Soft Budget Constraints as a Risk Sharing Arrangement in an Economic Federation

Erica Lindahl and Andreas Westermark
SOFT BUDGET CONSTRAINTS AS A RISK SHARING ARRANGEMENT
IN AN ECONOMIC FEDERATION

ERICA LINDAHL AND ANDREAS WESTERMARK
Abstract

We analyze a model where the federal government provides risk sharing arrangements to municipalities investing in a local public good. The risk sharing arrangements are an income equalization system and a system allowing for a soft budget constraint, i.e., a bailout. Our main result is that a bailout system in combination with income equalization can be a more efficient risk sharing arrangement than an income equalization system only. Thus, the introduction of a bailout system is welfare improving.

Keywords: Bailout, Fiscal federalism

JEL Classification: H72, H77
1 Introduction

Investments in local public goods are mostly undertaken in order to improve the general welfare in the local economy. However, there is also costs associated with investments. Firstly, the direct financial cost and secondly the costs from distorting incentives through taxes. Another potential disadvantage with investments is that the return is uncertain. However, an important property of redistributive taxation, as pointed out by Varian (1980), is that if income contains a random component, then a system of redistributive taxation will contribute to reduce the variance of after tax income. Hence, taxation can serve as insurance. Another important aspect concerning investment is whether the investors face a soft budget constraint; see e.g. Kornai et al (2003) and Wildasin (1997). A soft budget constraint might lead to inefficient investment but it can also have insurance properties. A central government can impose a risk sharing arrangement, including a soft budget constraint, on local municipalities which is financed by distortionary taxes.

The aim of this paper is to study how the welfare in the municipalities is affected by a soft budget constraint. That is, can a system allowing for bailouts serve as a risk sharing arrangement? We show that at the margin, the introduction of a bailout system can be welfare improving.

This paper relates to (and tries to combine) three fields: (i) fiscal arrangement for risk sharing and (ii) soft budget constraints and iii) optimal taxation in fiscal federalism.

Fiscal arrangements for risk sharing have been examined both theoretically and empirically. Sala-I-Martín (1992) and Asdrubali et al (1996) analyze risk sharing among US states. They find that the capital market and the credit market appear to be the most important mechanism for risk sharing. However, both studies recognize the federal government as an important complement for risk sharing. At a theoretical level, risk-sharing arrangements have been studied by Persson and Tabellini (1996a and 1996b) in a political economy set up. Their result suggest a trade off between federal risk sharing and moral hazard and federal risk sharing and redistribution respectively. Both papers focus on political economy outcomes under different fiscal constitutions. Lockwood (1999) assumes that local public goods give rise to spill over effects and he analyses the central government’s trade off between providing insurance and offering direct corrective incentives for local public goods. Aronsson and Wikström (2003) show that revenue sharing is motivated by the desire to avoid risk. This paper modifies their setup in order to study more explicitly how the risk taking behavior in municipalities is affected by risk sharing. In addition, this paper analyzes another type of risk sharing: a bailout system. As far as we know, soft budget constraints have not been analyzed as a tool for improving the efficiency of a
risk sharing arrangement before this paper.

The literature about soft budget constraints, i.e., the bailout problem, has become large during the last years (Kornai et al. (2003); Akai and Silvia (2003); Goodspeed (2002); Wildasin (1997), Nobuo and Emilson (2003) and Qian and Roland (1998)). Except for Wildasin (1997) the general lesson from the theoretical studies of the effects of the existence of a soft budget constraint is that it gives rise to inefficient economic behavior by the potentially bailed out organization. The intuition is simple; if the executives of an organization expect to be bailed out in case of financial trouble, it will not take full economic responsibility for their business and thereby run it inefficiently. Wildasin (1997) uses another approach; the soft budget constraint can induce socially efficient provision of local public goods that produce spill over benefits. Without federal intervention, municipal investment is lower than the efficient level since municipalities do not take into account the positive horizontal externalities. Kornai et al. (2003) address the question why the budget constraint is soft. As the authors point out the precedent literature give the impression that hardness of the budget constraint is "good" and softness "bad". With this simple conclusion it is hard to understand that soft budget constraints are so widespread and recurrent. The soft budget constraint addressed in this paper is the one that might occur in economic federations. The potential problem with soft budget constraints in this context, is that, if the sub-national governments disrespect the fiscal discipline imposed by their budget constraint, the central government may intervene and rescue them, hence they face a soft budget.

Finally, this paper relates to the optimal taxation literature. In the optimal taxation literature, the government uses taxes to induce individuals to reveal productivity and work at a second best efficient level. In general, optimal taxation schedules are complex nonlinear tax schemes. Here, municipalities’ investment decision correspond to the agents’ labor choice in the standard optimal taxation literature, although the information structure is different in the model presented here. Bailout introduces some nonlinearity in the otherwise linear taxation scheme.

The model in this paper includes three levels of decision making: the individual, the municipal and the federal. Labour supply and private consumption is determined at the individual level, investments in local public goods at the municipal level and finally risk sharing arrangements are chosen by politicians at the federal level. Investments in local public goods and risk sharing arrangements are financed by local (municipal) and federal income taxation respectively, and hence labour supply is a function of local and federal taxes.

We assume that the expected value of an investment as well as the risk associated with the project is increasing with the level of investment. This is not an uncommon assumption. For example, this is the case in a model where an individual invests a specific amount in an asset.
with uncertain proportional return, e.g. the interest rate. Moreover, we assume that the gross value of an investment is positive but uncertain. Since individuals are assumed to be risk averse, the investment level is restrained by both risk aversion and decreasing returns.

Local taxes (investments are partially financed by local taxes) are decided \textit{ex post} the realization of the outcome of investments. That is, a municipality has to set the local tax rate such that it can finance the investment cost associated with the actual investment level it has decided \textit{ex ante} the realization of the outcome of the investment. This assumption seems reasonable; a municipality with a bad realization of the outcome of its investments is forced to raise the local tax rate in order to attain budget balance.

We assume a continuum of municipalities and that all municipalities are assumed to be identical \textit{ex ante} the realization of the outcome of the local investment decision. Further, the local shocks are assumed to be independent among municipalities. This means that the differences in the outcome of investments among municipalities cancel out on an aggregate level.

The central government provides risk sharing in two ways. Firstly, it provides an income equalization system of the type existing in many economic federation (common in the Nordic countries). That is, all municipalities receive an identical (equal for all) lump sum transfer from the federal government. Secondly, it provides a bailout system in which transfers are redistributed differently, depending on the outcome of local investments. That is, municipalities with a realized return below a certain critical value receive a bailout. All transfers are financed by federal taxation.

We find that an introduction of a bailout system within an income equalization system is welfare improving. This result stems from the risk avert agents and the fact that a bailout system can serve as a more efficient insurance (compared to an income equalization system). The reason is that a lump sum transfer that is the same for all municipalities in general does not equate marginal utility between municipalities. In particular, if marginal utility for consumption is larger in municipalities with a bad outcome, a bailout system reduces the difference between municipalities, improving welfare in the federation.

Section 2 presents the model and shows how a bailout system affects the decisions taken at the individual, municipal and federal level. Section 3 presents the results and finally section 4 discusses the results.

2 The Model

We analyze a model where decisions are taken on three levels: federal, municipal and individual. On the individual level, agents choose labour supply and consumption. On the municipal level,
municipalities use revenues to finance local investments $\gamma$. The reason why these investments are undertaken on a local level is that it is reasonable to assume that municipalities have an information advantage in determining investment needs. Finally, on the federal level, the government redistributes among municipalities using two methods: lump sum transfers $r$ and bailouts $\beta$. The size of the lump sum transfer is equal for all municipalities independently of the outcome of the investment. A bailout is only allowed for those municipalities in which the investment outcome is below a critical value. Both redistribution types are financed by a federal tax $\tau_c$ on labour income. By providing risk sharing arrangement in this way, the central government reduces the after tax variance of income. However, taxes implies distortions on labour supply. Hence, there is a trade off between providing a risk sharing arrangement (via taxes) and keeping tax distortions low.

The federation consists of a continuum of municipalities. Each municipality is populated by one representative immobile individual. We start by characterizing the individual level. The utility function of the resident in municipality $i$ is

$$u_i = u^i(c^i, l^i),$$

where $c^i$ is private consumption and $l^i$ is labour supply. The utility function possesses the usual properties, i.e., it is increasing in both its arguments and it is strictly concave. The individual faces the following budget constraint

$$w^i (1 - \tau_c - \tau_l) l^i = c^i$$

where $w^i$ is the wage, $\tau_c$ a federal tax and $\tau_l$ a local tax. The wage in turn equals

$$w^i = \bar{w} + h(\gamma) \left(1 + \xi^i\right).$$

The wage includes a non stochastic part $\bar{w}$ and a stochastic part associated with the investment level $\gamma$, e.g., infrastructure investment. Further, $h(\gamma)$ is a production function with the usual properties: $h'(\gamma) \geq 0$ and $h''(\gamma) \leq 0$ and $\xi^i$ is a local shock with $E[\xi^i] = 0$, $Var(\xi^i) = \sigma^2$. We also assume that the support of $\xi^i$ is bounded: $\xi^i \in [\underline{\xi}, \bar{\xi}]$. The lower bound is restricted to be $\underline{\xi} \geq -1$. Hence, the worst realization of $\xi^i$ implies a wage at least equal to the initial level $\bar{w}$. Thus, the investment cannot lead to a decrease in the wage level.\(^1\) Thus, gross revenues of investments are positive. Since investments are costly, both in terms of the monetary cost and

\(^1\)Another reason for assuming that $\xi \geq -1$ is the following. It seems reasonable that wages at least should be positive, i.e., $w^i \geq 0$. Then $\xi \geq -\frac{\bar{w}}{h'(\gamma)} - 1$. Note that this must hold for all investment levels $\gamma$. If the investment technology is unbounded like when $h(\gamma) = \gamma^\alpha$ for $\alpha < 1$ or $h(\gamma) = \ln \gamma$ then $\frac{\bar{w}}{h'(\gamma)}$ converge to zero for $\gamma$ large and we must have $\xi \geq -1$.\)}
the induced tax distortions, net revenues can of course be negative. The local shock is realized after the labor choice.

2.1 A Bailout system

To analyze a bailout system, all municipalities are for simplicity assumed to be ex ante identical, so they all choose the same level of investment. Hence, the ex post difference between municipalities is the realized value of the stochastic term $\varepsilon$. Consider the following bailout system. All municipalities are guaranteed a minimum return on investment equal to

$$w^i = \bar{w} + h(\gamma)(1 + \hat{\varepsilon}).$$  \hfill (4)

In case of a return below the guaranteed return, the government has to supply additional resources to investment $\gamma_c(\hat{\varepsilon}, \varepsilon)$, i.e., bail out the municipality. The value of the bailout, i.e., the increase in net return is

$$\beta(\hat{\varepsilon}, \varepsilon) = (1 - \tau_c)(w(\hat{\varepsilon}) - w(\varepsilon))I.$$  \hfill (5)

We assume that

$$\gamma_c(\hat{\varepsilon}, \varepsilon) = f(\beta(\hat{\varepsilon}, \varepsilon))$$  \hfill (6)

where $f$ is increasing, concave, $f(0) = 0$ and that $f(x) \leq x$, i.e., the investments required to achieve an income increase of $x$ is weakly smaller than $x$. This implies that all municipalities with a realized value of $\varepsilon < \hat{\varepsilon}$ are compensated with a bailout equal to $\beta(\hat{\varepsilon}, \varepsilon)$. The worst realization of the of $\varepsilon^i$ within a bailout system is thus equal to a wage level above the initial level. That is, the after tax variation in the investment outcome is reduced. Note that if the central government chooses $\hat{\varepsilon} = \bar{\varepsilon}$, we have an income equalization system as a special case within a bailout system. As in Bordignon et al, we assume that total tax revenues is unobservable by the federal government and hence cannot be used as a signal of investment success.

Here, the lump sum transfers $r$ to the municipalities is an ex ante subsidy to encourage investment and $\gamma_c(\hat{\varepsilon}, \varepsilon)$ an ex post subsidy. If it is optimal for the federal government to set $\gamma_c(\hat{\varepsilon}, \varepsilon) > 0$ we say that the budget constraint is soft. This is similar to how Kornai et al (2003) defines a soft budget constraint.

Systems allowing for bailouts are often characterized with asymmetric information. That is, the central government has not full information about the potential need for a bailout in the lower levels of the federation. In earlier literature about soft budget constraints (Kornai et al (2003); Akai and Silvia (2003); Qian and Roland (1998)), asymmetric information is an important ingredient to the predominant conclusion that soft budget are undesirable. The
reason is that asymmetric information concerning the need for a bailout gives rise to moral hazard. Translated to this model, this means that the representatives of the municipalities may not report the true value of the shock when asking for a bailout. We believe this is an important (necessary) part of the analysis of soft budget constraint. In this model, municipalities need not report the true value of $\varepsilon$. Thus, the central government face the risk of untruthful reporting of the municipalities. We assume that the central government monitors the municipalities that report a bailout to detect possible cheaters. Each municipality is monitored with probability $p_m$. In case of observing a misreport by a municipality, that municipality is punished. The cost of the punishment for the municipality is $c_p$. Monitoring is costly and the unit cost is $c_m$.

### 2.2 Individual level

Now, let us analyze the problem of the consumer. Note that we assume that the labour choice is made before return on investment is revealed, while the consumption decision is made after uncertainty is resolved. A motivation for this is that the labour-leisure choice is a more long run decision, since it depends on e.g., human capital levels.\(^2\)

Also, if the return realization of investment is low, local taxes has to be raised in order to balance the municipal budget constraint. The labour decision is made before the actual local tax rate is known. From the municipal resource constraint we can solve for the local tax rate as a function of investment, actual return, and average labour supply in the municipality $l^m$,

$$
\gamma = r + \tau_l (\bar{w} + h(\gamma)(1+\varepsilon)) l^m \iff \tau_l (\gamma, \varepsilon, l^m) = \frac{\gamma - r}{(\bar{w} + h(\gamma)(1+\varepsilon)) l^m}
$$

Agents maximize expected utility, subject to the budget constraint

$$
c^i (\varepsilon) = w^i (\varepsilon) (1 - \tau_c - \tau_l (\gamma, \varepsilon, l^m)) l^i.
$$

Thus, the problem of the consumer can be written as

$$
\max_{l^i} \int_{\tilde{\varepsilon}}^{\tilde{\varepsilon}} u_c (c^i (\tilde{\varepsilon}), l^i) dF (\varepsilon) + \int_{\tilde{\varepsilon}}^{\bar{\varepsilon}} u_l (c^i (\varepsilon), l^i) dF (\varepsilon).
$$

subject to equation (8). The first-order condition is

$$
\Lambda^i = \int_{\tilde{\varepsilon}}^{\bar{\varepsilon}} (u_c (c^i (\tilde{\varepsilon}), l^i) (w^i (\tilde{\varepsilon}) (1 - \tau_c - \tau_l (\gamma, \tilde{\varepsilon}, l^m)) + u_l (c^i (\tilde{\varepsilon}), l^i)) dF (\varepsilon)
$$

$$
+ \int_{\tilde{\varepsilon}}^{\bar{\varepsilon}} (u_c (c^i (\varepsilon), l^i) (w^i (\varepsilon) (1 - \tau_c - \tau_l (\gamma, \varepsilon, l^m)) + u_l (c^i (\varepsilon), l^i)) dF (\varepsilon) = 0,
$$

\(^2\)This assumption implies that all agents choose the same labour supply, which simplifies calculations. Changing the timing would probably not affect the results.
assuming an interior solution. The solution to this problem is denoted $l(\hat{\varepsilon}, \gamma, \tau_c)$. The following intermediate result is needed below.

**Result 1.** We have $\frac{\partial \Lambda_i}{\partial \hat{\varepsilon}} \bigg|_{\hat{\varepsilon} = \bar{\varepsilon}} = 0$ and hence

$$\frac{\partial l}{\partial \hat{\varepsilon}} \bigg|_{\hat{\varepsilon} = \bar{\varepsilon}} = -\frac{\partial \Lambda_i}{\partial \hat{\varepsilon}} \bigg|_{\hat{\varepsilon} = \bar{\varepsilon}} = 0.$$  \hspace{1cm} (11)

**Proof:** Studying the first order condition of the individual’s decision, one easily recognizes two effects of a small increase in $\hat{\varepsilon}$. First, the effects through the bounds cancel out. Second, a slight increase in $\hat{\varepsilon}$ from $\bar{\varepsilon}$ affects $w^i(\hat{\varepsilon})$ and $c^i(\hat{\varepsilon})$ in the first integral. However, since the integral is evaluated at $\hat{\varepsilon} = \bar{\varepsilon}$, bailout is a probability zero event. Hence, there is no effect on the first-order condition of a change in $\hat{\varepsilon}$ when $\hat{\varepsilon}$ is at its lowest value, and we have $\frac{\partial \Lambda_i}{\partial \hat{\varepsilon}} \bigg|_{\hat{\varepsilon} = \bar{\varepsilon}} = 0$. Then, a change in $\hat{\varepsilon}$ does not affect the labour choice either and we get $\frac{\partial l}{\partial \hat{\varepsilon}} \bigg|_{\hat{\varepsilon} = \bar{\varepsilon}} = 0$. 

2.3 Municipal level

As mentioned above, local politicians at the municipality level invest in local public goods in order to increase the general welfare in the municipality. The price of investment is normalized to one and the cost of investment is hence equal to the investment level $\gamma$. Investments are financed by local income taxation. We require that the budget is balanced, i.e.,

$$\gamma = \begin{cases} r + \tau_l \left( \bar{w} + h(\gamma (1 + \varepsilon)) l(\hat{\varepsilon}, \gamma, \tau_c) \right) & \text{if } \varepsilon \geq \hat{\varepsilon} \\ r + \tau_l \left( \bar{w} + h(\gamma (1 + \hat{\varepsilon})) l(\hat{\varepsilon}, \gamma, \tau_c) \right) & \text{if } \varepsilon < \hat{\varepsilon}. \end{cases}$$  \hspace{1cm} (12)

where $r$ is a lump sum transfer distributed to the municipality from the government. Note that the requirement of budget balance implies that the municipality in case of a bad realization is forced to increase the local tax rate *ex post* the realization of the shock to enable budget balance.

The local politicians maximize the expected utility of a representative voter, subject to the constraint above. Note that the constraint implicitly defines the local tax rate as a function of investment; $\tau_l(\gamma)$. This function is different from equation (7), since the municipality takes into account the effect of a change in investment on municipal labour supply $l^m$.

We do not know whether truthtelling is an equilibrium or not. However, we start by focusing on conditions that are required for truthful reporting. Substituting (12) into (8) gives

$$c^m(\varepsilon, \gamma) = w^i(\varepsilon) (1 - \tau_c) l(\hat{\varepsilon}, \gamma, \tau_c) - \gamma + r,$$  \hspace{1cm} (13)

i.e., the individual consumption as a function of $\varepsilon$ and the municipal choice of $\gamma$. The *ex post* equilibrium payoffs for the municipalities when reporting truthfully are when $\varepsilon \geq \hat{\varepsilon}$:

$$u(c^m(\varepsilon, \gamma), l(\hat{\varepsilon}, \gamma, \tau_c))$$  \hspace{1cm} (14)
and when $\varepsilon < \hat{\varepsilon}$

$$u(c^m(\hat{\varepsilon}, \gamma), l(\hat{\varepsilon}, \gamma, \tau_c))$$

Now consider untruthful reporting. There are two ways of gaining by misreporting. First, a municipality that would not receive a bailout when reporting the truth can report a lower value and receive a bailout. Second, a municipality that would receive a bailout when reporting the truth could report an even lower value of $\varepsilon$ to receive a larger bailout. We start with the former case and hence analyze municipalities that have a true $\varepsilon$ larger than $\hat{\varepsilon}$ but report an $\varepsilon$ lower that $\hat{\varepsilon}$. There are lots of possible deviations a municipality could make. However, since a lower reported $\varepsilon$ leads to a larger bailout, municipalities that deviate gets the largest gain when reporting the lowest possible value of $\varepsilon$, i.e., $\underline{\varepsilon}$. The value of such a report is $\beta(\hat{\varepsilon}, \underline{\varepsilon})$. Also, since marginal utility is decreasing, of all types $\varepsilon \geq \hat{\varepsilon}$, type $\hat{\varepsilon}$ has the largest gain in payoff by misreporting $\underline{\varepsilon}$. Hence, if type $\hat{\varepsilon}$ does not gain by reporting $\underline{\varepsilon}$ then no other misreport is profitable. To ensure truthtelling for types $\varepsilon \geq \hat{\varepsilon}$ it is enough to consider type $\hat{\varepsilon}$ when reporting $\underline{\varepsilon}$.

The payoff when announcing $\underline{\varepsilon}$ and not being detected for type $\hat{\varepsilon}$ is

$$u(c^m(\hat{\varepsilon}, \gamma) + \beta(\hat{\varepsilon}, \underline{\varepsilon}), l(\hat{\varepsilon}, \gamma, \tau_c))$$

The payoff for a municipality that misreports $\varepsilon$ and is monitored is then

$$u(c^m(\hat{\varepsilon}, \gamma), l(\hat{\varepsilon}, \gamma, \tau_c)) - c_p$$

To avoid misreporting we require that

$$u(c^m(\hat{\varepsilon}, \gamma), l(\hat{\varepsilon}, \gamma, \tau_c)) \geq u(c^m(\hat{\varepsilon}, \gamma) + \beta(\hat{\varepsilon}, \underline{\varepsilon}), l(\hat{\varepsilon}, \gamma, \tau_c)) - \frac{p_m}{1 - p_m} c_p$$

(18)

For municipalities that receive bailout when reporting the truth, i.e., has a realization $\varepsilon < \hat{\varepsilon}$, the no mimicking condition is identical to condition (18). \(^3\)

From above we know that municipalities truthfully report the value of $\varepsilon$, given the values of $c_p$, and $p_m$. Substituting the constraint (13) into the expected payoff of the representative individual, the problem of the municipality can be written as

$$\max_{\gamma} \int_{\underline{\varepsilon}}^{\hat{\varepsilon}} u(c^m(\hat{\varepsilon}, \gamma), l(\hat{\varepsilon}, \gamma, \tau_c)) \, dF(\varepsilon) + \int_{\hat{\varepsilon}}^{\hat{\varepsilon}} u(c^m(\varepsilon, \gamma), l(\hat{\varepsilon}, \gamma, \tau_c)) \, dF(\varepsilon)$$

(20)

\(^3\)To see this, note that the best possible misrepresentation is again to say the lowest possible $\varepsilon$ and get $\beta(\hat{\varepsilon}, \underline{\varepsilon})$. To avoid misreporting we require that

$$u(c^m(\hat{\varepsilon}, \gamma), l(\hat{\varepsilon}, \gamma, \tau_c)) \geq u(c^m(\varepsilon, \gamma) + \beta(\hat{\varepsilon}, \underline{\varepsilon}), l(\hat{\varepsilon}, \gamma, \tau_c)) - \frac{p_m}{1 - p_m} c_p$$

(19)

Note that the right hand side is increasing in $\varepsilon$. Thus, if the expression holds for the highest possible $\varepsilon = \hat{\varepsilon}$, then it holds for all $\varepsilon < \hat{\varepsilon}$. Also, when $\varepsilon = \hat{\varepsilon}$ the right hand side of the expression above is identical to the right hand side of (18) and hence we get exactly the same condition as condition (18) above.
To analyze the problem for the federal government, we need the following intermediate result that analyzes how a change in $\hat{\varepsilon}$ affects the investment level and the first-order condition $\Lambda^m$ of the municipality. A small change in $\hat{\varepsilon}$, evaluated at $\hat{\varepsilon} = \bar{\varepsilon}$ does not affect the investment level chosen in the municipalities.

**Result 2.** We have $\frac{\partial \Lambda^m}{\partial \hat{\varepsilon}} = 0$ and hence

$$\frac{\partial \gamma}{\partial \hat{\varepsilon}} = \frac{\partial \Lambda^m}{\partial \hat{\varepsilon}} = 0.$$ (21)

**Proof:** See the appendix.

The intuition for this is similar to the one corresponding to the effect of $\hat{\varepsilon}$ on the first order condition at the individual level.

### 2.4 Federal level

The central government redistributes income between municipalities. Specifically, we assume that the government transfers an amount $r$ to each municipality. This is financed by a linear income tax $\tau_c$. Then, when optimizing, all municipalities are perceived as identical by the government. The government maximizes the expected utility of a representative municipality.

Here, the government chooses the four parameters $\tau_c, r, c_p$ and $p_m$. We assume that the punishment cannot be set arbitrarily high. In particular, there is a largest possible punishment $\bar{c}_p$. Note that, since we focus on truth-telling equilibria, $c_p$ do not enter the equilibrium payoffs, since no punishments need to be administrated. However, the punishments has an indirect effect on the problem via the monitoring probability. A higher punishment leads to a lower required monitoring probability in order to avoid misreporting. This in turn leads to lower monitoring costs. This induces the government to set $c_p = \bar{c}_p$.

Now consider equilibrium monitoring costs. Given the optimal punishment $\bar{c}_p$, the central government chooses $p_m$ such that expression (18) holds with equality. When holding with equality expression (18) implicitly defines $p_m$ as a function of $\hat{\varepsilon}$;

$$p_m (\hat{\varepsilon}).$$ (22)

Note that we, by using condition (18), obviously have $p_m (\bar{\varepsilon}) = 0$. The total monitoring cost for the government is

$$C_m (\hat{\varepsilon}) = \int_{\bar{\varepsilon}}^{\hat{\varepsilon}} p_m (\xi) c_m dF (\xi)$$ (23)

Before setting up the problem of the government we describe some properties of the $C_m (\hat{\varepsilon})$ function.
Result 3. We have
\[
\frac{\partial C_m(\hat{\varepsilon})}{\partial \hat{\varepsilon}} = 0 \tag{24}
\]
\[
\frac{\partial^2 C_m(\hat{\varepsilon})}{\partial \hat{\varepsilon}^2} = 2(1 - \tau_c) u_c(\cdot) \frac{\partial w_i(\varepsilon)}{\partial \varepsilon} l(\hat{\varepsilon}, \gamma, \tau_c) \frac{c_m}{c_p} dF(\varepsilon)
\]

Proof: See the appendix.

The reason for this result is that, at the lowest possible realization, no municipalities need to be monitored. A slight increase in \( \hat{\varepsilon} \) from \( \varepsilon \) as above, still makes bailout a probability zero event. Also, since \( p_m(\varepsilon) = 0 \) the effect through the bound is zero. Thus, marginal costs at \( \varepsilon \) is zero. Since costs are positive for all possible \( \hat{\varepsilon} > \varepsilon \) from condition (18), marginal costs must be increasing at \( \varepsilon \), implying that the second derivative is positive.

Now, let us analyze government’s choice of \( \tau_c \) and \( r \). The government maximizes the payoff of a representative individual subject to the federal budget constraint
\[
r \leq \tau_c E[w^l(\hat{\varepsilon}, \gamma, \tau_c)] - \int_{\hat{\varepsilon}}^{\hat{\varepsilon}} f(\beta(\hat{\varepsilon}, \varepsilon)) dF(\varepsilon) - C_m(\hat{\varepsilon}) \tag{25}
\]
and the reaction function of the municipalities
\[
\gamma = \gamma(\tau_c, r) \tag{26}
\]
Substituting (25) into (13) gives
\[
c^\theta(\varepsilon) = w^l(\varepsilon)(1 - \tau_c) l(\hat{\varepsilon}, \gamma, \tau_c) - \gamma + \tau_c (\bar{w} + h(\gamma)) l(\hat{\varepsilon}, \gamma, \tau_c) - \int_{\hat{\varepsilon}}^{\hat{\varepsilon}} f(\beta(\hat{\varepsilon}, \varepsilon)) dF(\varepsilon) - C_m(\hat{\varepsilon}) \tag{27}
\]
i.e., the individual consumption as a function of \( \varepsilon \). The government solves
\[
\max_{\tau_c} \int_{\hat{\varepsilon}}^{\hat{\varepsilon}} u(c^\theta(\varepsilon), l(\hat{\varepsilon}, \gamma, \tau_c)) dF(\varepsilon) + \int_{\hat{\varepsilon}}^{\hat{\varepsilon}} u(c^\theta(\varepsilon), l(\hat{\varepsilon}, \gamma, \tau_c)) dF(\varepsilon) \tag{28}
\]
subject to \( \gamma = \gamma(\tau_c, r) \) and expression (27).

Before presenting the results, we sum up the timing of the model. Individuals choose labour supply \( l \) and consumption \( c \), taken all other parameters as given. Municipal politicians choose the investment level \( \gamma \) and the local tax rate \( \tau_m \) and take into account the effect on labour supply, but the federal instruments \( r, b \) and \( t_c \) are given. At the federal level the lump sum transfer \( r \), the bailout \( b \) and the federal tax rate \( t_c \) are chosen. The objective function of the federal level is defined as the sum of the individual municipalities. Thus, at the federal level, all the reaction actions on lower levels (effects on labour supply and investments) are taken into account when choosing \( r, b \) and \( t_c \). The timing of the model is illustrated in figure 1.

Figure 1. The timing of the model
3 Results

In order to investigate under what circumstances a bailout system is welfare improving, we investigate whether an increase in $\hat{\varepsilon}$ at the lowest possible level is welfare improving at the margin. We denote the first-order condition $\Lambda^g$ and the value function $V(\hat{\varepsilon})$. Before continuing, we note the following intermediate result.

Result 4. We have

$$\frac{\partial V(\hat{\varepsilon})}{\partial \hat{\varepsilon}} = 0$$

and

$$\frac{\partial \tau_c}{\partial \hat{\varepsilon}} \Big|_{\hat{\varepsilon} = \bar{\varepsilon}} = 0.$$  (29)

Proof: See the appendix.

The intuition is again similar to the one corresponding to the effect of $\hat{\varepsilon}$ on the first order condition at the individual level.

Obviously we can not conclude that we have a maximum when $\frac{\partial V(\hat{\varepsilon})}{\partial \hat{\varepsilon}} = 0$. To determine whether $V(\hat{\varepsilon})$ is larger than $V(\bar{\varepsilon})$ for $\hat{\varepsilon}$ close to $\bar{\varepsilon}$ we need to look at the second derivatives

$$\frac{\partial^2 V(\hat{\varepsilon})}{\partial \hat{\varepsilon}^2} = u_c(\epsilon^g(\hat{\varepsilon}), l(\hat{\varepsilon}, \gamma, \tau_c)) h(\gamma) (1 - \tau_c) l(\hat{\varepsilon}, \gamma, \tau_c) dF(\hat{\varepsilon})$$

$$- \int_{\hat{\varepsilon}}^{\bar{\varepsilon}} u_c(\epsilon^g(\hat{\zeta}), l(\hat{\zeta}, \gamma, \tau_c)) K(\hat{\zeta}) dF(\hat{\zeta})$$

where

$$K(\hat{\varepsilon}) = f'(0) (1 - \tau_c) (h(\gamma) l(\hat{\varepsilon}, \gamma, \tau_c)) dF(\varepsilon) + \frac{\partial^2 C_m(\varepsilon)}{\partial \varepsilon^2}$$

This expression is derived in A.1 in the appendix. Substituting in the expression for $\frac{\partial^2 C_m(\varepsilon)}{\partial \varepsilon^2}$ and evaluating at $\varepsilon = \bar{\varepsilon}$ gives

$$K(\hat{\varepsilon}) = f'(0) + h(\gamma) (1 - \tau_c) l(\bar{\varepsilon}, \gamma, \tau_c) 2u_c(\epsilon^g(\bar{\varepsilon}), l(\bar{\varepsilon}, \gamma, \tau_c)) \frac{c_m}{c_p}.$$
Thus the second derivative is positive if
\[ u_c(c^g(\varepsilon), l(\varepsilon, \gamma, \tau_c)) dF(\varepsilon) > f'(0) + 2u_c(c^g(\varepsilon), l(\varepsilon, \gamma, \tau_c)) \frac{c_m}{c_p} \int_{\varepsilon}^{\hat{\varepsilon}} u_c(c^g(\varepsilon), l(\varepsilon, \gamma, \tau_c)) dF(\varepsilon) \]
(31)

Note that, since \( f(0) = 0 \) and that \( f(x) \leq x \), we must have \( f'(0) \leq 1 \). Since individuals are risk averse, implying that marginal utilities are decreasing in consumption, and if \( \frac{c_m}{c_p} \) is small we have
\[ \frac{\partial^2 V(\hat{\varepsilon})}{\partial \hat{\varepsilon}^2} \bigg|_{\hat{\varepsilon} = \varepsilon} > 0. \]
(32)

From the intermediate results above the second derivative is positive at \( \hat{\varepsilon} = \varepsilon \) if \( \frac{c_m}{c_p} \) is small. Hence, we have a (strict) local minimum at that point. Furthermore, since the second derivative is positive, the first derivative is positive for \( \hat{\varepsilon} \) close to \( \varepsilon \) and it must be optimal to choose an \( \hat{\varepsilon} \) larger than \( \varepsilon \). That is, introducing a bailout system in this model is welfare improving.

**Proposition 1** The introduction of a bailout system is welfare improving at the margin, evaluated at \( \hat{\varepsilon} = \varepsilon \) if \( \frac{c_m}{c_p} \) is small.

The intuition is the following. The federal government is defined as the sum of the individual municipalities. The municipalities (and the individuals), in turn, optimize their utilities given the federal instruments. Thus, when studying the effect of introducing a bailout system in a situation without bailouts, we only have to consider the direct effects. 4

There are two direct effects of introducing a bailout in this model. Note first that the effect of a change in \( \hat{\varepsilon} \) on consumption is \( h(\gamma)(1 - \tau_c) l(\varepsilon, \gamma, \tau_c) \). Firstly, the effect via the incentive compatibility constraint. That is, the monitoring cost in relation to the punishment cost \( \frac{c_m}{c_p} \) matter. The monitoring cost is assumed to be fixed, hence the punishment has to be large. A large punishment is necessary in order to guarantee that the cost of a bailout system in an asymmetric information set up (monitoring is costly) not becomes too large.

Secondly, it is obvious that the risk aversion assumption is essential. Risk aversion is equivalent to the assumption of decreasing marginal utility. The benefit in utility terms of a bailout is received by those with the lowest outcome of investment \( u_c(c^g(\varepsilon), l(\varepsilon, \gamma, \tau_c)) dF(\varepsilon) \) and the corresponding cost is \( \int u_c(c^g(\varepsilon), l(\varepsilon, \gamma, \tau_c)) dF(\varepsilon) \). That is, redistribution is welfare improving when individuals are assumed to be risk averse and acting (investing) under uncertain conditions. A bailout system can serve as a redistributive instrument and thereby be welfare improving. Moreover, a bailout system is more efficient than an income equalization system in

---

4The lower levels have already made their optimal choices (we use the Envelope theorem). Further, since we evaluate at \( \hat{\varepsilon} = \varepsilon \), the vertical externalities cancel out.
this respect. The reason is mainly due to the fact that redistribution is more targeted within a bailout system, but also due to the fact that a bailout system internalize the two types of vertical externalities existing in the model. Remember, that the municipalities neither take into account that the local income tax affect the federal tax base via labour supply, nor that the investment level affect the federal tax base via an increase in the wage level.

4 Discussion

The model above shows that a bailout system, in addition to an income equalization system, can improve the risk sharing arrangement and thereby be welfare improving. Income equalization systems are common in the Nordic countries and are motivated mostly due to equity reasons. The model above illustrates an efficiency argument for an income equalization system. This is also shown in Aronsson and Wikström (2003) but in a different set up. The uncertainty in our model, as opposed to Aronsson and Wikström (2003), is associated with the actions taken by local politicians. The model also shows that the inclusion of a bailout system, in addition to an income equalization system, can provide an even more efficient risk sharing.

In this model, the introduction of a bailout system leads to an increase in welfare. This is an opposite conclusion to most models analyzing the bailout phenomena. The general lesson from bailouts is rather that bailouts lead to inefficient economic behavior by the bailed out organizations.

The inefficient behavior is (often) related to the assumption of asymmetric information between the bailed out organization and the organization allowing the bailout. We believe that asymmetric information is important also in an economic federation. It is reasonable to assume that municipalities have an information advantage in determining investment needs. If the federal government had access to the same information as the municipalities, the first-best investment level could be determined at the federal level. However, since this is not the case, investments are determined at the municipal level.

Due to risk aversion, the federal government provides risk sharing which has to be funded by an income tax. The positive federal income tax leads to vertical externalities, since municipalities do not take into account the effects of their tax and investment choices on the federal tax base, implying that local tax and investment choices are inefficient. In this paper, we show that introducing a bailout system, in addition to an income redistribution system, is welfare improving on the margin. The intuition for our result is that a bailout system serves as a better insurance. The reason is that a lump sum transfer (that is the same for all municipalities) does not equate marginal utility between municipalities. If marginal utility for consumption is
larger in municipalities with a bad outcome, a bailout system reduces the difference between municipalities, leading to higher welfare. For this result to be true, note that it is necessary that monitoring costs are small relative to the punishment the federation can impose on municipalities for misreporting the investment return.

Finally note that the result is related to models of optimal taxation. In such models, the government uses taxes to induce individuals to reveal productivity and work at a second best efficient level. The result in these models is that optimal taxation schedules are complex nonlinear tax schemes. Here, the investment choice of municipalities correspond to the agents’ labor choice in optimal taxation models. The consequence of introducing bailout in addition to a federal tax is that it introduces some nonlinearity in the otherwise linear taxation scheme.
5 References


Appendix

A  Some results

Before proving the results, we derive the derivative of $\Lambda^i$ with respect to $\hat{\varepsilon}$, using that

$$w^i(\hat{\varepsilon})(1 - \tau_c - \tau_l(\gamma, \hat{\varepsilon}, l_m)) = \frac{c^i(\hat{\varepsilon})}{\hat{\varepsilon}},$$

$$\frac{\partial \Lambda^i}{\partial \hat{\varepsilon}} = \int_{\hat{\varepsilon}} u_c(.) \left( h(\gamma)(1 - \tau_c - \tau_l(\gamma, \hat{\varepsilon}, l_m)) - w^i(\hat{\varepsilon}) \frac{\partial \tau_l(.)}{\partial \hat{\varepsilon}} \right) dF(\varepsilon) + \int_{\hat{\varepsilon}} \left( u_{cc}(.) \frac{c^i(\hat{\varepsilon})}{\hat{\varepsilon}} + u_{cd}(.) \right) \left( h(\gamma)(1 - \tau_c - \tau_l(\gamma, \hat{\varepsilon}, l_m)) - w^i(\hat{\varepsilon}) \frac{\partial \tau_l(.)}{\partial \hat{\varepsilon}} \right) \hat{\varepsilon} dF(\varepsilon)$$

where

$$\frac{\partial \tau_l(\gamma, \hat{\varepsilon}, l)}{\partial \hat{\varepsilon}} = -\frac{\gamma - r}{(\bar{w} + h(\gamma)(1 + \hat{\varepsilon}))^2} h(\gamma).$$

Result 2 is derived from the following. The first-order condition with respect to $\gamma$ is, letting $u_{cc}^m(\varepsilon, \hat{\varepsilon}) = u_c(c^m(\varepsilon, \gamma), l(\hat{\varepsilon}, \gamma, \tau_c))$,

$$\Lambda^m = \int_{\hat{\varepsilon}} u_{cc}^m(\varepsilon, \hat{\varepsilon})(h'(\gamma)(1 + \hat{\varepsilon})(1 - \tau_c) l(\hat{\varepsilon}, \gamma, \tau_c) - 1) dF(\varepsilon) \quad (34)$$

$$+ \int_{\hat{\varepsilon}} u_{cc}^m(\varepsilon, \hat{\varepsilon})(h'(\gamma)(1 + \hat{\varepsilon})(1 - \tau_c) l(\hat{\varepsilon}, \gamma, \tau_c) - 1) dF(\varepsilon) + \Lambda^l l_{\gamma}$$

$$+ \int_{\hat{\varepsilon}} u_{cc}^m(\varepsilon, \hat{\varepsilon}) \frac{\partial c^m(\varepsilon, \gamma)}{\partial \gamma} dF(\varepsilon) + \int_{\hat{\varepsilon}} u_{cc}^m(\varepsilon, \hat{\varepsilon}) \frac{\partial c^m(\varepsilon, \gamma)}{\partial \gamma} dF(\varepsilon)$$

where $\frac{\partial c^m(\varepsilon, \gamma)}{\partial \gamma} = \tau_l w^i(\hat{\varepsilon}) l_{\gamma}$. Note that it is the case that $\gamma$ affects $\tau_l$ through the labour supply. From the first-order condition of the agents $\Lambda^i = 0$, this term vanishes. Note that the implicitly defined tax function $\tau_l(\gamma)$ only is affected by a change in $\varepsilon$ for realizations below $\hat{\varepsilon}$. Using that $\Lambda^i = 0$, and that all effects through the bounds of the integral cancels out, the effect of a change in $\hat{\varepsilon}$ on $\Lambda^m$ is, letting $u_{cc}^m(\varepsilon, \hat{\varepsilon}) = u_c(c^m(\varepsilon, \gamma), l(\hat{\varepsilon}, \gamma, \tau_c))$,

$$\frac{\partial \Lambda^m}{\partial \hat{\varepsilon}} = \int_{\hat{\varepsilon}} u_{cc}^m(\varepsilon, \hat{\varepsilon}) h'(\gamma)(1 - \tau_c) l(\hat{\varepsilon}, \gamma, \tau_c) dF(\varepsilon) \quad (35)$$

$$+ \int_{\hat{\varepsilon}} u_{cc}^m(\varepsilon, \hat{\varepsilon})(h'(\gamma)(1 + \hat{\varepsilon})(1 - \tau_c) l(\hat{\varepsilon}, \gamma, \tau_c) h(\gamma)(1 - \tau_c) l(\hat{\varepsilon}, \gamma, \tau_c)) dF(\varepsilon)$$

$$+ \frac{\partial \Lambda^l}{\partial \hat{\varepsilon}} l_{\gamma} + \int_{\hat{\varepsilon}} \frac{d}{d\hat{\varepsilon}} \left( u_{cc}^m(\varepsilon, \hat{\varepsilon}) \left( \frac{\partial c^m(\varepsilon, \gamma)}{\partial \gamma} \right) \right) dF(\varepsilon)$$

17
Evaluating at \( \hat{\varepsilon} = \varepsilon \) and using that \( \frac{\partial \Lambda^i}{\partial \varepsilon} = 0 \) gives

\[
\frac{\partial \Lambda^m}{\partial \varepsilon} = 0. \quad (36)
\]

By using \( \frac{\partial \Lambda^m}{\partial \varepsilon} = 0 \), we also have

\[
\frac{\partial \gamma}{\partial \varepsilon} = -\frac{\partial \Lambda^m}{\partial \gamma}. \quad (37)
\]

A small change in \( \hat{\varepsilon} \), evaluated at \( \hat{\varepsilon} = \varepsilon \) does not affect the investment level chosen in the municipalities.\( \blacksquare \)

**Result 3.** The effect on the cost of a change in \( \hat{\varepsilon} \) is

\[
\frac{\partial C_m (\hat{\varepsilon})}{\partial \hat{\varepsilon}} = p_m (\hat{\varepsilon}) c_m dF (\hat{\varepsilon}) + \int_{\hat{\varepsilon}}^{\varepsilon} \frac{\partial p_m (\hat{\varepsilon})}{\partial \hat{\varepsilon}} c_m dF (\varepsilon) \quad (38)
\]

Evaluating at \( \varepsilon \) gives

\[
\frac{\partial C_m (\varepsilon)}{\partial \varepsilon} = p_m (\varepsilon) c_m dF (\varepsilon) \quad (39)
\]

Note that we must have \( p_m (\varepsilon) = 0 \), since when \( \hat{\varepsilon} = \varepsilon \) we get \( \beta (\varepsilon, \varepsilon) = 0 \) from expression 5 leading to \( p_m = 0 \) in 18. Thus, marginal cost is zero. Now, let us analyze the second derivative of \( C_m \):

\[
\frac{\partial^2 C_m (\hat{\varepsilon})}{\partial \hat{\varepsilon}^2} = 2 \frac{\partial p_m (\hat{\varepsilon})}{\partial \hat{\varepsilon}} c_m dF (\hat{\varepsilon}) + p_m (\hat{\varepsilon}) c_m \frac{d (dF (\hat{\varepsilon}))}{d \hat{\varepsilon}} + \int_{\hat{\varepsilon}}^{\varepsilon} \frac{\partial^2 p_m (\hat{\varepsilon})}{\partial \hat{\varepsilon}^2} c_m dF (\varepsilon) \quad (40)
\]

Evaluating at \( \varepsilon \) gives since \( p_m (\varepsilon) = 0 \),

\[
\frac{\partial^2 C_m (\varepsilon)}{\partial \varepsilon^2} = 2 \frac{\partial p_m (\varepsilon)}{\partial \varepsilon} c_m dF (\varepsilon) \quad (41)
\]

Using that the no mimicking condition 18 holds with equality and differentiating with respect to \( p_m \) and \( \hat{\varepsilon} \) gives , where the second equality uses that \( p_m (\varepsilon) = 0 \) and, from expression 5, that

\[
\frac{\partial^3}{\partial \varepsilon^3} (1 - \tau_c) \frac{\partial w^i (\hat{\varepsilon})}{\partial \varepsilon} l (\hat{\varepsilon}, \gamma, \tau_c), \quad (36)
\]

\[
\frac{\partial p_m (\hat{\varepsilon})}{\partial \hat{\varepsilon}} = -\frac{u_c (\cdot)}{1 - p_m} \frac{\partial^3}{\partial \varepsilon^3} (1 - \tau_c) \frac{\partial w^i (\hat{\varepsilon})}{\partial \varepsilon} l (\hat{\varepsilon}, \gamma, \tau_c). \quad (44)
\]

We get

\[
\frac{\partial^2 C_m (\varepsilon)}{\partial \varepsilon^2} = 2 (1 - \tau_c) u_c (\cdot) \frac{\partial w^i (\varepsilon)}{\partial \varepsilon} l (\varepsilon, \gamma, \tau_c) \frac{c_m}{c_p} dF (\varepsilon) \quad (45)
\]

\[\text{Note that the effects on bailout of a change in } \hat{\varepsilon} \text{ are}
\]

\[
(1 - \tau_c) \frac{\partial w^i (\hat{\varepsilon})}{\partial \hat{\varepsilon}} l (\hat{\varepsilon}, \gamma, \tau_c) - \frac{\partial \tau_c}{\partial \varepsilon} \left( w^i (\hat{\varepsilon}) - w^i (\varepsilon) \right) l (\hat{\varepsilon}, \gamma, \tau_c) + (1 - \tau_c) \left( w^i (\hat{\varepsilon}) - w^i (\varepsilon) \right) \frac{\partial l}{\partial \hat{\varepsilon}} \quad (42)
\]

Evaluating at \( \varepsilon \) gives

\[
(1 - \tau_c) \frac{\partial w^i (\varepsilon)}{\partial \varepsilon} l (\varepsilon, \gamma, \tau_c) \quad (43)
\]
Result 4 Finding $\frac{\partial \tau_c}{\partial \hat{\varepsilon}}$ at $\hat{\varepsilon} = \varepsilon$.

Letting

$$\beta (\hat{\varepsilon}, \varepsilon) = (1 - \tau_c) (w (\hat{\varepsilon}) l (\hat{\varepsilon}, \gamma, \tau_c) - w (\varepsilon) l (\hat{\varepsilon}, \gamma, \tau_c))$$

and defining

$$b_l (\hat{\varepsilon}) = (-w^l (\hat{\varepsilon}) + (\bar{w} + h (\gamma))) l (\cdot, \hat{\varepsilon}) + \int_{\hat{\varepsilon}}^{\bar{\varepsilon}} f' (\beta (\hat{\varepsilon}, \varepsilon)) (w (\hat{\varepsilon}) - w (\varepsilon)) l (\cdot, \hat{\varepsilon}) dF (\varepsilon)$$

$$b_h (\hat{\varepsilon}) = (-w^h (\hat{\varepsilon}) + (\bar{w} + h (\gamma))) l (\cdot, \hat{\varepsilon}) + \int_{\hat{\varepsilon}}^{\bar{\varepsilon}} f' (\beta (\hat{\varepsilon}, \varepsilon)) (w (\hat{\varepsilon}) - w (\varepsilon)) l (\cdot, \hat{\varepsilon}) dF (\varepsilon)$$

and

$$a_l (\hat{\varepsilon}) = w^l (\hat{\varepsilon}) (1 - \tau_c) + \tau_c (\bar{w} + h (\gamma)) - \int_{\hat{\varepsilon}}^{\bar{\varepsilon}} f' (\beta (\hat{\varepsilon}, \varepsilon)) (1 - \tau_c) (w (\hat{\varepsilon}) - w (\varepsilon)) dF (\varepsilon)$$

$$a_h (\hat{\varepsilon}) = w^h (\hat{\varepsilon}) (1 - \tau_c) + \tau_c (\bar{w} + h (\gamma)) - \int_{\hat{\varepsilon}}^{\bar{\varepsilon}} f' (\beta (\hat{\varepsilon}, \varepsilon)) (1 - \tau_c) (w (\hat{\varepsilon}) - w (\varepsilon)) dF (\varepsilon)$$

The effect of $\hat{\varepsilon}$ on the value $V (\hat{\varepsilon})$ and its derivative is, suppressing the arguments in the utility function, letting $u^g (\varepsilon, \hat{\varepsilon}) = u_c (c^g (\varepsilon), l (\hat{\varepsilon}, \gamma, \tau_c))$

$$\frac{\partial V (\hat{\varepsilon})}{\partial \hat{\varepsilon}} = \Lambda^g \frac{\partial \tau_c}{\partial \hat{\varepsilon}} + \int_{\hat{\varepsilon}}^{\bar{\varepsilon}} u^g_l (\hat{\varepsilon}, \hat{\varepsilon}) h (\gamma) (1 - \tau_c) l (\hat{\varepsilon}, \gamma, \tau_c) dF (\varepsilon)$$

$$- \int_{\hat{\varepsilon}}^{\bar{\varepsilon}} \left( \int_{\hat{\varepsilon}}^{\bar{\varepsilon}} f' (\beta (\hat{\varepsilon}, \varepsilon)) (1 - \tau_c) h (\gamma) l (\hat{\varepsilon}, \gamma, \tau_c) dF (\varepsilon) + \frac{\partial C_m (\hat{\varepsilon})}{\partial \hat{\varepsilon}} \right) dF (\varepsilon)$$

$$- \int_{\hat{\varepsilon}}^{\bar{\varepsilon}} u^g_h (\varepsilon, \hat{\varepsilon}) \left( \int_{\hat{\varepsilon}}^{\bar{\varepsilon}} f' (\beta (\hat{\varepsilon}, \varepsilon)) (1 - \tau_c) h (\gamma) l (\hat{\varepsilon}, \gamma, \tau_c) dF (\varepsilon) + \frac{\partial C_m (\hat{\varepsilon})}{\partial \hat{\varepsilon}} \right) dF (\varepsilon)$$

where

$$c^g (\varepsilon) = (\bar{w} + h (\gamma) (1 + \hat{\varepsilon})) (1 - \tau_c) l (\hat{\varepsilon}, \gamma, \tau_c) - \gamma + \tau_c (\bar{w} + h (\gamma)) l (\hat{\varepsilon}, \gamma, \tau_c)$$

$$- \int_{\hat{\varepsilon}}^{\bar{\varepsilon}} f (\beta (\hat{\varepsilon}, \varepsilon)) dF (\varepsilon) - C_m (\hat{\varepsilon})$$
the first-order condition for the federal government is,

\[
\Lambda^g = \int_{\hat{\varepsilon}}^{\bar{\varepsilon}} u^g_c (\hat{\varepsilon}, \hat{\varepsilon}) dF(\varepsilon) + \int_{\hat{\varepsilon}}^{\bar{\varepsilon}} u^g_h (\hat{\varepsilon}) dF(\varepsilon)
\]

\[
+ \int_{\hat{\varepsilon}}^{\bar{\varepsilon}} u^g_l (\hat{\varepsilon}) dF(\varepsilon) l_f + \int_{\hat{\varepsilon}}^{\bar{\varepsilon}} u^g (\varepsilon) dF(\varepsilon) l_f
\]

\[
+ \int_{\hat{\varepsilon}}^{\bar{\varepsilon}} u^g (\varepsilon) a_l (\hat{\varepsilon}) dF(\varepsilon) l_f + \int_{\hat{\varepsilon}}^{\bar{\varepsilon}} u^g (\varepsilon) a_h (\hat{\varepsilon}) dF(\varepsilon) l_f + \Lambda^{m0} \frac{\partial \gamma}{\partial \tau_c}
\]

\[
+ \int_{\hat{\varepsilon}}^{\bar{\varepsilon}} u^g (\varepsilon) \left( \tau_c - (1 - \tau_c) \right) \int_{\hat{\varepsilon}}^{\bar{\varepsilon}} f'(\beta(\hat{\varepsilon}, \varepsilon))(\hat{\varepsilon} - \varepsilon) dF(\varepsilon) dF(\varepsilon) (h'(\gamma) l(\varepsilon) + h(\gamma) l_f) \frac{\partial \gamma}{\partial \tau_c}
\]

Note that most of the effects through \( \gamma \) vanishes by using the first-order condition of the municipality and the effects through \( l \) vanishes (except in the revenue constraint) by using the first-order condition of the individual. The last two terms are the vertical externalities that municipalities do not take into account when choosing investment.

Consider the effects of a change in \( \hat{\varepsilon} \) on \( \Lambda^g \), i.e.,

\[
\frac{\partial \Lambda^g}{\partial \hat{\varepsilon}}.
\]

First, note that the effect of a change in \( \hat{\varepsilon} \) on \( \Lambda^g \) through the bounds of the integral cancels out. Furthermore, it is easy to see that the effect of a change in \( \hat{\varepsilon} \) on \( \int_{\hat{\varepsilon}}^{\bar{\varepsilon}} f(\beta(\hat{\varepsilon}, \varepsilon)) dF(\varepsilon) \) is zero.

Also, since \( \Lambda^{m0} = 0 \) and \( \frac{\partial \Lambda^m}{\partial \tau_c} = 0 \) the effects through the term \( \Lambda^{m0} \frac{\partial \gamma}{\partial \tau_c} \) are zero.

Finally, consider the effects through \( (h'(\gamma) l(\hat{\varepsilon}, \gamma, \tau_c) + h(\gamma) l_f) \frac{\partial \gamma}{\partial \tau_c} \), i.e., the effects working through the vertical externalities. The effects through this term are

\[
(\frac{h''(\gamma)}{l(\hat{\varepsilon}, \gamma, \tau_c) + h'(\gamma) l_f) \frac{\partial \gamma}{\partial \tau_c} + \left( h'(\gamma) \frac{\partial l}{\partial \varepsilon} + h(\gamma) \frac{\partial^2 l}{\partial \varepsilon \partial \gamma} \right) \frac{\partial \gamma}{\partial \tau_c} \right) \left( \frac{\partial^2 l}{\partial \tau_c \partial \varepsilon} \right)
\]

First, note that \( \frac{\partial l}{\partial \varepsilon} = 0 \), eliminating the first term. Second, we have

\[
\left. \frac{\partial l}{\partial \varepsilon} \right|_{\hat{\varepsilon} = \bar{\varepsilon}} = -\frac{\partial \Lambda^m}{\partial \gamma}
\]

Then

\[
\left. \frac{\partial^2 l}{\partial \varepsilon \partial \gamma} \right|_{\hat{\varepsilon} = \bar{\varepsilon}} = \frac{\partial \Lambda^m}{\partial \gamma} \frac{\partial^2 \Lambda^m}{\partial \gamma \partial \tau_c} - \frac{\partial^2 \Lambda^m}{\partial \gamma \partial l} \frac{\partial l}{\partial \varepsilon}.
\]
From result 1 we have $\frac{\partial \Lambda^i}{\partial \hat{\varepsilon}}$. Using the expression for $\frac{\partial \Lambda^i}{\partial \hat{\varepsilon}}$ in the proof of result one, it is easy to see that $\frac{\partial^2 \Lambda^i}{\partial \hat{\varepsilon} \partial \gamma} = 0$ implying that

$$\frac{\partial l}{\partial \hat{\varepsilon}} \bigg| _{\hat{\varepsilon} = \varepsilon} = 0,$$

$$\frac{\partial^2 l}{\partial \hat{\varepsilon} \partial \gamma} \bigg| _{\hat{\varepsilon} = \varepsilon} = 0,$$

eliminating the second term. Finally, we have

$$\frac{\partial \gamma}{\partial \hat{\varepsilon}} = \frac{\partial \Lambda^m}{\partial \hat{\varepsilon}}$$

and

$$\frac{\partial^2 \gamma}{\partial \tau_c \partial \hat{\varepsilon}} = \frac{\partial^2 \Lambda^m}{\partial \hat{\varepsilon} \partial \tau_c} - \frac{\partial^2 \Lambda^m}{\partial \hat{\varepsilon} \partial \tau_c},$$

(54)

Using expression (35) we have $\frac{\partial \Lambda^m}{\partial \hat{\varepsilon}} = 0$ and

$$\frac{\partial^2 \Lambda^m}{\partial \hat{\varepsilon} \partial \tau_c} = \frac{\partial \Lambda^m}{\partial \tau_c},$$

(55)

It is easy too see from 33 that $\frac{\partial^2 \Lambda^i}{\partial \hat{\varepsilon} \partial \tau_c} = 0$ when $\hat{\varepsilon} = \varepsilon$. Evaluating the expression above at $\hat{\varepsilon} = \varepsilon$ gives, using that $\frac{\partial \Lambda^i}{\partial \hat{\varepsilon}} = 0$,

$$\frac{\partial^2 \Lambda^m}{\partial \hat{\varepsilon} \partial \tau_c} = 0,$$

(56)

eliminating the last term.

Finally, all other effects on $\Lambda^g$ of a change in $\hat{\varepsilon}$ appear in terms where integration is from $\varepsilon$ to $\hat{\varepsilon}$ and are thus zero. Hence, we have

$$\frac{\partial \Lambda^g}{\partial \hat{\varepsilon}} = 0$$

(57)

and thus we get

$$\frac{\partial \tau_c}{\partial \hat{\varepsilon}} = -\frac{\partial \Lambda^g}{\partial \tau_c} = 0.$$

(58)
A.1 Deriving an expression for $\frac{\partial^2 V(\hat{\varepsilon})}{\partial \hat{\varepsilon}^2}$.

The second derivative with respect to $\hat{\varepsilon}$ is, using $l(.) = l(\hat{\varepsilon}, \gamma, \tau_c)$,

\[
\frac{\partial^2 V(\hat{\varepsilon})}{\partial \hat{\varepsilon}^2} = \Lambda^g \frac{\partial^2 \tau_c}{\partial \hat{\varepsilon}^2} + \left( \frac{\partial \Lambda^g}{\partial \tau_c} \frac{\partial \tau_c}{\partial \hat{\varepsilon}} + \frac{\partial \Lambda^g}{\partial \gamma} \frac{\partial \gamma}{\partial \hat{\varepsilon}} \right) \frac{\partial \tau_c}{\partial \hat{\varepsilon}} + \frac{\partial \Lambda^g}{\partial \gamma} \frac{\partial \gamma}{\partial \hiat{\varepsilon}} \frac{\partial \tau_c}{\partial \hat{\varepsilon}} + \int_{\hat{\varepsilon}} u_c(.) \left( -f'(0) (1 - \tau_c) (h(\gamma) l(.) \right) \right) dF(\hat{\varepsilon}) dF(\varepsilon)
\]

\[
+ u_c(.) h(\gamma) (1 - \tau_c) l(.) dF(\hat{\varepsilon}) + \int_{\hat{\varepsilon}} u_{cc}(.) (h(\gamma) (1 - \tau_c) l(.)^2 dF(\varepsilon)
\]

\[
+ \int_{\hat{\varepsilon}} u_{cc}(.) \left( -\int_{\hat{\varepsilon}} f'(\beta(\hat{\varepsilon}, \varepsilon)) (1 - \tau_c) (h(\gamma) l(.) \right) dF(\varepsilon) - \frac{\partial C_m(\hat{\varepsilon})}{\partial \hiat{\varepsilon}} \right) \frac{\partial \tau_c}{\partial \hat{\varepsilon}} \right) \frac{\partial \tau_c}{\partial \hat{\varepsilon}} + \int_{\hat{\varepsilon}} u_{cc}(.) \left( -\int_{\hat{\varepsilon}} f'(\beta(\hat{\varepsilon}, \varepsilon)) (1 - \tau_c) (h(\gamma) l(.) \right) dF(\varepsilon) - \frac{\partial C_m(\hat{\varepsilon})}{\partial \hiat{\varepsilon}} \right) \frac{\partial \tau_c}{\partial \hat{\varepsilon}} \right) \frac{\partial \tau_c}{\partial \hat{\varepsilon}} + \int_{\hat{\varepsilon}} u_{cc}(.) \left( -\frac{\partial^2 C_m(\hat{\varepsilon})}{\partial \hat{\varepsilon}^2} \right) dF(\varepsilon) + \int_{\hat{\varepsilon}} u_c(.) \left( -\frac{\partial^2 C_m(\hat{\varepsilon})}{\partial \hat{\varepsilon}^2} \right) dF(\varepsilon)
\]

Using that $\Lambda^g = 0, \frac{\partial \tau_c}{\partial \hat{\varepsilon}} = 0, \frac{\partial \gamma}{\partial \hat{\varepsilon}} = 0, \frac{\partial C_m(\hat{\varepsilon})}{\partial \hat{\varepsilon}} = 0$ and $\hat{\varepsilon} = \hat{\varepsilon}$ gives

\[
\frac{\partial^2 V(\hat{\varepsilon})}{\partial \hat{\varepsilon}^2} = u_c^g(\hat{\varepsilon}, \hat{\varepsilon}) h(\gamma) (1 - \tau_c) l(\hat{\varepsilon}, \gamma, \tau_c) dF(\hat{\varepsilon})
\]

\[
- \int_{\hat{\varepsilon}} u_c^g(\varepsilon, \hat{\varepsilon}) \left( f'(0) (1 - \tau_c) (h(\gamma) l(\hat{\varepsilon}, \gamma, \tau_c) \right) dF(\varepsilon) + \frac{\partial^2 C_m(\hat{\varepsilon})}{\partial \hat{\varepsilon}^2} \right) \frac{\partial \tau_c}{\partial \hat{\varepsilon}} \right) dF(\varepsilon).
\]
WORKING PAPERS*
Editor: Nils Gottfries


2005:11 Martin Ågren, Myopic Loss Aversion, the Equity Premium Puzzle, and GARCH. 34 pp.


* A list of papers in this series from earlier years will be sent on request by the department.


2006:3 Magnus Gustavsson and Henrik Jordahl, Inequality and Trust: Some Inequalities are More Harmful than Others. 29pp.


See also working papers published by the Office of Labour Market Policy Evaluation http://www.ifau.se/

ISSN 0284-2904