Eidgenössische Finanzverwaltung, Bundesgasse 3, CH-3003 Bern
Administration fédérale des finances, Bundesgasse 3, CH-3003 Berne
Amministrazione federale delle finanze, Bundesgasse 3, CH-3003 Berna
Swiss Federal Finance Administration, Bundesgasse 3, CH-3003 Bern

Doc. No.: ÖT/2003/4

Author: P.-A. Bruchez

Will the Swiss Fiscal Rule Lead to Stabilization of the Public Debt?

Working Paper – 26 August 2003

Abstract

The primary goal of the Swiss fiscal rule is to achieve a null average deficit (more exactly: the federal deficit as defined in the cash flow statement except that extraordinary expenditure and revenue are not taken into account). However, some concerns have been expressed that this rule will not yield a null average deficit (even by the narrow definition that we use here). The present paper computes a formula for the lower and upper bounds of the expected deficit which is valid under realistic assumptions and shows that the expected deficit is close enough to zero to conclude that the Swiss fiscal rule is well-conceived to attain its primary goal.

1. Introduction

Switzerland has recently introduced a fiscal rule aiming at stabilizing the level of the public federal debt[1] while avoiding a pro-cyclical fiscal policy. There have been some concerns[2] that the above fiscal rule may not lead to an average deficit equal to zero. This is important, because an average deficit far from zero would imply that the fiscal rule will not achieve its primary goal.

Computations with past data as well as with artificial data show that the mean deficit does not differ too much from zero[3]. These computations however are open to the criticism that future data may differ from past data, or that artificial data generated in a different way could yield a different result. Thus, while these computations give good reason to think that the mean deficit will not differ too much from zero, they do not offer a complete rigorous proof.

The present paper's aim is to compute lower and upper bounds to the expectation of the deficit. We will see that under plausible assumptions the expectation of the yearly deficit will be relatively small.

Section 2 will describe the fiscal rule in greater detail, and will indicate how a formula for the lower and upper bounds of the expectation of the deficit can be computed. Sections 3, 4 and 5 will move toward this formula in three successive steps. In section 6 some realistic numbers will be plugged into this formula in order to quantify these bounds. Section 7 will conclude.

---

[1] More precisely the fiscal rule applies to the cash flow statement (except that extraordinary expenditure and revenue are not taken into account). This implies that some important expenses such as those relative to the pension schemes of the employees of postal services and Federal administration are not taken into account.

[2] KOF (2003), « Gutachten zu ausgewählten Problemen der Schuldenbremse – Schlussbericht »

[3] Geier Alain (2003), « Application of the Swiss Fiscal Rule to Artificial Data », unpublished, Swiss Federal Administration.

2. The Swiss fiscal rule and the expected yearly fiscal deficit

After presenting the fiscal rule in greater detail (§2.1), we will indicate an approach for finding a formula for the lower and upper bounds of the expected deficit (§2.2).

2.1. The fiscal rule

According to the fiscal rule, an expenditure ceiling A_t is computed by the following formula: $A_t=k_t*R_t$, where R_t is the tax revenue at time t (forecasted at time t-1) and k_t is a coefficient reflecting business cycle conditions. This coefficient is equal to the real GDP's trend, y^T_t, divided by the real GDP y_t. During recessions k_t is larger than 1, and it is possible to spend more than the tax revenue. During booms k_t is smaller than 1 and the expenditure ceiling is lower than revenue. For the fiscal rule to attain its primary goal of null average deficit, the k coefficient has to be such that the average expenditure is equal to the average revenue.

y^T_t is computed by applying the Hodrick-Prescott (HP) filter with smoothing parameter=100 to the interval [1980 ; t+3] where the data from t and onward are expert forecasts. For reasons explained elsewhere[4] the Swiss Finance Administration proposes some small changes to this algorithm:

i) The filter will be applied to the rolling interval [t-23 ;t]

ii) The filter will be applied to $\ln(y_t)$ instead of y_t (thus $y^T_t \equiv e^{\text{trend of } \ln(yt)}$).

iii) The filter will be a modified HP filter for which y_t has less impact on y^T_t than with the HP filter (it is well known that the HP filter gives too much weight to the last data in calculating the trend; the modification tries to reduce this end point problem).

4 Bruchez Pierre-Alain (2003), « Réexamen du calcul du coefficient k », working paper, Swiss Federal Administration.

In the present paper we will focus on this new version of the rule. However, point (iii) will play no role in our argument (the only assumption concerning the filter will be that the trend at time t is computed as a weighted average of present or past values of the series, and that the filter is such that if the data follow a straight line then the trend is equal to the data).

2.2. An approach for computing the lower and upper bounds of the expected deficit

There is a rough argument which indicates that the expected value of the deficit should be null. But it suffers from caveats.

A rough argument for null average deficit

The trend computed by applying the HP filter ex-post passes through the middle of the data. More specifically: if the trend at time t is computed on the interval [a;b] with perfect knowledge of all data on [a;b] (and not only of previous data) then the sum of deviations from trend on [a;b] is null, that is $\sum_{t=a}^{b}(y_t - \mathrm{y}^{\mathrm{T}}{}_{\mathrm{t}}) = 0$, or $E(y_t - y^T{}_t)=0$ (where E is the expectation operator) which can be written $E(y_t)=E(y^T{}_t)$. Thus, $E(k_t) = E(\mathrm{y}^{\mathrm{T}}{}_{\mathrm{t}} / y_t) = E(\mathrm{y}^{\mathrm{T}}{}_{\mathrm{t}}) / E(y_t) = 1$ and $\mathrm{E(A)} = \mathrm{E(k * R)} = \mathrm{E(k)E(R)} = 1 * \mathrm{E(R)}$. This would imply E(A-R)=0. That is: the expected deficit is null.

Three caveats

There are three caveats to this argument:

i) The filter is not applied ex-post but recursively (that is, applied on a rolling interval). Thus it is not obvious that $E(y_t)=E(\mathrm{y}^{\mathrm{T}}{}_{\mathrm{t}})$.

ii) Even if $E(y_t)=E(y^T{}_t)$, it does not strictly follow that E(k)=1. The problem is that $E(y^T{}_t / y_t) \neq \dfrac{E(y^T{}_t)}{E(y_t)}$.

iii) Even if E(k)=1, it does not strictly follow that E(A-R)=0. The problem is that $E(k*R) \neq E(k)E(R)$.

Because of these caveats the above rough « proof » is false. However, these caveats can be fixed. Why? The basic intuition is that the rough argument would be correct if $y_t=y^T_t$ for all t. Thus, one can hope that it will be approximately correct for y_t not too far from y^T_t as is to be expected in the business cycle context where the output gap is small in comparison to output. The rest of the paper will present a rigorous proof along the lines of this rough argument. Besides more rigor, this will provide a quantitative statement about how far from zero the yearly expected deficit can be, and will make explicit the assumptions under which this result is correct (and we will see that these assumptions are realistic).

3. E[ln(k_t)]=0

The Hodrick Prescott filter is a common filter used to compute a trend. It is well known that when this filter is applied ex-post on an interval (that is: not only previous data but all data in the interval are known to compute the trend at any time in the interval), the sum (on that interval) of the gaps between actual data x_t and its trend x^T_t is exactly 0. Other filters share the same property[5]. The question arises whether this is still the case when the HP filter is applied recursively. Application of the following theorem yields $E[\ln(y_t) - \ln(y^T{}_t)] = 0$, and thus: E[ln($k_t$)]=0.

---

[5] Formally, it is easy to prove that this is the case for any (weighted) moving average filter if, when applied ex-post, it leaves constant data unchanged, and is such that the weight of x_m on the trend at time n is the same as the weight of x_n on the trend at time m for all m and n in the time interval on which the filter is applied ex-post (the modified HP filter does not satisfy this symmetry property exactly).

Theorem

If i) *the expectation of the first difference*[6] *of the actual data is constant*

ii) *the filter is a (weighted) moving average of the past and current data, and is such that the trend of a straight line is equal to this straight line.*

Then *the expectation of the gap between the actual data and the trend computed by recursive application of the filter is null.*

Proof

The proof uses the following lemma:

Lemma

If *the filter is a (weighted) moving average of the past and current data, and is such that the trend of a straight line is equal to this straight line,*

Then *the difference between the actual data and the trend can be written as a linear combination of the changes in the data from one period to the next, and the sum of these weights is null.*

Proof of the lemma

Let x_t be some data, and $x^T{}_t$ its recursive trend. Assume that $x^T{}_t = \sum_{j=0}^{N} w_j x_{t-j}$ and that the w_j are such that for $x_t=\alpha+\beta t$ the trend is $x^T{}_t=x_t$ (whatever α and β), then (the proof is given elsewhere[7]):

[6] Yearly changes, if the year is the time unit.

[7] Bruchez Pierre-Alain (2003), « A Modification of the HP Filter Aiming at Reducing the End-Point Bias », working paper, Swiss Federal Administration.

$x_t - x^T{}_t = \sum_{j=0}^{N-1}\left[\tilde{w}_j\left(x_{t-j} - x_{t-j-1}\right)\right]$, where $\tilde{w}_j = \sum_{i=j+1}^{N} w_i$, and has the property that $\sum_{j=0}^{N-1}\tilde{w}_j = 0$.

The proof of the theorem follows directly from the assumption that the expectation of the first difference of the actual data is constant which implies $E\left(x_{t-j} - x_{t-j-1}\right) = \Omega\,\forall t$ where Ω is a constant) and the lemma[8]. Formally:

$$\left.\begin{array}{l} x_t - x^T{}_t = \sum_{j=0}^{N-1}\left[\tilde{w}_j\left(x_{t-j} - x_{t-j-1}\right)\right](Lemma) \\ E\left(x_{t-j} - x_{t-j-1}\right) = \Omega\,\forall t\,(stationary\ first\ diff.) \end{array}\right\} \Rightarrow E(x_t - x^T{}_t) = \sum_{j=0}^{N-1}\left[\tilde{w}_j E\left(x_{t-j} - x_{t-j-1}\right)\right] = \Omega\sum_{j=0}^{N-1}\left[\tilde{w}_j\right] = 0$$

Discussion of the assumptions of the theorem

Assumption (ii) is a condition satisfied by the modified HP filter as well as by the usual HP filter (the HP filter as applied in the fiscal rule uses not only past and current data but also forecasts; but the theorem could be generalized to this case).

The only real assumption is that the first difference of $\ln(y_t / y^T{}_t)$ is stationary, which seems to be realistic.

---

[8] Note that if the first difference is not stationary, then the result does not hold. For example if the expected first difference is positive but decreasing, then the weights $\tilde{w}$ corresponding to recent expected first differences (these weights can be shown to be positive) will have less impact than would be the case with the stationary first difference, and weights corresponding to earlier expected first differences (these weights are negatives or close to zero) will have more weight. Thus, $E(x_t - x^T{}_t)$ will be negative in this case. It is easy to provide examples in which the trend is always above actual data.

4. E(k) $\approx$ 1+0.5*var[ln(k)]

In the preceding section we have proved that the expectation of the gap between the data and the trend is null. But since $k_t = y_t^T/y_t$ we need a statement about their ratio instead of their difference. The solution is to apply the filter on ln(GDP). Let $x_t \equiv \ln(y_t)$ and define $y^T{}_t \equiv e^{x^T{}_t}$, then according to the previous section: $E[\ln(y_t{}^T) - \ln(y_t)] = E[\ln(y_t{}^T / y_t)] = 0$. We are interested in $E[y_t{}^T / y_t] = E[e^{\ln(y_t{}^T / y_t)}]$. Notice that, since in general $E[e^x] \neq e^{E(x)}$, we cannot conclude directly that $E[k_t] = 1$. The following theorem gives the approximation $E(k) \approx 1 + \frac{Var[\ln(k)]}{2}$.

Theorem

If *i)* $E[\ln(k_t)] = 0$

ii) the density of k is such that $\sum_{j=3} \frac{E[\ln^j(k)]}{j!} << 1 + \frac{Var[\ln(k)]}{2}$

Then $E(k) \approx 1 + \frac{Var(k)}{2}$

Proof

The following proof is written for a random variable x instead of ln(k) to stress the fact that it does not depend on the definition of k as long as the assumptions of the theorem are satisfied.

$$E(e^x) = E(\sum_{j=0}^{\infty} \frac{x^j}{j!}) = E(1 + x + \frac{x^2}{2!} + \sum_{j=3}^{\infty} \frac{x^j}{j!}) = 1 + E(x) + \frac{E(x^2)}{2!} + \sum_{j=3}^{\infty} \frac{E(x^j)}{j!}$$
$$= 1 + E(x) + \frac{E(x^2) - E^2(x)}{2!} + \frac{E^2(x)}{2!} + \sum_{j=3}^{\infty} \frac{E(x^j)}{j!}$$

$$= 1 + E(x) + \frac{Var(x)}{2!} + \frac{E^2(x)}{2!} + \sum_{j=3}^{\infty} \frac{E(x^j)}{j!}$$

$$= 1 + \frac{Var(x)}{2!} + \sum_{j=3}^{\infty} \frac{E(x^j)}{j!} \quad if \ E(x) = 0$$

$$\approx 1 + \frac{Var(x)}{2!} \qquad if \sum_{j=3}^{\infty} \frac{E(x^j)}{j!} << 1 + \frac{Var(x)}{2!}$$

For the rest of the paper, σ^2 will stand for Var[ln(k)].

Discussion of the assumptions of the theorem

Assumption (i) is satisfied if the assumptions of the theorem in §3 are satisfied.

Assumption (ii) is satisfied when x is always close enough to 0. In this case $\sum_{j=3}^{\infty} \frac{x^j}{j!}$ is small, and so must be its expectation (notice that for x=0.1, $\sum_{j=3}^{\infty} \frac{x^j}{j!} < 2*10^{-4}$). But even when x can take arbitrarily large values this assumption is still satisfied as long as the probability of getting large values decreases fast enough. This is for example the case if ln(k) is normally distributed. The appendix gives examples of common density functions satisfying assumption (ii), and presents a density, which does not satisfy this assumption (but this density does not seem pertinent for business cycle analysis).

5. Lower and upper bounds for $\frac{E(kR-R)}{E(R)}$

Knowing that $E[k_t] \approx 1 + \frac{\sigma^2}{2}$ we may want to conclude that:

$$\frac{E(kR-R)}{E(R)} = \frac{E(k)E(R) - E(R)}{E(R)} = E(k) - 1 = (1 + \frac{\sigma^2}{2}) - 1 = \frac{\sigma^2}{2}$$

If σ is small, this would imply that the expected deficit as a fraction of the expected revenues is even much smaller. But there is a problem with this computation: $E(k*R) \neq E(k)E(R)$. However, there are limits to how much different E(k*R) can be from E(k)E(R). The following theorem uses the fact that the correlation between k and R is larger or equal to –1 (like any correlation coefficient) and smaller or equal to 0 (a positive correlation would imply that the tax revenue is larger during recessions) in order to get lower and upper bounds on the expected deficit.

Theorem

If *i) the correlation between k and R is not positive*

Then $$[1-\frac{\sigma(k)}{E(k)}\frac{\sigma(R)}{E(R)}]E(k)-1 \leq \frac{E(kR-R)}{E(R)} \leq E(k)-1$$

If moreover

ii) the first difference of ln(k) is stationary

iii) the probability that k takes values far from 1 goes sufficiently quickly toward 0 (as the distance to 1 increases) for the following to hold:

iii.a) $$\sum_{j=3}\frac{E[\ln^j(k)]}{j!} << 1+\frac{Var[\ln(k)]}{2}$$

iii.b) $$Var[k] \approx Var[\ln(k)]$$

Then $$\boxed{-\frac{\sigma^2}{2}-\sigma[\frac{\sigma(R)}{E(R)}-\sigma] \leq \frac{E(kR-R)}{E(R)} \leq \frac{\sigma^2}{2}} \qquad (1)$$

where σ^2=Var[ln(k)]

Proof

The correlation between k and R, like any correlation coefficient, is between –1 and 1:

$$-1 \leq \frac{E(k*R) - E(k)E(R)}{\sigma(k)\sigma(R)} \leq 1$$

Where σ(k) and σ(R) are the standard deviations of k and R.

Using simple algebra this implies that:

$$[1 - \frac{\sigma(k)}{E(k)}\frac{\sigma(R)}{E(R)}]E(k) - 1 \leq \frac{E(kR - R)}{E(R)} \leq [1 + \frac{\sigma(k)}{E(k)}\frac{\sigma(R)}{E(R)}]E(k) - 1$$

If assumption (i) is satisfied, we can even be more restrictive:

$$[1 - \frac{\sigma(k)}{E(k)}\frac{\sigma(R)}{E(R)}]E(k) - 1 \leq \frac{E(kR - R)}{E(R)} \leq E(k) - 1$$

If assumption (ii) is satisfied, then §3 has shown that E[ln(k)]=0. If assumption (iii.a) is also satisfied, then §4 has shown that $E(k) \approx 1 + \frac{Var[\ln(k)]}{2}$. If assumption (iii.b) is also satisfied, then the above inequalities become:

$$-\frac{\sigma^2}{2} - \sigma[\frac{\sigma(R)}{E(R)} - \sigma] \leq \frac{E(kR - R)}{E(R)} \leq \frac{\sigma^2}{2}$$

where σ is the standard deviation of ln(k).

Discussion of the assumptions of the theorem

Assumption (i) should be satisfied since a positive correlation between k and R would imply that the tax revenue is larger during recessions (remember that k is small during booms).

Assumption (ii) as already been discussed (§3) and seems realistic.

Assumption (iii.a) has also already been discussed (§4) and is satisfied for common density functions.

Assumption (iii.b) could be derived by assuming that the values that ln(k) can take are close enough to zero for the linear approximation of the exponential around zero to be sufficient. But we want to stick with the second order approximation of the exponential. The reason we need to be so precise is that we have [E(k)-1]*E(R) as upper bound for the expected fiscal deficit and that E(R) is very high ($\approx$Sfr 50 billion). If we linearize the exponential we get $E(k) \approx 1$. But a small difference between E(k) and 1 already has a sizeable impact on the upper bound (for example E(k)-1=10^{-3} will imply an upper bound of Sfr 50 million, which cannot be neglected). With our second order approximation we get E(k)-1=σ^2 which is of order 10^{-4} if σ is of order 10^{-2}, which is satisfying. Thus we do not have a formal derivation of the (iii.b) assumption from a weaker (or less obscure) assumption. However, the appendix shows for three common density functions that $Var[e^x] \approx Var(x) * [1 + \gamma Var(x)]$ with γ not too large compared to 1. In such a case, the term $Var^2(x)$ can be neglected.

Further comment

1) The algebraic formula (1) is independent of the GDP elasticity of R, but the value of the parameter $\frac{\sigma(R)}{E(R)}$, and thus the value of the lower bound, may depend on this elasticity.

2) For $\frac{\sigma(R)}{E(R)} \approx \sigma$, the inequalities $-\frac{\sigma^2}{2} - \sigma[\frac{\sigma(R)}{E(R)} - \sigma] \leq \frac{E(kR-R)}{E(R)} \leq \frac{\sigma^2}{2}$ are symmetrical: $-\frac{\sigma^2}{2} \leq \frac{E(kR-R)}{E(R)} \leq \frac{\sigma^2}{2}$

6. Numerical example

In order to compute the lower and upper bounds we have to plug values for σ and $\frac{\sigma(R)}{E(R)}$ into $-\frac{\sigma^2}{2} - \sigma[\frac{\sigma(R)}{E(R)} - \sigma] \leq \frac{E(kR-R)}{E(R)} \leq \frac{\sigma^2}{2}$. We will first propose some numbers, then give the tools for a sensibility analysis.

The expected yearly fiscal balance will be a smaller deficit than Sfr 40 million and a smaller surplus than Sfr 80 million

Strictly speaking, we would need several occurrences of the data for a given year in order to compute σ and $\frac{\sigma(R)}{E(R)}$. Since we do not have this, we will use data for different years instead. Recursive application of the modified HP(100) filter on ln(GDP) in interval [t-23;t] for t=1971 to 2002 yields σ smaller than 4%. Computation of $\frac{\sigma(R)}{E(R)}$ is a bit more complicated since R is not stationary. Computing on the interval [1950:2002] σ(R) as the standard deviation of deviation from trend, and dividing by the average of R, gives for $\frac{\sigma(R)}{E(R)}$ a value smaller than 6%[9].

Assuming σ=4% and $\frac{\sigma(R)}{E(R)}$=6% yields:

$$-16*10^{-4} \leq \frac{E(kR-R)}{E(R)} \leq 8*10^{-4}$$

Assuming that E(R) is around Sfr 50 billion, it follows that:

$$\boxed{-80\,million \leq E(kR-R) \leq 40\,million}$$

Thus the expectation of the deficits (which should be understood as an average over a long period) is relatively close to 0. However, in any given year the deficit might be much larger (we have not computed the variance of the deficit).

Notice that although we have used some data, our result is much more robust than what we would get by simply computing the average deficit with historical (or artificial) data. We have found that the details of the data is not important: what

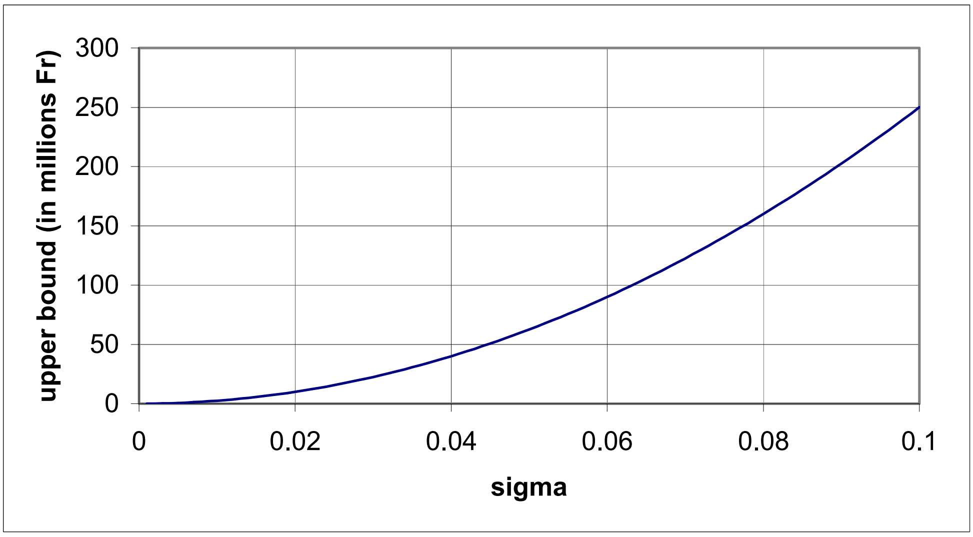
[9] Since this value is computed with the actual data for R, it indirectly takes account of the GDP elasticity of R (whatever the actual value of this elasticity may be).

counts is only σ and $\frac{\sigma(R)}{E(R)}$. We can also roughly assess the reliability of the numbers we have chosen. For example σ=4% means approximately that the average of the absolute value of the output gap is 4% of GDP, which seems reasonable.

Sensibility analysis

The following graphs present upper and lower bounds of the expected yearly deficit for various values of σ and $\frac{\sigma(R)}{E(R)}$.

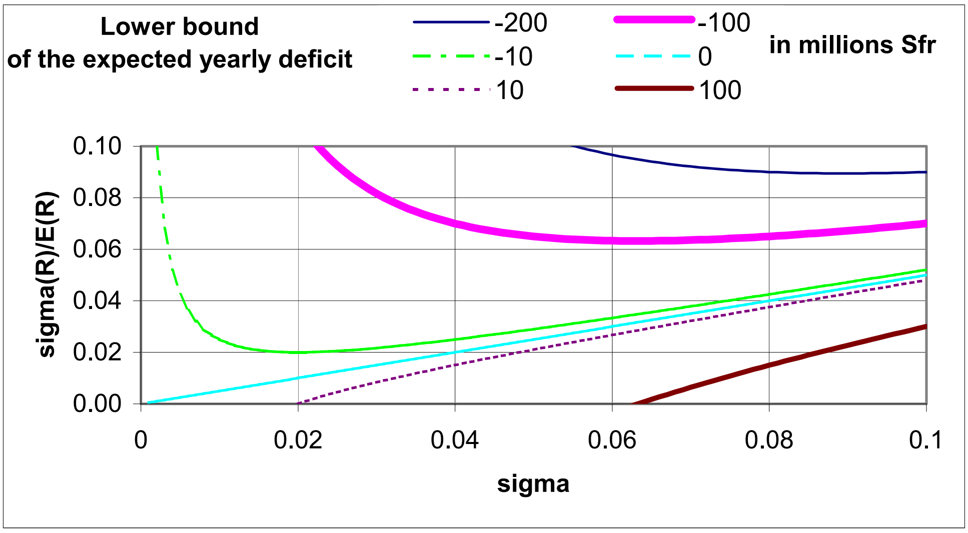
Upper bound of the expected yearly deficits in millions Sfr



The expected yearly deficit is smaller than Sfr 100 million if sigma is smaller than 6%.

Lower bound of the expected yearly deficits in millions Sfr

In the parameter space (σ;$\frac{\sigma(R)}{E(R)}$), the following graph presents lines made of all the points leading to the same expected deficit.



The lower bound is not a larger yearly deficit or surplus than Sfr 100 million if both σ and $\frac{\sigma(R)}{E(R)}$ are smaller than 6%.

7. Conclusion

The present paper has shown some theoretical results. In particular, expected deviation from trend is zero even when the trend is computed recursively (as long as the first difference of the data is stationary). We have also found lower and upper bounds of the expected yearly deficit. Under realistic assumptions these bounds are close enough to zero to conclude that the Swiss fiscal rule is well-conceived in this respect to attain its primary goal: long term stabilization of the debt level.

A natural continuation of this work could be to compute the lower and upper bounds for the variance of the deficit in order to know how far from its expectation this deficit is likely to be.

Appendix E[e^x] and Var[e^x]

$E[e^x] \approx (1+\frac{Var(x)}{2})$ and $Var[e^x] \approx Var(x)$ when E(x)=0 and Var(x) is small

If f and the density function of x are known, then E[f(x)] and Var[f(x)] can be computed exactly. For example for f=e^x we have (an approximation for σ small is also computed):

i) if half the values of x are at μ-δ and half at μ+δ, then δ is in fact the standard deviation σ of x, and we have

$$E[e^x] = \frac{e^{\mu-\sigma}+e^{\mu+\sigma}}{2} = \frac{e^{\mu}(e^{-\sigma}+e^{+\sigma})}{2} = e^{\mu}\sum_{j=0}^{\infty}\frac{\sigma^{2j}}{(2j)!} \approx e^{\mu}(1+\frac{\sigma^2}{2})$$

$$Var[e^x] = E(e^{2x}) - E^2(x) = e^{2\mu}\left[\frac{e^{2(-\sigma)}+e^{2(+\sigma)}}{4} - \frac{1}{2}\right] \approx e^{2\mu}\sigma^2(1+\frac{1}{3}\sigma^2)$$

ii) if x is uniformly distributed on [μ-δ;μ +δ], its variance is $\sigma^2=\delta^2/3$ and

$$E[e^x] = \frac{1}{2\delta}\left(e^{\mu+\delta}-e^{\mu-\delta}\right) = e^{\mu}\,\frac{e^{\delta}-e^{-\delta}}{2\delta} = e^{\mu}\sum_{j=0}^{\infty}\frac{\delta^{2j}}{(2j+1)!} \approx e^{\mu}(1+\frac{\delta^2}{6}) = e^{\mu}(1+\frac{\sigma^2}{2})$$

$$Var[e^x] = E(e^{2x}) - E^2(x) = \frac{e^{2\mu}}{(2\delta)^2}\left[\delta(e^{2\delta}-e^{-2\delta})-(e^{2\delta}+e^{-2\delta})+2\right] \approx e^{2\mu}\sigma^2(1+\frac{4}{5}\sigma^2)$$

iii) if x follows a N(μ,σ^2), then e^x follows a lognormal and

$$E[e^x] = e^{\mu+\sigma^2/2} \approx e^{\mu}(1+\frac{\sigma^2}{2})$$

$$Var[e^x] = (e^{\sigma^2}-1)e^{2\mu+\sigma^2} \approx e^{2\mu}\sigma^2(1+\frac{3}{2}\sigma^2)$$

Thus for f(x)=e^x and σ small, the approximation $E[e^x] \approx e^{\mu}(1+\frac{\sigma^2}{2})$ seems to be valid for a wide variety of density functions. Moreover, if we neglect the terms of order σ^4 or higher we get $Var[e^x] \approx e^{2\mu}\sigma^2$. On this basis we will assume that $E[e^x] \approx (1+\frac{Var(x)}{2})$ and $Var[e^x] \approx Var(x)$ when E(x)=0 and Var(x) is small. We have rigorously proved these approximations for three specific density functions only. We conjecture however that it will still be true for a wide variety of density functions[10].

Upper bound for E[f(x)] and Var[f(x)] when f is convex

The following lemma illustrates a direction that could be taken in order to compute bounds that could be proven to be robust for a wide variety of density functions.

Let f be a convex function. Then Jensen's inequality gives a lower bound of $E[f(x)]$: $E[f(x)] \geq f(E(x))$. Let's find an upper bound.

Call $\mu \equiv E(x)$. Assume f to be twice differentiable and defined everywhere on [μ-δ;μ+δ] where δ is a constant. If x takes all its values on [μ-δ;μ +δ] then it can be shown[11] that E[f(x)] will be the largest if half the values of x are at μ- δ and the other half at μ + δ. Thus:

$$\frac{f(\mu-\delta)+f(\mu+\delta)}{2} \geq E[f(x)]$$

---

10 It is possible to construct a density function for which these approximations are not correct. For example, it is possible to get an arbitrary large value for E[e^x] and Var[e^x] if the weight of x is sufficiently focused on x=μ with an arbitrary small mass being arbitrarily far from μ. This is however a very special density function according to which the x values are constant except in rare circumstances where the deviation is arbitrarily large. This does not seem relevant for business cycle analysis.

11 Think of density as masses placed on a lever (the density function corresponds to the linear density of these masses). A mean preserving transformation of this density would keep the center of gravity at the same place. Thus any movement of one mass to the right must be compensated by corresponding movements of mass(es) to the left. Noticing that by assumption the first derivative of f is increasing, it can easily be seen that any decrease of E[f(x)] due to movements to the left is more than compensated by the increase due to movement to the right. Thus, this transformation of the density increases E[f(x)]. This transformation can be repeated until the density given by probability ½ for x to take value μ- δ and ½ to take μ+ δ is attained.

For the special case of $f(x)=e^x$,

$$\frac{f(\mu-\delta)+f(\mu+\delta)}{2}=\frac{e^{\mu-\delta}+e^{\mu+\delta}}{2}=e^{\mu}\,\frac{(e^{-\delta}+e^{\delta})}{2}=e^{\mu}\sum_{j=0}^{\infty}\frac{\delta^{2j}}{(2j)!}\approx e^{\mu}(1+\frac{\delta^2}{2})$$

The approximation is valid for δ small, whatever the density function as long as its support is included in [μ-δ;μ +δ].

It can also be shown that this distribution gives the maximal variance (among densities defined on a closed interval with expectation in the middle of the interval, the one which has all weights equally distributed on the two extremities has maximal variance; this can be proved by applying the previous result to $f(x)=x^2$).

Thus:

$$[\frac{f(\mu-\delta)-f(\mu+\delta)}{2}]^2 \geq Var[f(x)]$$

For the special case of $f(x)=e^x$:

$$[\frac{f(\mu-\delta)-f(\mu+\delta)}{2}]^2 = e^{2\mu}[\frac{e^{-\delta}-e^{\delta}}{2}]^2 \approx e^{2\mu}\delta^2$$