Evaluation of Skattefunn. Referee report

Erik Fjærli,

07.01.2018

Skattefunn is meant to be a universal, rights-based and low-threshold tax-incentive scheme for stimulation of R&D in the business sector. It is designed to decrease the costs of R& D investments. According to the commission from the Ministry of finance, the evaluation of Skattefunn by Samfunnsøkonomisk Analyse AS shall answer the following questions:

1. Is the aid measure aimed at a well-defined objective of common interest?
2. Is the aid well designed to deliver the objective of common?
3. Are the distortions of competition and effect on trade limited?

The report has 10 chapters, and its conclusions very much rests on chapters 4, 5, 8 and 9. My comments are organized according to these three questions above.

Question 1. Is the aid measure aimed at a well-defined objective of common interest?

Chapter two in the report provides descriptive statistics and self-reported measures from survey data that support the view that Skattefunn indeed targets the intended goals.

The report documents that the Skattefunn scheme is widely utilized and has become the largest public support scheme for private R&D investments in Norway. The report further argues that Skattefunn addresses a well-defined objective of common interest.

I agree that Skattefunn in a broad sense can be said to be aimed at an objective of common interest (to increase R&D in the business sector). However, it is questionable how well-defined this objective is. The authors point out that "Because of the various positive externalities following investments in R&D, these investments conducted by the business sector will be less than what is socially optimal." (My underlining). This is elaborated in chapter 2.1:

"(…) markets left on their own will probably generate less innovation than would be desirable from society's point of view. This could be due to risk aversion and the fact that knowledge is not completely excludable. R&D investments are subject to the classical free rider problem of public goods. R&D investments are characterized by high risk and having high start-up costs, but relatively low marginal costs [Referee's comment: I suppose the authors meant variable costs]. Typically, these investments also face a low probability of success, but provide significant proceeds if succeeding.

When a firm succeeds in developing ideas, these can be easily copied and utilized by other firms. Non-excludability discourages firms to invest in research since the returns to investment will not entirely accrue to the firm. It is therefore well-recognized in economic literature that the broader economic effect of R&D investments exceeds the private economic effects (Arrow, 1962).

Furthermore, it is often difficult for firms to obtain funding for innovation projects in the private market. The information possessed by the firm and the investor is typically highly asymmetric, causing higher risk."

So, the report points at three different reasons that could justify public intervention: Knowledge spillovers, "risk aversion" and capital constraints caused by asymmetric information.

> **Comment:** Even if it is a frequently used argument, I disagree that the term *risk aversion* should be used to describe a market failure or as a problem by itself. In short; risk is a bad thing (for a given expected return, safer projects should be preferred over riskier projects) and risk aversion is not a bad thing in itself. Risk and risk aversion boils down to the question

of pricing risk correctly, i.e., determining the market-clearing risk premium. With well-functioning markets for risk sharing, public policy should neither encourage nor discourage private risk-taking, and a modern economy like Norway has rich opportunities for diversification and risk-sharing. What *can* be a problem for firms, however, is under-diversification caused by capital constraints (Leland, and Pyle, 1977). Bittler et al. (2005) show that active owners typically may have a large portion of wealth tied up in their company. Due to a poorly diversified portfolio they may seek to counteract this risk exposure by choosing a more cautious investment profile. Caggese (2012) finds similar results.

Hall and Lerner (2010) point to a similar effect due to moral hazard that arises when the goals of the management are not aligned with those of the owners. In the context of R&D and innovation, moral hazard problems may manifest themselves as the management of a company having strong aversion to default risk and being reluctant to invest in long-term, risky development projects - even when this is in the owners' interest.

In conclusion, there are three possible market failures in question: (1) externalities like knowledge spillovers, (2) capital constraints and (3) prudence, that is, the management or entrepreneurs are unwilling to accept (more) risk or are unable to diversify risk, both leading to prudent behavior.

As stated in the report's chapter 1.1.1: "To contribute to a common objective, the scheme must address a market failure" (p.1) and all three types of market failure may justify the use of tax payer's money on R&D subsidies. However, the task of aiming at "an objective of common interest" is not that well-defined, as we have *different* market failures that likely each require a *different* measure to be dealt with efficiently. It is not clear that a subsidy like Skattefunn is equally well suited for all types of market failure listed in the report. For example, as shown by Stiglitz and Weiss (1981), a risky loan can stimulate firms to undertake more risky projects. By contrast, a subsidy in the form of a grant (like Skattefunn) does not affect the choice between projects with different risks but is well suited to reduce the wedge between social and private returns to R&D.

I would like to see a more critical and stringent discussion of the arguments underlying public R&D support, including

1. Which of the market failures are most likely to be in effect among the Skattefunn clientele? Are there indications that the liquidity aspect is very important for many Skattefunn users? And previous evaluations suggest that Skattefunn-innovations typically are "new to the firm, but not new to the market" (i.e. small spillover effects)? Could the authors elaborate a bit on the market failure question? Provide some descriptive statistics on financial strength variables, use references to statements from Skattefunn users (the surveys) and the recommendations in NOU 2000:7? Etc.
2. Is it an established fact that it is so hard to find private funding for R&D? I would like to see references to empirical studies of the significance of capital rationing for entrepreneurs in general and particularly for firms doing intangible R&D investments.
3. Spillover effects as discussed in Arrow (1962) (and others), are well known *theoretical* results, but I would like to see references to empirical studies related to the gap between social and private return on R&D and knowledge spillover effects.
4. Skattefunn has built in a strong incentive for cooperation with R&D institutions. By intuition, one might think that such cooperation ensures spillover effects beyond effects in single (and small?) firms operating alone. However, this topic appears to be overlooked by the evaluators. As far as I remember, the Skattefunn administration has information on "Skattefunn-clusters", like the SINTEF-NTNU environment. In my opinion, the evaluation of additionality would benefit from a discussion of benefits from research cooperation through Skattefunn, as well as empirical assessments of its importance for activities in universities and research institutes. (This could possibly be done as a part of chapter 4). I'm aware that cooperation with universities and with other firms is examined in chapter 5, tables 5.2 – 5.4 (why not in chapter 4, too?), but only in terms of effects within the Skattefunn firm. The evaluation would benefit from a look at input- and in particular output additionality in the university sector.

Point 2 and 3 are particularly important when seen in the context of rent-seeking. Both the recipients and administrators of R&D subsidies may have common interest in the maintaining of various R&D support schemes. Therefore, the benefit to the society of R&D subsidies must be reasonable and well documented. Referring to the first paragraph on p.5 (ch 2.1), the fact that "everybody else does it" is of course not a sufficient argument for public R&D support.

Useful early references are Cohen et al (1987), Zoltan and Audretsch (1988) and Cohen et al (1989). Furthermore, Lerner (2009), chapter three, discusses and emphasizes the role of new, small firms for innovation (as opposed to Schumpeter's view that large firms have an advantage in innovation).

Question 2: Is the aid well designed to deliver the objective of common?

Chapter 4 in the report examines how Skattefunn affects R&D (input additionality) while chapter 5 analyzes how R&D intensity affects innovation (both for Skattefunn firms and non-Skattefunn firms) and how R&D capital affects productivity (for Skattefunn firms, firms with other types of support and others). Chapter 6 examines what types of R&D is stimulated by Skattefunn.

The evaluation report is well structured and can almost be regarded as a CDM-inspired approach where R&D activities are described by a sequential model in which some exogenous factors, *X,* affect

the probability of undertaking R&D (as well as its scale), which (with some probability) leads to innovation and in turn increased productivity (Crepon *et al.* ,1998). Below I have some specific comments and questions to the chosen methodology.

Chapter 4 Input additionality of Skattefunn

The input additionality analysis uses similar methods as previous studies by Statistics Norway (Hægeland and Møen, 2007, henceforth HM2007) and Mohnen et al. (2017) and thus must be said to be knowledge-based and solidly rooted in state-of-the-art methodology. The analysis also expands the Mohnen approach, by combining difference-in-difference (DID) with matching. These are well-known methods, too, and similar to the approach used by Statistics Norway (2016).

Chapter 4.3 Estimation of input additionality by discontinuity approach

The evaluation report uses "(…) a 'discontinuity approach', but instead of looking at introduction of the scheme, we evaluate effect of an *increase in the limit* of tax-deductible R&D expenditures in 2009" (p. 36). This approach gives positive, but lower results than HM2007. The report claims that this confirms that only firms with R&D expenditures under the cap are stimulated to do more R&D.

> **Comment:** When using a "discontinuity approach" to evaluate effects, one usually tries to secure comparability of the control- and treatment groups by narrowing the sample down to firms right above and right below the point of discontinuity or cap in the scheme. This will normally imply a trade off since it may cause a loss of observations. Besides, as pointed out by HM2007, the more one narrows the sample to improve the comparability of the control and treatment groups, the more likely it is that firms are misclassified.
>
> In HM2007 the Skattefunn cap is 4 million and the authors compare firms above and below the cap. Using DID, HM2007 find that firms below the cap responds positively to the introduction of Skattefunn, compared to firms above the cap.
>
> HM2007 examines the implications of their wide bandwidth (R&D expenses ranging from 0 to 4 million), using sensitivity analysis as shown by their tables 6.10 and 6.11 (pp. 37-38). They find that their results are driven by firms with R&D well below the cap (less than 1 million). Since HM2007 examines the effect of the *introduction* of the scheme, their findings do not weaken the result that the introduction of the scheme had an additionality effect, but it may perhaps raise some doubt about the size of their estimated "bang-for the-buck" (BFTB).
>
> The bandwidth trade off problem is more serious in the analysis by Samfunnsøkonomisk Analyse, since firms with very low R&D are irrelevant for the *marginal effect of increasing* the CAP (as opposed to HM2007, which looks at the introduction of the scheme). Strictly speaking, the expansion of the cap from 4 to 5.5 million does not change marginal incentives for firms with R&D below 4 million.

The increase in the cap gives an incentive for increased R&D after 2009 only if initial R&D is between 4 and 5.5 million.
I would like to see regressions like the two reported in table 4.5, column 2 and 4, but replacing the dummy for "below 5.5 million" with a dummy for X < R&D < 5.5 million, with trial X-values from 4 down to 3, 2, 1. Also, I would like to see regressions similar to table 6.10 and 6.11 in HM2007. If these regressions turn out to show the same results as HM2007 (i.e. if the effects are driven by firms with initially very low R&D), then the conclusions in the current version of the report should be modified.

(This is a rather important issue, since it concerns the general question of additionality of the Skattefunn scheme, in addition to the specific question of the benefit of the expansions of the scheme after 2009.)

Other comments:

- Figure 4.3 show a sharp *reduction* in R&D expenditures for firms above the cap in 2009 (caused by the financial crisis?). How could this affect your regression results? Are the results driven by reductions in R&D in the control group above the cap rather than increases in the treatment group?
- I find the discussion in chapter 4.1 (Self-reported input additionality) and 4.2 (Data on R&D expenditures) a bit difficult to follow. To clarify the different data sources, the meaning of the variables and their uses it would help to see a table with variable names, their definitions and where they come from.

Chapter 4.4 Estimation of input additionality by generalized difference-in-difference approach combined with matching

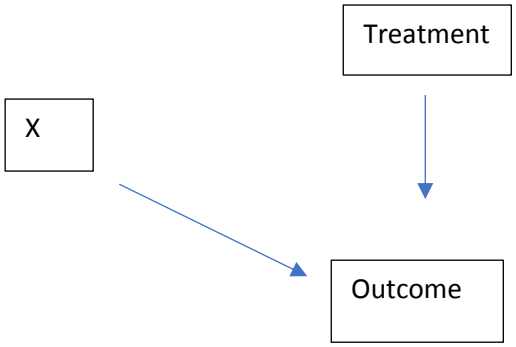
The DID approach shows that the use of Skattefunn is associated with increased R&D compared to a control group, which is reassuring. However, from there to concluding that there is a causal relationship is not obvious, and the methodology used in the report is not fully convincing. The cause of my doubts is the fact that Skattefunn is a universal, rights-based and low-threshold R&D subsidy. Why all firms that do R&D don't make use of it is a puzzle. One explanation could be that there are fixed application costs, implying that firms with constantly low R&D in nominal terms don't find it worthwhile to apply for Skattefunn while, on the other hand, firms with higher (and possibly growing) R&D expenses may use it more regularly. This could possibly produce similar results as the findings in the report, even if there were no additionality. However, even with this modification there should be little doubt that Skattefunn benefits the "right" type of firms, i.e. firms that indeed have a high degree of R&D (chapter 4) and innovations (chapter 5).

I also have some comments to the report's use of propensity score matching, which I find unsatisfactory motivated and a bit ad hoc.

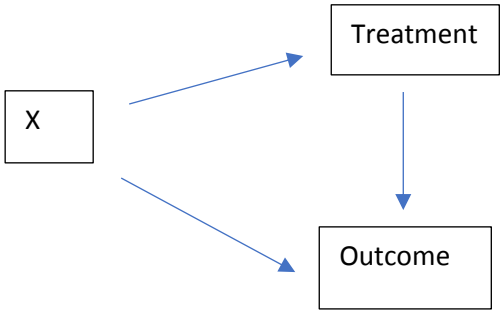
1. Why use (propensity score) matching (PSM)? The report states that it uses matching "to construct the control group of firms that is as much as possible comparable with Skattefunn firms (given observable characteristics)." This is at best an incomplete motivation for PSM, as multivariate models normally are sufficient to deal with heterogeneous and uneven distribution of covariates, if this is the only problem. The advantage of PSM is to control for *selection bias* that may occur when some (unobservable) factor affects both the selection into treatment and the outcome, which makes it difficult to identify the effect of the treatment, as shown in fig. 1 below.

Fig 1.

Multivariate models



Exact matching
Propensity score matching



Is there a realistic concern for this analysis that firms that apply for Skattefunn also are better innovators? Why? One question which may arise when using regression adjustment with propensity scores, is what is gained by using PSM rather than performing a regression adjustment with all the covariates used to estimate the propensity score included in the model. Another question is related to the choice of matching variables.

2. Choice of matching variables: When choosing Variables for Propensity Scores, one should

• include: Variables that are theoretically related to treatment and outcome and correlated with unmeasured confounders and

• not include: Variables hypothesized to be associated with treatment but *not* with outcome or variables that may be affected by the treatment.

The matching procedure used in the report does consider the importance of using matching variables that are independent of treatment (i.e., predetermined), but fails to fulfill the other requirements: There is no hypothesizing around the theoretical relation between the matching variables, the likelihood of treatment and the outcome, and there is no discussion around what the unmeasured confounders might be.

Also, the results from the matching regression are not shown, nor are the estimated effects of the matching variables on the outcome in the main regression table 4.9 (“in order to save space”, p. 58). The latter is particularly inadequate.

3. Another question related to the use of PSM is whether we might *lose* something when using the propensity score: Table 4.7 and 4.8 show that the matching implies that 44 percent, respective 58 percent of the Skattefunn firms are omitted from the regressions, and 94, respective 95 percent of the not-Skattefunn firms are omitted. This low range of common support means loss of information and it may cause bias.

Indeed, taking a closer look at the tables 4.7 and 4.8, we see that large Skattefunn-firms are selected out of the analysis. This is likely to cause an upward bias in the effect estimates, since the DID's are measured as percentages and large firms should be expected to have high initial R&D compared to the Skattefunn cap. (Inter alia, this is one reason why it is important to see the estimated coefficients for the matching variables used in the main regressions underlying table 4.9).

4. Tables 4.7 and 4.8 reveal another interesting fact, not commented by the authors: Skattefunn firms are on average larger, have a much higher share of high-skilled workers and are less cash-rich than other firms. In my view, the report would gain from a closer analysis of what determines use of Skattefunn. Indeed, as mentioned earlier, why shouldn't all firms that do R&D apply for Skattefunn-support? After all this is a universal, right-based and low-threshold support scheme? Furthermore; the fact that the Skattefunn-firms have a high share of high-skilled workers could be taken as circumstantial evidence of additionality in the sense that the money is not wasted on firms with low R&D abilities.

5. Panel data regressions reported in table 4.9; which method is used? (Still using fixed effects? Stata XTREG procedure?)

6. Finally (and apart from the question of using PSM or not): The analysis in chapter 4.4 follows closely the procedure in Cappelen et al (2016). However, Cappelen et al study schemes that are not universal (apart from Skattefunn), while this report considers Skattefunn only. As pointed out by HM2007, with the scheme being universal it is very challenging to construct a valid control group:

"Schemes and measures are often general in nature. This creates great challenges for evaluators. The more general the schemes, i.e. the more equally similar firms are treated in the scheme, the more complicated is the evaluation. The reason is that a higher degree of "generality" or "equal treatment" brings us further away from the ideal evaluation setting. When all comparable firms either have access to the scheme or not, it is impossible to construct a control group telling us anything about the counterfactual situation." (HM2007, p. 17)

As mentioned earlier it remains an open question why some firms do R&D without making use of Skattefunn. It would be interesting to see a regression with reversed causality, analyzing the decision to apply for Skattefunn, checking if fixed application costs and decreasing average cost of Skattefunn-participation could drive low-R&D firms out.

Chapter 5 Output-additionality of Skattefunn

Tables 5.2 – 5.4 summarize the results of effects on innovations. An obvious problem is the fact that these are self-reported data, but still I believe that the significant effect of the variable **d_SKF** (Skattefunn participation in each sub-period) is a strong and confident result.

> **Comment:** The overall impression I have from the regression results is this: The sample consists of two groups: One with low R&D intensity and little innovations and another with relatively high and persistent R&D activity and who (because of this) uses Skattefunn often (more than once during the entire observation period). This is indicated by the fact that $\mathbf{Y_{t-1}}$ is in the regressions serve as a substitute for the variable **SKF-firm** and the variable "**Additional R&D**" being a substitute for the variable **d_SKF**. In the light of this (or interpreting data in this manner), it is not surprising that the treatment group shows a higher propensity to innovate (also see my comments regarding PSM in chapter 5).
>
> I will emphasize that my caution regarding identification of causal effects in chapters 4 and 5 does *not* necessarily mean that the report is too benevolent when answering the evaluation question 2. However, the authors perhaps make it a bit difficult for themselves when focusing so much on proving causal effects. The report could possibly give a bit weaker (and more credible) "yes" to the question whether the scheme is well-designed by also pointing out that the users of Skattefunn indeed are R&D-intensive, do innovate and indeed face significantly lower R&D costs when they receive Skattefunn subsidies.
>
> **Other comments:** I believe that a reference to Klemetsen et al (2016) would be appropriate in chapter 5.1.3.

Chapter 8 Misuse of Skattefunn

The evaluators have analyzed the potential for misuse of SkatteFUNN based on selected indicators in collaboration with the Tax Administration. They find examples of what can be interpreted as tax motivated misuse of the scheme, but conclude that there is no reason to believe that it is extensive.

I notice that figure 8.5 in the report (p.114) show a sharp decline in the ratios of R&D personnel costs to total wage costs after 2007, when the templates for wage costs and ceilings for total annual hour per employee was introduced (presumably due to the advices from the previous evaluation by Statistics Norway and the findings in Fjærli, 2007). From 2009, it increases again, in line with the increases in the Skattefunn cap.

> **Comment:** Would it be possible to provide a table or graph with the cost per R&D employee or per total employees in Skattefunn firms over the same period? This could give the Tax authorities and the Ministry of Finance a clue regarding the effects of introducing such templates and limitations to project costs.

According to the report, the Tax administration has conducted an analysis of 20 randomly selected users of SkatteFUNN with relatively large R&D projects. The main finding was that the required project accounts were not detailed enough. None of the firms (!)had conducted a timetable completely in line with the regulations. It was not possible to utilize the project accounts to conclude whether expenses in fact where related to the R&D project. This is in line with the findings in Fjærli (2007), who proposed standardized reporting forms to improve the quality of Skattefunn cost reporting.

Comment: It should be possible to implement stricter requirements with respect to reporting/accounting, even if this means increased administrative costs for the users of Skattefunn? Is this among the recommendations from the evaluator?

The report states that "To some extent, however, one must accept misuse as a cost to support schemes intending to attract many firms. This is particularly so when, as is the case with SkatteFUNN, control routines and administrative expenditures are kept at a low level".

Comment: I agree that some misuse is inevitable, and will add: The fact that Skattefunn is founded on tax law triggers off certain standards that may be stricter than standards in other support schemes like for example Innovation Norway or the Norwegian Research Council. It is not obvious that what tax authorities will label as misuse is regarded as such in support schemes where funding is based on thorough evaluation of applicants and applications in advance.

Other comments: The evaluators expect to receive data from the Tax administration related to misuse of SkatteFUNN later, enabling them to link the information on tax deduction for each firm with accounting information and the R&D reporting to Statistics Norway. Indeed, it is important that the evaluators follow up this. Apart from that, I find that the report's handling of the misuse topic is satisfactory.

Question 3. Are the distortions of competition and effect on trade limited?

This question is addressed in chapter 9, summarized by:

"We have assessed potential distortive effects of SkatteFUNN on competition and trade. These may be positive, as well as negative. First of all, SkatteFUNN is neutral by design. Being a rights-based scheme, there is no selection bias related to receiving SkatteFUNN. Neutrality is achieved along most other domestic dimensions, including geographic location, industry, ownership, result, and subject of research We do find evidence of a slight favouring of SMEs, as intended, which arguably has a positive impact on competition as it reduces the entry barriers and counteracts the bias towards large firms by other available R&D schemes (…)" and "(…) To the extent that there are distortive effects of SkatteFUNN they are probably applicable also to most of the other member states having similar arrangements, levelling out the distortions".

Comment: This way of arguing is a formalistic one, or "in the eye of the law", rather than "economic" or "factual". As such I suppose that EU regulators will be satisfied by the conclusions in chapter 9 and I have no objections neither to the report's reasoning nor its conclusions, given the background for it (i.e., within the scope of a mandatory evaluation according to certain criteria). Having said that, I will, however, point out that Skattefunn does distort competition - in favor of R&D-intensive firms, cf. descriptive statistics in chapter 2.7.1. This is intentional, and in line with EU regulations, but owners and employees in low-tech industries may not necessarily agree that this is to the best for society (see my earlier comments related to the evaluation's question 1).

Concluding remarks

According to the evaluator, "The main objective of the evaluation report (…) has been to assess to what extent Skattefunn increases R&D investments in the Norwegian business sector, especially among small and medium sized enterprises." As revealed in chapter 4 and 5, they operationalize this task as identification of causal effects of SKF on increase in R&D (input- and output additionality). Apart from my specific comments and questions to the methods used, I find that the report are doing almost as best as possible to rule out possible alternative explanations behind their findings, by using state-of-the-art econometric procedures. However, in my opinion the report could have been more distant or humble with respect to the question of causality, which I find too superficially treated in the current version.

On the other hand, I also think that the report's conclusions could accentuate some of the findings in chapter 2.7; in the current version of the report the descriptive statistics in 2.7 is not well integrated. I would highlight that

- Skattefunn-firms are R&D-intensive
- Skattefunn-firms are innovative
- Skattefunn-firms have more human capital
- A large fraction of Skattefunn-firms comes from R&D industries like Science & technology, ICT and Advanced manufacturing industries
- SME's and young firms are well represented
- A large fraction of Skattefunn-firms also has qualified for support from other R&D-schemes

Finally, minor comments:

- In my view, there is too much use of graphs and figures in the report. I encourage more use of tables, eventually in an appendix. Also, some of the figures I don't find very informative – consider other graphical presentation or preferably a table (see fig. 4.7).
- In chapters 2.7.1, 2.7.2, 2.7.3 and 2.7.4 the distribution of Skattefunn firms should be compared to the distributions in the entire population of firms. After all, it would not be very informative if the distributions of Skattefunn-firms simply reflect the distribution of the entire population of firms.
- I have discovered at least one case of copy-paste (from HM2007). Please check the manuscript and remember to use quotation marks when you quote.

To conclude: In sum I agree that despite some shortcomings, the report properly documents that it is likely that "Skattefunn increases R&D investments in the Norwegian business sector and is well designed to deliver this objective".

References

Bitler, M.P., T.J. Moskowitz and A. Vissing-Jørgensen (2005): Testing Agency Theory with Entrepreneur Effort and Wealth, Journal of Finance, 60(2), 539–576.

Caggese, A. (2012): Entrepreneurial risk, investment and innovation, Journal of Financial Economics, 106(2), 287–307.

Cohen, W.M., R. C. Levin and D.C. Mowery (1987): Firm Size and R&D Intensity: A Re-examination, Journal of Industrial Economics 35. 543-63.

Cohen, W.M. and R. C. Levin (1989): Empirical Studies of Innovation and Market Structure, chapter 18 in Handbook of Industrial Organization, vol.2 (New York: North-Holland, 1989).

Hall, B.H. and J. Lerner (2010): The Financing of R&D and Innovation, Chapter 14 in (Hall, B.H. og N. Rosenberg) Eds.: Handbook of the Economics of Innovation, Vol. 1, North-Holland Publ. Comp., Amsterdam.

Klemetsen, M. E., Bye, B. and Raknerud, A. (2016), Can Direct Regulations Spur Innovations in Environmental Technologies? A Study on Firm-Level Patenting. Scand. J. of Economics.

Leland, H.E. and D.H. Pyle (1977): Informational asymmetries, financial structure, and financial intermediation, Journal of Finance, 32(2), 371–387.

Lerner, J. Boulevard of Broken Dreams: Why Public Efforts to Boost Entrepreneurship and Venture Capital Have Failed – and What to Do About It. Princeton Univ. Press, 2009.

Schumpeter, J. (1942): Capitalism, Socialism, and Democracy, Harper and Row, New York.

Stiglitz, J.E. and A. Weiss (1981): Credit Rationing in Markets with Imperfect Information, The American Economic Review, 71(3), 393–410.

Zoltan, J.A. and D.B. Audretsch (1988): Innovation in Large and Small Firms: An Empirical Analysis. American Economic Review 78, 678-90.