Designing feasible and effective health plan payments in countries with data availability constraints

Abstract

Risk equalization is a fundamental tool in health plan payment in many countries. Data availability often constrains the feasible models. This paper proposes, implements and quantifies the gains of a risk equalization scheme which incorporates risk sharing in a data-poor context. Risk sharing relies on total spending data likely available for purposes of payment, potentially increasing feasibility of an effective payment design. To examine incentives for risk selection, alternative models are evaluated in terms of fit at individual, insurer, and group level. Using Chile's private health insurance market as case study, we show that modest amount of risk sharing greatly improves fit even in simple demographic-based risk equalization. Expanding the model's formula to include morbidity-based adjustors and risk sharing redirects compensations at insurer level and reduces opportunity to engage in profitable risk selection at group level. Our emphasis on feasibility may make alternatives proposed attractive to countries facing data-availability constraints.

Key words: risk equalization; diagnostic groups; risk sharing; performance fit; risk selection.

1. Introduction

A number of high and middle-income countries rely on individual health insurance markets to provide health insurance to their residents. Left unregulated, private markets lead to high prices for the sick and provide no subsidy relief to the poor. While market-driven prices have efficiency properties, equity is sacrificed (Van de Ven & Ellis, 2000). Policy based on principles of regulated competition address this dilemma by standardized health plans, open enrolment requirements, and by disconnecting individuals' premiums from individuals' risks through community-rated premiums. One downside of community rating is that it creates profitable and unprofitable risk groups, resulting in incentives for risk selection. Although direct selection is often prohibited by open enrolment regulations, plans may take other actions to attract good or deter bad risks, such as limiting access to services used by groups undercompensated in the plan payment system (Van Kleef, Schut, & Van de Ven, 2018). To mitigate these incentives, countries have adopted forms of risk equalization, which, based on empirical models of predicted expenditures, redistribute revenues to/from insurers that have above/below average risks (McGuire & Van Kleef, 2018a).

One of the guiding principles for the design of risk-equalization models is feasibility.[1] The availability of data on which to base risk equalization differs radically across countries. Even in some high-income countries (e.g. Switzerland) diagnostic data are not available for use in risk equalization. Data limitations tend to be more severe in middle-income countries. To be feasible in a particular institutional context, design of an effective risk equalization system must accommodate the local reality regarding data availability. This is certainly the case in our application: the regulated private health insurance market in Chile. In this paper, we design a feasible and effective risk equalization model for Chile. The proposed design may have value in the large number of national markets where data for risk equalization are limited.

Facing feasibility constraints imposed by data availability (e.g. privacy or data collection regulations, undeveloped registration systems), countries have taken different approaches to risk equalization. A cell-based approach to risk equalization, wherein individuals are classified

[1] Other guiding principles include appropriateness of incentives (e.g. selection by insurers) lack of gameability, and clinical meaningfulness. For a general discussion, see (Ellis, Martins, & Rose, 2018).

according to basic demographic variables, such as age, gender, and location, is applied in Australia (Paolucci, Sequeira, Fouda, & Matthews, 2018), Chile (Velasco, Henriquez, & Paolucci, 2018), Israel (Brammli-Greenberg, Glazer, & Shmueli, 2018) and Colombia (Bauhoff, et al., 2018). Switzerland has also for some years relied on a cell-based approach without diagnostic information, but has recently added morbidity information based on pharmacy claims (Beck, Kauer, McGuire, & Schmid, 2020). Countries with sophisticated systems such as the Netherlands (Van Kleef, Eijkenaar, Van Vliet, & Van de Ven, 2018), Germany (Wasem, Buchner, Lux, & Schillo, 2018) or the Medicare private market in the U.S (Pope, et al., 2011) have dedicated substantial resources to data collection and research over many years to bring the systems to where they are today. Other countries can benefit from the intellectual advances embodied in these systems, applied to settings with sometimes severe data limitations.

In addition to statistical models of prediction based on risk adjustor variables, risk sharing (the transfer of some responsibility for costs from a plan to the regulator or the overall insurance market) is a tool for health plan payment (McGuire & Van Kleef, 2018b). Forms of risk sharing are used, among other countries, in the Marketplaces in the U.S. (Layton, Ndikumana, & Shepard, 2018), Switzerland (Schmid, Beck, & Kauer, 2018) and embedded in the plan payment system in Ireland (Armstrong, 2018) and Australia (Paolucci, Sequeira, Fouda, & Matthews, 2018).[2] Researchers have recently found that a modest amount of risk sharing in Israel would substantially reduce incentives for risk selection (Brammli-Greenberg, Glazer, & Waitzberg, 2019). Risk sharing, like risk equalization, brings plan revenues more in line with plan spending, and may be particularly effective for high-spending individuals with spending levels that are difficult to predict with available data. Risk sharing relies on total spending data to partially determine plan payment, data which are likely to be readily available for purposes of insurer payment, potentially increasing the feasibility of an effective payment system design (Ellis, Martins, & Rose, 2018; McGuire & Van Kleef, 2018b). Estimation of risk equalization

[2] Israel and Colombia have forms of risk sharing through a high cost account where selected health conditions receive supplemental payments (Brammli-Greenberg, Glazer, & Shmueli, 2018; Bauhoff, et al., 2018). A form of risk sharing based on variables defined on past spending is in use in The Netherlands (Van Kleef & Van Vliet, 2012). Risk sharing in the form of high-cost group dummy variables and "residual-based reinsurance" have been studied for Germany. Residual-based reinsurance pays on the gap between spending and risk-adjusted payment rather than on spending per se (Schillo, Lux, Wasem, & Buchner, 2016). Residual-based reinsurance makes more effective use of reinsurance funds, but the gain from using residual-based reinsurance is less when the predictive model itself is not as powerful. The reason is that with simpler models, the difference between actual and predicted spending is not so great, so the incremental contribution of residual-based reinsurance over conventional reinsurance is less, as here in the case of Chile.

weights can easily incorporate the presence of risk sharing, as we do in this paper.

A relevant consideration that arises with the implementation of risk sharing vis-à-vis the sophistication of the risk equalization formula (e.g. addition of morbidity-based risk adjusters) is incentives for insurer efficiency (i.e. cost-control). As risk sharing will pay insurers for part of their realized spending, cost control incentives are diminished. The development of adjusters which compensate insurers for those high-spending individuals (e.g. diagnosis-based adjusters) would address this problem. Nevertheless, this approach involves additional considerations. Perverse incentives, such as upcoding can be a threat to the payment system (Geruso & Layton, 2020) and consistency with those factors that are desired in cross-subsidies. As such, a trade-off between feasibility, effectiveness and efficiency emerges.

Using Chile as a case study, this paper proposes, implements and quantifies the gains of a risk equalization scheme which incorporates risk sharing. Relative to current practice, our approach is novel as it adapts research and experience from countries with sophisticated (and data-driven) risk equalization formulas (e.g. The Netherlands, U.S. Medicare) to a data-poor setting befitting the health insurance markets in numerous countries around the world where regulation of health insurance markets must take into account data scarcity.

Using an administrative dataset that covers all privately insured in Chile, we model health insurer expenditures (i.e. health insurance claims net of copayments) for two benefit packages (what we will refer to as the "Current Plan" for GES services, and a "Universal Plan," as proposed by the current reforms). Various specifications of risk-adjusters and thresholds of risk sharing are tested for model fit and risk selection opportunities at the insurer and group levels. The implications for transfers of risk-equalization payments across insurers are calculated. We find that a modest amount of risk sharing greatly improves fit even in a simple demographic-based risk equalization. Expanding the model's formula to include morbidity-based risk adjustors and risk sharing redirects over- and under-compensation at the insurer level and reduces the opportunity to engage in profitable risk selection at the group level.

Currently, the Chilean health care system allows choice between public (Fonasa) and private insurers (Isapres). As part of the regulatory framework that governs the competitive private insurance market, risk equalization is present for the benefit package referred to as "GES

services,"[3] which is comprised of a set of 85 healthcare conditions[4] and services that are by law guaranteed in terms of access, opportunity, quality and financial protection. In response to ongoing concerns about efficiency and fairness, e.g. risk selection through discrimination on gender, age and pre-existing conditions, lack of transparency due to product differentiation and a proliferation of plans (Velasco, Henriquez, & Paolucci, 2018), among other concerns, the Chilean government presented to Congress a reform (Ministerio de Salud, 2019)[5] that aims to restructure the regulatory framework in the competitive private health insurance market. Among the main changes proposed are an expanded minimum benefit package, referred to as a "Universal Plan" (in Spanish, "Plan de Salud Universal"), partially risk-rated premiums, a risk equalization mechanism, yet to be developed, for the new benefit package.

The remainder of the paper is structured as follows: Section 2 introduces the Chilean healthcare system; Section 3 presents the data and specifications of the risk-equalization models; Section 4 presents the results; and, finally, Section 5 concludes and discusses the key policy recommendations.

2. Chilean healthcare system and risk equalization

Mandatory health insurance in Chile is mainly provided by two parallel components: Fonasa (*Fondo Nacional de Salud*), the public option (one public insurer), and Isapres (*Instituciones de Salud Previsional*), the private option (several private insurers). Together, Fonasa and Isapres covered 91.4% of the population in 2020 (15.1 million - or 78.9% of the population, and 3.4 million or 12.5% of the population, respectively).[6] The mandatory health insurance law in Chile states that workers must allocate 7% of their gross salary (with a cap for the maximum salary to which the contribution is applied) to an insurer of their preference to obtain health coverage. Nonetheless, the two components operate under different regulatory arrangements in relation to coverage, pricing, enrolment, and contracting with healthcare providers.

[3] GES services are part of both Fonasa and Isapres, but the risk equalization scheme only applies to private insurance.

[4] https://www.minsal.cl/plan-auge-85/

[5] The private insurance act - currently under discussion, together with the past version of the project (Ministerio de Salud, 2011)- is part of a broader health financing reform that also addresses the public insurer.

[6] The number of insured in each scheme was taken from Fonasa website and Superintendence of Health website for 2020. The percentage distribution is from the CASEN survey.

Fonasa has a standard benefit package, and individual copayments are related to income and type of provider (public, or one of the three types of private providers). Its funding comes from the mandatory salary contributions (7%) and a state subsidy based on tax revenues. Fonasa covers all family members of those paying salary contributions as well as destitute people.

The Isapres sector consists currently of 9 private insurers, six of which are open to the public and the other three are closed and limited to members of certain associations.[7] The largest private insurer accounts for 21.2% of the market, while the four largest private insurers together account for 82.5% of the total private insurance market. Isapres can offer as many plans as they wish and, currently, are allowed to use underwriting to select risks (no open enrolment) by means of a health status declaration form at the time of enrolment. Insurers use this declaration to determine if the individual is allowed to enrol[8]. In 2020, there were around 4,600 health plans offered, which differed in terms of benefits, financial coverage, providers and premiums. Premiums to enrollees in the private system include a risk-rated component based on age, and dependent status, as well as a component that is community-rated at the insurer level and dedicated to GES services finance. Premiums relate to the mandatory 7% in the following way: the contribution could or could not suffice to pay for the premium of the plan. In the former, it could even generate an excess which is saved by the individual in a health savings account with the insurer, and in the latter, s/he would have to pay on an absolute amount in pesos on top. These supplemental contributions averages 3% of gross salary.

In 2005, a risk equalization scheme for private insurance was introduced, restricted to the mandatory GES services (only about 8% of total private health expenditure in 2017, as individuals can opt-out of this coverage).[9] The purpose of the scheme was to deter risk selection incentives potentially arising from the establishment of an insurer-level community-rated price for GES services (i.e. every enrollee in the same insurer pays the same price for GES services).[10] In practice, risk equalization works through a virtual fund that is managed by the

[7] Not open to individuals who do not belong to those firms, which cover less than 3% of total Isapres beneficiaries. In what follows, we aggregate the closed Isapres into a single group.

[8] Despite this feature not affecting enrolment within plans in the market, Atal (2019) provides evidence of the welfare loss resulting from lock-in of high-risk individuals in their insurance plans.

9 Receiving treatment through the GES services plan entails certain restrictions similar to a "managed care" model. Individuals in Isapres can opt out to stay with their usual doctor or provider, and get coverage through other parts of their insurance plan.

[10] Recent simulations on the utilization of GES services have shown that the public sector insurance overall draws an adverse selection of the risks (Pardo, 2019).

Superintendent of Health. It is based on the cell method that uses age (18 groups) and gender (male and female), with the risk equalization payment equal to the average insurer expenditures for each cell in a prior period. No diagnostic information is included in the payment formula. Details of operation of the risk equalization scheme can be found in (Velasco, Henriquez, & Paolucci, 2018).

To date, few studies have assessed the performance of the health plan payment system in Chile. One evaluation called attention to the lack of important variables available for payment (e.g. morbidity risk-adjusters) (Ellis, Ibern, & Wasem, 2008). Empirical analysis (Henriquez, Velasco, Mentzakis, & Paolucci, 2016) using data from 2013 documented the poor predictive power (i.e. an R-squared of less than 1 percent) of the current risk-equalization mechanism.

3. Data and methods

3.1 Data

Data come from the Superintendent of Health and include plan and enrollee expenditures on health services in 2017 among all enrollees in one of the 12 Isapres[11], 3,661,280 individuals. The data contain information on socio-demographics (age, gender), clinical characteristics of inpatient care (including number of days hospitalized and principal diagnosis of hospitalization). Health expenditures are grouped into different categories which allow us to identify GES services and total expenditures.

3.2 Model specification and risk adjusters

We use a form of concurrent risk equalization, where 2017 health care utilization is modelled on same-year individual characteristics.[12] Specifically, two different health care expenditure outcomes are examined, each representing a different coverage package/plan:

[11] During 2017, there where 6 closed Isapres.

[12] Concurrent risk equalization uses information from the prediction period (year t), reducing data burden. This form is used for practical reasons in the US Marketplaces because of the high rates of turnover in the Marketplaces. Data on health care utilization in the current year is of course more predictive of spending in the current year than data from a previous period. This higher predictiveness comes at a cost in terms of reducing insurers' incentives to control health care costs. Prospective and retrospective models share some of the main trade-offs then the sophisticated (e.g. morbidity based) vis-à-vis risk sharing risk equalization formulas, regarding effectiveness and efficiency. All these models can be seen as steppingstones to developing models that achieve the right balance in

1) Current Plan: GES services expenditure claims net of copayments (i.e. expenditures that the insurers reimburse).[13]
2) Universal Plan: an expanded health care plan that considers all healthcare expenditures[14] claims (including GES[15]) net of copayments.

Policy applications of risk equalization are universally based on least-squares linear[16] regressions, taking advantage of the large sample sizes typically involved, low computational demands, and ease of implementation (Ellis, Martins, & Rose, 2018). We also used least-squared regression,[17] beginning with the current risk equalization formula where only age and gender are included. This comprises Model A where for each gender (i.e. Males/Females) age is divided into 18 classes (i.e. 0-1, 2-4, 5-9, 10-14, 15-19, 20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, 60-64, 65-69, 70-74, 75-79, 80+) resulting in 35 gender/age estimable parameters with Males 35-39 years-old set as the reference category in the estimation. We add clinical diagnostic information (79 categories, to be described) and city of residence (Santiago – reference category, Antofagasta y Calama, Viña del Mar/Valparaiso/Con-Cón, Concepción/Talcahuano, Temuco, Other) to Model A to obtain Model B.

Diagnostic information is available for hospital admissions based on the ICD-10 system. Our data include over 6,800 distinct codes. No diagnostic grouping methodology is currently being used or discussed in Chile. To collapse diagnostic codes into a manageable number of groups, we group three-digit ICD-9 codes into 78 clinically more homogeneous groups following (Ash, Porell, Gruenberg, & Beiser, 1989) (see Table A6 in Appendix). To apply this grouping to our data, we cross-walk diagnostic codes from ICD-9 (original format used for coding) to the ICD-10 system,18 retaining the 78 diagnostic subgroups with one additional residual group for admissions without a match to one of the 78 categories. Note that individuals may have more

light of the regulator's objectives. See (Ellis, Martins, & Rose, 2018) for further discussion on the advantages and disadvantages of concurrent risk equalization.

[13] In the implementation of risk equalization, the Superintendence of Health uses tariffs net of copayment that assimilate the costs of the services and frequencies that are estimated for the matters. Here we use market prices and administrative level frequencies as recorded.

[14] Some forms of expenditure, such as mental, dental and pharmaceuticals, among others, are less common, as they are generally not covered, or only partially covered in the plans.

[15] In the 2019 health insurance act sent to Congress, the Current Plan is separate from the Universal Plan.

[16] Some use weighted models.

[17] More flexible linear and non-linear models were also tested (i.e. GLM, Poisson, Zero-inflate Poisson, two-part) with OLS performing similarly or better and hence analysis proceed with OLS. For further discussion on risk equalization functional forms and model specifications see (Ellis, Martins, & Rose, 2018). Researchers are exploring new methodologies, while OLS is still the prevalence (Rose, Bergquist, & Layton, 2017).

[18] http://www.nber.org/data/icd9-icd-10-cm-and-pcs-crosswalk-general-equivalence-mapping.html

than one diagnostic indicator, if they have multiple admissions or multiple diagnoses on a single admission are indicated.[19] Alternative grouping methods could at a later stage be compared to the one we apply here (Juhnke, Bethge, & Muhlbacher, 2016).

Partial year enrollees (i.e. deaths, opting out of private insurance to the public system or other) are accommodated by annualizing their expenditures and weighting observations by the months of enrolment (Ellis, Martins, & Rose, 2018).

3.3 Risk sharing

We consider risk sharing in the form of reinsurance with two attachment points (thresholds above which reinsurance covers some percentage of spending) chosen based on the distribution of spending in our sample and corresponding to levels used in other private health insurance markets: USD $25,000 (e.g. some states in U.S.) (Layton, Ndikumana, & Shepard, 2018) and USD $50,000 (e.g. Australia) (Paolucci, Sequeira, Fouda, & Matthews, 2018).[20] When insurers are not responsible for some costs, incentives for restraining spending are reduced. Generally, in reinsurance systems, the insurer retains some share of the risk to preserve incentives for cost control. Here we select 65% as the share of spending over the attachment point to be covered by reinsurance.[21]

Risk sharing is integrated into the regression model by constructing a variable that takes the value of zero if the claim is below the threshold and, if above, the value of the costs over the threshold (e.g. for a threshold of $50, the variable would take the value of zero for a claim of $25 and the value of $25 for a claim of $75). In addition, the regression model restricts the parameter of the risk sharing variable to 0.65 (McGuire & Van Kleef, 2018b).[22] Formula (1) depicts the integration of risk equalization and risk sharing in our application, where y_i reflects

[19] In another example of data limitations, for admissions where more than one diagnosis is listed, it is not possible to distinguish which of the diagnoses is primary. 0.3% of the enrollees have more than 1 diagnostic code.

[20] We conduct sensitivity analysis with respect to the thresholds, which are available in Appendix Table A1.

[21] Such a share mirrors the empirical average demand-side cost-sharing level of individuals under the current system. In general terms, if a greater (lesser) percentage of funds over the attachment point is included, predictions should come closer (further) to actual spending, increasing (decreasing) fit measures and reducing (increasing) residual spending measures.

[22] Practically, this is estimated using a constraint regression in STATA v15, by implementing the command cnsreg. Alternatively and equivalently, estimation could be done in a two-step process, first pulling out reinsurance payments from plan obligations and then estimating weights. Choice between the one-step regression with the constrained coefficient and pulling out the reinsurance before estimation can be made on practical grounds.

actual spending for individual i (for the two dependent variables), x_i is the vector of risk adjustors (e.g. Model A and B), and T will take up the values of our thresholds (e.g. $25,000 and $50,000).

$$y_i = \beta x_i + 0.65\{y_i \geq T\}(y_i - T) + \epsilon_i \quad (1)$$

The purpose of integrating risk sharing at the estimation phase is to optimize the risk equalization weights on the demographic and disease-based risk adjustors accounting for the presence of reinsurance. Intuitively, if very high spending for some disease is to be covered by reinsurance, it is inefficient to "pay twice" with a very high weight on that predictor. A further advantage of the proposed approach is that the R^2 measure of fit is equivalent to the "payment system fit" used to evaluate fit in payment models including risk equalization and risk sharing.[23]

3.4 Model performance

In total, we estimate twelve models (i.e. two health plans – Current Plan and Universal Plan, two specifications - Model A (age-gender classes) and B (adds diagnostic categories and city of residence to Model A), and three risk sharing thresholds - no risk sharing as the baseline, $50,000 and $25,000 thresholds) with fit of payments to insurer expenditures examined across all models.[24] The full data set is randomly split into two equal parts with one half used for parameter estimation (i.e. training dataset) and the other half for prediction and evaluation of model performance (i.e. test dataset)[25] to avoid overfitting. Goodness-of-fit and predictive power is assessed through the adjusted R-Square (R^2) and Cumming's prediction measure (CPM) (Van Veen, Van Kleef, Van de Ven, & Van Vliet, 2015). The adjusted R^2 measures how close the data is to the predicted values, adjusted for the number of predictors in the model. The measure is based on squared errors, which weigh large errors more than small errors, making it sensitive to variance in expenses and outliers in the data. CPM captures predictive

[23] Payment system fit is a generalization of an R^2 for describing fit of a payment system at an individual level. It simply substitutes payment for predicted value in the formula for share of explained variance. Typically, it is evaluated using simulation methods. Here, the integrated estimation produces payment system fit as the R^2 in the regression output. For discussion, see (McGuire & Van Kleef, 2018b).

[24] Because the reform considers some clinical information in the risk equalization formula, we omit the models that include risk sharing under Model A of risk equalization. Results for these are available in the Appendix Tables A1-A5.

[25] 1,830,640 individuals in each sample.

accuracy in absolute terms, that is how far the observed values are from predictions.[26] The CPM is less sensitive to large outliers from the predicted values. Higher values of both the adjusted R^2 and CPM indicate better fit of the risk-equalization model.

3.5 Risk selection measures based on residual spending

Two risk selection measures are examined based on residual spending (i.e. the difference between what the insurer spends and what the insurer would be paid by the risk-equalization model) (Park & Basu, 2018). First, for insurer *k* we compute the difference between the average predicted expenditure (i.e. risk-adjusted), $\overline{\hat{y}_k}$, and average actual expenditure, $\overline{y_k}$, (i.e. $\overline{\hat{y}_k} - \overline{y_k}$) (Van de Ven, Van Vliet, & Van Kleef, 2016). The difference between predicted and actual describes the financial result for an insurer. A negative sign implies the insurer is undercompensated (predicted falls short of actual) and conversely, a positive sign means the insurer is overcompensated. If, after risk equalization, average residual expenses differ from zero, selection incentives may be present on variables not in the risk equalization model (which might be observed or unobserved).[27] In any case, implied transfers among insurers is one set of results of interest to policy makers and to the insurers themselves.

The second measure is based on average residual spending for groups defined according to two sets of categorical variables: days of hospitalization (0 days, 1 to 5 days, 6 to 10 days, 10 to 30 days and +30 days) and number of comorbidities (total diagnostic groups recorded during a year; 0 comorbidities, 1, 2 and +3). The difference between the average predicted spending per group *j*, $\overline{\hat{y}_j}$, and the actual spending of that group $\overline{y_j}$ (i.e. $\overline{\hat{y}_j} - \overline{y_j}$), measures how profitable (and therefore attractive) they are for the insurer (Van Kleef, Eijkenaar, & Van Vliet, 2019). Such an exercise identifies specific groups for which the formula systematically under- or over-pays, creating risk-selection incentives for insurers.

For each measure we first present results in the absence of risk equalization (i.e. mean actual expenditure to mean average overall expenditure, - community-rating) followed by the addition

[26] The CPM = 1 – (Mean Absolute Prediction Error) / (Mean Absolute Deviation from Average).
[27] The measure needs to be interpreted with caution. Differences in efficiency between insurers might also explain the differences between average predicted and average actual spending. See (Van de Ven, Van Vliet, & Van Kleef, 2016) for discussion.

of risk equalization, and the different risk sharing thresholds to the pure risk equalization model.

3.6 Risk equalization payments

Risk equalization redistributes funds among insurers. To convert risk-adjusted expenditures into risk-equalization payments (*REP*), the average risk-adjusted expenditure across all insurers is subtracted from the risk-adjusted expenditure for enrollees of insurer k. $REP_k = \sum_i (\hat{y}_i - \bar{\hat{y}}), \forall\ k\ \in K,$ where $\hat{y}_i$ is the predicted expenditure by an individual enrolled in k and $\bar{\hat{y}}$ is the population average predicted expenditure.[28] These payments capture how much an insurer receives/contributes from/to the risk-equalization fund. The payment scheme adopted is zero-sum, reallocating funds from insurers with low predicted risk to those with high predicted risk.

Generally, as risk equalization improves, risks will be predicted more accurately, and more funds will be reallocated among insurers. If the scheme is doing nothing (i.e. predicting the mean only), there would be no transfers among funds. The higher the absolute values of these payments, the more the scheme is picking up systematic differences in need across plans.

3.7 Incentives for cost control

One of the purposes of capitation-based payments to health insurers is to convey incentives for insurers to control costs. Basing capitation payments on variables outside the control of providers or insurers, such as age, gender, or residence, do not interfere with cost control incentives, whereas risk adjustors based on health care encounters (the source of morbidity data) or spending do affect incentives. There is no consensus about how to assess the effect of payment systems on incentives (Geruso & McGuire, 2016). We comment on incentive effects of alternative models based on the share of people and the share of spending affected by risk sharing. We do not assess the incentive effects of the morbidity-based adjustors, not because these adjustors do not introduce incentives, but because there is no readily available method for assessing how incentives are affected by such adjustors.

[28] With an ordinary least squares regression, the average of predictions equals the average plan obligations.

4. Results

4.1 Sample descriptive statistics

The Current Plan has an average insurer spending of USD 94 (standard deviation: USD 1,564) while the Universal Plan has an average of USD 812 (standard deviation: USD 5,043).[29] In Figure 1, important features of the spending distributions are depicted. The figure shows that only 9.1% of individuals made a claim in the Current Plan (second vertical dotted line), while many more people (84.7%) did so for the Universal Plan (first vertical dotted line). In addition, it shows that spending in both plans is highly right-skewed, as the 95th percentile of spending for the Current Plan is around $161, rising to $1,512 for the 99th percentile. For the Universal Plan, the 95th percentile is $3,145, and the 99th percentile is $10,427.

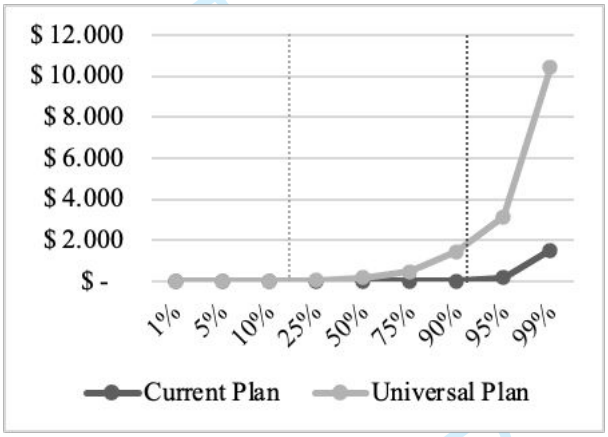

Figure 1: Distribution of the dependent variables by percentiles
Note: The vertical lines represent the percentage of claimants with zero-expenditure, at 90,9% for the Current Plan, and at 15,3% for the Universal Plan.

Males comprise 54.3% of the population. Most individuals are less than 59 years old (90.5%) with the largest age groups being 25-29 and 30-34 (around 10% each). 58.8% of the population resides in Santiago (Chile's capital). Figure 2 presents mean expenditures across the gender-age distribution for both health plans. For the Current Plan, males in nearly all age groups have higher expenditures than females. This pattern is reversed during women's child-bearing age for the Universal Plan highlighting the partial coverage for maternity related services within GES (i.e. only specific conditions, such as premature delivery, are covered). Moreover, in the

[29] Chilean pesos were converted to USD using an exchange rate of 0.001525 dollars per peso.

Universal Plan, annual expenditures are large for infants 0-1 years old, decreasing steadily until the 10-14 age group. Women starting at the age of 15 have larger expenditure than males, up to age 64, where the male overtakes the female average. The male-female gap reaches its maximum in the age group 80+ where males have USD 1,300 higher expenditure on average.

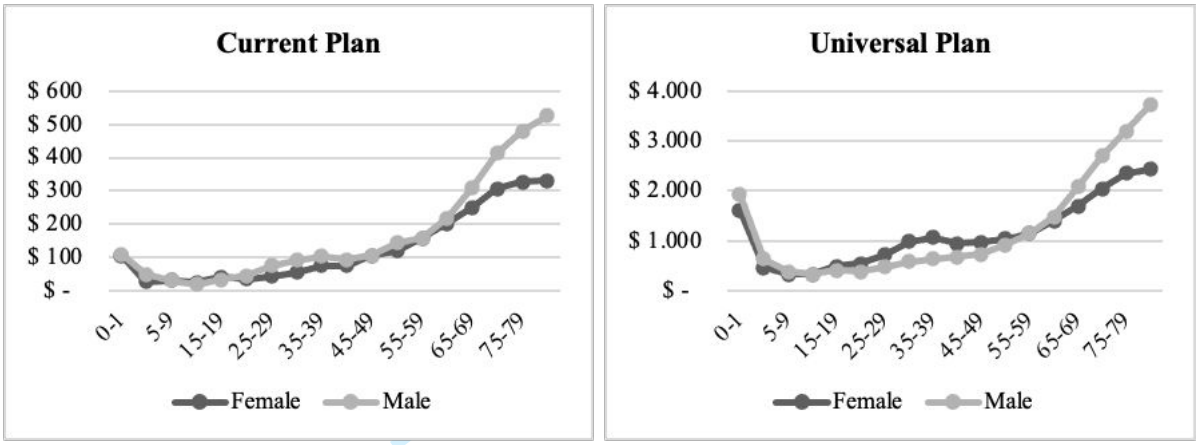

Figure 2. Expenditure distribution by gender and age for the two health coverage plans.

4.2 Incentives for Cost Control

High-cost risk sharing will only affect high spenders. Table 1 shows that in both plans, few people are touched by reinsurance, with a higher proportion touched by the lower threshold. For example, a threshold of $25,000 in the Current Plan touches only 0.03% of the population (3 in 10,000). For both plans, the $50,000 threshold touches about one-third of the people compared to the lower threshold. The share of funds (sum of costs above the threshold divided by total expenses under each plan and multiplied by 0.65) affected by risk sharing ranges from 5.9% to 2.9% for the Current Plan and 7.8% to 4.6% for the Universal Plan.

Table 1: Shares of individuals and funds above the risk sharing thresholds for the two health coverage plans

	Current Plan		Universal Plan	
	# of Individuals (% of Total obs.)	Funds out of Total	# of Individuals (% of Total obs.)	Funds out of Total
Threshold $25,000	1,149 (0.03%)	5.9%	9,141 (0.27%)	7.8%
Threshold $50,000	317 (0.01%)	2.9%	3,359 (0.10%)	4.6%

As noted above, no single metric is available to assess the effects of risk sharing on incentives for cost control. An approximate measure of the effect on incentives is the share of funds affected by reinsurance. This measure is exact in a proportional risk sharing scheme (where uniform risk sharing applies over the full range of spending) and is approximately true with

non-linear risk sharing as here. Based on this, incentives for cost control are reduced by this form of risk sharing in the range of 3-8%.

4.3 Performance measures (fit) across risk-equalization formulas

Model performance in terms of fit (Table 2) shows that with risk equalization only, fit is poor for the Current Plan (GES services), with an R^2 of 0.2% for Model A. This model represents the regression version of the risk equalization mechanism currently in place (i.e. cell-based). R^2 rises to only 3.5% with a full set of risk adjusters including morbidity indicators (Model B). Fit is better for the Universal Plan with corresponding values of 1.0% to 18.4%, for Models A and B, respectively.

Inclusion of risk sharing greatly improves fit. Such changes are significant even for the basic demographic model (Model A). With the $25,000 threshold in the Current Plan, R^2 jumps from 0.2% to 61.0%, and the Universal Plan from 1% to 68.2%. In Model B fit also increases, and in the Current Plan goes from 3.5% to 63.2% in terms of adjusted R^2, while for the Universal Plan from 18.4% to 78.4%. Fit measured by CPM improves as well, but not to the same degree, with values changing from 9.4% to 17.4%, and 31.2 % to 41.6%, for each plan respectively. In general terms, the improvement of the R^2 measure can be partly attributed to the nature of the indicator, which is based on squared differences between revenue and highly skewed costs at the individual level. Squaring weights enlarges differences more, and it is exactly the large differences that are targeted by our risk-sharing. As expected, raising the threshold for risk sharing to $50,000 decreases fit somewhat (i.e. having less expenditures shared), with, for example, R^2 falling to 51.4% and 57.8% in Model A, and 54.1% and 70.5% for Model B - in each plan, respectively.

Notably, these results for fit measures are high by international standards, for two reasons. First, in most cases, researchers report only an R^2 from a regression, which does not capture any fit contributed by risk sharing. When added fit contributed by risk sharing is taken into account, fit of any system with risk sharing will be higher than the R^2 from a regression. In a recent paper on Switzerland, for example, Beck et al. (2020) find that when the fit effect of Swiss risk sharing is taken into account, the overall fit of the payment system rises from just above 20% as measured by the R^2 from a regression to 57%. Substantial improvements in fit have also been observed in work by McGuire, Schillo & Van Kleef (2020), and McGuire, Zink

and Rose (2021). The second reason for the big jump in fit in our case is the extreme skewness of Chilean health care spending, which may be more skewed than in other countries, implying that a very larger portion of the total variation in spending is concentrated among the high spenders.

Table 2: Results of statistical performance indicators of the demographic (Model A) and diagnostic model (Model B) for the two health coverage plans

		No RS		RS $50k		RS $25k	
Specification		% R^2 adj.	% CPM	% R^2 adj.	% CPM	% R^2 adj.	% CPM
Current Plan							
	Model A	0.2	3.6	51.4	8.8	61.0	12.0
	Model B	3.5	9.4	54.1	14.4	63.2	17.4
Universal Plan							
	Model A	1.0	4.2	57.8	11.5	68.2	15.6
	Model B	18.4	31.2	70.5	38.0	78.4	41.6

Note: RS = Risk Sharing

4.4 Residual spending

In the previous section, we demonstrated the power of risk sharing, even in a data-poor context, when only risk adjusters such as those in Model A (age and gender) are available. In what follows, given that in our case study, it is feasible to include some form of diagnoses, the subsequent analysis focuses on these models (Model B). The results of Model A with and without the inclusion of risk sharing, for both Current Plan and Universal Plan can be found in the Appendix (Tables A2, A4, & A5).

a. Insurer-level

Table 3 presents residual spending at the insurer level for the Current Plan with and without risk sharing. Insurers are ranked by their average expenditure in the respective plan. Table A2 in the Appendix presents results for the Current Plan and Universal Plan under Model A. Financial results give dollars per person that an insurer would be over- or under-compensated. First, without risk equalization and with community rating, closed Isapres (shown as one group) and Isapres 2, 3, 5 and 6 would be undercompensated, while Isapres 1 and 4 would be overcompensated. Risk equalization redistributes funds. Adjusting for diagnostic information (Model B) closed Isapres and Isapre 3 are the most affected in a positive way. In comparison to the no risk equalization case, all other Isapres are less well-compensated.

Table 3: Residual spending for the Current Plan, Model B

		Financial Result ($ per person)			
	% of total	No RE	No RS	RS $50k	RS $25k
Closed Isapres	2.7	-25.6	6.2	3.3	1.7
Isapre 3	4.4	-11.3	20.4	17.9	16.3
Isapre 5	20.8	-8.9	0.7	-0.2	-0.3
Isapre 2	20.6	-8.2	-11.6	-11.6	-11.1
Isapre 6	20.2	-1.3	-2.5	1.4	3.2
Isapre 1	17.9	13.3	3.9	3.4	2.6
Isapre 4	13.4	19.5	7.2	5.1	3.3
Total	3,396,919				

Note: RE = Risk Equalization; RS = Risk Sharing
Insurers are ranked by their average expenditures in the Current Plan.

Adding risk sharing to Model B tends to shrink residual spending towards zero with the exception of Isapre 2 where there is little change in any of the payment alternatives from the no risk equalization case, and with the exception of the small Isapre 3 which benefits from both risk equalization and risk sharing.

Following practice in the literature on risk equalization, we do not test the statistical significance of the results in Table 3 (or in results tables below). We acknowledge that the particular redistributions observed will be sensitive to where high-spending outliers happen to appear in any one year. The main point of the Table is that addition of risk sharing for those high outliers compresses the win/loss discrepancies for the industry as a whole.

Table 4 presents residual spending for the Universal Plan. Compared to the Current Plan, a different pattern of over and undercompensated reflects the different nature of the expenses captured in each package. Moreover, residual spending seems more pronounced in absolute terms in comparison to the Current Plan simply in part because the level of spending is larger. With no risk equalization, under and over-compensations ranges from -$580 to +$212 dollars per person. Conversely, Model B which includes diagnostic information, significantly reduces these figures to a range of -$415 to +$137 dollars per person. In all payment models, the closed Isapres and the small Isapres 3 would see the largest net negative payments. Augmenting models with risk sharing has little effect on the magnitudes of Isapres residual spending in the Universal Plan.

Table 4: Residual spending for the Universal Plan, Model B

	Financial Result ($ per person)			
	No RE	No RS	RS $50k	RS $25k

Closed Isapres	-580.8	-415.0	-431.9	-428.3
Isapre 3	-448.6	-131.8	-129.1	-123.6
Isapre 1	-108.5	-159.8	-143.7	-133.8
Isapre 2	-33.9	-34.2	-40.6	-42.3
Isapre 5	22.6	128.3	123.7	119.3
Isapre 4	105.4	-13.4	-6.4	-3.6
Isapre 6	212.2	137.0	131.1	124.9

Note: RE = Risk Equalization; RS = Risk Sharing

Insurers are ranked by their average expenditures in the Universal Plan

b. Group-level

The main goal of risk equalization is to address risk selection incentives. The inclusion of high-cost risk sharing can be expected to have more of an impact on selection incentives for high-cost groups. Table 5 presents residual spending according to days of hospitalization and comorbidities under the Universal Plan (the corresponding results for the Current Plan can be found in the Appendix Table A3. Table A4 in the Appendix contains the results for the Current Plan and Universal Plan under Model A). Results are presented in terms of dollars per person that an insurer would be over- or under-compensated for a particular group, and therefore, indicates how alternative payment models impact in those group-level profits and losses.

One factor affecting total payments is the size of the fund (see Table 3 above for the share of each Isapres in the total population). In the absence of risk equalization, low spenders, such as those with no hospitalization or no comorbidities are very profitable for the insurers, while individuals with hospitalization or comorbidities exhibit considerable losses with a clear gradient as the number of days or comorbidities increase. The addition of risk equalization, in both of the groups considered, reduces losses (and gains). Bigger changes are attributed to the Model B which includes diagnostic information. Still, some large under and overpayment remain after Model B risk equalization. For example, while under-compensation for those with +30 days of hospitalization falls from -$54,920, it remains at -$37,177, indicating a very large under-payment for this group. Similarly, for those with +3 comorbidities underpayment falls from -$69,641 to -$24,151, but a significant underpayment remains for this group. Overall, risk equalization eliminates the over-payment for healthy groups (those with no hospitalization and no comorbidities), and reduces underpayment for sicker groups, but large underpayments remain for sicker individuals.

High-cost risk sharing further reduces underpayments for members of very sick groups. Those with +30 days of hospitalization exhibit a consistent reduction in the under-payments from Model B to the inclusion of the highest risk sharing threshold. In the case of the comorbidity groups, underpayment for the very sickest group, those with 3+ comorbidities, falls significantly from -$24,151 to -$5,833. Moreover, the average over and underpayment moves consistently towards zero for all four groups.

Interestingly, the groups with short hospitalizations and just one comorbidity flip from being underpaid with risk equalization only to being overpaid (though to a lower absolute amount) with risk sharing. This is likely a consequence of the risk equalization model in which diagnoses are drawn from hospital discharges, and payment weights for a particular hospital discharge will tend to overpay those with a short stay (given that diagnosis) and underpay those for a long stay. Notably, risk sharing reduces the overpayments for these groups (by decreasing the weight on the morbidity indicators).

Table 5: Financial revenues of the Universal Plan, Model B

		Financial Result ($ per person)			
	% of total	No RE	No RS	RS $50k	RS $25k
Days of hospitalization					
0 days	92.8	421.8	-25.7	-12.4	-5.2
1 to 5 days	6.2	-3,218.8	1,625.1	1,203.4	946.7
6 to 10 days	0.6	-9,180.4	-335.0	-1,242.2	-1,525.2
10 to 30 days	0.3	-20,629.3	-7,639.8	-7,653.6	-6,489.8
+30 days	0.1	-54,920.1	-37,176.9	-23,314.6	-17,989.8
Comorbidities					
0	95.6	291.5	-20.4	-12.2	-8.2
1	4.0	-5,207.3	759.6	466.5	314.0
2	0.3	-21,202.5	-3,208.9	-2,180.2	-1,519.7
+3	0.0	-69,641.8	-24,151.3	-10,603.8	-5,833.3

Note: RE = Risk Equalization; RS = Risk Sharing

4.5 Risk equalization payments

Finally, Table 6 presents information on total (not per-person) risk equalization payments each insurer, on net, would contribute or receive from the fund. Figures correspond to the differences between risk-adjusted and average across-all-insurer risk-adjusted expenditure for the two plans under Model B. A negative sign of the difference indicates an insurer would contribute to, whereas a positive sign implies that an insurer would receive from, the fund. (Appendix Table A5 contains the results for Model A).

Payment models have significant implications for individual Isapres. Under the Current Plan, Isapre 1 would contribute an amount of $5.7 million under Model B with no risk sharing, increasing to $6.5 million with risk sharing with the lower threshold (RS $25k). By contrast, Isapres 5 would receive around $6.8 million under model B with no risk sharing, a value which decreases slightly under risk sharing. Small differences are observed between the risk sharing thresholds, except for Isapres 6, which nearly doubles what it would receive from the fund under the $25,000 threshold. As a reminder, in the presence of risk sharing, the resulting inter-plan transfers in any one year will be affected by the very high-cost cases appearing in a plan.

Table 6: Risk equalization payments (in $millions), Model B

	Current Plan			Universal Plan		
	No RS	RS $50k	RS $25k	No RS	RS $50k	RS $25k
Closed Isapres	2.9	2.6	2.5	15.2	14.0	13.6
Isapre 1	-5.7	-6.0	- 6.5	-31.2	-15.4	-21.4
Isapre 2	-2.4	-2.4	- 2.0	-0.2	-5.8	-4.7
Isapre 3	4.7	4.4	4.1	47.3	48.6	47.7
Isapre 4	-5.6	-6.6	-7.3	-54.0	-49.6	-50.9
Isapre 5	6.8	6.1	6.1	74.6	68.3	71.3
Isapre 6	-0.8	1.8	3.1	-51.7	-60.0	-55.8
Total	0	0	0	0	0	0

Note: RE = Risk Equalization; RS = Risk Sharing

Turning to the Universal Plan, we note that the absolute level of the equalization payments is much larger than for the Current Plan. Also, payment positions (who contributes and receives) change (observed in signs reversals through the models for a specific insurer), are observed when moving from the Current Plan to the expanded Universal Plan, particularly for Isapres 6 in the models that include risk sharing.

Moving to risk sharing, some Isapres will either reduce (i.e. Isapres 1, Isapres 4) or increase (i.e. Isapres 2, Isapres 6) the amounts they contribute, while others will get lower compensations (i.e. closed Isapres, Isapres 5), implying that high-risk enrollees are unevenly distributed among the insurers.

5. Conclusion and discussion

This paper constructs a novel risk equalization scheme, adding new risk adjusters and incorporating risk sharing, using data from all privately insured in Chile. Comparisons are

made for two benefit packages, the Current Plan, composed of "GES services" and a version of a Universal Plan, an expanded package covering a comprehensive set of health care expenses, as contained in current reforms. The paper proposes, estimates, and evaluates a feasible model adaptable to settings where information and institutional capacities are in development, as is the case of Chile. Performance with respect to goodness of fit and risk selection incentives based on residual spending for insurers and population groups are assessed. We use these results to evaluate and comment on the performance of the alternative insurer payment models.

Adding morbidity adjustors to the risk equalization formula improved fit measures. Nevertheless, the biggest improvements were obtained by the addition of high-cost risk sharing in the form of reinsurance for both packages, even in the demographic model. In our data, a higher reinsurance threshold, decreasing the share of funds devoted to reinsurance, does not substantially reduce fit, implying that in evaluating the tradeoff between better fit and decreased incentives for cost control, the preferred option for Chile is likely to be the higher threshold of $50k. Still, the country should consider improving its data collection so that over time, morbidity-based adjustors could do more of the work and rely less on risk sharing. Availability of morbidity-based adjustors still leaves the regulator with the need to evaluate the tradeoff between improvements in fit and the adverse incentives associated with either form of risk equalization. Risk sharing can reduce incentives for cost control. Morbidity-based adjustors can introduce incentives for additional treatments and for "upcoding" diagnoses.

High-cost risk sharing may be particularly useful in countries and sectors where diagnostic data are unreliable or not available universally. In our context in Chile, setting relatively high thresholds for risk sharing accomplished significant improvements in overall fit affecting only a small share of spending and an even smaller share of the population. Cost control incentives for the large bulk of health care would be unaffected by this selectively targeted high-cost risk sharing. Thresholds chosen here are for illustrative purposes. Ideally, the threshold would be customized for the particular country application by observing some of the statistics we present: share of total funds and share of individuals affected. The regulator should therefore examine the local spending distribution and the number of individuals in different points of the distribution. Performing evaluations of the values being considered using the fit metrics proposed here could provide insights to the effects of the thresholds.

At the insurer level, some over/under compensation remained after improved risk equalization. Some of the over/under-compensation may be due to relative efficiency of insurers at providing care for persons of comparable risk. While we cannot rule this out, given the present structure of the market (i.e. where underwriting is allowed and there is a high degree of product differentiation (Velasco, Henriquez, & Paolucci, 2018)), risk selection based on personal characteristics not in the model is the more likely explanation.

Possibly the most interesting set of results are the ones relative to risk selection incentives at the group level. Improvements in the formula have a direct impact on the projected compensation received by high risks groups. We find that adding morbidity adjustors supplemented by high-cost risk sharing significantly reduce the opportunity to exploit individual patient information and engage in profitable risk selection.

Insurers are interested in the consequences alternative payment models have for them. Although the models we simulated are all balanced budget in aggregate, risk equalization creates winners and losers in relation to an equal payment to all insurers independent of enrollee risk characteristics. The demographic-only model is less effective at equalizing risk, therefore implying less redistribution among insurers. More effective risk equalization reallocates more of the funds. This indicates that the risk equalization is doing more work, that is, redistributing funds to those insurers whose risk profile is higher, and away from low-risk insurers. Our results on insurer redistribution should be interpreted as providing advantages to the entire insurance sector in the form of protection from the financial effects of high-cost outliers, rather than implying particular insurer winners and losers from our policy options.

While our approach to developing a payment model for settings in which morbidity data are limited may have general applications, the implications of our results for the specific Chilean context are also significant. Our paper shows that the current formula for GES services is inadequate, and if the same formula were applied to the more comprehensive plan, significant over and under-compensation for different risk groups would result, creating strong incentives to insurers to engage in risk selection. Making use of the morbidity data that are available when supplemented by high-cost risk sharing can significantly reduce incentives for selection.

Chile's ongoing discussions over reforms to its healthcare system, have had some consensus that a better risk equalization method is necessary to avoid undesired behaviors by insurers

related to risk selection. Policymakers have concerns over the feasibility of the implementation of an effective scheme with the available information. Our research implies that it is possible, in the Chilean context, to implement a system that is effective in diminishing incentives for risk selection.

References

Armstrong, J. (2018). Health Plan Payment in Ireland. In T. McGuire, & R. van Kleef, *Risk Adjustment, Risk Sharing and Premium Regulations in Health Insurance Markets: Theory and Practice.* Elsevier.

Ash, A., Porell, F., Gruenberg, L. S., & Beiser, A. (1989). Adjusting Medicare Capitation Payments Using Prior Hospitalization Data. *Health Care Financing Review*, 17-29.

Atal, J.P., (2019). Lock-in Dynamic Health Insurance Contracts: Evidence from Chile. *PIER Working Paper 19-020.* University of Pennsylvania.

Bauhoff, S., Rodriguez-Bernate, I., Gopffarth, D., Guerrero, R., Galindo-Henriquez, I., & Nates, F. (2018). Health Plan Payment in Colombia. In T. McGuire, & R. van Kleef, *Risk Adjustment, Risk Sharing and Premium Regulations in Health Insurance Markets: Theory and Practice.* Elsevier.

Beck, K., Kauer, L., McGuire, T., and Schmid, C. (2020). Improving Risk Equalization in Switzerland: Effects of Alternative Reform Proposals on Reallocating Public Subsidies for Hospitals, *Health Policy*, 124: 1363-1376.

Brammli-Greenberg, S., Glazer, J., & Shmueli, A. (2018). Regulated Competition and Health Plan Payment Under the National Health Insurance Law in Israel - the Unfinished Story. In T. McGuire, & R. van Kleef, *Risk Adjustment, Risk Sharing and Premium Regulation in Health Insurance Markets: Theory and Practice.* Elsevier.

Brammli-Greenberg, S., Glazer, J., & Waitzberg, R. (2019). Modest Risk-Sharing Significantly Reduces Health Plans' Incentives For Service Distortion. *The European Journal of Health Economics*, 1359-1374.

Ellis, R., Ibern, P., & Wasem, J. V. (2008). *Panel de Expertos Para la Evaluación del Fondo de Compensación Solidario Entre Isapres.* Ministerio de Salud.

Ellis, R., Martins, B., & Rose, S. (2018). Risk Adjustment for Health Plan Payment. In T. McGuire, & R. van Kleef, *Risk Adjustment, Risk Sharing and Premium Regulations in Health Insurance Markets: Theory and Practice*. Elsevier.

Geruso, M., & McGuire, T. (2016). Tradeoffs in the Design of Health Plan Payment Systems: Fit, Power and Balance. *Journal of Health Economics*, 1-19.

Geruso, M., & Layton, T. (2020). Upcoding: Evidence from Medicare on Squishy Risk Adjustment. *Journal of Political Economy,* 984-1026.

Henriquez, J., Velasco, C., Mentzakis, E., & Paolucci, F. (2016). Más Equidad y Eficiencia en Isapres: Evaluación y Propuestas al Mecanismo de Compensación de Riesgos. *Debates de Política Pública 18, Centro de Estudios Públicos*.

Juhnke, C., Bethge, S., & Muhlbacher, A. (2016) A Review on Methods of Risk Adjustment and their Use in Integrated Heathcare Systems. *International Journal of Integrated Care*, 1-18.

Layton, T., Ndikumana, A., & Shepard, M. (2018). Health Plan Payment in Medicaid Managed Care: A Hybrid Model of Regulated Competition. In T. McGuire, & R. van Kleef, *Risk Adjustment, Risk Sharing and Premium Regulations in Health Insurance Markets: Theory and Practice.* Elsevier.

McGuire, T., Schillo, S., & Van Kleef, R. (2020). Reinsurance, Repayments and Risk Adjustment in Individual Health Insurance. Germany, the Netherlands and the US Marketplaces. *American Journal of Health Economics.*

McGuire, T., & Van Kleef, R. (2018a). *Risk Adjustment, Risk Sharing and Premium Regulations in Health Insurance Markets: Theory and Practice.* Elsevier.

McGuire, T., & Van Kleef, R. (2018b). Risk sharing. In T. McGuire, & R. van Kleef, *Risk Adjustment, Risk Sharing and Premium Regulations in Health Insurance Markets: Theory and Practice.* Elsevier.

McGuire, T., Zink, A., & Rose, S. (2021). Improving the Performance of Risk Adjustment Systems. *American Journal of Health Economics*.

Ministerio de Salud. (2011, December 7). Boletín N 8105-11 Modifica el Sistema Privado de Salud, incorporando un Plan Garantizado.

Ministerio de Salud. (2019, April 22). Formula Indicación Sustitutiva al Proyecto de Ley que Modifica el Sistema Privado de Salud, Incorporando un Plan Garantizado (Boletín N 8105-11).

Paolucci, F., Sequeira, A., Fouda, A., & Matthews, A. (2018). Health Plan Payment in Australia. In T. McGuire, & R. van Kleef, *Risk Adjustment, Risk Sharing and Premium Regulation in Health Insurance Markets: Theory and Practice*. Elsevier.

Pardo, C. (2019). Health Care Reform, Adverse Selection and Health Insurance Choice. *Journal of Health Economics*, 67.

Park, S., & Basu, A. (2018). Alternative Evaluation Metrics for Risk Adjustment Methods. *Health Economics*, 984-1010.

Pope, G., Kautter, J., Ingber, J., Freeman, S., Sekar, R., & Newhart, C. (2011). *Evaluation of the CMS-HCC Risk Adjustment Model. In Final Report, RTI Project Number 0209853.006.* RTI International.

Rose, S., Bergquist, S., & Layton, T. (2017). Computational Health Economics For Indentification of Unprofitable Health Care Enrollees. *Biostatistics*, 682-694.

Schillo, S., Lux, G., Wasem, J., & Buchner, F. (2016). High Cost Pool or High Cost Groups-How to Handle the High(est) Cost Cases in Risk Adjustment Mechanism? *Health Policy,* 141-147.

Schmid, C., Beck, K., & Kauer, L. (2018). Health Plan Payment in Switzerland. In T. McGuire, & R. van Kleef, *Risk Adjustment, Risk Sharing and Premium Regulations in Health Insurance Markets: Theory and Practice.* Elsevier.

Van de Ven, W., & Ellis, R. (2000). Risk Adjustment in Competitive Health Plan Markets. *Handbook of Health Economics.* Elsevier.

Van de Ven, W., Van Vliet, R., & Van Kleef, R. (2016). How Can the Regulator Show Evidence of (no) Risk Selection in Health Insurance Markets? Conceptual Framework and Empirical Evidence. *European Journal of Health Economics,* 167-180.

Van Kleef, R., & Van Vliet, R. (2012). Improving Risk Equalization Using High Cost as a Health Indicator. *Medical Care*, 140-144.

Van Kleef, R., Eijkenaar, F., & Van Vliet, R. (2019). Selection Incentives for Health Insurers in the Presence of Sophisticated Risk Adjustment. *Medical Care Research and Review,* 1-22.

Van Kleef, R., Eijkenaar, F., Van Vliet, R., & Van de Ven, W. (2018). Health Plan Payment in the Netherlands. In T. McGuire, & R. Van Kleef, *Risk Adjustment, Risk Sharing and Premium Regulations in Health Insurance Markets: Theory and Practice.* Elsevier.

Van Kleef, R., Schut, F., & Van de Ven, W. (2018). Premium Regulation, Risk Equalization, Risk Sharing and Subsidies: Effects on Affordability and Efficiency. In T. McGuire, & R. van Kleef, *Risk Adjustment, Risk Sharing and Premium Regulations in Health Insurance Markets: Theory and Practice.* Elsevier.

Van Veen, S., Van Kleef, R., Van de Ven, W., & Van Vliet, R. (2015). Is There One Measure-of-Fit That Fits All? A Taxonomy and Review of Measures-of-Fit for Risk-Equalization Models. *Medical Care Research and Review, 72*(2), 220–243.

Velasco, C., Henriquez, J., & Paolucci, F. (2018). Health Plan Payment in Chile. In T. McGuire, & R. van Kleef, *Risk Adjustment, Risk Sharing and Premium Regulations in Health Insurance Markets: Theory and Practice.* Elsevier.

Wasem, J., Buchner, F., Lux, G., & Schillo, S. (2018). Health Plan Payment in Germany. In T. McGuire, & R. van Kleef, *Risk Adjustment, Risk Sharing and Premium Regulations in Health Insurance Markets: Theory and Practice.* Elsevier.

Appendix

Table A1: Results of statistical performance indicators on alternative thresholds

		RS $5k		RS $10k	
Specification		% R^2 adj.	% CPM	% R^2 adj.	% CPM
Current Plan					
	Model A	80.0	30.5	72.7	20.4
	Model B	81.0	34.6	74.3	25.1
Universal Plan					
	Model A	82.2	31.6	77.5	23.3
	Model B	87.5	53.5	84.7	47.6

Note: RE = Risk Equalization; RS = Risk Sharing

Table A2: Residual spending for the Current Plan and Universal Plan, Model A

	Financial Result ($ per person)					
	Current Plan			Universal Plan		
	No RS	RS $50k	RS $25k	No RS	RS $50k	RS $25k
Closed Isapres	24.2	20.8	18.4	-276.1	-308.8	-316.2
Isapre 1	10.6	9.5	8.4	-103.4	-92.2	-86.0
Isapre 2	-11.0	-11.0	-10.5	-45.0	-49.8	-50.6
Isapre 3	0.9	-0.4	-1.1	-371.6	-345.9	-325.5
Isapre 4	7.1	5.2	3.6	29.7	31.7	31.8
Isapre 5	-7.1	-7.7	-7.3	25.8	29.3	30.8
Isapre 6	1.0	4.5	6.2	208.5	197.3	187.6

Note: RE = Risk Equalization; RS = Risk Sharing
The table shows residual spending at the insurer level for the Current Plan and Universal under Model A and with the two thresholds of risk sharing. Financial results give dollars per person that an insurer would be over- or under-compensated.

Table A3: Financial revenues of Current Plan, Model B

			Financial Result ($ per person)		
	Freq (% of total)	No RE	No RS	RS $50k	RS $25k
Days of hospitalization					
0 days	3,152,042 (92.8)	36.4	-0.2	-0.2	-0.1
1 to 5 days	210,043 (6.2)	-153.3	169.4	149.5	133.5
6 to 10 days	19,577 (0.6)	-903.0	-86.7	-127.1	-145.8
10 to 30 days	10,619 (0.3)	-2,347.1	-658.2	-721.5	-679.3
+30 days	4,638 (0.1)	-8,619.1	-5,612.2	-4,454.1	-3,784.4
Comorbidities					
0	3,249,010 (95.6)	26.6	0.0	0.0	0.1
1	137,556 (4.0)	-483.1	-2.4	-3.1	-4.7
2	10,113 (0.3)	-1,866.0	23.0	23.8	36.3
+3	239 (0.0)	-5,083.4	79.2	159.0	187.5

Note: RE = Risk Equalization; RS = Risk Sharing
The table shows financial by days of hospitalization and comorbidities under the Current Plan for Model B. Financial results group-level profits and losses.

Table A4: Financial revenues of the Current Plan and Universal Plan, Model A

	Financial Result ($ per person)					
	Current Plan			Universal Plan		
	No RS	RS $50k	RS $25k	No RS	RS $50k	RS $25k
Days of hospitalization						
0 days	34.6	32.4	30.3	403	371.6	350.1

1 to 5 days	-136.9	-137.5	-135.9	-3,024.9	-3,032	-3,022.7
6 to 10 days	-844.3	-835.3	-803.2	-8,704.5	-8,538	-8,139.7
10 to 30 days	-2,263.5	-2,218.9	-2,065.4	-20,019.9	-18,211.9	-15,837
+30 days	-8,532.5	-7,169.1	-6,298.1	-54,329.9	-37,497.3	-30,401.2
Comorbidities						
0	25.1	23.4	21.7	279	252	234.1
1	-451.8	-423.6	-395.5	-4,954.4	-4,621.9	-4,383.3
2	-1,784.4	-1,646.6	-1,492.9	-20,618.4	-16,984.7	-14,709.3
+3	-4,946.8	-4,465.5	-4,024.8	-68,762.3	-47,558.3	-38,038.2

Note: RE = Risk Equalization; RS = Risk Sharing
The table shows financial by days of hospitalization and comorbidities under the Current Plan for Model B. Financial results group-level profits and losses.

Table A5: Risk equalization payments (in $millions), Model A

	Current Plan			Universal Plan		
	No RS	RS $50k	RS $25k	No RS	RS $50k	RS $25k
Closed Isapres	4.6	4.2	4.0	27.9	24.9	24.2
Isapre 1	-1.7	-2.3	-2.9	3.1	9.9	13.6
Isapre 2	-2.0	-1.9	-1.6	-7.8	-11.1	-11.6
Isapre 3	1.8	1.6	1.5	11.5	15.3	18.4
Isapre 4	-5.6	-6.5	-7.2	-34.4	-33.5	-33.5
Isapre 5	1.3	0.88	1.1	2.2	4.7	5.8
Isapre 6	1.6	4.0	5.2	-2.5	-10.3	-16.9
Total	0	0	0	0	0	0

Note: RE = Risk Equalization; RS = Risk Sharing
The table shows payments each insurer, on net, would contribute or receive from the fund.

Table A6: Description of 78 diagnostic subgroups.

Short	Name	Icd-9-cm code	Number
GENi	Genital prolapse	618	1
PRGa	Pregnancy-related problems	630-676, 760-779	2
GENh	Disorders of breast	610-611	3
DIGh	Appendicitis, excluding other disease	540-542	4
NRVf	Diseases of ear and mastoid processes	380-389	5
GENg	Diseases of male genitalia, except prostate	603-609	6
DIGI	Cholelithiasis and other disorders of gallbladder and biliary tract	574-576	7
DIGe	Hernia of abdominal cavity	550-553	8
RSPc	Other diseases of respiratory tract	470-478	9
NEOe	Benign neoplasms	210-229	10
GENf	Prostate disorders	600-602	11
NRVb	Cataract	366	12
NRVd	Disorders of the peripheral nervous system	350-359	13
GENc	Urethral stricture	598	14
GENj	Other diseases of female genital tract, excluding prolapse and abnormal bleeding	617, 619-629	15
NRVe	Disorders of the eye and adnexa, except for NRVb and NRVg	360-364, 367-379	16
NRVg	Glaucoma	365	17
GENd	Hydronephrosis, calculus of kidney and ureter, other disorders of kidney and ureter, calculus of lower urinary tract	591-594	18
INJa	Injuries involving fractures and dislocations	800-839	19
MSKd	Rheumatism, excluding the back and polymyalgia rheumatism	726-729	20
MSKc	Dorsopathies, except for inflammatory spondylopathies	721-724	21
GENb	Kidney infections	590	22
INJd	Superficial injury, contusions, effect of object entering through orifice	910-919, 920-924, 930-939	23
SYMa	General symptoms	780	24
INJc	Intracranial injury (excluding skull fracture, internal injury, open wound, injury to blood and vessels, late effects of injuries, poisonings, toxic effect and external causes, crushing injury, burns, injury to nerves, spinal cord, unspecified injuries	850-909, 925-929, 949-949, 959-959	25
DIGa	Diseases of oral cavity, salivary glands, and jaws	520-529	26
RSPb	Acute respiratory infections, except bronchitis	460-465	27
DIGf	Noninfective enteritis and colitis	555-556, 558	28
CNGa	Congenital anomalies	740-759	29
INJb	Sprains and strains of joints and muscles	840-848	30

NEOc	Malignant neoplasm of bone, skin, cartilage, soft tissue, male and female breasts, skin melanoma and carcinoma	170, 171, 173, 174, 175, 172	31
DIGg	Intestinal obstruction (nonherniated), diverticula of intestine, peritonitis and other disorders of intestine or peritoneum	560-562, 567-569	32
GENe	Disorders of urethra and urinary tract, urinary symptoms and nonspecific finding on urine examination	595-597, 599, 788, 791	33
DIGb	Gastric, duodenal, peptic and gastrojejunal ulcer, diseases of the esophagus, gastrointestinal hemorrhage	530-534, 578	34
CRCh	Diseases of veins and lymphatics and diseases of circulatory system, anal fissures, fistulae and anal or rectal abscess	451-459, 565-566	35
INGa	Infectious diseases except those in INFb	001-139	36
DIGc	Gastritis and duodenitis and other disorders of stomach and duodenum, functional digestive disorders, digestive symptoms	535-537, 564, 787	37
ENDe	Diseases of other endocrine glands, disorders of the thyroid gland, nutritional and metabolic symptoms, nonspecific findings on blood examinations	251-254, 256, 257, 259, 240-246, 783, 790	38
CRCe	Cerebrovascular disease	430-438	39
SYMc	Respiratory symptoms	786	40
RSPd	Pneumonia and influenza	480-487	41
SKNa	Diseases of the skin and subcutaneous tissue, symptoms involving skin	680-709, 782	42
SYMd	Symptoms involving abdomen and pelvis	789	43
MSKa	Diseases of connective tissue, rheumatoid arthritis and inflammatory polyarthropathies, osteoarthrosis and like disorders rheumatic fever and rheumatic heart diseases, polyarthritis and like conditions, inflammatory spondylopathies, polymyalgia rheumatica, symptoms of nervous and musculoskeletal system	710, 714, 715, 390-398, 446, 720, 725, 781	44
MSKb	Various arthropathies, disorder and derangement of joints, osteopathies, chondroplasties and acquired musculoskeletal deformities	711-713, 716, 717-719, 730-739	45
GENk	Inflammatory diseases of female pelvic organs	614-616	46
CRCb	Ischemic heart disease, diseases of pulmonary circulation, cardiovascular symptoms	410-414, 415-417, 785	47
CRCg	Aortic aneurysm, other aneurysms, arterial embolism and thrombosis	441, 442, 444	48
MNTb	Neurotic, personality and sexual disorders, alcohol and drugs dependence and abuse, other personality disorders, non-psychotic mental disorders and mental retardation	300-302, 303-305, 306-319	49
SYMe	Nonspecific abnormal findings, other ill-defined and unknown causes of morbidity and mortality	792-796, 799	50
EVLa	Supplementary classification of factors influencing health status	V01-V82	51
RSPa	Acute bronchitis and bronchiolitis	466	52

MNTa	Organic psychotic conditions, other psychoses, senility without psychosis	290-294, 295-299, 797	53
DIGj	Diseases of pancreas, intestinal malabsorption	577579	54
CRCa	Hypertensive disease	401-405	55
CRCf	Atherosclerosis, other vascular disease, disorders of arteries, arterioles and capillaries, vascular insufficiency of intestine	440, 443, 447-448, 557	56
ENDb	Nutritional deficiencies, metabolic disorders, obesity and immune disorders, disorders of blood and blood-forming organs, adrenal and polyglandular disorders	260.269, 270-275, 277, 278, 279, 280-289, 255, 258	57
NEOf	Carcinoma in situ in skin and on other unspecified sites, neoplasms of uncertain behavior unspecified nature, malignant neoplasm of the lip	232, 234, 235-239, 140	58
ENDc	Disorder of fluid, electrolyte and acid base balance	276	59
CRCc	Other forms of heart diseases, except heart failure	420-427, 429	60
INJf	Toxic effects of nonmedical substances, unspecified effects of external causes	990-995	61
NRVa	Other disorders of the central nervous system	340-349	62
INJe	Poisoning by drugs, medicines and biological substances	960-979	63
NRVc	Hereditary and degenerative diseases of central nervous system	330-337	64
INJg	Complications of medical care not elsewhere classified	996-999	65
SYMb	Symptoms involving head and neck	784	66
ENDa	Diabetes mellitus	250	67
NEOa	Malignant neoplasm of digestive organs and peritoneum, carcinoma of digestive organs	150-159, 230	68
NEOd	Malignant neoplasm of genitourinary organs	179-189	69
DIGh	Liver disorders and diseases	570-573	70
INFb	Various infectious and parasitic diseases	013, 038, 045-049, 070, 093-095, 112, 114-116, 135, 320-326	71
ENDd	Other and unspecified anemia's	285	72
RSPe	Chronic obstructive pulmonary disease and like conditions, pneumoconiosis and other lung diseases due to external agents	490-496,500-508	73
RSPf	Other diseases of the respiratory system	510-519	74
CRCd	Heart failure	428	75
NEOg	Malignant neoplasm of oral cavity and pharynx, malignant neoplasm of unspecified sites, malignant neoplasms of lymphatic and hematopoietic tissue	141-149, 190-199, 200-208	76
NEOb	Malignant neoplasm of respiratory and intrathoracic organs	160-165	77
GENa	Nephritis, nephrotic syndrome, and nephrosis	580-589	78
Note: 91,81% Doesn't have ICD-10 code. 37% of codes weren't matched using 4-digit ICD 10.			

Dear Joan Schmit, Editor, and Senior Editor:

Thank you for taking the time to carefully read our paper. We are grateful for your comments and for the opportunity to revise this paper following your guidance. We have made the following changes (summarized):

1) Revised the abstract and introduction to clarify the contribution. We have also moved up the discussion of risk equalization/risk sharing tradeoffs to a paragraph in the introduction.
2) The results section contains results related to the demographic risk equalization model (Model A) and simplified the tables. We have added the remaining models to the appendix and have more clearly described them.
3) We have addressed all the other descriptive points and minor edits.

Below, you can find our detailed responses.

Senior Editor
Comments to the Author:

Dear Josefa, Marica, Thomas, Emmanouil, and Francesco,

Thank you for submitting your paper "Designing feasible and effective health plan payments in countries with data availability constrains" for consideration for the special issue on health-insurer decision making at the Journal of Risk and Uncertainty. I have received a thorough review of the paper by an expert reviewer with familiarity of this work and have read the paper myself. I'm happy to say that both the reviewer and I believe this paper is a good fit for the JRI generally and in particular for the special issue. The reviewer has given a detailed set of suggestions for edits, which are essentially all expositional and quite doable. Here is my specific guidance on edits to strengthen the message of the paper building on what the reviewer has written:

1. The reviewer gives suggestions for the introduction and I would highlight in particular the comments about bringing up more discussion on risk equalization schemes, risk sharing and their tradeoffs into the introduction. I think this would help the paper considerably.

This is an important point. We have added to the introduction the following paragraph: “A relevant consideration that arises with the implementation of risk sharing vis-à-vis the sophistication of the risk equalization formula (e.g., addition of morbidity-based risk adjusters) is incentives for insurer efficiency (i.e., cost-control). As risk sharing will pay insurers for part of their realized spending, cost control incentives are diminished. The development of adjusters which compensate insurers for those high-spending individuals (e.g., diagnosis-based adjusters) would address this problem. Nevertheless, this approach involves additional considerations. Perverse incentives, such as upcoding can be a threat to the payment system (Geruso & Layton, 2020) and consistency with

those factors that are desired in cross-subsidies. As such, a trade-off between feasibility, effectiveness and efficiency emerges."

2. The paper does not clearly enough spell out the nature of the contribution in terms of the proposed approach to risk equalization. There needs to be a clearer statement in the introduction along the lines of "Relative to current practice (or maybe theoretical literature?) our proposed approach is novel because..." My understanding from reading the paper is that standard practice in data-poor regions is to use the cell-based method (Model A in your setup). In more data-rich environments, they use morbidity-based adjustment systems (i.e., individual characteristics). There is a question of whether to use concurrent-year characteristics or prior-year characteristics in these schemes with the tradeoff being that current-year increases fit but dampens insurer incentives to control costs. In practice in this paper due to the "data poor" environment you use concurrent year (Model B), which you also note is how it is done in US Marketplaces. So my read is that your general approach to Model B is not conceptually novel and is similar to risk equalization schemes in US Marketplaces, right? What seems to me to be the real contribution here is not so much the development of a "new risk equalization scheme" but rather the quantification of the large value of having modest risk sharing *given that* the risk equalization scheme cannot be so sophisticated or data rich. I strongly encourage the authors to do a little work for the abstract and introduction (building also on the reviewer's feedback here) to help nail down this contribution in a clearer way for readers.

We agree with this point. Indeed, "the quantification of the large value of having modest risk sharing *given that* the risk equalization scheme cannot be so sophisticated or data rich" is the main contribution of the paper. We have elaborated on this point by modifying the abstract in the following way: "This paper proposes, implements and quantifies the gains of a risk equalization scheme which incorporates risk sharing in a data-poor context." and "Using Chile's private health insurance market as a case study, we show that modest amount of risk sharing greatly improves fit even in a simple demographic-based risk equalization."

We have clarified the contribution in the introduction as well by modifying the following paragraph: "Using Chile as a case study, this paper proposes, implements and quantifies the gains of a risk equalization scheme which incorporates risk sharing. Relative to current practice, our approach is novel as it adapts research and experience from countries with sophisticated (and data-driven) risk equalization formulas (e.g. The Netherlands, U.S. Medicare) to a data-poor setting befitting the health insurance markets in numerous countries around the world where regulation of health insurance markets must take into account data scarcity. "

3. Related to point 2, I wondered why in your results (e.g., Table 2 and on) you do not show how risk sharing improves the system under the more constrained Model A? It feels like the really big point of this paper is that some level of risk sharing can

significantly improve the risk equalization scheme, wouldn't that point also be worth drawing out in the context of the more simplified risk equalization approach too?

We agree. We have added the requested results (Model A with the two risk sharing thresholds) into Table 2 and discuss them at that point in the paper. With this, we show the power of risk sharing, the main contribution of our paper. In what follows, given that we are talking about the Chilean context where it is feasible to include some form of diagnoses – as we do in the paper-, for the subsequent analysis we focus on what Chile can do. This translates into not including all combinations of the estimations (12 models in total – packages, models, and thresholds-) but showing only some of these in the main body of the text.

We include the remaining results for the readers to review in the appendix. We have made this explicit at the beginning of section 4.4 by adding "In the previous section, we demonstrated the power of risk sharing, even in a data-poor context, when only risk adjusters such as those in Model A (age and gender) are available. In what follows, given that in our case study, it is feasible to include some form of diagnoses, the subsequent analysis focuses on these models (Model B). The results of Model A with and without the inclusion of risk sharing, for both Current Plan and Universal Plan can be found in the Appendix (Tables A1, A3, and A4)."

4. The reviewer highlighted that you might think about consolidating and presenting some of the results in Tables 3, 4 & 6 differently since most readers won't be familiar with the names of the different insurance plans in Chile. I think that's a good idea and wonder if there is a way to show these results that sorts things from "insurers with highest-cost enrollment pools" to those with lowest-cost that would be more effective.

We recognize that the tables do not appeal to the reader. We have made several changes to Tables 3, 4 in an attempt to make them easier to digest. First, we have trimmed the frequency and left the % of total and total insured or total in the groups (Table 3 and Table 5). We have narrowed down what we are showing to reflect our decision of showing only Model B in the main body. Insurers have been sorted in terms of their average spending in their pool for the respective plan in study.

I encourage you to consider all of the Reviewer's comments as you do your revision. Ultimately, though, we will leave a lot up to your discretion and hope you see these comments as constructive feedback that we hope makes the paper more effective for JRI readers. I likely will not need to send the paper back to the reviewer after your edits and anticipate being able to read the final version myself to just check that all looks good going forward.

Reviewer(s)' Comments to Author:

Reviewer: 1

Comments to the Author
Referee Report JRI-Jan-21-004
“Designing feasible and effective health plan payments in countries with data availability constraints”

This paper uses 2017 data on the universe of privately insured in Chile to simulate the performance and redistributional consequences of alternative risk equalization and risk sharing schemes. The authors run different model variants for two existing health care plans and assess their performance using the standard measures (i) adjusted R2 and (ii) Cumming’s Prediction measure (CPM). The also calculate and discuss redistribution between insurers as well as over- and undercompensation at the group level (patients with hospital days + comorbidities). Their findings show that, especially in countries with limited data availability, a modest amount of risk sharing greatly improves fit and reduces insurer incentives for risk selection at the group level.

The authors study a topic of high relevance and potentially high policy impact in an international context. The paper is well written and structured, concise and conveys new and important findings for the reader. I consider it a great fit for the JRI special issue. I hope that the following comments and suggestions help to further improve the quality of the manuscript.

1. I felt like the authors should introduce general discussions on the pros and cons of refined risk adjustment already in the Introduction. For example, I enjoyed section 3.7 but believe this should be relegated to the Introduction, maybe also with aspects of footnotes 8 (final sentence) and footnote 23. In general, most readers will not be familiar with details of risk adjustment. The authors should see this as an opportunity to give them a general Introduction into the topic and its relevance.

We thank the reviewer for this comment. We have added a paragraph to reflect the pros and cons of refined risk adjustment in the introduction (see response to senior editor). Regarding the comments made in old footnote 8 (now 12), about the timing of estimation (e.g. concurrent/retrospective/prospective discussion), this is outside the scope of our paper, but we add in a note the following: “Prospective and retrospective models share some of the main trade-offs then the sophisticated vis-à-vis risk sharing risk equalization formulas, regarding effectiveness and efficiency. All these models can be seen as steppingstones to developing models that achieve the right balance in light of the regulator’s objectives. See (Ellis, Martins, & Rose, 2018) for further discussion on the advantages and disadvantages of concurrent risk equalization.”

2. Introduction: Is it just my impression or do most countries start with a simple cell-

based approach and then incorporate more refined morbidity measures over time? Following up on the previous point, could you maybe briefly summarize whether there is expert consensus on what adjusters to not include in risk adjustment, e.g., administrative costs? Likewise, what type of "personal characteristics not in the model" are the prime candidates for active or passive insurer risk selection? What does the literature say on the topic?

We thank the reviewer for this comment. With regard to the characteristics to be included in the model, we have added a note (new footnote 1) on this topic. The paper focuses on the feasibility, but there are other guiding principles to the design of risk equalization: "Other guiding principles include appropriateness of incentives (e.g., selection by insurers) lack of gameability, and clinical meaningfulness. For a general discussion, see (Ellis, Martins, & Rose, 2018)."

3. Introduction: the authors could potentially broaden the reference list to appeal to a wider audience. They are the experts here and one of the authors was editor of a JHE special issue on the topic, so I defer to their best judgement on how to implement this suggestion. For example, I would cite work by Juan Pablo Atal who is an expert on the health insurance market in Chile.

We have added further references in the paper to some very recent work along the lines of ours, and have included a relevant reference from Atal.

4. Section 2: I got confused a couple of times by the institutional details. First, the authors write that insurers use underwriting to select risks. How exactly is that done and priced and what are the average loadings on the risk factors age, gender, and dependent status (not region)? Is the risk rated premium then linked to these "mandatory salary contributions" or is it an absolute premium in pesos? Could you provide a simplified table or graph on this? What are the implications of this risk rating for the potential to select risks, and also for under- and overcompensated care at the group level? Do we have an evidence for active selection by insurers in Chile?

We thank the reviewer for this comment. There are many details in the workings of the Chilean healthcare system. We have provided the following clarification to the questions posed.

- First, the authors write that insurers use underwriting to select risks. How exactly is that done and priced and what are the average loadings on the risk factors age, gender, and dependent status (not region)?

We have added: "Underwriting is done by means of a health status declaration form at the time of enrolment. Insurers use this declaration to determine if the individual is allowed to enrol."

After consideration, we decided that some details would overly burden the paper. We provide the reader with references to Velasco et al (2018) where more detail can be found.

- Is the risk rated premium then linked to these “mandatory salary contributions” or is it an absolute premium in pesos?

We have clarified this point by adding: “Premiums relate to the mandatory 7% in the following way: the contribution could or could not suffice to pay for the premium of the plan. In the former, it could even generate an excess which is saved by the individual in a health savings account with the insurer, and in the latter, s/he would have to pay on an absolute amount in pesos on top. These supplemental contributions average 3% of gross salary. “

- What are the implications of this risk rating for the potential to select risks, and also for under- and overcompensated care at the group level? Do we have an evidence for active selection by insurers in Chile?

There is active selection as people are currently rejected from the Isapres market. But, there is no official number to determine the extent that this occurs. Related to this point, Cid (2011) shows the difference in average risk between private insurers and the public option (Fonasa) which is substantial. (Cid, Camilo. 2011. “Problemas y desafíos del seguro de salud y su financiamiento en Chile: el cuestionamiento a las ISAPRE y la solución funcional”. Temas de la agenda pública 49. Centro de Políticas Públicas UC. Diciembre.)

5. Could you provide more explanations what the mandatory GES services are, and why they only account for 8% of total private expenditures? What are the remaining 92%?

We have added new footnote 9 “Receiving treatment through the GES services plan entails certain restrictions similar to a “managed care” model. Individuals in Isapres can opt out to stay with their usual doctor or provider, and get coverage through other parts of their insurance plan.”

Am I correct that you have the total spending amount per enrollee and year but only diagnostic information for hospital admissions (without associated claims)? In general, I would add a descriptive statistics table to the Appendix that lists all the variables that you see and describe in 3.1.

We have clarified this point by modifying Section 3.1 and adding “Health expenditures are grouped into different categories which allow us to identify GES services and total spending.”. In Section 4.1. the descriptive statistics of all the variables that we use in our models are described.

6. Section 3.2. I stumbled across this detail and was wondering what the implications are for the model. My understanding is that people with multiple admissions and diagnoses simply appear in several of the 79 clinical categories, that is, while Model A and the gender-age cells are mutually exclusive, the clinical information in Model B is

not. How many enrollees with multiple diagnostics or admissions do we have in the data?

The answer to the latter question has been added to footnote 19.

“Alternative grouping methods could at a later stage be compared to the one we apply here (Juhnke, Bethge, & Muhlbacher, 2016).”

Could you run this as a robustness check and show it in the Appendix along with robustness checks that uses alternative thresholds, e.g. $5000 and $10,000?

We have added these robustness checks in Appendix Table A1 and noted in footnote 20.

7. Most readers not familiar with Chile will find Tables 3, 4, and 6 less exciting. I would streamline the discussion and move the tables to an Appendix. Maybe the results could be shown graphically and normalized; that is, as a reader, I would be interested in how much % of overall spending is redistributed between insurers based on method x and y, and how much % of overall insurer revenues for single insurers that comprises. Doesn’t the amount of redistribution also give us a strong indication of selection into plans/insurers? Does the literature have anything to say on this?

We recognized the Tables to be less exciting. This is addressed in an alternative take on the tables. See response to Senior Editor. In addition, regarding the redistribution comment, the reviewer is right. The larger the package, the stronger the gains of risk selecting for insurers. The amount of redistribution tells you how much of risk differences are being picked up by the risk equalization system.

8. I find the Conclusion strong. However, could you specify and discuss the aspect below a bit more? Maybe develop it to some policy guidance or at least a discussion on trade-offs? If a policymaker from country x came to you and would ask “what threshold should I set”, what would you reply?

“Thresholds chosen here are for illustrative purposes. Ideally, the threshold would be customized for the particular country application by observing some of the statistics we present: share of total funds and share of individuals affected.”

We thank the reviewer for this comment. Indeed, it’s a very policy relevant question. We have added the following concrete policy guidance and tradeoff analysis to the above statement in quotes: “The regulator should therefore examine the local spending distribution and the number of individuals in different points of the distribution. Performing evaluations of the values being considered using the fit metrics proposed here could provide insights to the effects of the thresholds.”

Style and Format:

9. The quotation style is sometimes awkward for my taste. For example, when you write “see (Smith, 2020)” I would set it direct quotes. Also, I would always put indirect quotes at the end of sentences.

Addressed where relevant.

10. The references should be consistently formatted. Sometimes the first letter is capitalized, sometimes not.

This has been fixed and now first letters are all capitalized.

11. The notes to the tables are note extensive enough, especially those in the Appendix. The tables should be self-contained. Further, the Appendix tables are not well integrated into a discussion in the main text. Most Table AX do not even appear as reference in the main text.

Addressed.

Manuscript Number: JHE_2020_125
Designing feasible and effective health plan payments in countries with data availability constraints

Dear,
Thank you for submitting your manuscript to Journal of Health Economics.

I regret to inform you that I have decided not to proceed with your manuscript based on the comments of the reviewers. This was a difficult decision because each reviewer identified this as an important and interesting topic and identified relatively minor concerns with the paper. The reviewers, however, were also in agreement that they thought the manuscript did not clear the bar for JHE - it was not adequately innovative with respect to health economics. They did emphasize, however, that the manuscript makes an important contribution to applied health policy. They recommended specific journals including Health Economics, the Journal or Risk and Insurance, and health services research journals as alternative outlets.

As you know, this took much longer than I had hoped, and, once again, I sincerely apologize for that. I ended up with three reviews because each of the reviewers I initially contacted near the beginning of 2020 ultimately provided comments as did a "back-up" reviewer I contacted after our last correspondence. I hope these comments are useful as you update your paper and submit it to another journal. I want to emphasize again that all three reviewers were enthusiastic about your research and the primary concern was an issue of fit for JHE.

Please refer to the comments listed at the end of this letter and the attached for more detail.

We appreciate you submitting your manuscript to Journal of Health Economics and thank you for giving us the opportunity to consider your work.

Kind regards,
Editor
Journal of Health Economics

Prior Review Comments Journal of Risk and Insurance

This document contains the summary of the comments of reviewers and our responses, which have shaped the present version of the original manuscript "Designing feasible and effective health plan payments in countries with data availability constrains". The manuscript was previously submitted to the Journal of Health Economics (only prior submission).

Reviewer 1

Several comments were made on the threshold chosen for risk sharing including what should regulators consider when choosing a threshold, sensitivity analysis for the threshold, and the fit improvement when adding risk sharing, both in relation to the goodness of fit measures and in relation to the international context.

We have included further discussion on decision-making behind the threshold in the conclusion (see page 21) pointing out some statistics that can be used to choose the attachment point. Upon further revision of the estimations, we have conducted some sensitivity analysis (see Footnote 20), which are available upon request. In Footnote 21 we have added a conceptual discussion on the improvements in fit due to risk sharing and contextualized upon our results in the end of page 15- start of page 16). We have added a paragraph on the international context too (see page 16).

A second comment on the diagnostic groups used in the estimation was made, in particular regarding other alternative diagnostic groupings that could be used, about the effects and measurement of the incentives that diagnostic groups have, and on the tradeoff between collection of information for diagnostic groupings vs adding risk sharing.

We have replied by pointing out that there are no other diagnostic groupings available or being discussed in the context and have included a reference to Juhnke et al 2016 where there is a list of alternative models that could also be used (see page 9). In addition, we have included the reasons why we are not assessing incentives in this paper (see page 13), and about the trade off in page 21.

Last, it was suggested to add to the paper those models previously omitted. We have done so in the appendix.

We have addressed nearly small comments regarding typos and stylistic suggestions.

Reviewer 2

A shared comment with reviewer 1 about the risk sharing threshold and the distribution of spending, as well as the international context of our results.

See our Reply to Reviewer 1.

We have better discussed the definitions of R-squared and CPM.

Last, the reviewer asks to contextualize our revenue results (Table 5) with respect to financial indicators of the insurer.

We have added a better discussion on this topic in page 18.

Reviewer 3

A shared comment with reviewer 1 about incentives of the risk equalization and the diagnostic coding was made. Refer to our reply in Reviewer 1.

Its was suggested by the reviewer to add the mathematical form of the risk sharing equation.

We have done so in Formula (1).

Last, there is a comment about quantification of uncertainty.

We have not included standard error or p values as this is not standard practice when measuring over/under compensations, while the coding (which is made available upon request), can be used to easily derive standard deviations of our parameters (see page 17)

We have addressed nearly small comments regarding typos and stylistic suggestions.