical drugs for treatment of the nervous system as indicators for mental diseases. These potential new REF adjusters are derived from the claims data in the Agis sickness fund administration. If added to the REF adjusters that are already included in the 2004 Dutch REF model, it shows that they substantially reduce the gap between REF predicted costs and normative costs. Following equation (2.6), the weighted average of the absolute deviations of REF predicted costs from normative costs for subgroups of S-type adjusters equals 157. Given that this number would be 687 in the absence of any system of risk-adjusted premium subsidies (1 – 157/687) x 100% = 77.1% of effective cross-subsidies can be achieved after adding the aforementioned REF adjusters to the 2004 Dutch REF equation. Remember that this figure equals 71.2% if these new REF adjusters are not added. Thus, adding PMCGs, MDCGs and MHCGs will lead to a substantial improvement of the risk-adjusted premium subsidies; however, some room for improvement still remains.

As a second example, a specific ex-post risk sharing scheme is tested as a supplement to incomplete and/or imperfect REF adjusters. An analogue of the 2004 Dutch risk sharing scheme is applied: a 100% retrospective reimbursement of production-independent observed hospital costs and a 90% retrospective reimbursement of observed production-dependent hospital costs, medical specialty costs and costs of other health care services above a threshold of € 12,500. In this study, this turns out to close the gap between REF predicted costs and normative costs to a large extent. The weighted average of the deviations of REF predicted costs from normative costs for subgroups of S-type adjusters equals 141 in this case. Therefore, (1 – 141/687) x 100% = 79.5% of the intended cross-subsidies can be achieved as by the aforementioned ex-post risk-sharing arrangement as a supplement to the 2004 Dutch REF equation. Ex-post risk sharing will also remain an important supplement to incomplete and/or imperfect adjusters, even after the PMCGs, MDCGs and MPCGs are added to the REF equation. It should be noted that implementation of an ex-post arrangement in practice reduces the incentives for efficiency and therefore introduces a tradeoff between the improvement of the cross-subsidies and efficiency.

As a third example, REF weights are derived in a GLM framework under the assumption of a Gamma error distribution and a log link between REF predicted costs and the REF adjusters. The discrepancy between REF predicted costs and normative costs is substantially removed for almost all subgroups defined by the

S-type adjusters; albeit this reduction is smaller than under the aforementioned variants of the 2004 Dutch REF equation. The weighted average of the deviations of REF predicted costs from normative costs for subgroups of S-type adjusters equals 172. Therefore, (1 – 172/687) x 100% = 75.0% of effective cross-subsidies can be achieved by changing the statistical specification of the REF model.

Throughout this study, it is assumed that there is a periodic open enrollment requirement for a specified benefit package. Insurers are free to rate their premiums, and a system of risk-adjusted premium subsidies is organized by a sponsor to safeguard financial access to coverage for high-risk insured. Given that none of the specifications of the REF model induces cross-subsidies that fully capture cost variation caused by S-type risk factors, the sponsor may be tempted to impose premium rate regulations in order to safeguard financial access to coverage for high-risk insured anyhow.

Premiums are direct payments which insured people pay to their insurers and their rates can be regulated by the sponsor. Premium rate regulation can take several forms: community-rating per insurer, a ban on certain rating factors or rate-banding (by class). Community-rating is the most extreme variant of rate regulation, because all insured people enrolled with the same insurer are obliged to pay the same premium for a specified benefit package irrespective of their individual risk profiles. However, although premium rate restrictions are intended to create implicit cross-subsidies for cost variation caused by S-type risk factors alone, these may also induce cross-subsidies for cost variation caused by N-type risk factors, which is by definition in conflict with the policy goals of the sponsor. Given the theoretical framework developed in this study, it is possible to determine the specific amount of S-type and N-type cost variation that is implicitly cross-subsidized across subgroups in case of premium rate regulation.

In Chapter Eight, it is shown that for most of the subgroups defined by self-reported prior medical utilization and self-reported health status, diseases and conditions, the cost variation that is cross-subsidized under community-rating is largely caused by S-type risk factors. However, for the subgroups that can be defined by the number of years that survey respondents belong to the 25% of insured people with the highest total expenses within each year prior to 2002, the predictable profits and losses appear to be mainly caused by N-type risk factors. It follows that for some subgroups of enrollees, premium rate restrictions create implicit cross-subsidies which are largely in accordance with the policy goals of the sponsor; however, for other subgroups these cross-subsidies are not intended by the sponsor.

Since 2006, Dutch insurers have been allowed to risk-rate premiums across the twelve provinces under the Health Insurance Act. From the results in Chapter Eight

it appears that, given the specification of the 2004 Dutch REF equation, the premium rate variation across the twelve Dutch provinces must be mainly attributed to N-type risk factors. This appears to be even more so if the REF weights are purged from the bias caused by N-type cost variation during the estimation phase. These results justify the decision of the Dutch government to allow premium rates to differ across the twelve Dutch provinces; S-type cost variation appears to be already adequately subsidized by the Dutch REF equation. It should be noted that insured people of only one Dutch insurer are included in the study sample, therefore this research result cannot be treated as representative of all Dutch provinces.

The answer to the third research question of this study is that alternative specifications of the 2004 Dutch REF model can substantially improve the cross-subsidies. Up to 79.5% of the cross-subsidization as intended by the Dutch government can be achieved by application of an ex-post cost sharing arrangement; however, this arrangement reduces the incentives for efficiency at the same time. This drawback does not exist if new adjusters such as PMCGs, MDCGs and MPCGs are added to the 2004 set of REF adjusters; or if a multiplicative instead of additive specification of the REF equation is applied. In these latter cases 77.1% and 75.0% of the effective cross-subsidies can be achieved respectively.

To the extent that the improved cross-subsidies do not yet fully meet the policy goals of the sponsor, premium rate restrictions may induce the necessary cross-subsidization among the subgroups defined by the S-type risk factors. However, for other subgroups of insured people, these rate restrictions may also induce undesired cross-subsidization for N-type cost variation. At the same time, they also create incentives for selection which may have several adverse effects on quality of care, affordability of coverage for high-risk insured people and efficiency in the production of care (Van de Ven, Van Vliet and Lamers 2004). Application of premium rate restrictions as a supplement to risk-adjusted premium subsidies must therefore be necessary and proportional to achieve the stated policy goals; it also demands a careful tradeoff with the incentives for selection (and their possibly adverse effects) which are induced at the same time.

An answer to the central question

The answer to the central question of this study is that the REF adjusters in the 2004 Dutch REF equation generate cross-subsidies up to 71.2% of what is desired given the policy goals of the Dutch government. This achievement can be improved to 79.5% by application of an ex-post risk sharing arrangement, at the expense of the incentives for efficiency. As an alternative strategy to improve the cross-subsidies from the 2004 Dutch REF equation, new adjusters such as

PMCGs, MDCGs and MPCGs appear to be good candidates, and additional research on a multiplicative instead of additive specification of the REF equation is recommended. Premium rate regulation should preferably only hold up to the extent that it is necessary as a supplement to the risk-adjusted premium subsidies in order to meet the specific policy goals of the sponsor; because it creates incentives for selection at the same time. Finally, the use of adjusted REF weights are recommended instead of the original REF weights corresponding to the REF adjusters in any specification of the REF equation, although doing this at the expense of cross-subsidization for S-type cost variation must be avoided as much as possible.

9.2 Discussion

General policy recommendations

The model specifications in Chapters Six, Seven and Eight are used to illustrate the use of the test procedure developed in this study. The theoretical framework developed in this study may be applied to test any risk equalization model that a sponsor has implemented in a competitive health insurance market for the effectiveness of the cross-subsidies they induce in practice. In general, this theoretical framework can also be applied to other sectors where indirect standardization of large populations is guided by a normative decision rule, rather than merely being a statistical exercise.

The first policy recommendation is that the sponsor should always make an explicit choice about the specific categorization of the S-type and the N-type risk factors when implementing a risk equalization model. Given such an explicit choice, the theoretical framework developed in this study can be applied in order to determine the extent to which a REF model safeguards financial access to coverage for high-risk insured people.

The second policy recommendation is to apply the theoretical framework developed in this study on a regular basis to a national sample of insured people to test for the effectiveness of the cross-subsidies, and to check whether the adjusted REF weights should be used instead of the original REF weights in case of imperfect REF adjusters in practice.

The third policy recommendation is to improve cross-subsidies for the insured population with functional impairments by the development of REF adjusters based on utilization of physiotherapy, medical devices and pharmaceutical drugs for mental diseases. The latter category of REF adjusters may prove especially valuable if – contrary to this study – the cost definition also includes mental health care costs.

The fourth policy recommendation is to include an ex-post cost-sharing arrangement in the REF model as a supplement to an incomplete set of REF adjusters. From this study, it appears that ex-post risk sharing increases financial access to coverage for those at high risk even more than the implementation of additional REF adjusters based on the utilization of physiotherapy, medical devices and pharmaceutical drugs for mental diseases. However, ex-post risk sharing also reduces the incentive for efficiency. Thus, there exists a tradeoff between the effectiveness of the cross-subsidies and the incentives of efficiency with ex-post risk sharing.

The fifth policy recommendation is to allow premiums to be risk rated for cost variation caused by N-type risk factors. Premium rate restrictions may hold with respect to S-type risk factors, given that cost variation caused by S-type risk factors is cross-subsidized by the REF model and/or ex-post risk sharing. However, by definition, the (implicit) cross-subsidies that are induced by the premium rate restrictions are not supposed to compensate for cost variation caused by N-type risk factors. Furthermore, premium rate restrictions create incentives for selection which may have adverse effects on quality, affordability and efficiency. The undesirable compensation for N-type cost variation and (the possibly adverse effects of) the incentives for selection can be avoided if premiums are allowed to be risk-rated for cost variation caused by N-type risk factors.

If the sponsor desires cross-subsidies for cost variation caused by the S-type risk factors alone – for example, age, gender and health status – then a proportional implementation of premium rate restrictions would be a ban on the rating factors that are measures of these S-type risk factors (including the REF adjusters), but not on all other rating factors that are measures of the N-type risk factors as is done when requiring community rating. A less stringent alternative to community rating per insurer, per product might be rate-banding. Given the opportunity of rate-banding, insurers will then indicate which relevant measures of the S-type risk factors should actually be added to the REF equation in the next few years. Community rating per premium risk group or class can be implemented by the sponsor in order to protect the consumers against too strong premium increases in case of new rating factors. In any case, the question remains: why does the Dutch government still forbid insurers to rate their premiums based on measures of N-type risk factors, such as, regional input prices, practice style and propensity for consumption?

Country-specific policy recommendations

In the Netherlands, under the 2006 Health Insurance Act, government is legally obliged to undertake a scientific evaluation of the risk equalization system by a panel of international experts in 2008 and 2011 (MoHWS 2005, page 26). The theoretical

framework developed in this study can be applied to test for the effectiveness of the cross-subsidies given a stratified sample of the total Dutch population for this purpose. Such a sample can be found in the Permanent Survey of Living Conditions (POLS) that is conducted yearly by the Dutch National Bureau of Statistics (CBS). The 'Health Status' module in this survey is largely similar to the 2001 Agis Health Survey, and can be applied to derive normative costs for a national study sample of insured people (the SF-12 is used instead of the SF-36).[152] In 2008, it is expected that curative mental health care will be included in the basic benefits package, and the relative importance of the four mental SF-36 scales in the normative equation is expected to differ from that presented in this study.

In Switzerland, the REF adjusters are age, gender and the canton in which an enrollee lives, whereas ex-post risk sharing is not applied as a supplement (Beck et al. 2003, Van de Ven et al. 2007). In 1996 it was decided by law that the model specification would remain unaltered for a period of 10 years, yet in 2004 the law changed, and the risk equalization system was prolonged until 2010 (Bundesrat 2004, Bundesrat 2005). In 2006 the first chamber voted in favor of legislation to include prior hospitalization in hospitals and nursing homes (if admitted for at least three days) as a REF adjuster in the Swiss REF equation (Ständerat 2006, Art. 18a, Abs. 2, page 76). Furthermore, risk equalization is accepted as being permanent. Additional (yet unknown) health indicators are accepted as a long run option and the REF weights are going to be calculated prospectively instead of retrospectively in the future. The National Council did not make a decision after a hearing in May 2006; however, the National Council is expected to vote in favor of this legislation in 2007. Before the end of 2010, the Swiss government must make up their minds with respect to the inclusion of other morbidity-related risk adjusters in the REF equation (for example, PCGs), the implementation of a high-risk sharing arrangement or even the option to abolish the risk equalization system entirely (BBI 2004 4259, page 4273). The theoretical framework developed in this study can be of value in this delicate Swiss debate, as it clearly defines the extent to which cross-subsidies are needed to reduce the market mechanism problems that the Swiss regulators face (Van de Ven et al. 2007). Appendix A6.1 and Section 7.2 show the consequences of these decisions based on the Dutch sample used in this study. Note that the Swiss definition of the benefits package also includes (nursing) home care. It is expected that the specifications of the Dutch REF model

152. The regular frequency to repeat the test procedure should be bi-annually at least, because of sample size limitations that hold for the national survey. The net response to the 'Health Status' module in the yearly POLS survey is targeted to be about 10,000 respondents, whereas the 2001 Agis Health Survey contains 18,617 records. In general, the recommended size depends on the level of detail which is necessary to define the REF equation and the normative equation in this context.

presented in Chapters Six and Seven will induce less effective cross-subsidies if (home) care is included in the Dutch benefits package as well. Therefore, it is expected that the challenge to find an adequate specification will turn out to be at least as large as that for the Dutch REF model.

In Germany, the implementation of morbidity-related REF adjusters to capture the S-type risk factor health status in the REF equation was originally planned in 2007, but is now scheduled for 2009 and is expected to capture 50-80 conditions (Büchner and Wasem 2003, Bundesrat 2007). There will also be a move from an internal modality to channel the REF payments to an external modality, i.e. to modality B in Van de Ven et al. (2000, p. 324). In 2002 risk sharing was introduced for 60% of the costs above a threshold of more than 20,000 euros, thereby increasing the level of ex-post cost compensation from 0% to 4% (in 2006) (Van de Ven et al. 2007, Table 1). In 2003 the (voluntary registration for an) accredited so-called Disease-Management-Program (DMP) has been added as a REF adjuster to the German REF equation. Both the ex-post risk sharing and the DMPs are seen as temporary measures and will be abolished with the planned introduction of a risk equalization fund in 2009. In 2004, a group of international experts advised to use PCGs (RxGroups) and DCGs (HCC) for this purpose (IGES/Lauterbach/Wasem 2004); in 2006, the debate on the implementation of morbidity related REF adjusters was still ongoing (Schokkaert et al. 2006, Table 1). The two-tier insurance system for social and private health insurance will be maintained after 2009; however, sickness funds will be allowed to charge higher premiums for their members than the nation-wide premium based on expectations by the German government. Private health insurers will be confronted by open enrollment regulation for a specified benefits package that is equal to that for sickness funds. Furthermore, risk loading is not allowed when setting premiums. And the premium rates will also be capped, where the cap depends on the average premium rate in the sickness fund sector. There will be some form of risk pooling as a consequence of this premium rate regulation; however, the model specification is yet unknown. The effect of the PCGs and DCGs in the sickness fund sector and the effect of (the implementation of) a system of ex-post risk sharing in both social and private health insurance can be determined by application of the test procedure that is developed in this study.

In Israel, age is the only REF adjuster included in the REF equation. As of 2005, there are eleven instead of nine age subgroups. Gender cannot (yet) be made available because of feasibility problems and no ex-post risk sharing scheme is implemented (Shmueli, Chernichovsky and Zmora 2003). There has been a growing dissatisfaction with this formula. It is argued that children are overcompensated, while the elderly are under compensated, and arguments have been put forward

to include more risk adjusters (Van de Ven et al. 2007). In 2006, the debate on the implementation of morbidity related REF adjusters was still ongoing (Schokkaert et al. 2006, Table 1). Ability to work is not included as a REF adjuster because this is not a relevant measure of health status. Furthermore, socio-economic characteristics are not included as REF adjusters under the hypothesis that time price ("time to visit a doctor") is a more important determinant of observed cost variation than health status. However, based on the results presented in Table 6.7 it must be concluded that there is no good reason not to include ability to work or specific socio-economic characteristics in the Israeli REF equation, given that an adjustment of the REF weights can be applied to take care of cost variation caused by N-type risk factors such as the time price of self-employed insured people. In 2005 it was agreed that the Israeli REF equation would be updated after every four years, therefore a potential revision of their REF equation may be due in 2009. In the meantime, the theoretical framework developed in this study can be applied in Israel to determine the research results that apply to their specific situation. It should be noted that the Israeli formula is not based on individual claims data supplied by the sickness funds, but on a health services utilization survey and a hospitalizations data collection. Data on costs of drugs are not included at all.

In Belgium, the estimation of the REF weights has been based on individual data since 2002. The Belgian REF adjusters are age, sex, morbidity related and socio-economic variables; for example, indicators of chronic illness and disability categories. The information on DRGs and on pharmaceutical consumption is being collected but has not yet been implemented. However, there is a general consensus about the desirability to include it in the future risk adjustment model (Schokkaert et al. 2006). In the on-going debate about the categorization of the S-type and N-type risk factors, many observers keep arguing in favor of a risk adjustment formula which includes as many variables as possible, for example "number of days in the hospital" which is presented as an indicator of morbidity. It was decided not to include medical supply in the REF equation. This held sickness funds responsible for the regional cost variation that is caused by medical supply, although sickness funds do not have adequate instruments to influence the expenditures of their members (Van de Ven et al. 2007). The theoretical framework developed in this study can be used to evaluate the decision with respect to regional cost variation in the Belgian setting.

Since 2004, CMS in the USA has applied a "frailty" REF adjuster to finance PACE organizations. In the context of Medicare, this serves community populations with functional impairments for integrated care so that they can live in their own homes instead of being institutionalized. Due to several methodological problems of feasibility associated with the use of survey data for calculating risk equalization

payments, a "frailty" REF adjuster shall not be implemented program-wide in the CMS-HCC REF equation for Medicare Advantage plans in 2008 (CMS 2007, Attachment II, Section A). However, CMS announces that it will continue to explore ways to incorporate factors into the CMS-HCC REF equation that will better predict costs associated with the "frailty" of individual beneficiaries. From the results presented in this study, utilization of physiotherapy and medical devices may be used to overcome the feasibility problems that the CMS is currently faced with. In the meantime, the REF weights associated with the current REF adjusters can be adjusted in order for the cross-subsidies to better capture cost variation due to functional impairments.

Limitations of this study

The study sample used to illustrate the application of the theoretical framework developed in this study consists of those enrolled with Agis sickness fund in both 2001 and 2002. The results presented in this study may therefore not be representative of all Dutch regions, nor are they representative of the formerly private health insurance population which has also been insured under the Dutch Health Insurance Act since 2006. The cross-subsidies from the 2004 Dutch REF equation are not calculated for insured people younger than 16 years of age in this study, because the health survey was not conducted under this population.

The specification of the 2004 Dutch REF equation differs from its implementation in this study in some respects. The REF adjuster age defines ten-year instead of five-year classes of age, and interactions between the REF adjusters insurance eligibility and age are absent in this study. The cost definition includes production-independent hospital costs in this study; however, these costs are 95% reimbursed in Dutch practice up to 2006. Since 2006, about one third of hospital costs are defined as production-independent; they are effectively 100% reimbursed. Lastly, the proportional risk sharing (amongst insurers) and retrospective reimbursement scheme (between an individual insurer and the sponsor) which existed in 2004 with respect to production-dependent and medical specialty costs, is not applied in this study.

It may be possible that the S-type adjusters included in the normative equation in this study do not capture cost variation caused by S-type risk factors to the full extent. The range of health indicators may be expanded depending upon availability in future studies. The crux of the application of the theoretical framework in this study is that the current array of S-type adjusters is less limited than the set of REF adjusters which are used in practice. In this sense, the approach developed in this study produces a lower bound on the extent to which the REF equation induces the cross-subsidies that the sponsor desires. In other words, given the implementation of normative costs in this study, the performance scores of the REF models as

presented in Chapters Six, Seven and Eight will probably indicate a maximum for the extent to which the policy goals of the Dutch government are met.

The omitted variables approach to removing the bias from the imperfect REF weights appears to be rather limited given the set of N-type adjusters applied in this study. Results might change if another potentially broader set of N-type adjusters are applied.

Further research

The empirical results presented in this study apply to the 2002 Agis sickness fund insured people. Although the REF weights are not expected to change significantly for most REF adjusters, the exercise should be repeated for the total Dutch population of sickness fund enrollees, in order to be representative in this respect. In particular, the relative importance of the regional REF weights might change as a result.

Furthermore, the Dutch REF model holds with respect to all 16 million Dutch citizens under the 2006 Dutch Health Insurance Act. Therefore, the theoretical framework developed in this study is relevant to the total Dutch population since then and should be applied as such. The normative equation can be implemented based on information from national surveys which are currently available; for example, the 'Health Status' module in the POLS survey of the CBS.

Within a competitive health insurance market without risk equalization, premium rebates under the option of voluntary deductibles will reflect cost variation caused by S-type risk factors as a consequence of adverse selection (Van Kleef, Van de Ven and Van Vliet 2006). To some extent, this will still be the case if cross-subsidies are based on incomplete and/or imperfect REF adjusters. The extent to which this is the case can be determined under the approach developed in this study. Furthermore, the level of the voluntary deductible chosen by the insured people may be included in the REF equation as a proxy for health status. Although implementation of this proxy as a new REF adjuster may also induce undesired cross-subsidies for N-type cost variation, these effects can be explicitly weighed against each other by application of the theoretical framework developed in this study. An adjustment of the corresponding REF weight can be applied in order to avoid compensation for these N-type effects. Note that an analogous exercise is already performed with respect to insurance eligibility and region in the empirical part of this study.

The ultimate goal of managed competition is that insurers take up their role as prudent purchasers of health care. For this situation to occur, consumers must not only switch insurers as a consequence of their sensitivity to observed premium differences among insurers, but consumers must also be enabled to observe and

be sensitive to quality differences of health care delivery. Under the pressure of market competition, insurers are then forced to organize and purchase/provide health care according to the preferences of their enrollees. A more direct way to stimulate insurers to meet consumers' preferences, is to let the cross-subsidies depend on explicit measures of the quality of the health care that they contracted (in addition to compensation for S-type cost variation). The IOM (2006) recommends creating pools from a reduction in the base Medicare payments for each class of providers (hospitals, skilled nursing facilities, Medicare Advantage plans, dialysis facilities, home health agencies, and physicians). Initially, pay-for-performance programs should be designed such that providers are rewarded who achieve high performance and manage to improve performance significantly over time. Given that these indicators are often available for only a limited number of consumers, a natural approach seems to be to include these measures in the normative equation such that these are reflected in the adjusted REF weights.

A well-documented case of under provision necessary health care of social subgroups can be found in Shmueli (2000). This is for Arab insured people living in Israel. Table A8.3 of this study shows that REF predicted costs for first-generation immigrants are slightly above observed costs, but at the same time it is revealed that REF predicted costs are significantly lower than normative costs. This means that there is severe underutilization by first-generation immigrants. In this case there is no financial incentive for an insurer to tackle this problem of underutilization, because REF predicted costs are barely different from actual costs. This generates incentives for adverse selection against first-generation immigrants, yet removing the problem will lead to predictable losses for the subgroup of first-generation immigrants. Although this problem should probably be targeted by direct subsidies or educational programs, according to Schokkaert et al. (2006), an alternative approach may be to include first-generation immigrants as a REF adjuster in the REF equation and adjust the corresponding REF weight for N-type cost variation at the same time. This approach makes them preferred risks for insurers. Note however, that insurers must then also be allowed to adjust their premiums for the subgroup of first-generation immigrants in order to reduce the danger that the resulting predictable profits may be allocated for other purposes than combating underutilization.

In sum, the approach to risk equalization developed in this study can be applied in practice in several ways, and it is relevant to all countries with competitive health insurance markets. This approach is recommended to test for and improve the cross-subsidies of REF models in these cases.