Titel der Arbeit: Essays on the Effects of Asymmetric Information and Behavioral Biases

Schriftliche Promotionsleistung zur Erlangung des akademischen Grades Doctor rerum politicarum

vorgelegt und angenommen an der Fakultät für Wirtschaftswissenschaft der Otto-von-Guericke-Universität Magdeburg

Verfasser: Benedikt Mihm, M.Sc.

Geburtsdatum und –ort: 07.05.1982, Friedberg

Arbeit eingereicht am: 12.08.2016

Gutachter der schriftlichen Promotionsleistung:

Prof. Dr. Joachim Weimann

Prof. Dr. Andreas Knabe

Datum der Disputation: 15.02.2017

# Summary

This dissertation consists of four essays that analyze the effects of asymmetric information and behavioral biases. The dissertation start with two essays that use laboratory experiments to study how biases and asymmetric information affect two important economic policy issues: taxpayers' response to taxation and the likelihood of agreement in bargaining over the provision of a public good.

In Essay 1 "*Biased effects of taxes and subsidies on portfolio choices*" (Joint with Hagen Ackerman and Martin Fochmann) an experimental investigation is carried out to evaluate the impact of a tax perception bias in a choice problem under risk. In the experiment, subjects are confronted with a portfolio choice problem of how much to invest in a risk free asset and how much in a risky asset. There are four treatments that vary in regards to the form of government intervention on the risky asset. In three of the treatments the return on the risky asset is either subject to a tax, a subsidy, or a tax and a subsidy, whereas the return on the risk free asset is tax free. In the remaining treatment, the risky asset is neither taxed nor subsidies. Payoffs are chosen such that the net incomes are the same in all four treatments and the setup therefore has a clear theoretical prediction: payoff maximizing subjects should invest the same amount in the risky asset in each treatment. Any deviation from this prediction then indicates a perception bias.

The results of the experiment provide strong evidence of a bias, with the investment in the risky asset being significantly lower in all treatments with some form of government intervention. This main finding is confirmed in a range of variations of the baseline experiment indicating that the bias is quite robust to changes in the choice situation.

A laboratory experiment is well suited to studying the central question of the first essay. The controlled environment makes it possible to setup the taxpayer's problem so that the effect of a bias on outcomes can be carefully extracted. In the next essay the advantages that the controlled environment of the laboratory provides is again utilized to consider the interplay between payoff maximizing and non-payoff maximizing behavior in bargaining problem.

Essay 2 "*To commit or not to commit? An experimental investigation of pre-commitments in bargaining situations with asymmetric information*" (joint with Sönke Hoffmann and Joachim Weimann) is an experimental investigation of a negotiation model developed by Konrad and Thum (2014). The model sets out a bilateral bargaining problem over the provision of a public good. Asymmetric information plays an important role in the setup as it leads to the possibility of the bargaining outcomes being inefficient.

Specifically, the environment is characterized by a sequential game between two parties, A and B, in which both have private information about their contribution cost. Player A makes a transfer offer to B, which B can either accept or reject. Since A does not know B's cost, however, the transfer offer may be too low, leading B to reject the offer and the efficient outcome not being achieved. Analyzing two versions of the game, Konrad and Thum (2014) show that the likelihood of agreement is lower in a version of the game with a pre-commitment to the public good by A.

The experiment is setup to both test the model's main prediction and to evaluate how non-payoff-maximizing behavior affects its main result. To do so, the experiment is run in three settings, each allowing differing amounts of freedom for subjects to deviate from payoff maximizing behavior. The first setting is the closest to the original model. The second allows subjects to choose whether to contribute to the public good even if the transfer offer has been rejected. In the third setting pre-play communication between the subjects is allowed. The two versions of the game (one with and one without pre-commitment) were played in each setting to see if the main prediction of the model about the counterproductive effect of pre-commitment could be verified.

The prediction of the model is indeed shown to hold in all three setting even though significant evidence of non-payoff maximizing behavior was observed, with B players rejecting offers that were payoff maximizing and accepting offers that were not. Despite engaging in non-payoff maximizing behavior subjects therefore responded to the incentive as predicted by the model enough for the main result to nevertheless hold. Based on the results of the second essay it can therefore be concluded that both asymmetric information and non-payoff maximizing behavior are likely to have an effect on outcomes of bargaining problems in relation public goods.

Moving away from the laboratory, the next two essays develop theoretical models to analyze how biases and asymmetric information affect outcomes in another important policy area: sovereign debt. Essay 3 "*Mispricing of risk in a sovereign bond market with asymmetric information*" starts by considering the problem when there is no biased behavior but asymmetric information plays a key role. In the model, a government borrows from international investors but a commitment problem means that the government may not repay. The likelihood of repayment depends on which one of two types (high or low) the government is, with the type being the government's private information. This asymmetric information problem means that the setup can be modeled as signaling game. The government issues debt and the level of borrowing sends a signal to investors about the government's ability to repay. Observing the debt supply investors must in turn infer the likelihood of repayment from the amount of debt issued.

This sovereign debt signaling game has a *pooling equilibrium* in which both types of government borrow the same amount and the borrowing signal is uninformative about the government's type. As a result, the two types receive the same bond price and the bond price does not accurately reflect the government's likelihood of repayment. This disconnect arises even though the default risk is the only determinate of the bond price, so that in the absence of asymmetric information the bond price would fully reflect the default risk.

The results of the third essay highlight how important asymmetric information can be in sovereign bond markets and rationalize the mispricing of risk in bond markets in a setup in which everyone is payoff maximizing. In the next essay the government is again assumed to have private information about its likelihood of default but now a behavioral bias is also introduced.

Essay 4 "*Biased government borrowing and yardstick competition in a sovereign debt market*" develops a model in which government borrows from international investors and can be either safe or risky, with the risky government having a higher probability of default. The government's type is private information but the government may miss-perceive it own likelihood of default: a safe government may perceive itself to be risky, and vice versa. This miss-perception introduces a behavioral aspect into the framework as the government is not fully rational in terms of how much it borrows.

With asymmetric information and biased government borrowing, bond prices become partially disconnected from the default risk even in a *separating equilibrium* in which the two types send different borrowing signals. Since investors must take into account that the government has miss-perceived its own type, the debt levels are only partially informative and bond prices do not fully reflect default risk. In the fourth essay it is therefore the combination of asymmetric information and a behavioral bias that leads to risk mispricing. It is also shown that in a two country version of the model, correlations in the default risk across the two countries creates information spillovers making it either easier or harder to spot a government which has miss-perceived its own type.

Economics Letters 120 (2013) 23–26

Contents lists available at SciVerse ScienceDirect

Economics Letters

journal homepage: www.elsevier.com/locate/ecolet

# Biased effects of taxes and subsidies on portfolio choices

Hagen Ackermann[a], Martin Fochmann[b,*], Benedikt Mihm[a]

[a] Otto-von-Guericke University Magdeburg, Germany
[b] Leibniz University Hannover, Germany

## HIGHLIGHTS

- Experimental study on the effects of taxes and subsidies on portfolio choices.
- Four treatments with either no tax, a tax, a subsidy or a tax and a subsidy.
- Net payoffs identical in all treatments so investment level should be constant.
- Find a highly significant negative impact from both types of intervention.

ARTICLE INFO

*Article history:*
Received 12 November 2012
Received in revised form
12 March 2013
Accepted 22 March 2013
Available online 28 March 2013

*JEL classification:*
C91
D14
H24

*Keywords:*
Tax perception
Risk-taking behavior
Portfolio choice
Distorting taxation
Behavioral economics

ABSTRACT

We study how taxes and subsidies affect portfolio choices in a laboratory experiment. We find highly significant differences after intervention, even though the net income is identical in all our treatments and thus the decision pattern of investors should be constant. In particular, we observe that the willingness to invest in the risky asset decreases markedly when an income tax has to be paid or when a subsidy is paid. We investigate this result further in a range of variations of the baseline experiment and find our main result to be largely robust. However, as we reduce the number of states of nature the bias weakens considerably.



## 1. Introduction

In a recent experiment, Fochmann et al. (2012) find that a tax perception bias influences risk-taking behavior when subjects are able to offset losses from their taxable base. In this paper, we investigate whether a perception bias also has an effect in a more general investment problem with different types of government intervention. We look at the effects of both subsidies and taxes on portfolio choices in a laboratory experiment to see how they influence the choice between risky and risk-free assets. We find that imposing a tax and paying a subsidy both have a highly significant negative effect on the willingness to invest in a risky asset.

This paper adds to a small but growing literature on the effect of biases from government intervention. Chetty et al. (2009), for example, find that consumption decisions are influenced by the salience of sales taxes and show that the resulting distortions may have important welfare effects. Sausgruber and Tyran (2011) also find that biased tax perception can have an impact on welfare in the context of voting decisions. Gamage et al. (2010), Djanali and Sheehan-Connor (2012), and Fochmann et al. (forthcoming) observe that labor market decisions are distorted by a biased tax perception. Our contribution to this literature is twofold: (1) we shed further light on the effect of government intervention on investment decision and (2) we are to our knowledge the first to analyze the effect of subsidy perception on risk-taking.

## 2. Experimental design and hypothesis

In our setting, subjects have to decide on the composition of an asset portfolio in different choice situations. At the beginning

* Correspondence to: Faculty of Economics and Management, University of Hannover, Königsworther Platz 1, D-30167 Hannover, Germany. Tel.: +49 391 6712158.
*E-mail address:* fochmann@tax.uni-hannover.de (M. Fochmann).

0165-1765/$ – see front matter 
http://dx.doi.org/10.1016/j.econlet.2013.03.038

**Table 1**
Returns of risky asset A and risk-free asset B (example).

| State of nature | Risky asset A | | | | | | | | | | | | | Risk-free asset B |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | No subsidy/tax | Subsidy | | | | Tax | | | | Subsidy–tax | | | | No subsidy/tax, subsidy, tax, subsidy–tax |
| | | Gross | Subsidy | Tax | Net | Gross | Subsidy | Tax | Net | Gross | Subsidy | Tax | Net | |
| 1 | 1.000 | 0.667 | 0.333 | – | 1.000 | 2.000 | – | 1.000 | 1.000 | 1.333 | 0.667 | 1.000 | 1.000 | 1.300 |
| 2 | 1.100 | 0.733 | 0.367 | – | 1.100 | 2.200 | – | 1.100 | 1.100 | 1.467 | 0.733 | 1.100 | 1.100 | 1.300 |
| 3 | 1.200 | 0.800 | 0.400 | – | 1.200 | 2.400 | – | 1.200 | 1.200 | 1.600 | 0.800 | 1.200 | 1.200 | 1.300 |
| 4 | 1.300 | 0.867 | 0.433 | – | 1.300 | 2.600 | – | 1.300 | 1.300 | 1.733 | 0.867 | 1.300 | 1.300 | 1.300 |
| 5 | 1.400 | 0.933 | 0.467 | – | 1.400 | 2.800 | – | 1.400 | 1.400 | 1.867 | 0.933 | 1.400 | 1.400 | 1.300 |
| 6 | 1.500 | 1.000 | 0.500 | – | 1.500 | 3.000 | – | 1.500 | 1.500 | 2.000 | 1.000 | 1.500 | 1.500 | 1.300 |
| 7 | 1.600 | 1.067 | 0.533 | – | 1.600 | 3.200 | – | 1.600 | 1.600 | 2.133 | 1.067 | 1.600 | 1.600 | 1.300 |
| 8 | 1.700 | 1.133 | 0.567 | – | 1.700 | 3.400 | – | 1.700 | 1.700 | 2.267 | 1.133 | 1.700 | 1.700 | 1.300 |
| Subsidy | No | 50% of gross return | | | | No | | | | 50% of gross return | | | | No |
| Tax | No | No | | | | 50% of gross return | | | | 50% of gross return plus subsidy | | | | No |

of each situation, each subject receives an endowment of 100 Lab-points where 1 Lab-point corresponds to 1 Euro cent. The participants' task is to spend their endowment on two investment alternatives: asset A and asset B. The price for one asset of either type is 1 Lab-point.

The return of asset A is risky and depends on the state of nature. Eight states are possible and each state occurs with an equal probability of $\frac{1}{8}$. The return of asset B is risk-free and is therefore equal in every state of nature. The returns of both assets are chosen in such a way that asset A does not dominate asset B in each state of nature, but the expected return of asset A exceeds the risk-free return of asset B. The subjects know the potential returns on both assets in each state of nature before they make their investment decision.

The experiment consists of four treatments in which the presence of a tax and a subsidy is varied. In the no subsidy/tax treatment, no tax is levied and no subsidy is paid. In the subsidy treatment, a subsidy of 50% of the gross return is paid for each asset A, but no tax is imposed. In the tax treatment, a tax with a rate of 50% is levied on the gross return of each asset A, but no subsidy is paid. In the subsidy–tax treatment, a subsidy of 50% of the gross return is paid for each asset A, but in addition a tax has to be paid. In this case, the tax is 50% of the sum of the gross return of asset A and the subsidy. In all four treatments, the returns of the risk-free asset B are neither taxed nor subsidized. Before subjects make their investment decision, they are informed about the tax and subsidy situation.

Although the gross returns of asset A are treated differently across the treatments, they are transformed in such a way that the net returns remain the same (see Table 1 for an example). This leads to identical investment settings in all four treatments and the decision pattern should therefore also be identical across the treatments. Our hypothesis is:

**Hypothesis.** Investment in the risky asset A and the risk-free asset B is identical in all four treatments.

In each treatment, we have five decision situations in which we vary both the potential returns of asset A and the return of asset B. Each subject participates in each treatment (within-subject design) and therefore makes 20 investment decisions in total. To avoid learning effects, the order of these 20 decision situations is completely randomized for each subject.[1] Since we are only interested in the treatment differences, the risk attitude of the subjects is not of importance for our analysis. Participants with

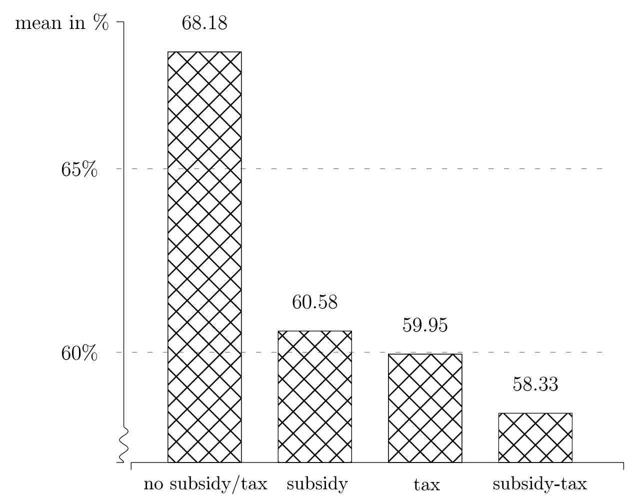



**Fig. 1.** Share of endowment invested in the risky asset A on average for each treatment (number of subjects: 119).

stable and unbiased preferences should follow the same decision pattern across the treatments independently of their individual attitude towards risk.

Despite the fact that we use a very simple setting, with simple tax and subsidy rates, several mechanisms are used to make sure subjects understand their decision environment. First, written instructions explain the calculation of the net returns in detail and provide one numerical example for each treatment. Second, each subject has to correctly solve one numerical example for each of the four treatments as a comprehension test. Third, subjects are provided with both a pocket calculator and a computerized "what-if"-calculator, which allows subjects to calculate their tax, subsidy, and net payoff at different investment levels in each decision situation.

All experiments were carried out at the computerized experimental laboratory at the Otto-von-Guericke University of Magdeburg (MaXLab) and were programmed with $z$-Tree (Fischbacher, 2007). To avoid income effects, we randomly selected five of the 20 decision situations to be paid in cash after the experiment was finished.

## 3. Results and discussion

### 3.1. Baseline experiment

Fig. 1 depicts the average share of endowment invested in the risky asset A for each treatment. In the no subsidy/tax treatment, subjects invested 68.18% of their endowment in asset A. Even though the net returns are identical in the other treatments, this

[1] This means that in each of the 20 rounds one of the five decision situations is randomly selected from any of the four treatments and presented to a subject instead of subjects receiving the choices in four blocks of five decision situations from the same treatment.

**Table 2**
Variation treatments.

| | Variation 1 | Variation 2 | Variation 3 | Variation 4 | Variation 5 |
|---|---|---|---|---|---|
| *Average share of endowment invested in the risky asset A (in %)* | | | | | |
| No subsidy/tax | 68.45 | 71.98 | 76.53 | 83.73 | 71.87 |
| Subsidy | 63.20 | 64.02 | 54.69 | 75.65 | 63.86 |
| Tax | 56.08 | 65.68 | 68.28 | 78.23 | 68.93 |
| Subsidy–tax | 55.68 | 62.87 | 65.39 | 75.20 | 67.67 |
| *Statistical comparison (p-value, two-tailed)*[a] | | | | | |
| No subsidy/tax vs. subsidy | 0.0589 | 0.0030 | <0.0001 | 0.0236 | 0.0019 |
| No subsidy/tax vs. tax | <0.0001 | 0.0439 | 0.0003 | 0.0143 | 0.0932 |
| No subsidy/tax vs. subsidy–tax | 0.0001 | 0.0234 | <0.0001 | 0.0007 | 0.1075 |
| Subsidy vs. tax | 0.0289 | 0.4971 | <0.0001 | 0.6799 | 0.1772 |
| Subsidy vs. subsidy–tax | 0.0088 | 0.6374 | 0.0001 | 0.3306 | 0.2752 |
| Tax vs. subsidy–tax | 0.5547 | 0.4520 | 0.459 | 0.3814 | 0.7562 |
| *No. of subjects* | 25 | 24 | 46[b] | 34 | 36 |

[a] The Wilcoxon signed-rank test is applied for the variation treatments 1, 2, 4, and 5 (treatments with within-subject design), the Mann–Whitney $U$ test for variation treatment 3 (treatment with between-subject design).
[b] 12, 10, 11, and 13 subjects participated in the no subsidy/tax, subsidy, tax, and subsidy–tax treatment, respectively.

share decreased markedly when a subsidy was paid (60.58%) or a tax had to be paid (59.95%). This effect intensified weakly when a subsidy was paid and a tax imposed simultaneously (58.33%). All differences are highly significant ($p < 0.001$, Wilcoxon signed-rank test, two-tailed) compared to the no subsidy/tax treatment. Our hypothesis is therefore rejected for all these comparisons. The difference between the subsidy and the subsidy–tax treatment is weakly significant ($p = 0.077$). However, we found no significant differences between the tax and subsidy–tax treatment or between the subsidy and tax treatment.

These findings are not only at odds with our hypothesis but also with a range of biases discussed in the literature. If subjects had tax aversion (Sussman and Olivola, 2011), tax affinity (Djanali and Sheehan-Connor, 2012), or gross payoff illusion (Fochmann et al., forthcoming) then the bias would have had a different sign in the tax treatment than it did in the subsidy treatment. Since a subsidy is essentially just a negative tax, subjects with tax aversion (affinity) would receive a lower (higher) utility in the tax treatment and a higher (lower) utility in the subsidy treatment when compared to the no subsidy/tax treatment. They would thus have invested less (more) when the risky asset was taxed and more (less) when it was subsidized. This is not what we observed.

Our pattern does not indicate gross payoff illusion either. Since the gross payoff was higher than the net payoff in the tax treatment and lower than the net payoff in the subsidy treatment, subjects with the illusion that their gross payoffs are relevant would not have reacted the same to both types of intervention. They would have been drawn to the higher gross payoff in the tax treatment and the lower gross payoff in the subsidy treatment. The fact that we observe a fall in investment in both treatments can therefore not be readily explained by any of these existing theories.

Given that our main result seems at odds with existing work we checked how robust it was by carrying out a range of variations of the baseline experiment. The results are shown in Table 2 and discussed in Section 3.2.

### 3.2. Variations of the baseline experiment

The tax and subsidy rate in the baseline experiment was deliberately chosen to be quite extreme (50%). To see whether this is important for our results we ran an experiment in which we used a much lower rate. In *variation* 1 we used a tax and subsidy rate of 5% while leaving everything else unchanged. Given that the difference between the net and gross payoffs was now very small we might have expected subjects to react less strongly to the subsidy and tax in variation 1 than they did in the baseline experiment. However, the results were very similar to those in our initial experiment with investment in the risky asset falling sharply under each type of intervention, although the difference between the no subsidy/tax and the subsidy treatment is now only weakly significant. Thus we have strong support for our main result even when the difference between net and gross payoffs has been drastically reduced.

One explanation consistent with the finding that investment in the risky asset fell under both types of intervention is that subjects have an aversion to computational complexity, which reduces their utility from an asset that has been subsidized/taxed. To test this idea we ran an experiment (*variation* 2) in which we subsidized and/or taxed the risk-free asset B instead of the risky asset A. If aversion to computing net payoffs explains our findings then we would expect the opposite results in this variation than we observed in the baseline experiment. However, the results were in fact very similar with a subsidy and/or tax on the risk-free asset also leading to a reduction in investment in the risky asset. Thus, our main result holds in variation 2 suggesting that aversion to computational complexity is not a fitting explanation.[2]

Even though the baseline experiment was set up to be as simple as possible the environment was nonetheless complex enough to suggest that this may be playing an important role. To test this we ran experiments in which we again subsidized and/or taxed the risky asset but simplified the choice environment. We did this in two ways. In *variation* 3 we ran an experiment using a between-subject design. This gave each subject 20 rounds in which they were confronted with just one type of intervention. Stabilizing the environment in this way provided subjects with a greater opportunity to figure out strategies for dealing with the complexity of the environment. In this variation, just as in the baseline experiment, investment in the risky asset fell significantly under each type of intervention, confirming our main result in this more stable environment.

A key difference between variation 3 and the baseline is that there is now a significantly greater reduction in the subsidy treatment than in the other two treatments with intervention. However, it is worth noting that this difference was only observed in early rounds. In the tax and the subsidy–tax treatments there was no trend in their difference to the no subsidy/tax treatment over the 20 rounds. In the subsidy treatment, however, the difference to

[2] A further reason to doubt the computational complexity explanation is that our results are driven largely by subjects investing less in the risky asset under intervention (this made up on average 71% of the reduction) rather than subjects moving away from it completely. This intensive margin of reaction is harder to rationalize using computational complexity.

the no subsidy/tax treatment was much higher in early rounds and gradually fell to being of similar magnitude to the bias observed in the other two treatments. In the last five rounds, for example, the difference between the subsidy treatment and the other two treatments with intervention is no longer significant at the 10% level. But the difference between the no subsidy/tax treatment and the subsidy ($p = 0.0009$), the tax ($p = 0.0336$), and the subsidy–tax ($p = 0.0059$) treatments is still significant.

The second way in which we reduced the complexity of the environment was to reduce the number of states of nature. In *variation* 4 we reduced the states from eight to four and in *variation* 5 we reduced them to two. Investment in the risky asset again fell in all treatments with intervention in both these variations. While the difference between the treatments with and without intervention were smaller in variation 4 than in the baseline experiment they continue to be significant at the 5% level. With two states, however, the difference between the no subsidy/tax treatment and the tax and the subsidy–tax treatment are no longer significant at this level. Thus reducing the complexity along this dimension weakened the bias considerably.

## 4. Conclusion

The baseline experiment together with our five variations shows that the finding that investment in a risky asset falls in the presence of a tax and/or a subsidy is quite robust. This behavior is not consistent with theories such as tax aversion, tax affinity or gross payoff illusion, which would predict that tax and subsidy biases would have different signs. Further, our results do not appear to be driven by an aversion to computational complexity since investment in the risky asset also falls if we subsidize and/or tax the risk-free asset. However, reducing the complexity of the environment by reducing the number of states does seem to affect the strength of the bias. This indicates that the extent to which government intervention biases risk-taking behavior may fall with the complexity of the environment in which the intervention takes place.

## References

Chetty, R., Looney, A., Kroft, K., 2009. Salience and taxation: theory and evidence. The American Economic Review 99 (4), 1145–1177.
Djanali, I., Sheehan-Connor, D., 2012. Tax affinity hypothesis: do we really hate paying taxes? Journal of Economic Psychology 33 (4), 758–775.
Fischbacher, U., 2007. *z*-tree: zurich toolbox for ready-made economic experiments. Experimental Economics 10 (2), 171–178.
Fochmann, M., Kiesewetter, D., Sadrieh, A., 2012. Investment behavior and the biased perception of limited loss deduction in income taxation. Journal of Economic Behavior & Organization 81 (1), 230–242.
Fochmann, M., Weimann, J., Blaufus, K., Hundsdoerfer, J., Kiesewetter, D., Net wage illusion in a real effort experiment. Scandinavian Journal of Economics (forthcoming).
Gamage, D., Hayashi, A., Nakamura, B.K., 2010. Experimental evidence of tax framing effects on the work/leisure decision. SSRN Working Paper.
Sausgruber, R., Tyran, J.-R., 2011. Are we taxing ourselves? How deliberation and experience shape voting on taxes. Journal of Public Economics 95 (1), 164–176.
Sussman, A.B., Olivola, C.Y., 2011. Axe the tax: taxes are disliked more than equivalent costs. Journal of Marketing Research 48 (Special Issue 2011), S91–S101.

Journal of Public Economics 121 (2015) 95–105

Contents lists available at ScienceDirect

Journal of Public Economics

journal homepage: www.elsevier.com/locate/jpube

# To commit or not to commit? An experimental investigation of pre-commitments in bargaining situations with asymmetric information

Sönke Hoffmann *, Benedikt Mihm, Joachim Weimann

*Faculty of Economics and Management, Otto-von-Guericke University, Magdeburg, Germany*

ARTICLE INFO

*Article history:*
Received 2 August 2013
Received in revised form 8 October 2014
Accepted 13 November 2014
Available online 18 November 2014

*Keywords:*
Pre-commitments
Bargaining
Asymmetric information
Social preferences

ABSTRACT

In a recent paper Konrad and Thum (2014) present a model that shows that unilateral pre-commitment reduces the likelihood of agreement in bilateral negotiations over the provision of a public good when parties have private information over their contribution costs. We test the model in a laboratory experiment paying particular attention to how behavioral motivations other than payoff-maximization affect the strength of the model's result. We find that the result is no longer statistically significant when we allow for non-payoff-maximizing behavior at each stage of the game. Introducing communication has an interesting effect as it influences different forms of non-payoff-maximizing behavior asymmetrically and leads to the model's result again becoming significant. All in all, we find strong experimental support for Konrad and Thum's model even though we observe considerable amounts of non-payoff-maximizing behavior that is not accounted for in the original model.


## 1. Introduction

Bargaining over the private provision of public goods may lead to inefficient outcomes if parties have incomplete information or contracts are not enforceable. The literature has largely concentrated on the enforcement problem. Work on international negotiations on environmental regulations, for example, has paid particular attention to the enforcement problem because no common institution exists in this setting and the enforcement of contracts is thus difficult (Carraro and Siniscalco, 1993; Barrett, 1998).

In a recent paper Konrad and Thum (2014) focus instead on the problems that arise in a bargaining environment with asymmetric information. Their model (referred to as KT-model henceforth) assumes the enforcement problem is resolved and examines bargaining over contributions to a public good when parties are privately informed about their cost of provision.

Under asymmetric information bargaining outcomes will generally be inefficient as negotiations can break down with a positive probability even when mutually beneficial agreements are possible (Meyerson and Satterthwaite, 1983). It is well known that in markets for private goods the inefficiency disappears as the number of traders increases and the market becomes large (Gresik and Satterthwaite, 1989). However, Rob (1989) showed that even this asymptotic efficiency does not hold for public goods and thus under asymmetric information negotiations over the private provision of a public good are unlikely to ever achieve an efficient solution.

The question remains, however, how large the inefficiencies will be and under what kind of negotiation rules the likelihood of negotiation breakdown, and thus the inefficiency, can be minimized. In particular, it is unclear if prior commitments by one party have a positive influence on the prospects for achieving more efficient outcomes. The KT-model makes an important contribution to the literature on the private provision of public goods by investigating this issue in a non-cooperative game setting.

The role of prior commitments is highly relevant. The EU, for example, seems to view pre-committing to environmental damage prevention as an act that sets a good example for others and that will motivate others to follow suit. The KT-model, however, states the exact opposite. Comparing the equilibria of two sequential bargaining games – one with commitment and one without – the authors show that the probability for successful cooperation is strictly lower when one party has contributed to the public good before bargaining takes place. This result obviously has strong political implications.

Our paper is an experimental investigation of the findings of the KT-model. In addition to a direct experimental verification of the model our experiment focuses on the potential for the bargaining situation modeled by Konrad and Thum to be influenced by various motives that deviate from payoff-maximization and which could thus affect the results of the model. Inequality aversion, for example, might prevent players from payoff-maximizing if payoff differences are sufficiently large (Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000). Direct and

* Corresponding author.
*E-mail addresses:* sohoffma@ovgu.de (S. Hoffmann), mihm@ovgu.de (B. Mihm), joachim.weimann@ovgu.de (J. Weimann).

http://dx.doi.org/10.1016/j.jpubeco.2014.11.005
0047-2727/

indirect reciprocity (Falk and Fischbacher, 2006; Nowak and Siegmund, 2005) may also be a factor in the presence of pre-commitments.

Results from a world-wide survey of people involved in international climate policy indicate that fairness and equity considerations can play a significant role in climate negotiations (Lange et al., 2007, 2010). Since the KT-model's results are particularly relevant for climate negotiations it is thus important to investigate how the model performs in a bargaining environment in which real subjects may harbor such behavioral motivations. A laboratory setting is the ideal venue to explore this extension as the experiment can control for the amount of freedom subjects have to deviate from payoff-maximization.

In order to find out how behavioral motivations other than payoff-maximization affect the results of the KT-model the experiment is designed to be carried out in three settings. Each of the three settings has one treatment with pre-commitment and one without. The first setting is intended to be a direct assessment of the KT-model as it most closely follows the basic assumptions of the original theory, i.e., payoff-maximizing behavior and common knowledge. Technically, one subgame of the KT-model (standard prisoners' dilemma) is replaced by the corresponding Nash payoffs and thus players are forced to behave in a payoff-maximizing way in the final stage of the game. In our experiment we found that in this reference setting cooperation took place twice as often in the treatment without pre-commitment when compared to the treatment with pre-commitment (referred to as cooperation gap henceforth).

In the second setting the entire prisoners' dilemma is introduced to ascertain whether the KT-model is affected by giving subjects additional room to behave in non-payoff-maximizing ways, and if so, whether the cooperation gap persists. We found that the gap did persist in our experiment but became considerably smaller.

In the third setting the KT-model is pushed even further away from its original assumptions through the introduction of pre-play communication between the bargaining parties. There are two motivations for this extension. First, the experimental literature on the provision of public goods has shown that communication between subjects increases the level of cooperation even if communication is cheap talk (Brosig et al., 2003; Valley et al., 1998). It is still unclear, however, what effect communication has in environments with or without pre-commitment. Second, it is an artificial assumption that bargaining over the provision of public goods takes place without communication between the parties involved. It is thus important for the external validity of the KT-model to check whether or not it is communication proof. In fact, in our experiment we observed that with communication there was a strong increase in success rates in both the pre-commitment and no pre-commitment treatments but at the same time the cooperation gap again opened significantly.

The remainder of this paper is structured as follows. The next section outlines the KT-model as it was implemented in our experiment. In section three we specify the experimental procedure. Section four contains our main results, and in the final section five we discuss our findings.

## 2. The KT-model

The KT-model encompasses two variants of a sequential bargaining game, one with pre-commitment and one without. We start with the more general version without pre-commitment.[1]

[1] As the original model is too general to be directly implemented in the laboratory, some basic assumptions of the model had to be slightly adjusted. In particular, the KT-model applies to continuous random variables following arbitrary probability distributions that have a positive inverse hazard rate. In our experiments we use integer variables scaled by factor 10 and a uniform distribution of random variables. Therefore, our presentation of the major results is slightly different compared to the original paper. However, our modification is just a special case of the original theory.

Two players $i \in \{A, B\}$ negotiate over the provision of a public good $e = e_A + e_B$, where $e_A$ and $e_B$ denote the contribution of players $A$ and $B$ respectively. Both players can either make a contribution ($e_i = 10$) or not ($e_i = 0$). If player $i$ decides to contribute, his cost of contribution is $10 + c_i$ with $c_i \in \{1, 2, \ldots, 9\}$. The cost parameter $c_i$ is private information of player $i$ and is randomly drawn from a uniform distribution. In the bargaining process, player $A$ can offer a transfer $t \in \{-10, -9, \ldots, 9, 10\}$ to player $B$. If $t > 0$ the transfer goes from $A$ to $B$ which means that $A$ pays a price to $B$, if $t < 0$ the transfer is a price $B$ pays to $A$.

The overall bargaining structure is characterized by a take it or leave it offer similar to the classic ultimatum game: Player $A$ proposes a transfer to $B$ which $B$ can accept or reject. If $B$ accepts then both players become obliged to contribute to the public good ($e_i = 10$). If the offer is rejected no transfer is paid and both players decide over their contributions independently. In this case both players are in a prisoners-dilemma and choosing not to contribute is their dominant strategy. Fig. 1 visualizes the sequential structure of the game without pre-commitment.

This version of the model is contrasted with a version in which $A$ makes a commitment before the game starts. Technically, this pre-commitment is modeled by fixing $e_A = 10$ throughout the whole game, which removes strategy $e_A = 0$ from the prisoners' dilemma in the last stage. Thus, player $A$ no longer decides about his contribution and this is common knowledge.

In both cases the payoffs of the players can be written as

$$\pi_A = e_B - c_A \frac{e_A}{10} - t \quad \text{and} \quad \pi_B = e_A - c_B \frac{e_B}{10} + t. \tag{1}$$

Under the assumption of payoff-maximization the KT-model has the following two results.

**Result 1** (Konrad and Thum, 2014). *The probability that A and B agree on a cooperative outcome is higher without pre-commitment for all possible $c_A$.*

**Result 2** (Konrad and Thum, 2014). *The unique perfect Bayesian equilibrium transfer is non-positive in the game without pre-commitment and strictly positive in the game with pre-commitment. Specifically, under the conditions implemented in the experiment the equilibrium transfers are given by $t^*_{nPC} = \min\{-\frac{c_A}{2}, -1\}$ in the game without pre-commitment and $t^*_{PC} = 5$ in the game with pre-commitment.*

The intuition behind these results is as follows. If player $A$ does not pre-commit before bargaining takes place then his gain from reaching an agreement is greater. To keep the chances of getting this gain realized $A$ has to bargain less aggressively which enhances the likelihood of cooperation relative to the game with pre-commitment.

Furthermore, if $A$ does not pre-commit then he can sell his willingness to cooperate to $B$. Player $A$ thus demands a price for cooperation

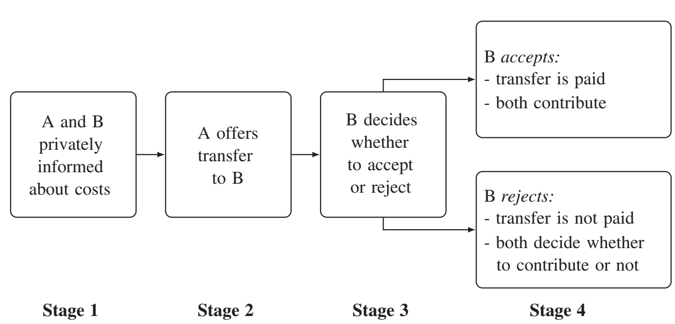



**Fig. 1.** Sequential structure of the game without pre-commitment.

and we have $t \leq 0$. If $A$ pre-commits, however, he needs to offer $B$ something to make him cooperate and we have $t > 0$.

## 3. The experiment

### 3.1. Background

One purpose of the experiment is to detect how behavioral motivations other than payoff maximization affect the strength of the KT-model's predictions. Generally, $B$-players can deviate from payoff-maximizing behavior in two ways. They may either accept offers although they would be better off rejecting, or reject offers although they would be better off accepting. If observed, these patterns can, however, not be attributed uniquely to a specific behavioral motive. Both for non-payoff-maximizing rejections and acceptances several motives may explain the data.

A non-payoff-maximizing rejection may for example be the result of either negative (direct) reciprocity or inequality aversion. Direct reciprocity refers to a situation in which the first mover $A$ provides a trigger to the second mover $B$ and $B$ directly responds by a choice that affects $A$'s payoffs.[2] The case in which offers are rejected although accepting would be the payoff-maximizing choice can be attributed to negative direct reciprocity because it implies that the $B$-players are directly punishing the offers made by the $A$-players. Rejections which yield a lower payoff for $B$-players may also be due inequality aversion if the $B$-players deem the payoffs resulting from acceptance to be too unequal.

A non-payoff-maximizing acceptance may be the result of either positive (indirect) reciprocity or a willingness to cooperate. In the treatments with pre-commitments a non-payoff-maximizing acceptance may be driven by upstream reciprocity, which is a form of indirect reciprocity of the type "somebody else helped me and I help you" (Nowak and Roch, 2006). Since the introduction of a fixed pre-commitment puts the $B$-players in a dominant position it may trigger them to reward the $A$-players at a cost to themselves even though the pre-commitment was not voluntary.[3] In the pre-commitment treatments the acceptance of offers when rejecting would be the payoff-maximizing choice may also be due to a willingness to cooperate as $B$-players may be more interested in the common good than acting only in their own self-interest.

For $A$-players, offers that differ from the transfers predicted by the KT-model may also be due to number of behavioral motivations. In the PC game the acceptance of an equilibrium transfer results in an unequal payoff in favor of the $B$-player so an inequality averse $A$-player may make lower offers than predicted by theory. In the nPC-game equilibrium transfers result in unequal payoff in favor of the $A$-player leading an inequality averse $A$-player to make higher offers than predicted by theory. Offers that differ from the equilibrium transfers may also be due to the $A$-player taking into account that the $B$-player may be driven by one of the behavioral motivations such as reciprocity outlined above.

As several motives may explain non-payoff maximizing behavior we keep the discussion of how these motivations explain our results quite general throughout the paper.

The experiment is carried out in three settings that give subjects varying amounts of freedom to follow behavioral motivations other than payoff-maximization. In all settings we have one treatment with and one without pre-commitment.

- *Setting 1* is the reference setting designed to control for non-payoff-maximizing behavior. If an offer is rejected both players' contributions are fixed to the dominant strategy $e_i = 0$ to force them to play the prisoners' dilemma's Nash equilibrium. The two treatments in this setting are labeled PC_nD_nC (pre-commitment, no decision, no communication) and nPC_nD_nC (no pre-commitment, no decision, no communication).
- *Setting 2* gives both players more room for non-payoff-maximizing behavior because after a rejected transfer they are free to choose whether to contribute to the public good or not. We conjecture that having a choice in the final stage of the game may have a feedback effect on the previous stages. If subjects are completely selfish we should observe no such feedback and we should thus observe no difference between Settings 1 and 2. The two treatments in Setting 2 are labeled PC_D_nC (pre-commitment, decision, no communication) and nPC_D_nC (no pre-commitment, decision, no communication).
- *Setting 3* is the same as Setting 2 but includes a three minute pre-play chat using a chat-box integrated into the user interface. The opportunity for communication occurs in stage one of the game after the subjects learn their cost of provision. Written content is essentially unrestricted but any information that reveals a player's identity is prohibited. The conjecture is that subjects will behave more cooperatively in the communication treatments as has been observed in various experiments on public good provision (Brosig et al., 2003; Valley et al., 1998). The two treatments in this setting are labeled PC_D_C (pre-commitment, decision, communication) and nPC_D_C (no pre-commitment, decision, communication).

Table 1 gives an overview of the three pairs of treatments played in the three settings.

### 3.2. Experimental setup

Each of our three settings contains two treatments, one without pre-commitment and one with pre-commitment. In each of the six treatments we played six sessions with a different group of ten subjects respectively. Because each subject participated in exactly one specific session, subjects were also different across treatments (between-subject design). At the beginning of a session subjects were randomly selected into the fixed roles of either an $A$- or a $B$-player. Over the five rounds the game was played, we used a rotating matching scheme under complete anonymity, i.e. each of the five $A$-players was paired exactly once with each of the five $B$-players ("round robin"). As every single subject was well informed about this setup we assume that a player's current decision was made independently from another player's history of decisions. However, among a total of 150 observations per treatment (5 pairs × 5 rounds × 6 sessions) we considered subject specific observations dependent which results in 30

**Table 1**
Overview of all treatments played.

| Setting | Name | Pre-play commitment | Post-rejection choice | Pre-play communication |
|---|---|---|---|---|
| 1 | nPC_nD_nC | No | No | No |
| | PC_nD_nC | Yes | No | No |
| 2 | nPC_D_nC | No | Yes | No |
| | PC_D_nC | Yes | Yes | No |
| 3 | nPC_D_C | No | Yes | Yes |
| | PC_D_C | Yes | Yes | Yes |

[2] Even for direct reciprocity alone there exist several different modeling approaches. Intention based reciprocity models focus on what one player believes about the intention of the other player (Rabin, 1993; Falk and Fischbacher, 2006). In type based models people weigh monetary payoffs according to the perceived type of person they face (Levine, 1998). Emotion based models explain reciprocal behavior in terms of relative payoffs and the player's current emotional state (Cox et al., 2007).

[3] Note that upstream reciprocity originated from Evolutionary Game Theory which analyzes the evolution of populations given many repetitions of the game. The behavioral pattern we observe here is closest to the character of upstream reciprocity, even though our games were played one-shot.

independent observations per treatment (5 pairs × 6 sessions).[4] In total 360 subjects participated in the experiment.

After the roles had been fixed the subjects were given written instructions and sufficient time to read them.[5] Subjects were informed that their possible cost values ranged from 1 to 9 and were randomly selected with the same probability. Once drawn, we used the same cost values in all sessions to make sessions comparable. The experiment started with three practice rounds in which the subjects played against the computer. All participants were informed that the computer played payoff-maximizing strategies throughout the practice rounds and that these rounds were not payoff relevant. In the first two practice rounds subjects learned the computer's cost as well as their own. The third round simulated the actual game as each player was informed only about their own cost of provision.

At the start of the actual game the subjects were given an initial endowment to ensure that it was not possible for them to make a loss. Subjects were informed that they received an endowment that covered losses but not how high the endowment actually was.[6] Subjects were not paid a show up fee on top of this endowment. During the game subjects had access to on-screen tables that provided information about relevant payoffs.[7] These tables included their own payoffs conditional on their and the other players possible decisions, as well as the payoffs of the other player conditioned on the other player's possible costs. The purpose of this information was to make the game easier to follow and to minimize calculation effort.

The experiment was carried out at the experimental laboratory at the University of Magdeburg, Germany (MaXLab) and was programmed using z-Tree (Fischbacher, 2007). The sessions lasted on average 40 min. At the end of the experiment the payoffs of all five games were paid and the average earnings of the subjects were 10.35 euro.

## 4. Results

### 4.1. Model predictions versus experimental results

In this section we address the question of whether the theoretical predictions in Results 1 and 2 are compatible with the experimental data of our reference setting (Setting 1), and if so, whether these results still hold given a post-rejection decision (Setting 2) and communication (Setting 3). Table 2 summarizes the (aggregated) experimental data of each treatment. Player $A$'s behavior is captured by the average transfer $\varnothing t$ and the average deviation $\Delta t^*$ from the predicted equilibrium value. The basic behavior of player $B$ is described by $a$, the total number of acceptances and $r$, the total number of rejections. Dividing $a$ by the total number of decisions (150 in each treatment) gives the total acceptance rate that is plotted in Fig. 2 over all three settings.

Result 1 predicts a higher probability of successful negotiation if there is no pre-commitment. In the first setting of our experiment we observed that the acceptance rate was 72 % in the nPC-condition and 36 % in the PC-condition (cf. Fig. 2, Setting 1). Clearly, this difference is statistically significant ($p$-value < 0.001, $\chi^2$-test). Allowing subjects to have a post-rejection decision in the prisoners dilemma (Setting 2) led to the cooperation gap becoming much smaller (55 % vs. 47 %) and no longer being statistically significant ($p$-value $= 0.204$, $\chi^2$-test). Result 1 is, nevertheless, confirmed qualitatively. Finally, adding pre-play communication (Setting 3) increased the agreement rates with and without pre-commitment. However, the influence of communication was much stronger without pre-commitment such that the cooperation gap once again opened widely (91 % vs. 73 %) and the difference between the rates once again becomes statistically significant ($p$-value < 0.001, $\chi^2$-test).

Result 2 predicts equilibrium transfers $t^*_{nPC}(c_A) = -0.5c_A$ in the game without pre-commitment and $t^*_{PC}(c_A) = 5$ in the game with pre-commitment. To compare these predictions with our experiments we fitted the linear model $t = \alpha + \beta c_A$ in all six treatments using ordinary least squares with observations clustered by subject and robust variance estimates.

Table 3 shows that without pre-commitment the experimental data of Settings 1 and 2 are in line with the prediction. $H_0 : \alpha = 0$ could not be rejected at $p$ values larger than 0.4. The slope estimates are both negative and significantly different from zero and $H_0 : \beta = -0.5$ could not be rejected in Setting 2. With communication involved (Setting 3) the level parameter still corresponds to the prediction ($H_0 : \alpha = 0$, $p = 0.6$) but the transfer seems to become completely unaffected by the cost of Player $A$ ($H_0 : \beta = 0$, $p = 0.7$).

In the three PC-treatments we no longer observed any evident relationship between $t$ and $c_A$. At all cost levels the transfers spread over their maximum range with correlation coefficients between 0.01 and 0.1 and $H_0 : \beta = 0$ could not be rejected in all three settings. Basically, this observation is in line with the model which predicts that transfer is not affected by the cost level. However, the transfer varies a lot, and while positive, it seems to be lower on average than the predicted value of five (cf. Table 2).

All in all, we see strong experimental evidence in support of Results 1 and 2 of the model, even under conditions that may deviate from the original KT-model.

### 4.2. Behavioral analysis

We now address the question of how the interplay between different behavioral motivations resulted in the support for the KT-model's main predictions. The behavioral analysis focuses on non-payoff-maximizing behavior and requires a number of refinements of the measures $a$ and $r$. The variables in columns $5-7$ of Table 2 represent the number of cases for which the transfer-cost constellation made accepting generate a strictly higher payoff than rejecting ($a_{hyp}$), rejecting generate a strictly higher payoff than accepting ($r_{hyp}$), and both decisions generate equal payoffs ($i_{hyp}$). $a_{hyp}$ is split up into $a_{pmx}$, the number of times $B$-players indeed accepted a profitable offer and $r_{npmx}$, the number of times $B$-players rejected even though accepting was the payoff-maximizing choice. The total number of acceptances $a$ is the sum of $a_{pmx}$ $a_{npmx}$ and $a_{ind}$, where the latter is the number of $B$-players who accepted when rejecting and accepting provided the same payoff. Equivalent decompositions hold for $r_{hyp}$ and $r$.

Fig. 3 shows the number of non-payoff-maximizing acceptances $a_{npmx}$ (Fig. 3a) and rejections $r_{npmx}$ (Fig. 3b) of the $B$-players in all three settings. Without pre-commitment we did not find any indication of positive reciprocity or cooperative behavior in the $B$-players' behavior. Acceptances were clearly payoff driven as the white bars in Fig. 3a show no deviations from payoff-maximization in the nPC-treatments. This is quite different when it comes to the rejections. A considerable number of transfers were rejected although accepting would have provided a higher payoff (cf. white bars in Fig. 3b). This indicates the presence of negative reciprocity or inequality aversion among $B$-players.[8]

With pre-commitment the pattern was completely reversed: $B$'s rejection behavior was in line with payoff-maximization but the

[4] Because of the dependence of observations within an individual we calculated robust variance estimates using clustered individuals. Even if we clustered over sessions we would get essentially the same results.

[5] See example of instruction in Appendix C.

[6] We adjusted both players final earnings by making their endowments unequal, otherwise $A$ would have earned nearly nothing on average in the PC-treatments. Letting subjects know about unequal endowments before or during the five rounds played would have affected behavior differently in different treatments, therefore we did not inform subjects about these endowment adjustments.

[7] See screenshots in Appendix D.

[8] As the transfer offered by the $A$-player is essentially an ultimatum offer it is not surprising that $B$-players were willing to reject offers which would make them better off but were perceived as unfair. This kind of behavior is well known from ultimatum game experiments (Güth et al., 1982).

**Table 2**

Behavioral results of all treatments. (1) Behavior of $A$: $\varnothing\ t$: Average transfer offered (Lab dollars), $\Delta t^*$: Average difference between $t$ and the theoretical equilibrium value $t^*$ (Lab Dollars), *median t*: Median values of the transfers. (2) Expected behavior of $B$ (absolute numbers): $a^{hyp}$: hypothetical acceptance cases (payoff from accepting was strictly higher), $r^{hyp}$: hypothetical rejection cases ((expected) payoff from rejecting was strictly higher), $i^{hyp}$: hypothetical indifference cases (accepting and rejecting had same payoffs) (3) Actual behavior of $B$: $a$: total acceptances, $a_{pmx}$: acceptances when accepting was the payoff-maximizing choice, $a_{npmx}$: acceptances when accepting was not the payoff-maximizing choice, $a_{ind}$: acceptances when accepting and rejecting generated equal payoffs (indifference). All rejection variables have the equivalent interpretation given that $B$ rejected the offer. Column "Totals" sums up the respective acceptance and rejections such that $\omega = a + r$, respectively.

| | Behavior of players | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $A$ | | | $B$'s exp. beh. | | | $B$ accepted | | | | $B$ rejected | | | | Totals | | |
| Treatment | $\varnothing\ t$ | $\Delta t^*$ | median $t$ | $a_{hyp}$ | $r_{hyp}$ | $i_{hyp}$ | $a$ | $a_{pmx}$ | $a_{npmx}$ | $a_{ind}$ | $r$ | $r_{pmx}$ | $r_{npmx}$ | $r_{ind}$ | $\omega_{pmx}$ | $\omega_{npmx}$ | $\omega_{ind}$ |
| nPC_nD_nC | −1.69 | 0.71 | −2.00 | 129 | 9 | 12 | 108 | 104 | 0 | 4 | 42 | 9 | 25 | 8 | 113 | 25 | 12 |
| PC_nD_nC | 0.77 | −4.23 | 2.00 | 38 | 102 | 10 | 54 | 35 | 11 | 8 | 96 | 91 | 3 | 2 | 126 | 14 | 10 |
| nPC_D_nC | −1.88 | −0.52 | −1.00 | 122 | 20 | 8 | 82 | 81 | 1 | 0 | 68 | 19 | 41 | 8 | 100 | 42 | 8 |
| PC_D_nC | 1.98 | −3.02 | 3.00 | 49 | 88 | 13 | 70 | 49 | 10 | 11 | 80 | 78 | 0 | 2 | 127 | 10 | 13 |
| nPC_D_C | 0.22 | 2.62 | 0.00 | 147 | 2 | 1 | 137 | 136 | 1 | 0 | 13 | 1 | 11 | 1 | 137 | 12 | 1 |
| PC_D_C | −2.84 | −2.16 | −3.00 | 52 | 76 | 22 | 109 | 50 | 39 | 20 | 41 | 37 | 2 | 2 | 87 | 41 | 22 |

acceptance behavior was not. $A$'s disadvantageous position (relative to $B$'s) caused by the pre-commitment seems to have triggered $B$'s willingness to accept offers even though rejecting would have provided a higher payoff.

It is important to note that the non-payoff-maximizing behavior both in regard to acceptances and rejections work against Result 1 of the KT-model. The non-payoff-maximizing acceptances in the PC-treatments increased the amount of agreements in the games with pre-commitment and the non-payoff-maximizing rejections in the nPC-treatments reduced the amount of agreements in the games without pre-commitment. Given the overall characterization of the behavior of the $B$-players it is surprising that the predictions of the KT-model are nevertheless confirmed by our findings. To see why this is the case, we have to look more closely at the three settings and we have to take the behavior of the $A$-players into account.

### 4.2.1. Setting 1: No last stage decision, no communication

In our reference setting we observed the general behavioral pattern outlined above: Without pre-commitment acceptances were payoff maximizing and rejections were not (e.g. negative reciprocity), whereas with pre-commitment rejections were payoff-maximizing and acceptances were not (e.g. positive reciprocity). As negative reciprocity reduces the number acceptances and positive reciprocity increases it, both observed effects work against the Result 1 of the KT-model.

The reason that there were nevertheless significantly less agreements in the PC-treatment than in the nPC-treatment in this setting can be attributed to the fact the offers made by the $A$-players were not high enough in the PC-treatment. Only 38 of the 150 (= 25 %) offered transfers in the PC-treatment would have made the $B$-player better off by acceptance than by rejection, compared to 129 of the 150 (= 86 %) offers for which this would have been the case in the nPC-treatment. This difference is clearly significant ($p$-value < 0.001, $\chi^2$-test).

The KT-model predicts a transfer payment in the PC game of $t^*_{PC} = 5$ but the average transfer payments in the PC-treatment was only 0.77. Thus, behavioral motivations such as inequality aversion appear to have had a strong effect on the behavior of the $A$-players in the PC-treatment and, as result, the agreement rate fell steeply enough that the difference in the number of agreements reached in the PC-treatment and the nPC-treatment was so pronounced.

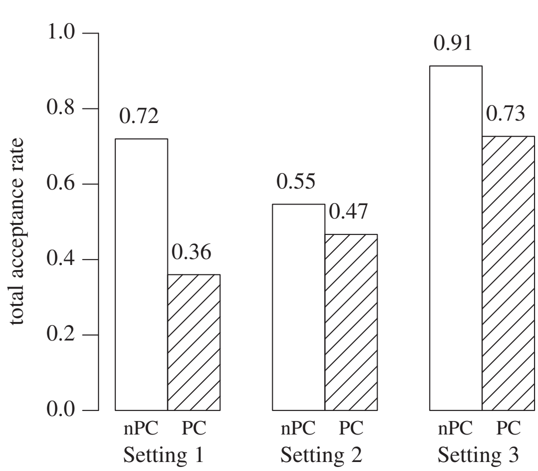



**Fig. 2.** Variations of the cooperation gap in three experimental settings.

### 4.2.2. Setting 2: Last stage decision, no communication

Under the PC-condition the behavior of the $B$-players did not change much between Setting 1 and Setting 2. Fig. 3 shows that the introduction of a contribution choice in the prisoners' dilemma made non-payoff-maximizing rejections change from $r_{npmx} = 3$ out of $a_{hyp} = 38$ (8 %) to $r_{npmx} = 0$ out of $a_{hyp} = 49$ (0 %) and the non-payoff-maximizing acceptances change from $a_{npmx} = 11$ out of $r_{hyp} = 102$ (11 %) to $a_{npmx} = 10$ out of $r_{hyp} = 88$ (11 %). Both changes are not significant ($p$-value > 0.15, $\chi^2$-test). Put differently, we still observed payoff-maximizing rejection behavior and moderate positive reciprocity or cooperative behavior from the $B$-players. The transfer behavior of the $A$-players did not change much either, as the $a_{hyp}$ values in Settings 1 and 2 are quite similar.[9]

The key difference between the two settings can be found in the nPC-condition. In Setting 1 the $B$-players rejected $r_{npmx} = 25$ out of $a_{hyp} = 129$ (= 19 %) advantageous offers made by $A$ and in Setting 2 this rate was 41 out of 122 (= 34 %). Thus, the tendency towards non-payoff-maximizing rejections by the $B$-players increased significantly ($p$-value = 0.01, $\chi^2$-test) when both players were free to choose their contribution in the prisoners' dilemma. This, in turn, made the number of agreements fall and the cooperation gap close.

This raises the question as to why $B$-players rejected offers even though the resulting payoff in the non-cooperative solution of the prisoners' dilemma was strictly lower than the safe payoff from accepting. In the treatment in Setting 2 without pre-commitment this type of behavior is only reasonable if the $B$-players expected that the $A$-players would not play their dominant strategy in the prisoners' dilemma game. If all players expected every other player to choose the non-cooperative strategy in the prisoners' dilemma then there should be no difference in the $r_{npmx}/a_{hyp}$ values between Settings 1 and 2. But as

[9] This behavior is not too surprising as $A$'s decision space is exactly the same in both settings when he has to pre-commit.

**Table 3**
Parameter estimates and $p$-values of the linear relationship between $t$ and $c_A$ for all six treatments.

| | Coeff. | | $p$-Values (t-test) | | |
|---|---|---|---|---|---|
| nPC | $\hat{\alpha}$ | $\hat{\beta}$ | $\alpha = 0$ | $\beta = 0$ | $\beta = -\frac{1}{2}$ |
| nPC_nD_nC | 0.002 | −0.35 | 0.996 | <0.001 | 0.025 |
| nPC_D_nC | 0.427 | −0.48 | 0.416 | <0.001 | 0.873 |
| nPC_D_C | 0.484 | −0.06 | 0.618 | 0.689 | 0.003 |
| PC | $\hat{\alpha}$ | $\hat{\beta}$ | $\alpha = 0$ | $\alpha = 5$ | $\beta = 0$ |
| PC_nD_nC | 2.31 | −0.32 | 0.056 | 0.028 | 0.124 |
| PC_D_nC | 3.02 | −0.22 | 0.007 | 0.070 | 0.274 |
| PC_D_C | 2.96 | −0.02 | 0.000 | 0.007 | 0.854 |

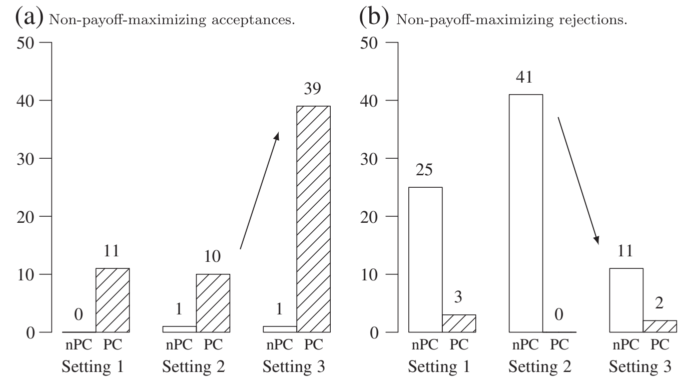



**Fig. 3.** Deviations from payoff-maximization over all treatments. Values above the bars represent the absolute number of $B$-players. The arrows indicate the effect of communication.

we did observe a difference it can be attributed to $B$'s expectation to exploit the other player in the final stage.

*4.2.3. Setting 3: Last stage decision, communication*

Adding pre-play communication in Setting 3 neither affected non-payoff-maximizing acceptances under the nPC-condition nor did it affect non-payoff-maximizing rejections under the PC-condition.[10] With communication $B$-players still did not reward $A$-players at cost to themselves when $A$ was not disadvantaged by pre-commitment, and $B$-players still did not punish $A$-players at cost to themselves when $A$ was disadvantaged by pre-commitment. However, the arrows shown in Fig. 3 indicate that communication did (i) increase the non-payoff-maximizing acceptances under PC ($a_{npmx}$ increased from 10 out of $r_{hyp} = 88$ ($= 11\,\%$) to 39 out of $r_{hyp} = 76$ ($= 51\,\%$)) and (ii) reduced the non-payoff-maximizing rejections under nPC ($r_{npmx}$ falls from 41 out of $a_{hyp} = 122$ ($= 34\,\%$) to 11 out of $a_{hyp} = 147$ ($= 7\,\%$)). Therefore, communication influences $B$ towards cooperation in two different ways: On the one hand it led $B$ to share in the disadvantage of $A$'s pre-commitment and accept non-profitable transfers ($a_{npmx}$ increases), on the other hand it neutralized the expectations $B$ had of exploiting $A$ that were present without communication in the nPC-condition ($r_{npmx}$ falls). The first effect works against Result 1 of the KT-model, the second one supports it.

If communication had no other effect we should have observed that the cooperation gap remained unchanged. The reason why the cooperation gap opened again in Setting 3 is that communication also affected the generosity of the $A$-players. Average transfers $\varnothing\, t$ as well as the number of offers $a_{hyp}$ that a payoff-maximizing actor would accept take their highest values in Setting 3.[11] This made the number of payoff-maximizing acceptances rise by $\Delta a_{pmx} = 55$ and the number of payoff-maximizing rejections change by $\Delta r_{pmx} = -18$ under the nPC-condition. In total this effect is stronger than under the PC-condition ($\Delta a_{pmx} = 1$ and $\Delta r_{pmx} = -41$), supporting Result 1 of the model.

*4.3. Estimation of treatment effects*

As a final piece of analysis we use our experimental data to estimate the logistic regression model

$$\Pr(Acc_B = 1) = \Lambda(\alpha + \beta_1 t + \beta_2 c_B + \gamma_1 D_1 + \ldots + \gamma_5 D_5) \tag{2}$$

in which $Acc_B$ (Did Player B accept? Yes = 1, No = 0) represents the binary outcome variable and $\Lambda(z)$ is the logistic link function. The success probability $\Pr(Acc_B = 1)$ is explained by the predictors $t$ (transfer offered by Player $A$) and $c_B$ (costs of Player $B$) and five treatment dummies $D_1$ to $D_5$. The Maximum-Likelihood fit of Eq. (2) and some technical detail on the estimated model's characteristics can be found in Appendix A.

Here, we start by using the fitted model to estimate the Average Marginal Effect (*AME*) of each of the predictors, which quantifies their average isolated effect on success probabilities when controlling for all other variables.[12] The first two rows of Table 4 show that the average effect of $B$'s costs on success probabilities is negative and significant ($z$-Test, $p$-value < 0.001), whereas for the transfer $t$ the *AME* is positive and significant. The remaining seven rows display the isolated effect of a treatment relative to a baseline both within (rows 3 - 5) and across (rows 6 - 9) the three settings. It is important to note that these comparisons differ conceptually from those carried out to quantify the cooperation gaps in Section 4.1. The cooperation gap is defined as the difference in the *acceptance rates* between two treatments. The acceptance rates themselves are not only influenced by whether there is a pre-commitment or not, but also by the transfers and costs that are present in each of the treatments being compared. Consequently, the evolution of *AME* values over the three settings draws a different picture than the observed cooperation gaps. In our reference setting the *AME* of pre-commitment is very strong ($AME = -0.51$) and becomes successively weaker under a post-rejection decision ($AME = -0.38$) and communication ($AME = -0.35$). In all three settings the isolated effect of a pre-commitment on success probabilities is negative and significantly different from zero.

The last four rows in Table 4 are across setting comparisons. We see that the introduction of the post-rejection decision has a significant negative effect without pre-commitment ($AME = -0.12$), whereas under pre-commitment this effect disappears ($AME = 0.02$, $p$-value $= 0.6$, $z$-Test). Finally, introducing communication has a positive and statistically significant effect without pre-commitment ($AME = 0.18$) and with a pre-commitment ($AME = 0.20$).

Clearly, the averages of the marginal effects are useful, but the aggregation of multidimensional data into a single number always eliminates information. For example, *AMEs* hardly allow for drawing conclusions about non-payoff-maximizing behavior or to identify those combinations of transfer $t$ and cost $c_B$ for which a treatment effect is the strongest. For that reason we also plotted the individual marginal treatment effects at every hypothetical combination of transfer $t$ and cost $c_B$. Fig. 4 visualizes this approach for each of the four across setting comparisons.

[10] See Appendix B for a detailed discussion of how communication was used by the subjects. It is worth noting here that the opportunity to communication was utilized extensively in both treatments although more often in the treatment without pre-commitment.

[11] Note that this effect of communication was so strong that even the $\varnothing\, t$ in the PC-condition became positive, which is difficult to explain in the context of the KT-model.

[12] See (Long and Freese, 2006) for details on the calculation of marginal effects of continuous and categorical predictors.

**Table 4**
Average marginal effects of continuous variables, within setting comparisons and between setting comparisons.

| | | | | | | 95 % conf. int. | |
|---|---|---|---|---|---|---|---|
| | | *AME* | Robust $\widehat{sd}$ | *z* | $P > \|z\|$ | Lower | Upper |
| | $c_B$ | −0.07 | 0.004 | −15.73 | <0.001 | −0.08 | −0.06 |
| | *t* | 0.07 | 0.005 | 12.47 | <0.001 | 0.06 | 0.08 |
| *Baseline* | *vs.* | | | | | | |
| nPC_nD_nC | PC_nD_nC | −0.51 | 0.049 | −10.41 | <0.001 | −0.61 | −0.42 |
| nPC_D_nC | PC_D_nC | −0.38 | 0.037 | −10.26 | <0.001 | −0.45 | −0.30 |
| nPC_D_C | PC_D_C | −0.35 | 0.052 | −6.85 | <0.001 | −0.45 | −0.25 |
| nPC_nD_nC | nPC_D_nC | −0.12 | 0.038 | −3.04 | 0.002 | −0.19 | −0.04 |
| PC_nD_nC | PC_D_nC | 0.02 | 0.041 | 0.52 | 0.6 | −0.06 | 0.10 |
| nPC_D_nC | nPC_D_C | 0.18 | 0.042 | 4.25 | <0.001 | 0.10 | 0.26 |
| PC_D_nC | PC_D_C | 0.20 | 0.044 | 4.56 | <0.001 | 0.12 | 0.29 |

Fig. 4a and b display the effect of a post-rejection decision with and without pre-commitment. Without pre-commitment we see a negative effect so introducing a post-rejection decision must have led some *B*-players to reject when they would have accepted otherwise. The solid black line represents all transfer/cost constellations for which *B* should be theoretically indifferent between accepting and rejecting. Above that line accepting results in higher payoffs, below it rejecting is more profitable. As the darkest shading (i.e. the strongest effect) is located strictly above that line a considerable number of *B*-player must have rejected even though it was not payoff-maximizing. This is in line with our previous conjecture that *B*-players began to develop expectations about exploiting the *A*-players in the prisoners' dilemma or simply wanted to punish *A* at cost to themselves.

In contrast to that, Fig. 4b shows that, when *A* had to pre-commit, the effect of a post-rejection decision was miniscule but positive and is located strictly below the indifference line.[13] Put differently, if there was an effect at all under pre-commitment, then it was one of positive reciprocity or willingness to cooperate, i.e. accepting non-payoff-maximizing offers.

Performing the same analysis for the effect of communication we observe a positive treatment effect from communication both without pre-commitment (Fig. 4c) and with pre-commitment (Fig. 4d). Communication must thus have made *B* players accept who would have rejected otherwise. Without pre-commitment those *B*-players who would have rejected without communication are those that should have accepted if they were payoff-maximizing because the darkest shading is located strictly above the indifference line. In other words, communication led to an increase in acceptance probabilities because it reduces non-payoff-maximizing rejections when no pre-commitment is involved. With pre-commitment those *B*-players who would have rejected without communication are those that should have rejected if they were payoff-maximizing because the darkest shading is located strictly below the indifference line. As communication made these players accept anyway, the increase in acceptance probabilities can be attributed to an increase in positive reciprocity or willingness to cooperate when pre-commitment is involved.

## 5. Discussion

Our experimental results lend considerable support to the main predictions of the KT-model, even though subjects in our experiment were influenced by behavioral motives that are not taken into account in the original model. The experiment revealed behavior consistent with a tendency towards cooperation, positive and negative reciprocity, and inequality aversion. All these traits can work against Result 1 of the model. Cooperative behavior and positive reciprocity can lead to more agreements under pre-commitment and negative reciprocity to less agreement without pre-commitment. So if subjects had been led more by reciprocity or social preferences than by payoff-maximizing behavior the theoretical results could well have failed to hold. Nevertheless, in all our comparisons we observed that a pre-commitment reduced the likelihood of an agreement being reached.

Communication between the players also had an interesting effect. As we have come to expect from past experiments communication led to subjects having a higher willingness to deviate from payoff-maximization and behave cooperatively. Given this, we expected the treatment with communication to strengthen the effects such as cooperative behavior and reciprocity that can work against the model's results. However, while communication did have the expected effect on the willingness to cooperate it also led to a sharp fall in negative reciprocity and on the attempts of the *B*-players to exploit the *A*-players in the treatment without pre-commitment. Combined this resulted in the number of agreement being reached being the highest of all treatments in the no pre-commitment treatment with communication. This in turn led to the difference between pre-commitment and no pre-commitment increasing with communication and the difference once again being statistically significant.

All in all, our experimental results provide important backing to the main conclusion of the KT-model regarding the potentially counterproductive effects of pre-commitments in climate negations. We have shown that even if negotiating parties are driven by behavioral motives outside the model the results of the model can still be expected hold. Moreover, we have shown that if the environment is extended to include the realistic feature of communication between parties the main conclusion still holds strong.

A possible critique of the KT-model – and thus our experimental investigation – is that the pre-commitment of the *A*-players is not voluntary. It could be argued that in the case of voluntary pre-commitments positive reciprocity would be more pronounced. There are good reasons for not testing this by running experiments with voluntary pre-commitments using the KT-model. First, given the specific bargaining situation, we cannot expect to observe any voluntary pre-commitment. If *A*-players have the choice between committing and not committing, they will certainly opt for not committing, even if they harbor some kind of other regarding preferences. The reason is that to pre-commit actually increases the inequality of final payoffs.

Therefore, an inequality averse player would also choose to not pre-commit and would offer a low price ($t = 0$) which ensures that the *B*-player will agree to cooperate and which leads to equal payoffs for both players. Consequently, there is no motivation discussed in the literature on other regarding preferences which would make pre-commitment a rational choice.

This goes in line with the fact that we rarely observe truly altruistic pre-commitments in reality. In the context of climate damage abatement, for example, the pre-commitment is usually sold as having the advantage of allowing countries that pre-commit to gain a competitive advantage in the development of cleaner technologies and not out of some altruistic consideration.

Furthermore, previous experiments on the effect of pre-commitments in which these were voluntary have shown little evidence for reciprocity which go beyond that which we observed in our experiment. In the literature on leadership in climate negotiations we have seen only moderate levels of reciprocal behavior towards pre-commitments (Werner Güth et al., 2007; Vittoria Levati et al., 2007; Gächter et al., 2012; Sturm and Weimann, 2008). In a sequential bargaining experiment Brosig et al. (2004) find that voluntary pre-commitments are greatly taken advantage of and in a current paper Heinrich and Weimann (2014) find that in dictator games in which recipients could choose

[13] Note that the shading is so bright that it may become invisible in some printouts. We kept this color coding anyway to maintain comparability to the other three figures.

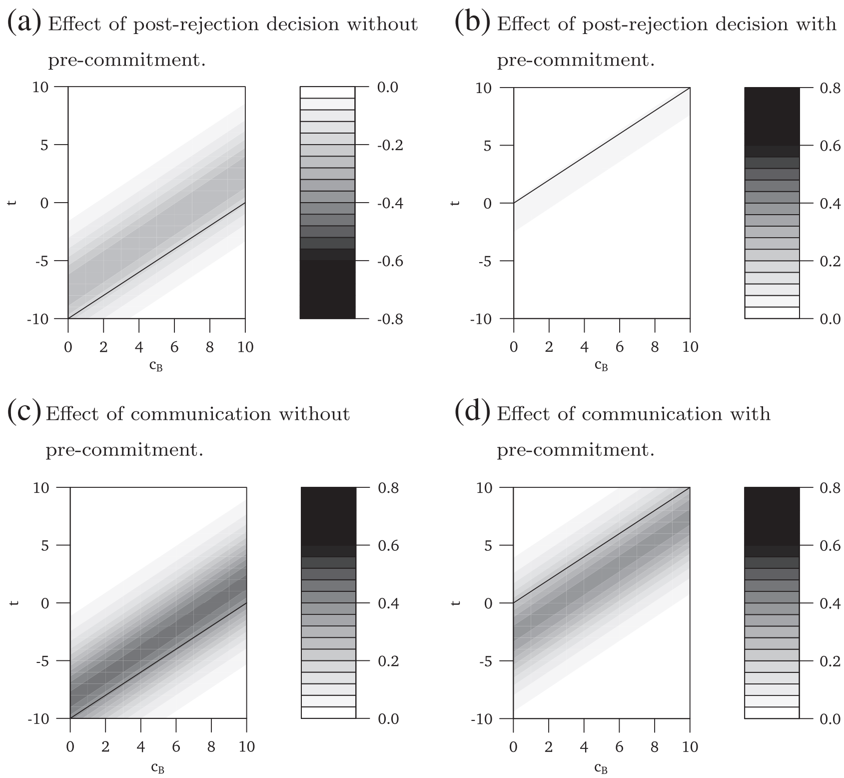



**Fig. 4.** Treatment effects on estimated acceptance probabilities over all transfers $t$ and costs $c_B$. Darker shadings represent stronger effects on the estimated acceptance probability. The black line represents theoretical $t/c_B$-combinations for which a payoff-maximizing $B$ should be indifferent between accepting and rejecting.

between different modified dictator games there was no reciprocity shown by dictators. Thus, there is no current experimental evidence that suggests a voluntary pre-commitment would lead to more reciprocity than that which we observed in our experiment.

## Acknowledgement

We are grateful to Annette Kirstein, Michael Kvasnicka, Hendrik Thiel, Florian Timme, the participants of the 2013 meeting of the Social Science Commission/Verein für Socialpolitik, the participants of the 2013 Economic Science Association World Meeting, and two anonymous reviewers for their valuable feedback.

## Appendix A. Logistic regression results

**Table A.5**
Logistic regression with nPC_nD_nC as baseline. Because of the dependence of observations within a subject we calculated robust variance estimates using the cluster() option in STATA. *Model characteristics*: (1) No multicollinearity: All variance inflation factors are $\ll 10$, (2) No perfect separation (3) Likelihood-Ratio test for overall fit: $\chi^2 = 153.04$, $p < 0.001$ (4) Predictive accuracy: In 84.5 % of all cases the observed 1 (0) was predicted at $p >= 0.5$ ($p < 0.5$).

| | $\hat{\beta}$ | Robust $\widehat{sd}$ | $z$ | $p$ | Odds ratio | 95 % conf. int. | |
|---|---|---|---|---|---|---|---|
| | | | | | | Lower | Upper |
| (Intercept) | 4.81 | 0.538 | 8.95 | <0.001 | – | – | – |
| Transfer t | 0.56 | 0.064 | 8.74 | <0.001 | 1.76 | 0.44 | 0.69 |
| Cost cB | −0.57 | 0.057 | −10.00 | <0.001 | 0.57 | −0.68 | −0.46 |
| PC_nD_nC | −4.26 | 0.670 | −6.35 | <0.001 | 0.01 | −5.57 | −2.94 |
| nPC_D_nC | −1.13 | 0.401 | −2.82 | 0.005 | 0.32 | −1.92 | −0.35 |
| PC_D_nC | −4.07 | 0.538 | −7.56 | <0.001 | 0.02 | −5.13 | −3.02 |
| nPC_D_C | 0.84 | 0.571 | 1.47 | 0.141 | 2.32 | −0.28 | 1.96 |
| PC_D_C | −2.49 | 0.547 | −4.56 | <0.001 | 0.08 | −3.56 | −1.42 |

## Appendix B. Analysis of communication

A closer look at how the communication was used in the two treatments can shed some light on the widening of the cooperation gap. Table B.6 shows how the chats were used in the two treatments with communication. As can be seen cost information was revealed in more cases in the nPC_D_C-treatment (71 %) than in the PC_D_C-treatment (42 %) with the share of truthful revelation being roughly the same between the two (60 % vs. 69 %). After there was a communication about costs there was a strong tendency to reach an agreement. Furthermore, the tendency was stronger in the nPC_D_C-treatment than in the PC_D_C-treatment (94 % vs. 72 %). Thus, the fact that both costs were revealed more often and that the tendency to reach agreement after a cost revelation was stronger in the nPC_D_C-treatment is consistent with the widening of the cooperation gap.

In the nPC_D_C-treatment 65 % of the subjects made a firm commitment to come to an agreement in the chat while in the nPC_D_C-treatment it was 53 %. Again, there was a strong tendency of offers to be accepted after a clear agreement was reached in the chat (97 % and 90 %). This difference between agreements reached in the chats is again consistent with a widening of the cooperation gap.

**Table B.6**
Communication in Setting 3.

| Cases in which… | nPC | PC |
|---|---|---|
| … *A* and *B* revealed own cost in chat | 214 (71.3%) | 127 (42.3%) |
| … revealed cost was true value | 129 (60.3%) | 87 (68.5%) |
| … *A* and *B* agreed after costs were revealed | 202 (94.3%) | 91 (71.7%) |
| … *A* and *B* agreed in chat | 194 (64.7%) | 160 (53.3%) |
| … *A* and *B* agreed in game | 189 (97.4 %) | 144 (90.0 %) |

## Appendix C. Instructions

The following text represents the instructions in the nPC_D_nC-treatment (translated from German):

OTTO VON GUERICKE UNIVERSITÄT MAGDEBURG FACULTY OF ECONOMICS AND MANAGEMENT

MaXLab Magdeburger Experimentallabor für Wirtschaftsforschung

# Rules and instructions, please read carefully!

1. **Before the experiment** switch-off your mobile phone. Read carefully the instuctions below and inform the experimenter by a show of hands if you have questions.
2. **During the experiment** you are neither allowed to talk nor to leave your place. We let you know when the experiment is finished. If your screen does not respond to your entry immediately, an other player has not decided yet. Please be patient.
3. **After the experiment** you get your payoff. Please remain seated until you are called.

**General procedure**: Before the actual experiments you will play 3 practice rounds against the computer. These practice rounds are not payoff relevant. The purpose of these rounds is to familiarize yourself with the experimental environment. The computers behavior will not be arbitrary but fully payoff-maximizing. In order that you are fully able to comprehend the computers actions you will be given the computers private cost information in the first two practice rounds. This information will be unknown to you both in the third practice round as well as in the actual game.

In the following 5 rounds you will play the actual experiment in which you will play each other participant once. At the start you will receive an endowment which can grow or shrink depending on your actions in each of the rounds. Your final payoff after the 5 rounds will certainly be positive. There are a total of 10 participants, 5 of which will be given the role of an $A$ player and 5 the role of a $B$ player. In each round the first decision will always be made by $A$ and the second by $B$. The $B$ player will thus be able to observe the action taken by the $A$ player before making a decision. Designated roles will stay the same throughout the experiment.

**Content of the experiment**: In each of the rounds you and the other player can contribute some effort $e$ towards the provision of a common good. The size of $e$ is measured in lab dollars with 5 lab dollars being worth 1 euro. The choice $e = 10$ means that you will contribute the effort, and $e = 0$ means you will not contribute the effort. Other values of $e$ are not possible. You and the other player will receive the sum of both your contributions. That is, 0 (if neither of you contributed), 10 (if one of you contributed), 20 (if you both contributed). This payoff will be reduced by your cost of contributing. The cost will be zero if you do not contribute ($e = 0$) and $10 + c$ if you do contribute ($e = 10$). The whole number $c$ is randomly selected and lies between 1 and 9, with each of the nine values being equally probable. Before each round each player will find out their own cost but not the cost of the other participant.

The $A$ player can offer the $B$ player a transfer payment which can take any whole number between −10 and +10. Positive values mean there is a payment from $A$ to $B$ and negative values mean there is a payment from $B$ to $A$. Player $B$ can either accept or reject the transfer offer. If the $B$ player accepts the transfer payment will be made and both players are required to contribute the effort $e = 10$ (you will then no longer be able to decide over this freely as $e = 10$ will be set automatically). If $B$ rejects the offer then both players can choose freely whether to contribute or not, i.e., you can either choose $e = 10$ or $e = 0$.

**Experimental timeline:**:

- **Stage 1:** Player $A$ receives the cost information $10 + c_A$ and offers B a transfer $-10 \le t \le 10$
- **Stage 2:** Player $B$ receives the cost information $10 + c_B$ and decides whether to accept the offer or not.
- **Acceptance, Stage 3a:** $e = 10$ is set for both players.
- **Rejection, Stage 3b:** Both player decide between $e = 10$ and $e = 0$, without knowing the decision of the other player.

After Stage 3 the current round ends and you are informed about your current payoff. Payoffs from previous rounds are logged and made available to you in each round.

**Payoffs**: Lab-dollars $\pi$ in a round are determined as follows (**red** for **Player A**, **blue** for **Player B**):

- If $B$ **accepts** (Stage 3a):

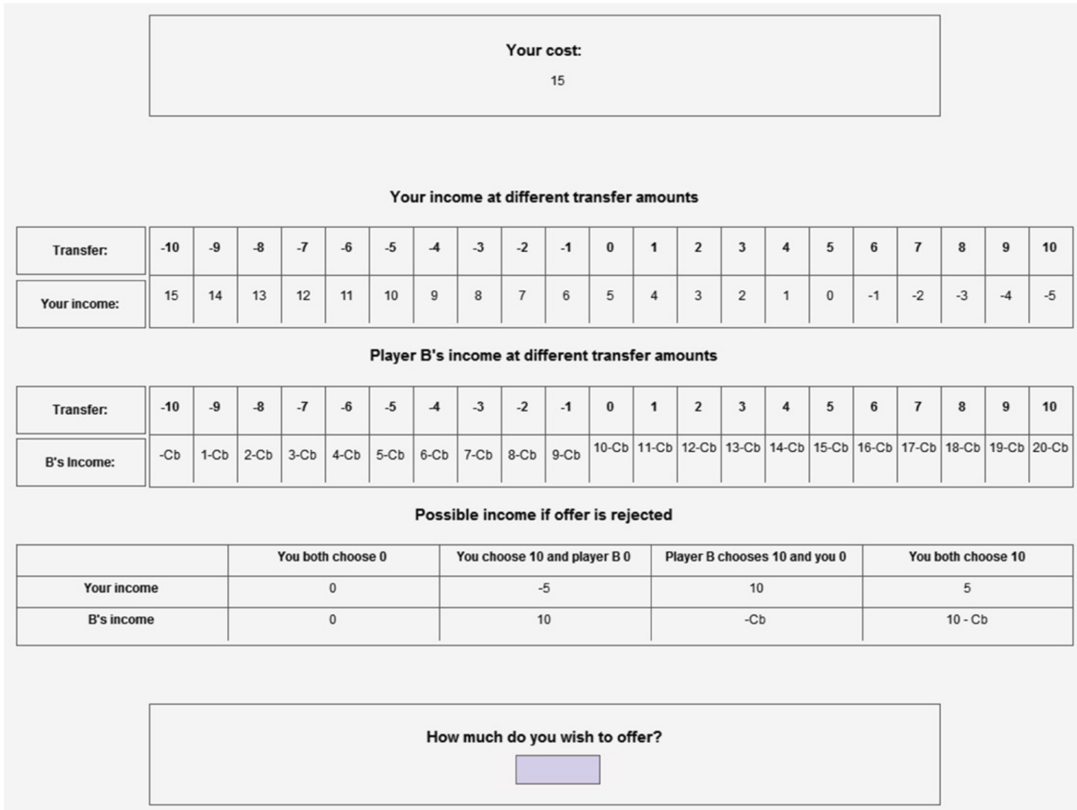
| | gross payoff | − | costs | ± | transfer |
|---|---|---|---|---|---|
| $\pi_A =$ | **20** | − | $(10 + c_A)$ | − | **t** |
| $\pi_B =$ | **20** | − | $(10 + c_B)$ | + | **t** |

- If $B$ **rejects** (Stage 3b):

| | | B | | | |
|---|---|---|---|---|---|
| | effort | $e_B = 0$ | | $e_B = 10$ | |
| | $e_A = 0$ | **0** | **0** | **10** | $-c_B$ |
| A | $e_A = 10$ | $-c_A$ | **10** | $10 - c_A$ | $10 - c_B$ |

ID: NPC-NC

## Appendix D. Screenshots



**Your cost:**
15

**Your income at different transfer amounts**

| Transfer: | -10 | -9 | -8 | -7 | -6 | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Your income: | 15 | 14 | 13 | 12 | 11 | 10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 | -1 | -2 | -3 | -4 | -5 |

**Player B's income at different transfer amounts**

| Transfer: | -10 | -9 | -8 | -7 | -6 | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| B's Income: | -Cb | 1-Cb | 2-Cb | 3-Cb | 4-Cb | 5-Cb | 6-Cb | 7-Cb | 8-Cb | 9-Cb | 10-Cb | 11-Cb | 12-Cb | 13-Cb | 14-Cb | 15-Cb | 16-Cb | 17-Cb | 18-Cb | 19-Cb | 20-Cb |

**Possible income if offer is rejected**

| | You both choose 0 | You choose 10 and player B 0 | Player B chooses 10 and you 0 | You both choose 10 |
|---|---|---|---|---|
| Your income | 0 | -5 | 10 | 5 |
| B's income | 0 | 10 | -Cb | 10 - Cb |

**How much do you wish to offer?**

**Fig. D.5.** Screen of player *A* in Setting 2 and nPC.

**Your cost:**
17

**Player A's offer:**
4

**Your income if you accept**
7

**Player A's income at different cost if you accept the offer**

| A's possible costs: | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| A's income: | 5 | 4 | 3 | 2 | 1 | 0 | -1 | -2 | -3 |

Possible income if offer is rejected

| | You both choose 0 | You choose10 and player A 0 | Player A chooses 10 and you 0 | You both choose 10 |
|---|---|---|---|---|
| Your income | 0 | -7 | 10 | 3 |
| A's income | 0 | 10 | -Ca | 10 - Ca |

**Do you accept the offer?**
Accept the offer
Reject the offer

**Fig. D.6.** Screen of player *B* in Setting 2 and nPC.

## References

Barrett, S., 1998. On the theory and diplomacy of environmental treaty-making. Environ. Resour. Econ. 11, 317–333.

Bolton, G.E., Ockenfels, A., 2000. ERC: A theory of equity, reciprocity, and competition. Am. Econ. Rev. 90, 166–193.

Brosig, J., Weimann, J., Ockenfels, A., 2003. The effect of communication media on cooperation. Ger. Econ. Rev. 4 (2), 217–241.

Brosig, J., Weimann, J., Yang, C.L., 2004. Communication, reputation, and punishment in sequential bargaining experiments. J. Inst. Theor. Econ. 160 (4), 576–606.

Carraro, C., Siniscalco, D., 1993. Strategies for the international protection of the environment. J. Public Econ. 52 (3), 309–328.

Cox, J.C., Friedman, D., Gjerstad, S., 2007. A tractable model of reciprocity and fairness. Games Econ. Behav. 59, 17–45.

Falk, A., Fischbacher, U., 2006. A theory of reciprocity. Games Econ. Behav. 54 (2), 293–315.

Fehr, E., Schmidt, K.M., 1999. A theory of fairness, competition, and cooperation. Q. J. Econ. 114 (3), 817–868.

Fischbacher, U., 2007. z-Tree: Zurich toolbox for ready-made economic experiments. Exp. Econ. 10 (2), 171–178 (June).

Gächter, S., Nosenzo, D., Renner, E., Sefton, M., 2012. Who makes a good leader? Social preferences and leading-by-example. Econ. Inq. 50, 867–879.

Gresik, T.A., Satterthwaite, M.A., 1989. The rate at which a simple market converges to efficiency as the number of traders increases: An asymptotic result for optimal trade mechanisms. J. Econ. Theory 48, 304–332.

Güth, W., Schmittberger, R., Schwarze, B., 1982. An experimental analysis of ultimatum bargaining. J. Econ. Behav. Organ. 3 (4), 367–388.

Güth, W., Levati, M.V., Sutter, M., Van Der Heijden, E., 2007. Leading by example with and without exclusion power in voluntary contribution experiments. J. Public Econ. 91 (5), 1023–1042.

Heinrich, T., Weimann, J., 2014. A note on reciprocity and modified dictator games. Econ. Lett. 121, 202–205.

Konrad, K., Thum, M., 2014. Climate policy negotiations with incomplete information. Economica http://dx.doi.org/10.1111/ecca.12065 (Wiley Early view (Online version)).

Lange, A., Vogt, C., Ziegler, A., 2007. On the importance of equity in international climate policy: an empirical analysis. Energy Econ. 29, 545–562.

Lange, A., Löschel, A., Vogt, C., Ziegler, A., 2010. On the self-interested use of equity in international climate negotiations. Eur. Econ. Rev. 54, 359–375.

Levati, M.V., Sutter, M., Van der Heijden, E., 2007. Leading by example in a public goods experiment with heterogeneity and incomplete information. J. Confl. Resolut. 51, 793–818.

Levine, D.K., 1998. Modeling altruism and spitefulness in experiments. Rev. Econ. Dyn. 1, 593–622.

Long, J.S., Freese, J., 2006. Regression Models for Categorical Dependent Variables Using Stata, 2 edition. Stata Press.

Meyerson, R.B., Satterthwaite, M.A., 1983. Efficient mechanisms for bilateral trading. J. Econ. Theory 29, 265–281.

Nowak, M.A., Roch, S., 2006. Upstream reciprocity and the evolution of gratitude. Proc. R. Soc. 274, 605–609.

Nowak, M.A., Siegmund, K., 2005. Evolution of indirect reciprocity. Nature 437, 1291–1298.

Rabin, M., 1993. Incorporating fairness into game theory and economics. Am. Econ. Rev. 83, 1281–1302.

Rob, R., 1989. Pollution claim settlements under private information. J. Econ. Theory 47, 307–333.

Sturm, B., Weimann, J., 2008. Unilateral emissions abatement: an experiment. Environmental Economics, Experimental Methods. Routledge, pp. 157–183.

Valley, K.L., Moag, J., Bazerman, M.H., 1998. A matter of trust: effects of communication on the efficiency and distribution of outcomes. J. Econ. Behav. Organ. 34, 211–238.

# Mispricing of Risk in Sovereign Bond Markets with Asymmetric Information

*Benedikt Mihm*
Otto-von-Guericke University Magdeburg

**Abstract.** *The likelihood that a government will repay its sovereign debt depends both on the amount of debt it issues and on the government's future ability to repay. Whilst the former is publicly observable, the government may have more information about the latter than investors. This paper shows that this asymmetric information problem impairs the market's ability to differentiate economies according to their fiscal sustainability, and can lead to a disconnect between bond prices and default risk. The model can help rationalise the behaviour of Eurozone bond prices prior to the recent European sovereign debt crisis.*

**JEL classification:** *E60, F34, G15.*

**Keywords:** Default risk; sovereign debt; political uncertainty; signalling.

## 1. INTRODUCTION

A puzzling feature of the recent European sovereign debt crisis is how poorly bond prices prior to the crisis predicted which countries were set to be engulfed by it. In the period between 1999 and 2007, the spreads in yields between Eurozone countries were close to zero, appearing to indicate that the default risk of the countries was perceived to be essentially the same. However, the crisis revealed that Eurozone countries had in fact differed greatly in their levels of fiscal sustainability with the crisis pushing some countries close to bankruptcy, whilst others continued on as normal.

The purpose of this paper is to show that the mispricing of risk can be an equilibrium outcome in sovereign bond markets with asymmetric information. The paper presents a model of a sovereign bond market in which countries can have the same equilibrium bond price even if the risk that the countries' governments will default on their debt obligations differ.

There are a total of three market frictions in the model. First, governments excessively discount the future due to uncertainty about re-election as in Alesina and Tabellini (1990). This *political distortion* results in a tendency of governments to accumulate positive levels of debt even in the absence of any consumption smoothing motive. Second, there is a lack of commitment on the part of the government to repaying its debt. This *commitment problem* introduces the possibility of default, an option a government will make use of as long as the benefit outweighs the cost. As a result, default is a potential equilibrium outcome as is typical in the willingness-to-pay literature introduced by Eaton and Gersovitz (1981), and means investors must take default probabilities into account when making their bond demand decision. Third, a government has more information about the fundamentals of their economy than investors. This *information asymmetry* means that investors must infer the fundamentals of an economy from its government's bond supply decision. However, governments may have no interest in having investor correctly infer the fundamentals. This strategic consideration introduces a signalling problem in the spirit of Spence (1973) to the framework, as a government needs to take into account that its debt issuance provides a signal to investors about its economy's fundamentals.

In this environment, bond prices can become disconnected from default risks in equilibrium, even though default risk is the only determinant of the bond price. This disconnect occurs because a government whose economy has poor fundamentals has a higher default probability and this in turn should reduce the price it gets for its bonds. If these fundamentals are unknown to investors, however, the government will have an incentive to mimic the bond supply of a government of an economy with better fundamentals in order to get a higher bond price. Since the only observable decision to investors is the government's bond supply, investors will take this mimicking incentive into account upon observing a government's bond supply. Hence, there is a possible pooling equilibrium in which two types of governments get the same price for their bonds even though the fundamentals of their economies – and thus default risks – differ. In this equilibrium, the bond supply of a government tells investors nothing about the type of government they face and the market fails to differentiate according to fiscal sustainability.

The model's pooling equilibrium is consistent with some of the pricing behaviour observed in the Eurozone sovereign bond market in the past decade. Prior to the European debt crisis Eurozone bond prices were remarkably similar with average 10-year-yield spreads relative to German bonds contained in range of little over 50 basis points in the period between 1999 and 2007 (Arghyrou and Kontonikas, 2011). Yet, it is now clear that the low spread in yields masked large underlying differences in the fiscal sustainability of Eurozone members with countries such as Greece and Portugal much worse effected by the 2007 financial crisis than countries such as Germany, France or the Netherlands. The model's pooling equilibrium is consistent with this chain of events since it also features countries with differing levels of fiscal sustainability being able to receive the same prices for their bonds.

## 2. RELATED LITERATURE

The model is related to three stands of the literature. First, it is related to the literature that introduces default risk into a framework in which a deficit bias arises due to political uncertainty as in Alesina and Tabellini (1990). Since the political uncertainty literature has focused on the tendency of advanced economies to accumulate debt it has not typically considered default, which is a rare event in such countries. Recent exceptions are Cuadra and Sapriza (2008) who quantify the role of political distortions in emerging markets, and Beetsma and Mavromatis (2014) who look at the role of debt mutualisation in a model with sovereign default. The contribution of my model is to add a hidden information friction to the problem and analyse the potential for this to lead to the risk of a government defaulting on its debt to be incorrectly priced.

Second, it relates to the literature which combines sovereign default risk modelled using the willingness-to-pay approach introduced by Eaton and Gersovitz (1981) with an asymmetric information problem. As in Eaton (1996), Alfaro and Kanczuk (2005) and Sandleris (2008), investors in my model are uncertain about the type of government from which they are purchasing bonds. Most closely related to mine is the work of Sandleris (2008) who also considers an environment in which a government knows more about the fundamentals of its economy than investors do. But whilst Sandleris (2008) analyses the cost of default in terms of the signal it sends about fundamentals, I consider the case in which the cost of default is exogenous but the government learns its fundamentals before it issues its debt and the debt issuance itself may act as a signal. This difference in the timeline allows me to tie the asymmetry of information to the way default probabilities are reflected in the price of bonds.

As in Corsetti and Dedola (2011) a sovereign's risk premium in the model, therefore, reflects other factors than only a country's fundamentals. Corsetti and Dedola (2011), however, analyse the interaction between fundamentals and confidence in driving fluctuations in risk premia, whereas I consider the role asymmetric information plays in the risk premium only partially reflecting a country's type.

Third, the signalling framework is related to the work of Spence (1973), whose seminal contribution on signalling in environments with asymmetric information has inspired a vast literature (Riley, 2001). The idea in my model is that a government takes into account the signalling effect of it bond supply when deciding how much debt to accumulate. This strategic incentive hampers investors in their attempts to distinguish between types and leads to the potential for bond price equalisation in equilibrium.

## 3. THE MODEL

### *3.1. Environment*

Consider a small open economy, whose government can borrow resources by selling bonds to investors on an international sovereign bond market. Time lasts for two periods ($t = 1,2$). The productive efficiency of the economy is determined by a parameter $\theta$ which can take one of two possible values $\theta_H > \theta_L > 0$, reflecting some underlying fundamentals. Specifically, an economy's period $t$ income is given by $y_t = \theta y$, where $y$ is an exogenously given endowment. The economy's fundamentals are determined by nature at the start of the game. With a probability $\lambda \in (0,1)$ the economy will have good fundamentals $\theta_H$ and with probability $1 - \lambda$ it will have poor fundamentals $\theta_L$. In what follows if a government is in charge of an economy with fundamentals $\theta_H$ it will be referred to as an $H$-government and if it is in charge of an economy with fundamentals $\theta_L$ it will be referred to as an $L$-government.

Uncertainty about re-election means that the government excessively discounts the future. At the start of period, one the government in the economy is controlled by a political party which cares only about the consumption of its own constituents. The party is re-elected at the end of the period with an exogenous constant probability $\pi \in (0,1)$. Letting $g_t \geq 0$ be the period $t$ consumption of a public good for the party's constituents, the utility function of the government of an economy of type $i \in \{L,H\}$ can be written as:

$$U_i = g_{1i} + \pi E[V_{2i}(g_{2i})], \tag{1}$$

where $V_{2i}$ is a function which represents that period two utility depends on a default option that the government has and which is defined later.[1]

Since the re-election probability is less than one, the government cares more about period one consumption than it does about consumption in period two. The political uncertainty, therefore, creates an incentive to issue debt, and since this incentive implies that the equilibrium bond supply will always be positive the terms bond supply and debt will be used interchangeably throughout the paper.

The government has no endowment available in period one so must finance its period one consumption by issuing one period real bonds $b_i$ at price $q_i$. These bonds are purchased by international investors on an international sovereign bond market with no legal enforcement to ensure debt repayment. This assumption reflects the relevant institutional structure for the borrowing of advanced economies, which are the focus of this paper.

In period two, the government receives a stochastic endowment $y$ with density $f(y)$. To keep things simple, $y$ is assumed to be uniformly distributed on [0,1]. The resource constraints in the two periods if there is no default are then given by:

$$g_{1i} = q_i b_i \tag{2}$$

$$g_{2i} = \theta_i y - b_i. \tag{3}$$

### 3.2. Timeline

The timing is as follows (Figure 1). At the start of period one the government observes the fundamentals $\theta_i$ and then makes its bond supply decision $b_i$. After the government's bond supply decision investors observe $b_i$ and decide whether to purchase the government's bonds or invest in another asset with a risk-free rate of return $r_f$. Investors do not know the country's fundamentals $\theta_i$ when making the bond demand decision. This simple asymmetry of information reflects that the fundamentals of an economy are likely to depend on things – such as structural reforms or government investment – that a government may know more about than investors. Since investors do not know $\theta_i$ they must form beliefs about $\theta_i$ based on $b_i$. At the end of period one the party in power faces an election and is re-elected with exogenous probability $\pi$.

1. The default decision will introduce non-linearity into the model, which is why to keep things tractable the consumption utility itself is linear.

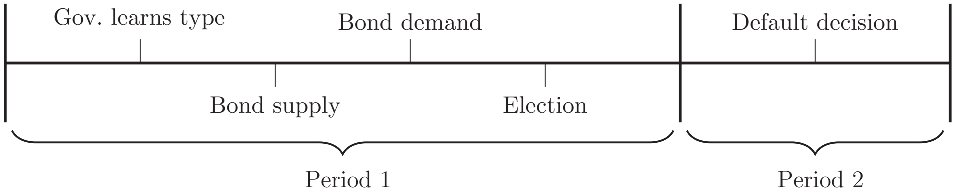



**Figure 1** Timeline.

In period two the government makes a default decision as the outcome of a trade-off between the cost and benefit of doing so. For a given debt level $b_i$ the probability that an *H*-governments will default is lower than it is for an *L*-government. It is about these default probabilities the investors must form beliefs.

### *3.3. Default decision*

In period two the government must decide on its debt repayment. The fact that there is no legal enforcement to ensure that a government repays its debt means that it has a default option. The benefit of using this default option is that resources can be used for consumption by the government that would otherwise be transferred to international investors who do not enter into the government's utility function.

There is a cost associated with a default as well, however. The cost comes in the form of a direct output cost that reduces period two resources to $\gamma\theta y$ after default with $\gamma \in [0,1)$. The default cost captures – in a reduced form way – the many potential disruptions which may arise from a default and thus have a negative effect on output.[2]

A government of type *i* will then have period two resources contingent on its default decision given by:

$$V_{2i}(g_{2i}) = \begin{cases} \theta_i y - b_i & \text{if there is no default} \\ \gamma\theta_i y & \text{if there is a default.} \end{cases}$$

For a given level of bonds issued in period one and the realisation of the endowment, the government thus solves:

$$V_{2i} = \max\{\theta_i y - b_i, \gamma\theta_i y\}.$$

If the left term in the brackets is greater than the right term, then the government will not default and vice versa. Thus, for a given debt level $b_i$ we can derive a threshold for the endowment $y^*$ such that $y < y^*$ leads to default and $y \geq y^*$ leads to repayment. The threshold value is given by:

$$y^* = \frac{b_i}{\theta_i(1-\gamma)}.$$

The default probability function $\delta_i(b_i, \theta_i, \gamma)$ then represents the likelihood that a government will default when it has a debt level of $b_i$, fundamentals $\theta_i$ and keeps

2. There has been a long debate about what exactly the costs of default are (Hatchondo *et al.*, 2007). Given the simple two period structure of this model an output cost is used since, for example, being barred from future borrowing would not be effective.

a share of output $\gamma$ if it chooses to write off its debts. The default probability function can be written as:

$$\delta_i(b_i, \theta_i, \gamma) = pr\{y < y^*\} = F\left[\frac{b_i}{\theta_i(1-\gamma)}\right], \tag{4}$$

where $F[y]$ is the CDF of the endowment $y$.

The following proposition states how the default probability function relates to the debt level $b_i$, the fundamentals $\theta_i$ and the share of output the economy has left after default $\gamma$.

**Proposition 1** The default probability $\delta_i(b_i, \theta_i, \gamma)$ is increasing in $b_i$ and $\gamma$ and decreasing in $\theta_i$.

*Proof.* See Appendix A.■

### 3.4. Investors' bond demand

At the end of period one the price the government receives for its bonds is determined on a competitive international sovereign bond market with no legal enforcement to ensure debt repayment. Risk-neutral international investors observe the bond supply decision of the government and decide to either invest in the government's bonds or to invest in an riskless outside option that pays the exogenous and constant international risk-free interest rate $r_f$. Whilst investors can fully observe the bond supply of the government they do not know the fundamentals of government's economy when they make this investment decision. They must thus infer from the bond supply what the fundamentals may be.

When observing the bond supply decision of the government, investors form beliefs about the fundamentals $\theta_i$ of the economy. Since the default decision affects period two utility the bond supply will act as a signal to investors based upon which they can differentiate between types according to their likelihood of default. These beliefs are given by the conditional probability $\mu(\theta_i|b)$. If upon observing the bond supply decision investors know with certainty that they face a government of type $i$ then they have beliefs given by $\mu(\theta_i|b) = 1$. If investors can infer nothing about the government's type from the bond supply, then the investors' posterior beliefs are equal to their priors so that $\mu(\theta_H|b) = \lambda$ and $\mu(\theta_L|b) = 1 - \lambda$.

For investors to make zero profits requires a no arbitrage condition so that the expected rate of return investors receive from investing in the government's bonds is equal to the rate of return investors receive from investing in the risk-free outside option. A lender who lends $q$ units to the government in period one receives one unit in period two if the government does not default and zero if it does. From the perspective of the investors they will thus get a return of $1/q_i$ unit per unit lent to a government of type $i$ with probability $1 - \delta_i(b_i, \theta_i, \gamma)$ and zero with probability $\delta_i(b_i, \theta_i, \gamma)$. The break even no arbitrage condition is then:

$$1+r_f=\frac{\mu(\theta_L|b)[1-\delta_L(b_L,\theta_L,\gamma)]+\mu(\theta_H|b)[1-\delta_H(b_H,\theta_H,\gamma)]}{q_i}. \tag{5}$$

To simplify notation and without loss of generality the risk-free interest rate is set to $r_f = 0$. Rearranging the no arbitrage condition (5) then pins down the price function for a government of type $i$ as:

$$q_i(b_i,\theta_i,\gamma)=\mu(\theta_L|b)[1-\delta_L(b_L,\theta_L,\gamma)]+\mu(\theta_H|b)[1-\delta_H(b_H,\theta_H,\gamma)]. \tag{6}$$

Since the price function is exclusively determined by the default probability of the government its key properties follow directly from the key properties of the default probability stated in Proposition 1. That is:

**Corollary 1.** The bond price function $q_i(b_i,\theta_i,\gamma)$ is decreasing in $b_i$ and $\gamma$ and increasing in $\theta_i$.

*Proof.* Follows from Proposition 1.■

It follows from Corollary 1 that an *H*-government is less tightly constrained in its debt issuance than an *L*-government. The fact that an *H*-government has a lower likelihood of defaulting means that it receives a higher price for its bonds at any given level of debt. Thus, having a high level of debt has a lower negative effect on the bond price of an *H*-government than it does for an *L*-government, and the *H*-government will choose a higher level of debt *ceteris paribus*.

### *3.5. Government's bond supply*

At the beginning of period one a government of type $i$ observes the fundamentals of its economy and then decides how many bonds to issue, taking bond demand as given. That is, it wishes to maximise the following intertemporal utility function:

$$U_i(b_i,q_i,\gamma)=q_i(b_i,\theta_i,\gamma)b_i+\pi\left[\int_{\frac{b_i}{\theta_i(1-\gamma)}}^{1}(\theta_i y-b_i)f(y)dy+\int_0^{\frac{b_i}{\theta_i(1-\gamma)}}\gamma\theta_i y f(y)dy\right]. \tag{7}$$

Solving the integral in (7) and using the price function (6) the intertemporal utility of a government of type $i$ that faces its own price function can be written as:[3]

$$U_i(b_i,q_i,\gamma)=\left[b_i-\frac{b_i^2}{\theta_i(1-\gamma)}\right]+\pi\left[\frac{\theta_i}{2}+\frac{b_i^2}{2\theta_i(1-\gamma)}-b_i\right]. \tag{8}$$

Utility representation (8) makes clear the intertemporal trade-offs the government faces when making its bond supply decision. Debt increases period one utility and reduces period two utility by a linear term, representing the usual trade-off when transferring resources between periods.

3. The explicit derivations for this section are given in Appendix B.

More important for the purposes of this paper, however, are the two quadratic terms that represent the trade-off created by the interaction between the government and investors on the sovereign bond market. This trade-off works as follows. In period two the government has the option to default on its debt. The benefit of having this default option is represented by the quadratic term in the second set of brackets and arises due to the possibility the government has of reducing the resources it must transfer to investors. A government that has a high debt level or whose economy has poor fundamentals, and which thus has a higher default probability, will be more likely to receive this benefit from having the opportunity to default in period two.

However, the higher period two benefit of having the default option for a government with a higher default probability is reflected in a tighter period one price function. This cost of having the default option is represented by the quadratic term in the first set of brackets. A government with a higher debt level or whose economy has poor fundamentals receives a lower price for the bonds and this reduces its period one utility, as it reduces the amount of resources it can transfer between periods for each bond sold. This constraint is a market disciplining mechanism at work. The market forces a government to trade-off its period two benefit of having the option to default with a cost of default via the bond price.

If a government of type $i$ does not face its own price function but rather the price function of a government of type $j \in \{L,H\}$, then intertemporal utility function is:

$$U_i(b_i, q_j, \gamma) = \left[b_i - \frac{b_i^2}{\theta_j(1-\gamma)}\right] + \pi\left[\frac{\theta_i}{2} + \frac{b_i^2}{2\theta_i(1-\gamma)} - b_i\right]. \tag{9}$$

Thus, a government of type $i$ will be better off facing the price function of a government of type $j$ whenever $\theta_j > \theta_i$. Since the price function is increasing – and thus the cost of default is falling – in $\theta$, a government will always be strictly better off being seen as being in charge of an economy with better fundamentals than it actually has.

## 4. EQUILIBRIUM SOVEREIGN RISK PRICING

### *4.1. Perfect information risk pricing*

Under perfect information both the government and investors know the fundamentals of a government's economy when making their decisions. The perfect information equilibrium serves as the benchmark for analysing the effect that asymmetric information has on equilibrium outcomes. The perfect information equilibrium of the model is defined as follows:

Definition 1. A *perfect information equilibrium* is a bond supply $b_i^p$, a bond price $q_i^p$ and a default probability $\delta_i^p$ such that:

- The bond supply $b_i^p$ is the solution to the government's optimisation problem when it takes the bond price function as given.

- The bond price $q_i^p$ reflects default probability $\delta_i^p$ and is consistent with investors' expected zero profits.

The perfect information equilibrium triple $(b_i^p, q_i^p, \delta_i^p)$ for a government of type $i$ is found by maximising the government's intertemporal utility function (8), i.e. maximising the government's utility, given it faces its own price function. This maximisation gives the perfect information equilibrium debt level $b_i^p$ for a government of type $i$. The equilibrium bond supply is shown in Appendix A. A feature of the equilibrium bond supply worth noting is that it is decreasing in the re-election probability $\pi$ regardless of type. This negative relationship represents the fact that a government that has a lower probability of being in power in the next period has a higher willingness to accumulate debt as there is chance it will not have to bear the cost of debt repayment.

Plugging $b_i^p$ into the default probability function then gives an equilibrium default probability $\delta_i^f$. Finally, plugging $\delta_i^p$ into the price function then gives $q_i^p$. As shown in Appendix A, the default probability is also decreasing in the re-election probability as a government that has a lower probability of staying in power will issue more debt and thus have a higher probability default. As a result, the equilibrium bond price a government receives will be lower if it is more likely to exit power and thus the equilibrium price takes into account the incentive effects of the political uncertainty.

The following proposition now outlines the equilibrium outcomes under perfect information:

**Proposition 2.** There is a unique perfect information equilibrium. Moreover, in this perfect information equilibrium the debt level of an $H$-government is higher than that of an $L$-government, both types receive the same equilibrium bond price and have the same equilibrium default probabilities. That is:

- $b_H^p > b_L^p,\ q_H^p = q_L^p,\ \delta_H^p = \delta_L^p.$

*Proof.* See Appendix A.■

Under perfect information an $H$-government faces a more relaxed pricing function than an $L$-government. This results in an $L$-government issuing a lower debt level than an $H$-government in equilibrium since investors see it as less credit worthy and its ability to transfer resources between periods is more tightly constrained. Again, this is the market disciplining mechanism at work. By forcing a government with a higher *ex-ante* default probability to issue less debt the market is able to perfectly equalise the *ex-post* equilibrium default probabilities across types. This in turn results in the equilibrium bond price of both types of government being equal as the default probability determines the equilibrium bond price.

The perfect information equilibrium has two important features. (1) A government that is more credit worthy is able to issue a higher level of debt in equilibrium. This outcome is what one would hope for as the market differentiates governments in terms of their fiscal sustainability. (2) The equilibrium bond price of both types of debt is the same, as the default probabilities of the two types of government are equalised. Again, this outcome is what one would hope for in a well-functioning bond market. Since the default probability is the only determinant of the bond price, we should be able to infer from an equal bond price that countries are equally fiscally sustainable.

### *4.2. Risk mispricing under asymmetric information*

When the fundamentals of an economy are not observable to the investors they will have to infer from the bond supply decision of the government what the fundamentals of the economy are. Since a government's utility depends on its likelihood of default and the likelihood of default depends on the level of debt, the bond supply acts as a signal to investors based upon which they can update beliefs regarding a government's type. A Perfect Bayesian equilibrium for this setting is defined as follows:

Definition 2. A *Perfect Bayesian equilibrium* is a set of bond supply decisions $b_i^*$, bond prices $q_i^*$, default probabilities $\delta_i^*$ and $\delta_j^*$ for $i \neq j$, and investors beliefs $\mu(\theta_i|b)$ such that:

- The bond supplies $b_i^*$ are optimal given the beliefs of investors.
- The bond prices $q_i^*$ reflect default probabilities $\delta_i^*$, and when necessary $\delta_j^*$, and are consistent with investors' expected zero profits.
- Beliefs $\mu(\theta_i|b)$ are consistent with Bayes rule.

The first part of the definition requires that the bond supply decisions fulfil the relevant incentive compatibility constraints. For the second part, the bond demand must be such that either a government faces its own price function or when necessary a price function which reflects a weighted average of the default probabilities of the two possible types. The condition that prices can also reflect the default probability of the other type of government differs from the condition under perfect information. However, it is a direct result of the information asymmetry and is essential for the zero profit condition to also hold.

The third part requires that investors' beliefs be consistent with Bayes rule on the equilibrium path. There will thus be one of the following two types of equilibria. (1) A *separating equilibrium* in which different types of government make different bond supply decisions and reveal their type perfectly, i.e. in which $b_H^* \neq b_L^*$ and $\mu(\theta_H|b_H^*) = 1$ and $\mu(\theta_L|b_L^*) = 1$. (2) A *pooling equilibrium* in which both types of government make the same bond supply decision and investors' posterior beliefs are equal to their priors, i.e. we have $b^* = b_H^* = b_L^*$ and $\mu(\theta_H|b^*) = \lambda$ and $\mu(\theta_L|b^*) = 1 - \lambda$.

The first key result is that since the perfect information equilibrium debt levels are such that $b_H^p > b_L^p$ they cannot be an equilibrium outcome under asymmetric information. This result follows from the following lemma:

**Lemma 1.** Let $b_H^*$ be a possible equilibrium bond supply for an *H*-government and $b_L^*$ a possible equilibrium bond supply for an *L*-government then under asymmetric information we must have $b_H^* \leq b_L^*$.

*Proof.* See Appendix A.■

The intuition behind Lemma 1 is as follows. Both types of government prefer to face an *H*-government's more relaxed price function than the price function of an *L*-government. Under asymmetric information an *L*-government thus has an incentive to mimic the bond supply decision of an *H*-government so it can face its preferred price function. Furthermore, for any $b_H^* > b_L^*$ if an *H*-government prefers to choose $b_H^*$ and face the *H*-government's price function so does an *L*-government and the *L*-government will have an incentive to mimic. The only way for there to be an equilibrium in which $b_H^* \neq b_L^*$ is for $b_H^*$ to be low enough for an *L*-government to prefer to choose a higher debt level and face its own price function, than to mimic and face the *H*-government's price function.

Lemma 1 requires that in any separating equilibrium an *H*-government must have a lower bond supply than an *L*-government. It also follows that in such an equilibrium an *H*-government will have a lower default probability than an *L*-government and receive a higher equilibrium bond price. In this type of separating equilibrium, there is no bond price equalisation as there is under perfect information. However, there is no risk mispricing either since prices accurately reflect the default probabilities of the different types. The following proposition formalises this point:

**Proposition 3.** Under asymmetric information, an equilibrium in which an *H*-government has a different debt level than an *L*-government will not feature risk mispricing as the *H*-government receives a different bond price than the *L*-government reflecting their different default probabilities.

*Proof.* See discussion above.■

A condition for the existence of a separating equilibrium is derived in Appendix B. However, for the purposes of the analysis here the pooling equilibrium is of more interest as it does exhibit risk mispricing. In the pooling equilibrium, both types of governments supply the same debt level and the investors can infer nothing about the fundamentals of an economy by observing the bond supply decision of its government. They will thus offer a government a pricing function which is a weighted sum of the default probabilities of the two types, with the weights given by the investors' priors. The following additional assumption is used to derive the existence of pooling equilibria:

**Assumption 1.** $\lambda > \pi/2$.

Assumption 1 is required for the following reason. As mentioned above the re-election probability $\pi$ affects the incentive to accumulate debt regardless of type. Moreover, this incentive effect is present both if a government faces its own price function and if it faces the pooled price function. However, a low probability $\lambda$ of there being $H$-governments reduces the incentive to accumulate debt only in a pooling equilibrium as it increases the cost of higher debt via a more restrictive pooled price function. If $\lambda$ is very low it can happen that the lowest level of debt at which an $L$-government is still just willing to be in a pooling equilibrium is greater than the highest level of debt at which an $H$-government is willing to be in a pooling equilibrium. Assumption 1 holding is sufficient but not necessary to rule this out.

The following proposition now establishes the existence of pooling equilibria and is the main result of the paper:

**Proposition 4.** Under Assumption 1 there exists an equilibrium bond supply $b^*$ under asymmetric information that is chosen by both an $H$-government and an $L$-government. In this equilibrium, both types of government have the same debt level, the same bond price but have different default probabilities. That is:

- $b_H^* = b_L^*$, $q_H^* = q_L^*$, $\delta_H^* < \delta_L^*$.

*Proof.* See Appendix A.■

As under perfect information, the equilibrium bond price is the same for both types of government. However, unlike under perfect information the equalisation of bond prices does not result from the fact the both types have the same equilibrium default probabilities but rather that both types have the same equilibrium debt level.

This outcome represents two central distortions due to the information asymmetry in the market. (1) The market's disciplining mechanism does not work to allow a more fiscally sustainable government to enjoy a higher level of debt. Since investors must take the $L$-government's mimicking incentive into account when buying bonds, the $H$-government's better fiscal position is not rewarded with a less tightly constrained debt issuance. (2) The default risk is not accurately reflected in the bond price. Thus, under asymmetric information observing that two countries have the same bond price does not tell us that they have the same level of fiscal sustainability.

## 5. CONCLUDING REMARKS

This paper presented a model which can explain how governments with different levels of fiscal sustainability can nonetheless have the same bond price. The model brings into focus the role that asymmetric information has on sovereign bond markets. The two central conclusions are that if asymmetric information plays a role then (1) the ability of the market to discipline the debt accumulation of governments may be hampered, and (2) bond prices may tell us less than we would hope about the fiscal position of an economy.

Throughout the paper, the assumption of linear consumption utility was used to allow for tractable analytical solutions. However, the main result of the paper should extend to adding some risk aversion into the government's preferences. Concavity in utility will *ceteris paribus* reduce a government's utility at all debt levels, independent of the price function it faces. However, how utility depends on which price function the government faces should not be greatly affected. Specifically, for any type of utility function both the *H*-government and the *L*-government will prefer facing the pooled price function than facing the *L*-government's price function at any given debt level. Thus, if off equilibrium path beliefs are such that deviations from equilibrium are seen by investors as coming from an *L*-government then both types will prefer to be in pooling equilibrium at any given bond supply. For there to be pooling equilibra there needs to be a range of bond supplies at which both types simultaneously prefer to be in the pooling equilibrium. In the linear case there is a large range of pooling equilibria for most parameter constellations, and while introducing risk aversion will alter this range it is unlikely that it will eliminate pooling equilibria entirely. [4]

A potential avenue for future research would be to analyse whether asymmetric information has additional effects in a richer model of sovereign debt. One way to extend the model would be to add domestic investors. This would allow the asymmetric information problem to be tied to a domestic liquidity crisis of the type analysed, for example, by Brutti (2011). Furthermore, adding institutional investors such as banks as in Bolton and Jeanne (2011) would allow the effect of asymmetric information on contagion in the banking system to be analysed. This could potentially add an addition spillover from economies with poor fundamentals to economies with good fundamentals in my framework. Contagion has been a much discussed problem during the European sovereign debt crisis so it would be interesting to consider what role asymmetric information may have played.

4. Extending the model to *CRRA* utility with various degrees of relative risk aversion can be done numerically. At high values of relative risk aversion utility can become negative for all debt levels. When parameters are such that utility is positive for some debt levels then pooling equilibria appear to almost always exist.

## APPENDIX A. PROOFS

### A.1. PROOF OF PROPOSITION 1

*Proof.* Given $\gamma \sim U[0,1]$ we have:

$$\delta(b_i, \theta_i, \gamma) = \frac{b_i}{\theta_i(1-\gamma)}.$$

Therefore:

$$\frac{\partial \delta(b_i, \theta_i, \gamma)}{\partial b_i} = \frac{1}{\theta_i(1-\gamma)} > 0$$
$$\frac{\partial \delta(b_i, \theta_i, \gamma)}{\partial \gamma} = \frac{b_i}{\theta_i(1-\gamma)^2} > 0$$
$$\frac{\partial \delta(b_i, \theta_i, \gamma)}{\partial \theta_i} = -\frac{b_i}{\theta_i^2(1-\gamma)} < 0. \qquad \blacksquare$$

### A.2. PROOF OF PROPOSITION 2

*Proof.* Maximising (8) with respect to $b_i$ gives the government's optimal bond supply under perfect information:

$$\begin{aligned} \frac{\partial U_i(b_i, q_i, \gamma)}{\partial b_i} &= 1 - 2\frac{b_i^p}{\theta_i(1-\gamma)} + \pi \frac{b_i^p}{\theta_i(1-\gamma)} - \pi = 0 \\ \Leftrightarrow b_i^p &= \left(\frac{1-\pi}{2-\pi}\right)\theta_i(1-\gamma), \end{aligned} \tag{A.1}$$

which is increasing in $\theta_i$ since:

$$\frac{\partial b_i^p}{\partial \theta_i} = \left(\frac{1-\pi}{2-\pi}\right)(1-\gamma) > 0.$$

Thus, since $\theta_H > \theta_L$, we have $b_H^p > b_L^p$.

Note also that:

$$\frac{\partial b_i^p}{\partial \pi} = \frac{-1}{(2-\pi)^2}\theta_i(1-\gamma) < 0,$$

so that the debt level is a negative function of the default probability.

Plugging (A.1) into $\delta_i^p = b_i^p/[\theta_i(1-\gamma)]$, we get $\delta_i^p = (1-\pi)/(2-\pi)$, which is independent of $\theta$. Thus, $\delta_H^p = \delta_L^p$. This in turn implies $q_H^p = q_L^p$. $\blacksquare$

### A.3. PROOF OF LEMMA 1

The proof starts by deriving the incentive compatibility constraints. Using utility function representations (8) and (9) we can derive four utility functions.

1. The utility function of an $L$-government that faces its own price function:

$$\begin{aligned} U_L(b_L, q_L, \gamma) &= \left[b_L - \frac{b_L^2}{\theta_L(1-\gamma)}\right] + \pi\left[\frac{\theta_L}{2} + \frac{b_L^2}{2\theta_L(1-\gamma)} - b_L\right] \\ &= \pi\frac{\theta_L}{2} + (1-\pi)b_L - \left[\frac{2-\pi}{2\theta_L(1-\gamma)}\right]b_L^2. \end{aligned} \tag{A.2}$$

2. The utility function of an $H$-government that faces its own price function:

$$U_H(b_H, q_H, \gamma) = \pi\frac{\theta_H}{2} + (1-\pi)b_H - \left[\frac{2-\pi}{2\theta_H(1-\gamma)}\right]b_H^2. \tag{A.3}$$

3. The utility function of an $L$-government that faces an $H$-government's price function:

$$\begin{aligned} U_L(b_L, q_H, \gamma) &= \left[b_L - \frac{b_L^2}{\theta_H(1-\gamma)}\right] + \pi\left[\frac{\theta_L}{2} + \frac{b_L^2}{2\theta_L(1-\gamma)} - b_L\right] \\ &= \pi\frac{\theta_L}{2} + (1-\pi)b_L - \left[\frac{2\theta_L - \pi\theta_H}{2\theta_L\theta_H(1-\gamma)}\right]b_L^2. \end{aligned} \tag{A.4}$$

4. The utility function of an $H$-government that faces an $L$-government's price function:

$$U_H(b_H, q_L, \gamma) = \pi\frac{\theta_H}{2} + (1-\pi)b_H - \left[\frac{2\theta_H - \pi\theta_L}{2\theta_L\theta_H(1-\gamma)}\right]b_H^2. \tag{A.5}$$

Let $b_L^*$ be a possible equilibrium bond supply for an $L$-government and $b_H^*$ be a possible equilibrium bond supply for an $H$-government. Then for $b_L^*$ to be incentive compatible for an $L$-government requires that (A.2) at $b_L^*$ be preferred to (A.4) at $b_H^*$. That is:

$$(1-\pi)b_L^* - \left[\frac{2-\pi}{2\theta_L(1-\gamma)}\right]b_L^{*2} \geq (1-\pi)b_H^* - \left[\frac{2\theta_L - \pi\theta_H}{2\theta_L\theta_H(1-\gamma)}\right]b_H^{*2}. \tag{A.6}$$

Likewise, for $b_H^*$ to be incentive compatible for an $H$-government requires that (A.3) at $b_H^*$ be preferred to (A.5) at $b_L^*$. That is:

$$(1-\pi)b_H^* - \left[\frac{2-\pi}{2\theta_H(1-\gamma)}\right]b_H^{*2} \geq (1-\pi)b_L^* - \left[\frac{2\theta_H - \pi\theta_L}{2\theta_L\theta_H(1-\gamma)}\right]b_L^{*2}. \tag{A.7}$$

(A.6) and (A.7) are the necessary incentive compatibility constraints.

Adding both sides of the incentive compatibility constraints and rearranging we get:

$$\begin{aligned} \left[\frac{2\theta_H - \pi\theta_L}{2\theta_L\theta_H(1-\gamma)}\right]b_L^{*2} - \left[\frac{2-\pi}{2\theta_L(1-\gamma)}\right]b_L^{*2} &\geq \left[\frac{2-\pi}{2\theta_H(1-\gamma)}\right]b_H^{*2} - \left[\frac{2\theta_L - \pi\theta_H}{2\theta_L\theta_H(1-\gamma)}\right]b_H^{*2} \\ \Leftrightarrow \left[\frac{\pi(\theta_H - \theta_L)}{2\theta_L\theta_H(1-\gamma)}\right]b_L^{*2} &\geq \left[\frac{\pi(\theta_H - \theta_L)}{2\theta_L\theta_H(1-\gamma)}\right]b_H^{*2} \\ \Leftrightarrow b_L^* &\geq b_H^*. \end{aligned}$$

■

## A.4. PROOF OF PROPOSITION 4

*Proof.* The proof starts by deriving the relevant utility functions. Let:

$$\begin{aligned}\bar{q}(b,\theta,\gamma) =& \lambda[1-\delta_H(b_H,\theta_H,\gamma)]+(1-\lambda)[1-\delta_L(b_L,\theta_L,\gamma)] \\ =& \lambda\left[1-\frac{b_H}{\theta_H(1-\gamma)}\right]+(1-\lambda)\left[1-\frac{b_L}{\theta_L(1-\gamma)}\right],\end{aligned} \tag{A.8}$$

denote the pooled price function. Using the pooled price function (A.8) in utility function representations (8) then allows a further two utility functions to be derived.

1. The utility function of an $L$-government that faces the pooled price function:

$$\begin{aligned}U_L(b,\bar{q},\gamma) =& \left[b-\lambda\frac{b^2}{\theta_H(1-\gamma)}-(1-\lambda)\frac{b^2}{\theta_L(1-\gamma)}\right]+\pi\left[\frac{\theta_L}{2}+\frac{b^2}{2\theta_L(1-\gamma)}-b\right] \\ =& \pi\frac{\theta_L}{2}+(1-\pi)b-\left[\frac{\lambda}{\theta_H(1-\gamma)}+\frac{1-\lambda}{\theta_L(1-\gamma)}-\frac{\pi}{2\theta_L(1-\gamma)}\right]b^2\end{aligned} \tag{A.9}$$

2. The utility function of an $H$-government that faces the pooled price function:

$$U_H(b,\bar{q},\gamma)=\pi\frac{\theta_H}{2}+(1-\pi)b-\left[\frac{\lambda}{\theta_H(1-\gamma)}+\frac{1-\lambda}{\theta_L(1-\gamma)}-\frac{\pi}{2\theta_H(1-\gamma)}\right]b^2 \tag{A.10}$$

In the equilibrium beliefs are such that $\mu(\theta_H|b^*)=\lambda$ and $\mu(\theta_L|b^*)=1-\lambda$. Beliefs off the equilibrium paths are not specified so I assume that $\mu(\theta_H|b)=0$ $\forall\ b\neq b^*$, so that any deviation from equilibrium is seen as coming from an $L$-government. Thus, utility functions (A.9) and (A.10) need to be compared to the utility function (A.2) of an $L$-government that faces its own price function and utility function (A.5) of an $H$-government facing an $L$-government's price function. Utility functions (A.2) and (A.5) are reproduced here for convenience:

$$U_L(b,q_L,\gamma)=\pi\frac{\theta_L}{2}+(1-\pi)b-\left[\frac{2-\pi}{2\theta_L(1-\gamma)}\right]b^2. \tag{A.11}$$

$$U_H(b,q_L,\gamma)=\pi\frac{\theta_H}{2}+(1-\pi)b-\left[\frac{2\theta_H-\pi\theta_L}{2\theta_L\theta_H(1-\gamma)}\right]b^2. \tag{A.12}$$

All of these utility functions are quadratic functions with the following important properties. (1) The intercept coefficients are all positive and differ only between the two types, i.e. $\pi\theta_H/2\neq\pi\theta_L/2$. (2) The linear coefficients are positive and do not differ between types, i.e. they are $(1-\pi)$ for all four utility functions. (3) The quadratic coefficients are negative for all the functions and differ across all four utility functions.

These three properties mean that all four utility functions have one unique maximum and that the relative position of the four maxima is determined by the differences in the quadratic coefficients.

For some debt $b^*$ to form a pooling equilibrium both types must prefer $b^*$ and facing the pooling price function (A.8) than any other possible debt level $b \neq b^*$ and facing the $L$-government's price function.

For an $H$-government the highest possible utility that can be reached facing an $L$-government's price function is the utility at the $b_H \neq b^*$ that maximises (A.12). Likewise, for an $L$-government the highest possible utility that can be reached facing its own price function is the utility at the $b_L \neq b^*$ that maximises (A.11). The following arguments – based on the properties of the utility functions outlined in above – insure that there exists a $b^*$ that provides both types with a higher utility than the maximum attainable utility when facing the $L$-government's price function, and which is thus a possible equilibrium:

- For an $H$-government the quadratic coefficient of the utility function (A.12) is greater than the quadratic coefficient of the utility function (A.10) since:

$$\frac{2\theta_H - \pi\theta_L}{2\theta_L\theta_H(1-\gamma)} > \frac{\lambda}{\theta_H(1-\gamma)} + \frac{1-\lambda}{\theta_L(1-\gamma)} - \frac{\pi}{2\theta_H(1-\gamma)},$$

holds as long as $\theta_H > \theta_L$. Since the quadratic coefficients are negative this means that (1) the function (A.12) is everywhere below the function (A.12), and (2) the function (A.12) reaches its maximum before the function (A.10).
Together this implies that there is a set of bond supplies $b^*$ that an $H$-government can choose when facing the pooled price function that make it strictly better off than the maximum possible utility it could achieve when facing the $L$-government's price function. Part of this set of bond supplies will be the bond supplies between the maxima of the utility functions (A.12) and (A.10).
- For an $L$-government the quadratic coefficient of the utility function (A.11) is greater than the quadratic coefficient of the utility function (A.9) since:

$$\frac{2-\pi}{2\theta_L(1-\gamma)} > \frac{\lambda}{\theta_H(1-\gamma)} + \frac{1-\lambda}{\theta_L(1-\gamma)} - \frac{\pi}{2\theta_L(1-\gamma)},$$

which holds as long as $\theta_H > \theta_L$. Thus, the function (A.11) is everywhere below the function (A.9) and the function (A.11) reaches its maximum before the function (A.9). This implies that there is a set of bond supplies $b^*$ that an $L$-government can choose when facing the pooled price function that makes it strictly better off than the maximum possible utility it could achieve when facing its own price function. Part of this set of bond supplies will be the bond supplies between the maxima of the utility functions (A.11) and (A.9).
- The set of bond supplies $b^*$ for which an $H$-government strictly prefers to face the pooled price function and for which the $L$-government strictly prefers to face the pooled price function will intersect. This occurs as long as the maximum of the function (A.11) occurs between the maxima of the functions

(A.12) and (A.10). This is ensured by the condition:

$$\frac{2\theta_H - \pi\theta_L}{2\theta_L\theta_H(1-\gamma)} > \frac{2-\pi}{2\theta_L(1-\gamma)} > \frac{\lambda}{\theta_H(1-\gamma)} + \frac{1-\lambda}{\theta_L(1-\gamma)} - \frac{\pi}{2\theta_H(1-\gamma)},$$

which holds as long as $\theta_H > \theta_L$ and $\lambda > \pi/2$, with the latter insuring that the second inequality does not reverse.

Finally, $b_H^* = b_L^*$ and $q_H^* = q_L^*$ must hold in any pooling equilibrium. $\delta_L^* = b^*/[\theta_L(1-\gamma)]$ and $\delta_H^* = b^*/[\theta_H(1-\gamma)]$ implies $\delta_H^* < \delta_L^*$ since $\theta_H > \theta_L$.■

# B. ADDITIONAL DETAILS

## B.1. REARRANGING THE GOVERNMENT'S INTERTEMPORAL UTILITY FUNCTION

The pricing function for a government of type $i$ is

$$\begin{aligned} q_i(b_i, \theta_i, \gamma) &= 1 - \delta_i(b_i, \theta_i, \gamma) \\ &= 1 - \frac{b_i}{\theta_i(1-\gamma)} \end{aligned}$$

Plugging this into (7) yields:

$$U_i(b_i, q_i, \gamma) = \left[1 - \frac{b_i}{\theta_i(1-\gamma)}\right] b_i + \pi\left[\int_{\frac{b_i}{\theta_i(1-\gamma)}}^{1} (\theta_i y - b_i) f(y) dy + \int_0^{\frac{b_i}{\theta_i(1-\gamma)}} \gamma\theta_i y f(y) dy\right]$$

Let $U_{2i}$ be the period two utility. Using $y \sim U[0,1]$ to solve the integral and rearranging gives:

$$\begin{aligned} U_{2i} &= \pi\left[\int_{\frac{b_i}{\theta_i(1-\gamma)}}^{1} (\theta_i y - b_i) dy + \int_0^{\frac{b_i}{\theta_i(1-\gamma)}} \gamma\theta_i y dy\right] \\ &= \pi\left\{\theta_i\left[\frac{y^2}{2}\right]_{\frac{b_i}{\theta_i(1-\gamma)}}^{1} - b_i[1]_{\frac{b_i}{\theta_i(1-\gamma)}}^{1} + \gamma\theta_i\left[\frac{y^2}{2}\right]_0^{\frac{b_i}{\theta_i(1-\gamma)}}\right\} \\ &= \pi\left\{\frac{\theta_i}{2} - \frac{\theta_i}{2}\left[\frac{b_i}{\theta_i(1-\gamma)}\right]^2 - b_i + \frac{b_i^2}{\theta_i(1-\gamma)} + \frac{\gamma\theta_i}{2}\left[\frac{b_i}{\theta_i(1-\gamma)}\right]^2\right\} \\ &= \pi\left\{\frac{\theta_i}{2} + \frac{b_i^2}{2\theta_i(1-\gamma)} - b_i\right\}. \end{aligned}$$

Thus we have:

$$\begin{aligned} U_i(b_i, q_i, \gamma) &= \left[1 - \frac{b_i}{\theta_i(1-\gamma)}\right] b_i + \pi\left[\frac{\theta_i}{2} + \frac{b_i^2}{2\theta_i(1-\gamma)} - b_i\right] \\ &= \pi\frac{\theta_i}{2} + (1-\pi) b_i - \left[\frac{2-\pi}{2\theta_i(1-\gamma)}\right] b_i^2 \end{aligned} \qquad (B.1)$$

If a government of type $i$ faces the price function of a government of type $j$ we have:

$$U_i(b_i, q_j, \gamma) = \left[1 - \frac{b_i}{\theta_j(1-\gamma)}\right] b_i + \pi\left[\frac{\theta_i}{2} + \frac{b_i^2}{2\theta_i(1-\gamma)} - b_i\right] \\ = \pi\frac{\theta_i}{2} + (1-\pi)b_i - \left[\frac{2\theta_i - \pi\theta_j}{2\theta_i\theta_j(1-\gamma)}\right] b_i^2 \tag{B.2}$$

## B.2. CONDITION FOR THE EXISTENCE OF A SEPARATING EQUILIBRIUM

In a separating equilibrium investors' posterior beliefs are $\mu(\theta_i|b_i^*) = 1$ after observing $b_i^*$. To show that such an equilibrium exists we need to check if there is an equilibrium for which $b_H^*$ and $b_L^*$ are optimal given the responses of investors and given the off equilibrium path beliefs given by $\mu(\theta_H|b) = 0 \ \forall \ b \neq b_i^*$.

For an $L$-government, the best response to being correctly perceived as such is to solve:

$$\max_b \pi\frac{\theta_L}{2} + (1-\pi)b - \left[\frac{2-\pi}{2\theta_L(1-\gamma)}\right] b^2,$$

since there is nothing to be gained from not choosing debt optimally given it faces its own pricing function. This leads to:

$$b_L^* = \left(\frac{1-\pi}{2-\pi}\right)\theta_L(1-\gamma). \tag{B.3}$$

By plugging (B.3) into the utility function of the $L$-government when it faces its own price function we can define the following incentive constraint that ensures an $L$-government prefers to choose (B.3) than deviating to the $H$-government's debt level $b_H^*$:

$$\frac{(1-\pi)^2}{2(2-\pi)}\theta_L(1-\gamma) \geq (1-\pi)b_H^* - \left[\frac{2\theta_L - \pi\theta_H}{2\theta_L\theta_H(1-\gamma)}\right] b_H^{*2}. \tag{B.4}$$

For an $H$-government not to deviate from $b_H^*$ it must be better off facing its own price function at that debt level than choosing another debt level but facing an $L$-government's price function. Again, the best an $H$-government can do if it faces an $L$-government's price function is to optimise. Thus, if an $H$-government deviates from $b_H^*$ knowing it will face the $L$-government's price function if it does so then it may as well optimise and solve:

$$\max_b \pi\frac{\theta_H}{2} + (1-\pi)b - \left[\frac{2\theta_H - \pi\theta_l}{2\theta_L\theta_H(1-\gamma)}\right] b^2,$$

Letting $b^*_{HL}$ be the $H$-government's optimal debt level if it faces an $L$-government's price function, we have:

$$b^*_{HL} = \left(\frac{1-\pi}{2\theta_H - \pi\theta_L}\right)\theta_L\theta_H(1-\gamma).$$

Plugging (B.5) into the utility function for an $H$-government that faces an $L$-government's price function, then allows the $H$-government's incentive compatibility constraint to be defined as:

$$(1-\pi)b^*_H - \left[\frac{2-\pi}{2\theta_H(1-\gamma)}\right]b^{*2}_H \geq \frac{(1-\pi)^2}{2(2\theta_H - \pi\theta_L)}\theta_L\theta_H(1-\gamma), \qquad (B.5)$$

which ensures that an $H$-government prefers to choose $b^*_H$ and face its own price function than deviating and choosing $b^*_{HL}$ and face the $L$-governments price function.

For there to exist a separating equilibrium the highest value of $b^*_H$ that still satisfies (B.4) must be higher than the lowest value of $b^*_H$ that still satisfies (B.5). If this is the case then there exists a set of debt levels $b^*_H$ for which the $L$-government prefers to choose $b^*_L$ and face its price function and the $H$-government prefers to choose $b^*_H$ and face its price function rather than $b^*_{HL}$ and face the $L$-government's price function.

For such a set to exist the lower of the two roots of a quadratic equation, given when incentive constraint (B.4) holds with equality, must be greater than the lower of the two roots of a quadratic equation, given when incentive constraint (B.5) holds with equality. This leads to the following condition:

$$\frac{1-\pi}{2\left(\frac{2\theta_L-\pi\theta_H}{2\theta_L\theta_H(1-\gamma)}\right)} - \frac{\sqrt{(1-\pi)^2 - 4\left(\frac{2\theta_L-\pi\theta_H}{2\theta_L\theta_H(1-\gamma)}\right)\left(\frac{(1-\pi)^2}{2(2-\pi)}\theta_L(1-\gamma)\right)}}{2\left(\frac{2\theta_L-\pi\theta_H}{2\theta_L\theta_H(1-\gamma)}\right)}$$
$$\geq \frac{1-\pi}{2\left(\frac{2-\pi}{2\theta_H(1-\gamma)}\right)} - \frac{\sqrt{(1-\pi)^2 - 4\left(\frac{2-\pi}{2\theta_H(1-\gamma)}\right)\left(\frac{(1-\pi)^2}{2(2\theta_H-\pi\theta_L)}\theta_L\theta_H(1-\gamma)\right)}}{2\left(\frac{2-\pi}{2\theta_H(1-\gamma)}\right)}$$
$$\Leftrightarrow \frac{1-\sqrt{\frac{2(\theta_H-\theta_L)}{2\theta_H-\pi\theta_H}}}{2\theta_L - \pi\theta_H} \geq \frac{1-\sqrt{\frac{2(\theta_H-\theta_L)}{2\theta_H-\pi\theta_L}}}{2\theta_L - \pi\theta_L}$$

The properties of this condition are difficult to prove analytically but it can shown computationally that when $\theta_H \in (\theta_L, 2\theta_L)$ and $\pi \in (0,1)$ the condition holds.

## ACKNOWLEDGEMENTS

I am grateful to Sönke Hoffmann, Maximilian Mihm, Hendrik Ritter, Joachim Weimann and an anonymous referee for helpful comments and suggestions.

Address for correspondence: Benedikt Mihm, Faculty of Economics and Management, Economic Policy, University of Magdeburg, Postbox 4120, D-39016 Magdeburg, Germany. Tel.: +49 391 6712158; fax: +49 391 67 12971; e-mail: mihm@ovgu.de

## REFERENCES

Alesina, A. and G. Tabellini (1990), 'A Positive Theory of Fiscal Deficits and Government Debt', *Review of Economic Studies* 57, 403–414.

Alfaro, L. and F. Kanczuk (2005), 'Sovereign Debt as a Contingent Claim: A Quantitative Approach', *Journal of International Economics* 65, 297–314.

Arghyrou, M. and A. Kontonikas (2011), 'Decomposing Euro-area Sovereign Spreads: Credit and Liquidity Risks'. European Economy - Economic Papers 436.

Beetsma, R. and K. Mavromatis (2014), 'An Analysis of Eurobonds', *Journal of International Money and Finance* 45, 91–111.

Bolton, P. and O. Jeanne (2011), 'Sovereign Defaults and Liquidity Crises', *IMF Economic Review* 59, 162–194.

Brutti, F. (2011), 'Sovereign Defaults and Liquidity Crises', *Journal of International Economics* 84, 65–72.

Corsetti, G. and L. Dedola (2011), *Fiscal Crises, Confidence and Default: A Bare-bones Model with Lessons for the Euro Area*, Unpublished manuscript, Cambridge University.

Cuadra, G. and H. Sapriza (2008), 'Sovereign Default, Interest Rates and Political Uncertainty in Emerging Markets', *Journal of International Economics* 76, 78–88.

Eaton, J. (1996). 'Sovereign Debt, Reputation, and Credit Terms', *International Journal of Finance and Economics* 1, 25–35.

Eaton, J. and Gersovitz, M. (1981), 'Debt with Potential Repudiation: Theoretical and Empirical Analysis', *Review of Economic Studies* 48, 289–309.

Hatchondo, J., L. Martinez and H. Sapriza (2007), 'The Economics of Sovereign Defaults', *Economic Quarterly* 93, 163–187.

Riley, J. (2001). 'Silver Signals: Twenty-five Years of Screening and Signaling', *Journal of Economic Literature* 41, 432–478.

Sandleris, G. (2008). 'Sovereign Defaults: Information, Investment and Credit', *Journal of International Economics* 76, 267–275.

Spence, M. (1973). 'Job Market Signaling', *Quarterly Journal of Economics* 87, 355–374.

# Biased government borrowing and yardstick competition in a sovereign debt market

Benedikt Mihm *
University of Magdeburg

July 18, 2016

**Abstract**

This paper develops a sovereign debt model in which governments are privately informed about their likelihood of default but can themselves have a biased perception of this likelihood. I show that in this setup government borrowing acts as a signal that is only partially informative about fundamental default probabilities, and bond prices do not necessarily reflect true credit risk. I also show that in a two country version of the model correlations between the two countries lead to a form of yardstick competition, and the borrowing decision of one government affects the bond price received by the other. Whether the information spillover increases or decreases the distortion created by the bias depends on the extent to which borrowing signals reinforce each other.

**Keywords:** Sovereign debt; behavioral bias; signaling; yardstick competition.

**JEL-Codes:** E60, F34

*Postal address: Faculty of Economics and Management, Chair in Economic Policy, University of Magdeburg, Postbox 4120, D-39016 Magdeburg, Germany. E-mail address: mihm@ovgu.de.

# 1 Introduction

The market for sovereign debt occupies a central role in the global financial system. Sovereign bonds are usually viewed as relatively safe and highly liquid, and one would generally expect well-functioning markets for such assets. However, there are also a number of peculiar features of the market for sovereign debt. First, while there are many buyers and sellers in the market, there is only a small number of issuers of sovereign debt, so that market power and strategic considerations may affect the interaction between issuers and other market participants. Second, sovereign debt is primarily sold as non collateralize bonds and legal enforcement of the debt contract is tough and uncertain. Third, the enforcement problem leads to credit risk but as default is an extremely rare event for most countries, estimating default likelihoods is difficult both for the issuers of debt and other market participants.

This paper develops a model to explore some of the effects of these features of the sovereign debt market on government borrowing and bond prices. In the model, a government borrows from risk neutral investors with a likelihood that it will not repay. The government can be either safe or risky depending on the likelihood of repayment with the government's type being private information. However, while the government knows more about its default probability than investors, the government can be biased in the sense that it miss-perceives its own likelihood of default: a risky government can perceive itself to be safe and vice versa. This biased perception captures that the rarity of defaults means that even better informed governments may make systematic mistakes in estimating there likelihood.

In the game between the issuer of sovereign debt and other market participants, I first show that there is a separating equilibrium in which a risky government borrows more than a safe government. The higher likelihood of default means that the risky government likes to borrow more, and a safe government can signal its type by choosing a low enough debt level. This separation can occur even though the safe government receives a higher price for its bonds and the risky government therefore has an incentive to be seen as safe. Importantly, while the equilibrium bond prices received by the two types differ, they only partially reflect the difference in default probabilities as investors take into account that the bond supply can come from a biased government. Moreover, while the government selects to receive its (perceived) correct bond price, the fact that it can be biased means that the price it receives can wrongly reflect its true type.

Empirical studies assessing the extent to which measures of government borrowing affect bond prices show mixed results but tend to show that higher borrowing leads to lower bond prices (Bernoth et al., 2012; D'Agostino and Ehrmann, 2014). A key channel leading to this result is likely the increased credit risk that comes directly from higher borrowing with more debt increasing the likelihood a government will not repay. In this paper, credit risk leads to pricing behavior that goes in the observed direction through an alternative channel that results from the signaling value of debt. Safe governments receive higher prices as they are able to signal their greater intent to repay though lower debt levels. The presence of the bias, however, means that although the signal can be successful, bond prices can incorrectly reflect true default probabilities.

It is interesting to consider what additional information investors might use to evaluate the credit worthiness of a government. Gande and Parsley (2005) suggests that one important source of information may be cross country comparisons. They show that news spillovers can have significant effects in sovereign bond markets, with credit rating changes in one country affecting the bond prices of other similar countries. In a two country version of my model a positive correlation in default probabilities can also generate information spillovers, as investors can infer information about the credit worthiness of one government from the bond supply of the other.[1] This feature introduces a form of *yardstick competition* to the sovereign debt context, as previously studied in the literature on industrial organization (Shleifer, 1985) and political economy (Besley and Case, 1995).

Intuitively, one might expect the information spillover to reduce the distortion created by the the bias, as investors have additional information to update beliefs. The second result of the paper is to show that this intuition is only partially correct. Specifically, when the governments of both countries signal the same type (either safe or risky) bond prices with yardstick competition are closer to those that would obtain under full information. If, on the other hand, the two governments signal that they are different types the bond prices with yardstick competition are further from those that would obtain under full information. The effect of yardstick competition on bond price therefore depends on whether signals reinforce each other or not.

The main results of the paper can be related to the puzzling behavior of bond markets in the Eurozone prior to the European sovereign debt crisis, where bond price differentials between euro countries were very small despite large differences in the state of public finances. One possible explanation for this pricing behavior is that the no bailout clause in the Eurozone was not credible, and investors therefore viewed sovereign debts by different Eurozone countries as more homogeneous than they actually were. While this argument is likely to be part of the explanation, it has some well-discussed limitations (see, .e.g., Honkapohja, 2013).

My paper suggests an addition consideration for the observed behavior based on information spillovers. When the government of one country signals that it considers itself safe with low borrowing and the government of another that it considers itself risky with high borrowing, then yardstick competition acts to narrow the bond price spread between them. Essentially, the safe government's borrowing behavior makes investors believe that the risky government might have overestimated its true default risk, and the risky government's behavior makes investors believe that the safe government might have underestimated its true default risk. As a result, when public finances differ the bond price spread between the countries is smaller with yardstick competition than it otherwise would be.

[1] The spillovers Gande and Parsley (2005) find do not result from debt but from credit ratings. Landon and Smith (2000) find evidence of spillovers from debt among Canadian provinces when creditworthiness is measured by credit ratings. Landon and Smith (2007) find no evidence of spillovers among Canadian providence from debt to bond prices. However, the amount of empirical work studying debt spillovers is limited, so the the role of debt spillovers remains an open question.

## 2 Related literature

The paper makes two contributions to the literature on sovereign borrowing. First, the model integrates the idea of biased signals into the sovereign debt context. Santos-Pinto (2012) studies the labor market signaling model of Spence (1973) when workers have biased beliefs about their cost of education and shows that the bias leads to wage compression. In addition to the different environment that I study, my work differs from Santos-Pinto (2012) as I extend my setup to a multiagent setting with correlated types in order to study information spillovers that play no role in his work. The wage compression in Santos-Pinto (2012) is similar to the risk mispricing that arises in my model, but the extension to multiple dependent agents allows me to introduce yardstick competition, which is the main focus of my work and which seems particularly relevant in the sovereign debt context as indicated by the work of Gande and Parsley (2005).

There are several recent sovereign debt papers that use unbiased signaling. Sandleris (2008) studies how the debt repayment decision can serve as a signal of fundamentals that are private information. Perez (2015) analyzes how debt maturity decisions can act as a signal for the likelihood of debt repayment. My paper is very different from Perez (2015) but shares with his that I also consider exogenous default probabilities and optimal debt is derived via concave preferences. Mihm (2015) looks at a framework with endogenous default and linear preferences, and shows that there are pooling equilibria in which differences in default probabilities are not reflected in bond prices. In Mihm (2015) prices do not accurately reflect differences in default probabilities because debt is completely uninformative about type, whereas here I look at a situation in which debt is partially informative but can still act as a credible signal.

The second contribution of the paper is to introduce a form of yardstick competition to a sovereign debt environment. Shleifer (1985) considers a framework in which regulators are uncertain about a monopoly's costs and use other monopolies with correlated costs as benchmarks to optimize regulation. More closely related, Besley and Case (1995) look at a political economy setup in which voters can learn addition information about the type of their government by observing the behavior of governments in other regions. Besley and Smart (2007) analyze a more theoretically formal setting and show that voter welfare can actually be lower when they use yardstick competition. Bordignon et al. (2004) use a very general type of correlation structure and show that the effect of yardstick competition on voter and government behavior depends, among other things, on the amount of correlation. I use the general form of correlation in Bordignon et al. (2004) to consider the effect of yardstick competition on bond prices. Like election probabilities in the voter's problem, bond prices in my model depend on beliefs. Observing borrowing behavior in one country provides information about the likelihood that the signal from the government in the other country is biased and there are spillovers from the borrowing in one country to the bond price of the other.

The remainder of the paper is organized as follows. Section 3 presents the basic one country version of the model. Section 4 analyzes the effect of biased borrowing on the equilibrium under asymmetric information. Section 5 then develops a two country version of the model to consider the effect of yardstick competition. Section 6 concludes. All proofs are in Appendix A

# 3 Model

## 3.1 The government

There is a small open economy that lasts for two periods. The consumption maximizing government can borrow resources by selling bonds to investors on an international sovereign bond market. A lack of commitment on the part of the government means that there is a probability $\theta \in [0,1]$ that it will default on its debts. This default probability depends on the government's type $i \in \{R, S\}$, which is private information. The government can be either risky ($i = R$) or safe ($i = S$), with $\theta_R > \theta_S = 0$ so that a risky government has positive probability of default and the safe government does not.

A government's type is determined by nature with a share $\lambda \in (0,1)$ of governments becoming a safe government ($S$-type) and a share $1-\lambda$ becoming a risky government ($R$-type). While the governments type is private information, the government can have biased beliefs about its own default probability: a share $\beta \in [0, 1-\lambda]$ of $R$-types believe themselves to be $S$-types and share $\alpha \in [0, \lambda]$ of $S$-types believe themselves to be $R$-types. The share of biased and unbiased governments is also determined by nature. Similar to the labor market model of Santos-Pinto (2012), a biased government neither knows that it is biased nor takes the fact that it might be biased into account.

The government receives an exogenous endowment $y$ in the second period so it will need to borrow to consume in the first period. It borrows by issuing one period real zero-coupon bonds $b_i$ at price $q(b_i)$. If the government defaults it receives an income $y_d > 0$. The expected utility of the government depends both on consumption in each period and on the state: default or not. I abstract from discounting and use log utility so that a government of type $i$'s expected utility is given by:

$$U_i[b_i, q(b_i)] \quad = \quad \ln[q(b_i)b_i] + (1-\theta_i)\ln[y-b_i] + \theta_i \ln[y_d]. \tag{1}$$

## 3.2 The bond market

The bond price $q(b_i)$ is determined by the demand decisions of risk neutral investors on a competitive international sovereign bond market. Investors can invest any amount in the government's bonds or in some other asset with an exogenously given risk-free interest rate $r = 0$. Since the government may default, the investors need to be compensated for any expected losses so that the bond price contains a default loss premium. Moreover, the private information and the potentially biased beliefs of a government about its own likelihood of default also influence the price of bonds. Investors know $\lambda$, $\alpha$, and $\beta$ but not the government's type or the government's beliefs about its own type. Therefore, when observing the bond supply $b_i$ of a government, the investors form beliefs about the government being safe $\mu(\theta_S|b_i)$ or risky $\mu(\theta_R|b_i) = 1 - \mu(\theta_S|b_i)$. A no arbitrage condition then pins down the bond price as:

$$q(b_i) = \mu(\theta_S|b_i)(1-\theta_S) + [1-\mu(\theta_S|b_i)](1-\theta_R),$$

where $\theta_S$ is zero but is left in the expression for clearer exposition.

# 4 Biased borrowing

In a separating equilibrium, the bond supply $b_S^* \neq b_R^*$ is only partially informative about the government's true type, as a government may chose $b_S$ although its actual default probability is $\theta_R$, and vice versa.[2] This uncertainty is reflected in the bond prices:

$$q(b_S) = \mu(\theta_S|b_S)(1-\theta_S) + [1-\mu(\theta_S|b_S)](1-\theta_R) \quad (2)$$
$$q(b_R) = \mu(\theta_S|b_R)(1-\theta_S) + [1-\mu(\theta_S|b_R)](1-\theta_R), \quad (3)$$

with $\mu(\theta_S|b_S) = \frac{\lambda-\alpha}{\lambda-\alpha+\beta}$, $\mu(\theta_S|b_R) = \frac{\alpha}{1-\lambda+\alpha-\beta}$, so the prices are a weighted average of the two default probabilities. In the absence of the bias, the separating bond price coincides with the full information bond price and both fully reflect the true default probabilities, i.e., $q(b_i) = 1-\theta_i$. The full information price, therefore, serve as a benchmark to evaluate the effect of the bias. The following proposition links the prices (2) and (3) to the full information prices:

**Proposition 1.** *With asymmetric information and biased types:*

*(i) A government that chooses bond supply $b_S$ receives a price $q(b_S)$ that is lower than the price it would receive under full information.*

*(ii) A government that chooses bond supply $b_R$ receives a price $q(b_R)$ that is higher than the price it would receive under full information.*

*Moreover, the higher is the likelihood of a government being biased, the further prices are from the full information prices.*

Proposition 1 has a straightforward intuition. Under full information bond prices fully reflect the true default probabilities. An $S$-type then receives a price $q(b_S) = 1-\theta_S$ and an $R$-type the price $q(b_R) = 1-\theta_R$, which are the upper and lower bound on prices respectively. Under asymmetric information and biased types a government that chooses $b_S$ receives the price (2) that is lower than under full information as it places a positive weight on $1-\theta_R$. A government who chooses $b_R$ receives the price (3) that is higher than under full information as it put a positive weight on $1-\theta_S$. As the bias disappears, the prices will converge to their upper and lower bounds.

Bond prices in a separating equilibrium with biased types have two key features. First, bond prices do not fully reflect the difference in default probabilities between types, and the bond price spread is narrower than it should be. The higher the likelihood of bias, the narrower the spread. Second, an $R$-type can receive an $S$-type's bond price and vice versa, so that bond prices do not always accurately reflect the true type.

To make the separating equilibrium interesting, it is necessary to limit the amount of bias, so that $q(b_S) \geq q(b_R)$ and it is the $R$-type and biased $S$-type that has an incentive to mimic. This is insured by the condition:

[2] The model will also have pooling equilibria in which the bond supply is completely uninformative and investors use the priors $\mu(\theta_S|b) = \lambda$. This paper concentrates exclusively on the separating equilibria.

$$\alpha + \beta \leq \lambda(1-\lambda), \tag{4}$$

which restricts the amount of bias and means that a bond supply $b_i$ is always more likely to have come from a government of type $i$ than from a government that mistakenly thinks it is type $i$. As the price the $S$-type and biased $R$-type then receives is greater than the price the $R$-type and biased $S$-type receives, the $R$-type and biased $S$-type may have an incentive to mimic to receive a higher price. However, the fact that the $R$-type and biased $S$-type have, or believe they have, a higher likelihood of default means they prefer a higher debt level at a given bond price. This increased incentive to borrow creates a *single-crossing* property, that states that the $R$-type and biased $S$-type will never choose a strictly lower debt level than the $S$-type and biased $R$-type. The following lemma formalizes the property:

**Lemma 1.** *Let $b_S^*$ be a possible equilibrium bond supply for the $S$-type and biased $R$-type, and $b_R^*$ be a possible equilibrium bond supply for the $R$-type and biased $S$-type. Then it must be that $b_S^* \leq b_R^*$.*

There are two qualitatively different types of separating equilibria in which $b_S^* \leq b_R^*$ and condition (4) holds:

**Definition 1.**

- Type I separating equilibrium*: The $R$-type and biased $S$-type maximize (1) subject to (3) and the $S$-type and biased $R$-type maximize (1) subject to (2).*
- Type II separating equilibrium*: The $R$-type and biased $S$-type maximize (1) subject to (3) and the $S$-type and biased $R$-type choose a debt level lower that that which maximizes (1) subject to (2).*

In the type I separating equilibrium the debt level chosen by the $S$-type and biased $R$-type when they maximize (1) subject to (2) is sufficiently low and the $R$-type and biased $S$-type would prefer to maximize (1) subject to (3) rather than mimic.[3] In the other type of separating equilibrium the debt level chosen by the $S$-type and biased $R$-type when they maximize (1) subject to (2) is not sufficiently low so that the $R$-type and biased $S$-type would prefer to mimic to receive the price (2). In this equilibrium the $R$-type and biased $S$-type must choose a lower debt level in order to separate.

The following proposition establishes the condition for the type of separating equilibrium:

**Proposition 2.** *A type I separating equilibrium occurs if:*

$$\frac{q(b_R)}{q(b_S)} \geq \frac{1}{(1-\theta_R)^{1-\theta_R}} \left(\frac{2-\theta_R}{2}\right)^{(2-\theta_R)}. \tag{5}$$

*Otherwise, separating equilibria will be of type II.*

[3] In the current setting the type I equilibrium coincide with the first best full information equilibrium. In general the two equilibria will not coincide so this paper will not evaluate borrowing relative to this benchmark.

Condition (5) is derived from the incentive compatibility constraint of the $R$-type and biased $S$-type. The condition reflects a simple intuition: when the bond price difference is small, the $R$-type and biased $S$-type have less incentive to mimic as they receive a relatively high price when they reveal themselves truthfully. Furthermore, the higher is the default probability $\theta_R$, the bigger the difference in the amount the $R$-type and biased $S$-type want to borrow relative to the $S$-type and biased $R$-type, and the less likely it is that $R$-types and biased $S$-types will want to mimic. Importantly, regardless of which of the two types of equilibria occurs, the $S$-types and biased $R$-types will have to choose a debt level low enough to dissuade the $R$-types and biased $S$-types from mimicking.

# 5 Yardstick competition

In this section, I extend the setup to include a second identical country. The two countries will be referred to as *home* and *foreign*, and the problem is considered from the perspective of the home country's government. The two countries are independent, except for the fact that the likelihood of the countries being either safe or risky is positively correlated, reflecting that a country's default probability can depend on fundamentals - such as productivity and financial conditions- that can be correlated with those in other similar countries.

Let the home country continue to be indexed by $i \in \{S, R\}$ and the foreign country be indexed by $j \in \{S, R\}$, and denoted with a prime. The joint probability of the two countries being safe or risky is then written as $p(\theta_i, \theta'_j)$ and the correlation is of the form $p(\theta_S, \theta'_S) = \sigma\lambda$, $p(\theta_S, \theta'_R) = p(\theta_R, \theta'_S) = (1-\sigma)\lambda$ and $p(\theta_R, \theta'_R) = 1 - \lambda(2-\sigma)$. The parameter $\sigma \in [\lambda, 1]$ then determines the degree of positive correlation.

Given that investors are risk neutral and there is unlimited funding, the countries are not in competition for funds. Rather, information spillovers occur because investors can infer important information about the type of one country's government by observing the behavior of the other country's government. Specifically, the beliefs of investors about the default probability of the home country $\mu(\theta_S|b_i, b_j)$ now also depend on the behavior in the foreign country and, as a result, so do the home bond prices. That is:

$$q(b_i, b'_j) = \mu(\theta_S|b_i, b'_j)(1-\theta_S) + [1 - \mu(\theta_S|b_i, b'_j)](1-\theta_R), \tag{6}$$

and there are now four possible bond prices $q(b_R, b'_S)$, $q(b_R, b'_R)$, $q(b_S, b'_S)$ and $q(b_S, b'_R)$.

Using Bayes' rule, and the fact that $b_i$ and $b'_j$ are independent conditioned on $\theta_S$, beliefs can be written as:

$$\mu(\theta_S|b_i, b'_j) = \frac{p(\theta_S)p(b_i|\theta_S)p(b'_j|\theta_S)}{p(\theta_S)p(b_i|\theta_S)p(b'_j|\theta_S) + p(\theta_R)p(b_i|\theta_R)p(b'_j|\theta_R)}.$$

When $p(b'_j|\theta_S) > p(b'_j|\theta_R)$ yardstick competition leads to more weight being put on $1 - \theta_S$, and vice versa.

The following proposition now links bond prices with yardstick competition to the benchmark full information prices:

**Proposition 3.** *Let $b_i$ be the home bond supply and $b'_j$ be the the foreign bond supply. Then under condition (4) and for a given amount of bias:*

*(i) When $i = j$, bond prices with yardstick competition are closer to the prices that would obtain under full information.*

*(ii) When $i \neq j$, bond prices with yardstick competition are further from the prices that would obtain under full information.*

*Moreover, the higher is the degree of correlation, the closer prices are to full information prices when $i = j$ and the further they are from full information prices when $i \neq j$.*

The intuition behind Proposition 3 is as follows. Condition (4) requires that the bond supplies $b_i$ and $b'_j$ are more likely to have been chosen because default probabilities are $\theta_i$ and $\theta'_j$, rather than because the governments are biased. When the governments in both countries signal the same type ($i = j$), yardstick competition then acts to reinforce the signals. The foreign bond supply $b'_j$ reassures investors that the home default probability is $\theta_i$, and the home bond supply $b_i$ reassures investors that the foreign default probability is $\theta'_j$. As a result, bond prices in both countries increase when $b_S$ and $b'_S$ are chosen and decrease when $b_R$ and $b'_R$ are chosen. It follows from Proposition 1 that prices with yardstick competition are then closer to the prices that would obtain under full information.

However, when the governments in the two countries send different signals from one another ($i \neq j$), investors become less assured the default probabilities are $\theta_i$ and $\theta'_j$ and more convinced that the governments are biased. As a result, when $b_S$ and $b'_R$ are chosen the home bond price is lower and the foreign bond price higher with yardstick competition. When $b_R$ and $b'_S$ are chosen the home bond price is higher and the foreign bond price lower with yardstick competition. It follows from Proposition 1 that this moves prices further from the full information prices. The higher the degree of correlation, the stronger the effect in either direction.

The consequences of yardstick competition are twofold. First, when signals reinforce each other, the distortion from the bias is reduced as prices move closer to the benchmark full information prices. When signals do not reinforce each other, the distortion is increased. Second, when the governments in the two countries send different signals, yardstick competition acts to narrow the bond price spread between the countries by increasing the price in one country and reducing the price in the other.

The following proposition establishes the condition for the type of separating equilibrium:

**Proposition 4.** *With yardstick competition a type I separating equilibrium occurs if:*

$$\frac{q(b_R, b'_S)^{\phi} q(b_R, b'_R)^{1-\phi}}{q(b_S, b'_S)^{\phi} q(b_S, b'_R)^{1-\phi}} \quad \geq \quad \frac{1}{(1-\theta_R)^{1-\theta_R}} \left( \frac{2-\theta_R}{2} \right)^{(2-\theta_R)}, \tag{7}$$

*with $\phi = p(b'_S|\theta_R)$, $1-\phi = p(b'_R|\theta_R)$. Otherwise, separating equilibria will be of type II.*

The condition (7) is again derived from the incentive compatibility constraint of the $R$-type and biased $S$-type. The difference to condition (5) is that $R$-type and biased

$S$-type can now receive two different prices if they choose $b_R$ and two when they choose $b_S$, depending on what the government in the foreign country does. After observing $\theta_R$ they form beliefs $\phi = p(b'_S|\theta_R)$ and $1 - \phi = p(b'_R|\theta_R)$ about the bond supply of the government in the foreign country.

# 6 Conclusion

This paper developed a simple model to consider the role of information frictions and behavioral bias in a sovereign bond market. The first result of the paper is to show that separating equilibrium can exist in which a government that is risky (or believes itself to be risky) prefers to borrow more and a safe government (or one believing itself to be safe) can separate to get a higher price for its bonds. When signals can be biased the resulting bond prices differ between safe and risky governments but the prices do not fully reflect the difference in credit risk. Moreover, a risky government can receive a safe governments bond price, and vice versa.

In a two country version of the model, a positive correlation between the two countries default probabilities leads to spillovers from the debt choice of one to the bond prices of the other. The effect this form of yardstick competition has on the distortion due to the bias depends on whether the signals from the two governments reinforce each other or not. When both government send the same signal then prices with yardstick are closer to those that obtain under full information. When the signals differ, prices with yardstick competition are further from the full information benchmark.

The model makes two predictions about how biased signaling can affect the interaction between bond prices and borrowing. First, bond prices react to debt levels in the expected fashion, with higher borrowing leading to lower bond prices. However, when the relationship arises due to biased signals, bond prices may not accurately reflect true credit risk. The possibility of biased signals, therefore, suggests caution needs to be taken in interpreting the pricing behavior as indicating that credit risk is accurately priced.

Second, yardstick competition can lead to higher (lower) borrowing in one country inducing lower (higher) bond prices in other similar countries. This is in line with what Gande and Parsley (2005) refer to as the common information effect - for which they find strong evidence- where a ratings event in one country signals a common trend and affects interest rates in other countries. In my framework, the signal comes directly from the borrowing decision made by the government in one country, which provides information about the default risk in other similar countries.

# A Proofs

## A.1 Proof of Proposition 1

*Proof.* Under full information, the bond supply $b_S$ results in beliefs $\mu(\theta_S|b_S) = 1$ and bond price $q(b_S) = 1 - \theta_S$. Under asymmetric information and $\beta \in (0, 1 - \lambda]$ and $\alpha \in (0, \lambda]$, beliefs are $\mu(\theta_S|b_S) \in [0, 1)$, and the bond price is $q(b_S) = \mu(\theta_S|b_S)(1 - \theta_S) + [1 - \mu(\theta_S|b_S)](1 - \theta_R) < 1 - \theta_S$.

With bond supply $b_R$ full information results in beliefs $\mu(\theta_S|b_R) = 0$ and price $q(b_R) = 1-\theta_R$. Under asymmetric information and biased beliefs, $\mu(\theta_S|b_R) \in (0,1]$ and the price is $q(b_R) = \mu(\theta_S|b_S)(1-\theta_S) + [1-\mu(\theta_S|b_S)](1-\theta_R) > 1-\theta_R$.

Finally, $\mu(\theta_S|b_S) = \frac{\lambda-\alpha}{\lambda-\alpha+\beta}$ is decreasing in $\alpha$ and $\beta$, whereas $\mu(\theta_S|b_R) = \frac{\alpha}{1-\lambda+\alpha-\beta}$ is increasing in $\alpha$ and $\beta$. Therefore, the higher $\alpha$ and $\beta$ the further are the prices from full information prices. □

## A.2 Proof of Lemma 1

*Proof.* We need:

$$U_S[b_S^*, q(b_S)] \geq U_S[b_R^*, q(b_R)] \tag{8}$$
$$U_R[b_R^*, q(b_R)] \geq U_R[b_S^*, q(b_S)]. \tag{9}$$

Using (1) conditions (8) and (9) can be written as:

$$\begin{aligned} \ln[q(b_S)b_S^*] + \ln[y - b_S^*] &\geq \ln[q(b_R)b_R^*] + \ln[y - b_R^*] \\ \ln[q(b_R)b_R^*] + (1-\theta_R)\ln[y - b_R^*] &\geq \ln[q(b_S)b_S^*] + (1-\theta_R)\ln[y - b_S^*]. \end{aligned}$$

Adding both sides:

$$\ln[y - b_S^*] + (1-\theta_R)\ln[y - b_R^*] \geq \ln[y - b_R^*] + (1-\theta_R)\ln[y - b_S^*],$$

so $b_S^* \leq b_R^*$. □

## A.3 Proof of Proposition 2

*Proof.* If the $R$-type and biased $S$-type maximize (1) subject to (3) the resulting bond supply is $b_R^* = \frac{y}{(2-\theta_R)}$. If the $S$-type and biased $R$-type maximize (1) subject to (2) the resulting bond supply is $b_S^* = \frac{y}{2}$. For this bond supply to be a separating equilibrium the $R$-type and biased $S$-type should not want to mimic. That is:

$$\ln\left[q(b_R)\frac{y}{(2-\theta_R)}\right] + (1-\theta_R)\ln\left[y - \frac{y}{(2-\theta_R)}\right] \geq \ln\left[q(b_S)\frac{y}{2}\right] + (1-\theta_R)\ln\left[y - \frac{y}{2}\right].$$

Solving yields condition (5):

$$\frac{q(b_R)}{q(b_S)} \geq \frac{1}{(1-\theta_R)^{1-\theta_R}}\left(\frac{2-\theta_R}{2}\right)^{(2-\theta_R)}.$$

If it holds there is a type I separating equilibrium.

When condition (5) does not hold then the $R$-type and biased $S$-type will want to mimic and $b_S^* = \frac{y}{2}$ no longer constituents a possible separating equilibrium. The $S$-type and biased $R$-type will need to choose a bond supply $b_S^* \leq \frac{y}{2}$ in order to signal that they are (or believe they are) safe. The resulting equilibrium is then a type II separating equilibrium. □

## A.4 Proof of Proposition 3

*Proof.* Using the law of total probability the relevant conditional probabilities can be written as:

$$
\begin{aligned}
p(b'_S|\theta_R) &= \frac{\frac{(\lambda-\alpha)}{\lambda}(1-\sigma)\lambda + \frac{\beta}{1-\lambda}[1-\lambda(2-\sigma)]}{1-\lambda} \\
p(b'_S|\theta_S) &= \frac{\frac{(\lambda-\alpha)}{\lambda}\sigma\lambda + \frac{\beta}{1-\lambda}(1-\sigma)\lambda}{\lambda} \\
p(b'_R|\theta_R) &= \frac{\frac{\alpha}{\lambda}(1-\sigma)\lambda + \frac{(1-\lambda-\beta)}{1-\lambda}[1-\lambda(2-\sigma)]}{1-\lambda} \\
p(b'_R|\theta_S) &= \frac{\frac{\alpha}{\lambda}\sigma\lambda + \frac{(1-\lambda-\beta)}{1-\lambda}(1-\sigma)\lambda}{\lambda}.
\end{aligned}
$$

$p(b'_S|\theta_S) \geq p(b'_S|\theta_R)$ when $\frac{\lambda-\alpha}{\lambda} \geq \frac{\beta}{1-\lambda}$ and $p(b'_R|\theta_R) \geq p(b'_R|\theta_S)$ when $\frac{1-\lambda-\beta}{1-\lambda} \geq \frac{\alpha}{\lambda}$, which are both the case due to condition (4). It follows that when $b_S$ and $b'_S$ are chosen more weight in the home bond price (6) is put on $1-\theta_S$ due to yardstick competition, and when $b_R$ and $b'_R$ are chosen more weight in (6) is put on $1-\theta_R$ due to yardstick competition. When $b_S$ and $b'_R$ are chosen less weight in (6) is put on $1-\theta_S$ due to yardstick competition, and when $b_R$ and $b'_S$ are chosen less weight in (6) is put on $1-\theta_R$ due to yardstick competition.

Finally, when $\frac{\lambda-\alpha}{\lambda} \geq \frac{\beta}{1-\lambda}$ then $p(b'_S|\theta_S)$ is increasing in $\sigma$ and $p(b'_S|\theta_R)$ is decreasing in $\sigma$. When $\frac{1-\lambda-\beta}{1-\lambda} \geq \frac{\alpha}{\lambda}$ then $p(b'_R|\theta_R)$ is increasing in $\sigma$ and $p(b'_R|\theta_S)$ decreasing in $\sigma$. □

## A.5 Proof of Proposition 4

*Proof.* The proof is similar to the one for Proposition 2 only now the incentive constraint is in term of expected payoffs:

$$
\begin{aligned}
& \phi \ln\left[q(b_R, b'_S)\frac{y}{(2-\theta_R)}\right] + (1-\phi)\ln\left[q(b_R, b'_S)\frac{y}{(2-\theta_R)}\right] + (1-\theta_R)\ln\left[y - \frac{y}{(2-\theta_R)}\right] \\
\geq \quad & \phi \ln\left[q(b_S, b'_S)\frac{y}{2}\right] + (1-\phi)\ln\left[q(b_S, b'_R)\frac{y}{2}\right] + (1-\theta_R)\ln\left[y - \frac{y}{2}\right],
\end{aligned}
$$

with $\phi = p(b'_S|\theta_R)$, $1-\phi = p(b'_R|\theta_R)$. Solving yields condition (7):

$$
\frac{q(b_R, b'_S)^{\phi} q(b_R, b'_R)^{1-\phi}}{q(b_S, b'_S)^{\phi} q(b_S, b'_R)^{1-\phi}} \quad \geq \quad \frac{1}{(1-\theta_R)^{1-\theta_R}} \left(\frac{2-\theta_R}{2}\right)^{(2-\theta_R)}.
$$

□

# References


Bernoth, K., von Hagen, J. and Schuknecht, L. (2012). 'Sovereign risk premiums in the european government bond market', *Journal of International Money and Finance* **31**.

Besley, T. and Case, A. (1995). ‘Incumbent behavior: Vote-seeking, tax-setting, and yardstick competition’, *The American Economic Review* **85**(1), 25–45.

Besley, T. and Smart, M. (2007). ‘Fiscal restraints and voter welfare’, *Journal of Public Economics* **91**, 755773.

Bordignon, M., Cerniglia, F. and Revelli, F. (2004). ‘Yardstick competition in intergovernmental relationships: theory and empirical predictions’, *Economics Letters* **83**, 325–333.

D’Agostino, A. and Ehrmann, M. (2014). ‘The pricing of g7 sovereign bond spreads - the times, they are a-changin’, *Journal of Banking and Finance* **47**, 155–176.

Gande, A. and Parsley, D. (2005). ‘News spillovers in the sovereign debt market’, *Journal of Financial Economics* **75**(3), 691–734.

Honkapohja, S. (2013). ‘Sovereign risk premiums in the european government bond market’, *Journal of Macroeconomics* **39**, 260–271.

Landon, S. and Smith, C. E. (2000). ‘Government debt spillovers and creditworthiness in a federation’, *Canadian Journal of Economics* **33**(3), 634–661.

Landon, S. and Smith, C. E. (2007). ‘Government debt spillovers in a monetary union’, *North American Journal of Economics and Finance* **18**, 135–154.

Mihm, B. (2015). ‘Mispricing of risk in sovereign bond markets with asymmetric information’, *German Economic Review* p. doi: 10.1111/geer.12068.

Perez, D. (2015). ‘Sovereign debt maturity structure under asymmetric information’, *Memeo* .

Sandleris, G. (2008). ‘Sovereign defaults: information, investment and credit’, *Journal of International Economics* **76**, 267–275.

Santos-Pinto, L. (2012). ‘Labor market signaling and self-confidence: Wage compression and the gender pay gap’, *Journal of Labor Economics* **30**(4), 873–914.

Shleifer, A. (1985). ‘A theory of yardstick competition’, *The RAND Journal of Economics* **16**(3), 319–327.

Spence, M. (1973). ‘Job market signaling’, *The Quarterly Journal of Economics* **87**, 355–374.