Presented at the Blue Sky II Forum on "What Indicators for Science, Technology and Innovation Policies in the 21st Century?", Ottawa, Canada, 25-27 Sept. 2006

What is missing in the analysis of input-output relationships of innovation processes?

Mark Knell and Svein Olav Nås

(NIFU STEP)

Background

1. It is widely recognized that innovation may result in significant improvements in firm performance (OECD, 2005). For this reason would expect to observe strong correlations between innovation inputs and economic results. One rationale of the Oslo Manual is to facilitate such analysis by operationalising a series of input- and intermediate innovation indicators. However, econometric estimations of the effects on economic performance of the Oslo Manual types of innovation inputs only explain small fractions of the observed variance. Why is this? Can we do something about it? Those are the core questions raised in the paper, with the aim of contributing to improved indicators and analysis.

2. One of the most important tasks facing empirical analyses of economic performance is the construction and interpretation of indicators for the generation and acquisition of knowledge. Early models developed by economists affiliated with the NBER incorporate a variable that captured the 'economically valuable technological knowledge', or what Zvi Griliches often called 'knowledge capital' and said very little about what knowledge is, or of how it becomes important for innovation and growth. More recent models make use of indicators created through innovation surveys but still lacks a proper understanding of how different forms of knowledge transforms into economic performance.

3. An important objective of the Oslo Manual has been to provide guidelines for creating new input and intermediate indicators and composing survey questionnaires. The manual was revised twice since the original edition issued in 1992, mainly to include the service sector and then to cover organizational and marketing innovations. The third edition of the Oslo Manual (OECD, 2005, p. 5) states that it is imperative to "capture the changes that affect firm performance and contribute to the accumulation of knowledge". This suggests that future questionnaires should include questions specifically targeted to the input-output relationships of the innovation process.

4. In the Community Innovation Survey (CIS) and similar surveys based on the Oslo Manual, firms are asked to give information on knowledge inputs into the innovation process including R&D expenditures within the firm, collaboration with other firms and organizations, and acquired from outside the firm. Other inputs include knowledge embodied in machinery, technology licenses, training, and market analysis. The survey also covers a variety of different sources of information (internal, suppliers, customers,

universities, etc.) as well as the ability to recognize hampering factors (funds, personnel, information, etc). Influences on innovation inputs include public and other financial support. Outputs include the introduction of new products and processes, the share of turnover due to new products, the effects of innovative activity on the enterprise, and the effects of organizational innovations on the enterprise.

5. In what follows the three aspects of the input-output relationship is addressed in turn: What are the main characteristics to take into account when modeling the relationships between inputs and outputs? To what degree are the relevant inputs factors identified and introduced in the estimations? What are the expected and identifiable results and effects of innovation from the perspectives of business firms and society?

Modeling the relationship between innovation and economic performance

6. Indicators derived from innovation surveys have been used in the most recent models developed by economists affiliated with the NBER. Early models focused mainly on the relationship between R&D activity and productivity growth within a production function framework (eg. Griliches, 1979; 1984; 1995; 1998). Pakes and Griliches (1984) developed a variant of this framework in which changes in knowledge capital, defined as the level of economically valuable technological knowledge, are unobservable, which allows for the inclusion of several interrelated innovation inputs. In this model, innovation inputs included only patents and R&D activity, but in an influential paper by Crépon, Duguet and Mairesse (1998), or CDM for short, indicators derived from French innovation survey were incorporated into the basic framework.

7. The CDM model combined a production function of knowledge relating R&D activity to patenting or innovative activities with economic performance as measured by labor productivity. It contains a system of three simultaneous equations where R&D activity and other factors generate new knowledge, which then propels innovation (output) and finally productivity growth. Other supply and demand factors as well as sectoral differences and unobserved heterogeneity are also included in the model to improve its explanatory power. A paper by Mohnen, Mairesse and Dagenais (2006) estimated this model using micro-aggregated data from the first Community Innovation Survey from Belgium, Denmark, Ireland, Germany, the Netherlands, Norway and Italy for 1992. Figure 1 describes the model using path analysis.

8. Variants of the CDM model have been estimated on data from France (Duguet, 2000), Germany and Sweden (Janz, Lööf and Peters, 2004), Germany (Peters, 2006), the Netherlands (van Leeuwen and Klomp, 2006), Chile (Benavente, 2006), Sweden (Lööf and Heshmati, 2002a, 2006), Norway and Finland (Lööf, Heshmati, Asplund and Nås, 2002)), China (Jefferson, Huamao, Xioajing and Xiaoyun, 2006), the UK (Criscuolo and Haskel, 2002), Canada (Mohnen and Therrien, 2003) and France, Germany, Spain and the UK (Griffith, et. al. 2005). Surveys by Hall and Mairesse (2006) and Mohnen (2006) show that there are some similarities across countries, but also that the differences indicate that there are some serious econometric problems. Nevertheless, some progress has been made in improving the estimation procedure and in matching the survey data with accounting data.

9. There are several econometric problems that the model should address. First, there are measurement issues in most of the models. Often we have only sales and not value added or material inputs. As a result the coefficients tend to be inflated because material inputs

are implicitly assumed as constant. Second, the model may contain as ‘identification’ problem. This problem appears because the inputs and outputs are chosen simultaneously. For example, innovation may improve productivity, but higher productivity growth can encourage innovation. Identification can be achieved by finding a suitable instrumental variable (see Arellano, 2004), in particular GMM methods applied to panel data (fixed-effects models can exacerbate the measurement errors). But choosing the wrong instrument can also cause additional problems. Third, the CDM model itself relies on the strong assumption that the four disturbance terms are not correlated. This makes it difficult to apply this model in the innovation context and it may not be appropriate for cross-section analysis, which is commonly used with innovation survey data because of the problem of comparability over time. Finally, the basic CDM model is linear, which means that non-linearity in the innovation system will be overlooked. Testing for nonlinearity can also be regarded as a test for heterogeneity, which is a problem plaguing these models.

10. Even though CIS and similar surveys have added significantly to our understanding of innovation processes, they also have limitations in terms of analyzing input-output relationships. Mohnen (2006) summarizes many of these problems mainly from a comparative perspective. We believe that when investigating the effects of innovation on performance the timing problems is essential. Basically the present CIS survey design identifies results (product or process innovations) obtained over three years combined with information about innovation efforts in the last of these three years. Alternatively, results obtained in the same year as the efforts (new to firm or market share of sales). To link the inputs and outputs it is necessary to make strong assumptions, for instance that current innovation efforts are representative of or a continuation of similar efforts in the past that actually lead to the results reported. Raymond et al (2006) show that this assumption is generally not confirmed. It would be necessary to construct a time series panel data to overcome the problem. The ideal would be to have balanced panels of innovation surveys with a reasonable frequency, preferably annually. Both frequency and balancing of panels are, however, resource demanding and will take time before usable data can be in place. Differences in lag structures between industries will also still represent a problem.

11. Alternatively innovation data can be supplemented with external performance data at the enterprise level, such as accounts. Such a solution helps in balancing the panels as all enterprises have to report results. Problems of lag structures would still remain. There are also problems with differences in accounting practices between countries. Lastly, the timing problem of including in the analysis intermediate results such as innovative sales is not solved this way. The latter would require linking frequent innovation surveys over several years. This makes it easier to apply GMM estimation techniques (Arellano, 2004).

12. A selectivity problem arises because only a fraction of firms engage in innovation in a given period, and their initial state varies and affects estimation. To correct for this problem a simultaneous equation system such as the CDM model can be used, with a Heckman selection equation identifying innovators. The information used for selection should be independent of the information in the input-output step of the relation. Presently too little information that can be used is collected for both innovators and non-innovators. To overcome this problem one option is to change survey questionnaires if it can be done without increasing the response burden too much. Another option is to gather the information that is needed from supplementary statistical sources. The latter is the only

possibility in the short run, and may constitute a permanent solution if matching can be done at the micro level.

13. The types of information that are relevant include age of the firms, some indicator of technological opportunities in the given industry and the composition of human resources in the firm compared to the industry average. Apart from technological opportunity all of these are available census-based information in many countries. Technological opportunity is difficult to observe directly, but may be approximated by for instance the numbers of patents or patent applications in an industry, or the R&D intensity. It is also related to the prevailing competition and market structures surrounding each firm. Such information is available in some cases, but usually limited to indicators such as share of national markets and export share of sales. In the increasingly open and free trading economies such information should include the full world markets, although it introduces significant measurement problems.

Understanding and measuring knowledge and learning

14. Knowledge and knowledge creation appears at the center of the NBER family of models, including the CDM model. Yet these models consider knowledge as an unobservable. The main reason is that we have only a rudimentary understanding of the factors that shape the rate, direction and effects that the learning process has on the creation of knowledge. Even among philosophers, the theory of knowledge remains controversial. Epistemological discussions have focus on analyzing the nature of knowledge, how it is acquired and how it relates to truth, belief, justification, and perception. For the early philosophers, including Plato, Aristotle, and Descartes, knowledge was purely objective in that it was impersonal, explicit and permanent. The Cartesian method, which was based on the search for certainty and universality through the use of conjecture, influenced the subsequent discussion on the limitations of knowledge and the meaning of experience. Adam Smith can be placed at the center of this discussion; both in terms of the philosophical debate and in defining how this epistemology can be incorporated into economic theory (cf. knowledge creation through the division of labor in chapters 1-3 of Book 1 and the assumption of known knowledge in chapters 5-7 of Book 1).

15. Perception, in the form of probabilistic statements, remained trivial in the theory of knowledge until the twentieth century. This changed when Merleau-Ponty argued for the primacy of perception over conjecture and experience, providing justification for moving away from the view that knowledge must be, at least, true and justified toward one where only beliefs need to be justified. Epistemology thus becomes ‘biological’ in the sense that knowledge must contain a tacit dimension as pointed out by M. Polanyi. This recognition of two different kinds of knowledge is closely related to Ryle’s distinction between knowing how (procedural knowledge), and knowing that, (descriptive knowledge). While tacit knowledge is difficult to define, it is important for evolutionary economic models and innovation policies based on these models.

16. Knowledge capital in economic models needs to become biological in the sense that it considers the tacit dimension. Discussions in the philosophy of knowledge suggest that knowledge appears in many different forms and localizations and interacts with other factors. In economics knowledge is usually poorly understood and it often appears in simple forms; in neoclassical theory usually reduced to information that can be transferred without transaction costs or time lags. The relevant forms of knowledge vary according to

the circumstances and over time, depending on the form of knowledge and the intended use. To be able to utilize knowledge, in any form, it must be learned by the relevant actors; firms and their management and employees, and customers. Knowledge forms are different when it comes to the time and effort it takes to learn and transfer them, the risk of copying, and the cost of utilizing them.

17. The tacitness, embeddedness and complementarity of knowledge to other factors are not currently satisfactorily dealt with in innovation surveys. Even though innovation surveys distinguish components of innovation activities and costs, the breakdown is still difficult to comprehend both for respondents and analysts. It mixes types of knowledge with sources of knowledge and methods of producing it, and leaves out questions about learning. Learning is necessary to be able to implement and utilize knowledge, which may not be trivial.

18. Innovation does not mean that only new, improved or recombined knowledge is relevant. Equally important is the utilization of the whole range of capabilities that the firms possess, embedded in the routines of the organization, the competences of the employees and the built in capabilities of machinery and equipment. Innovation surveys tend to address newly generated knowledge to a larger extent than existing knowledge. Some components of existing knowledge are included in innovation surveys, in particular when including recombination of existing knowledge in the definition. Parts of this problem can be handled by estimation techniques taking omitted variables and initial conditions into account. However, many of the components of existing competencies are interacting with innovation activities as a prerequisite for innovations to succeed also in economic terms.

19. In economic theory, knowledge is generally conceived as a stock, or knowledge capital. Knowledge is created through experience, learning or perception. Arrow (1962) emphasized that learning is the process of acquiring knowledge, skills, and values through study and experience. Learning by doing is one example of informal methods of acquiring knowledge and education is an example of a formal method. Knowledge often becomes obsolete or is forgotten, which implies that the stock of knowledge may also depreciate over time. The extent of this depreciation is difficult to assess, partly because it does not physically deteriorate (Boulding, 1966), but because the deterioration of its usefulness will be uneven and dependent on the generation of new knowledge (process of creative destruction). Since the outcome of innovation, and economic performance more generally, depends on the utilization of the entire stock of knowledge, and not only the new or modified parts, assessing the depreciation rate becomes important. In most studies, however, only the flow of new knowledge is usually taken into account, putting the knowledge stock aside as differences in initial conditions. To better understand differences in economic performance an assessment of differences in knowledge management between firms should be integrated in the analysis.

20. Firms operate in constant interactions with their surroundings, such as their customers and other firms in the value. In some cases even specialized suppliers delivering R&D or other kinds of knowledge services are involved. In other cases open sources of information that require no or very limited payments are consulted. Firms belonging to groups, nationally or internationally, or with other kinds of relationships, may exchange knowledge and information free of charge. As a result, large parts of access to relevant knowledge are not recorded, at least not in quantitative terms. To the degree this kind of

relations affects innovation output and firm performance they disturb the observed relationship between innovation inputs and results.

21. Innovation is known to be complex processes, to varying degrees involving a large array of different inputs. To be realized and generate economic returns, investments and activities that go far beyond the knowledge related activities of innovation are called for. This includes inputs such as production facilities, distribution systems and marketing, all of which may outperform innovation activities in magnitude. Without such complementary inputs no results will emerge, and the innovation becomes a failure and a loss. Part of policy debates circle around the problems of selecting and supporting promising innovations beyond the research and introduction stages. Information about such complementary investments linked to innovation is usually not available, thereby limiting our ability to identify reasons for success or failure.

What are the relevant output measures?

22. It is essential to be aware of the stage in the innovation process that is being addressed in the model. At the level of the firm it is possible to distinguish between different kinds of innovation outputs, for instance the discovery of a new active pharmaceutical component, or the introduction of a new product on the market. These are both only intermediary outputs for the firm, as we consider firms to be profit seekers so that it is the economic surplus generated over the relevant time horizon that matters most. This may take different forms such as annual net results, dividend to owners or shareholder value of the company. In this section we bring up some of the problems associated with measuring output.

23. The returns on investments in innovation are usually sought in the subsequent performance of the innovating firms. But innovators are only to a limited degree able to appropriate the full value of their innovations, or bear the costs of failures. Effects are found also outside the innovators, for instance among customers (consumer surplus), but also among competing or cooperating firms. For society as a whole it is the sum of the effects that are of interest. To legitimize public policy it is the social return that matters, not the private. Since such effects may be widespread, and sometimes the effects are the larger far from where the innovations originated, one must expect the social returns of successful innovations to be underestimated. There may also be a case where innovations with negative effects, such as pollution and other environmental concerns, also are underestimated.

24. Given that capitalistic enterprises are profit seekers and cost minimizers, investigating accounting data, to the degree that accounts reflect real performance measures, can identify economic returns. The main challenge becomes to identify the correlation between the concrete innovation activities and subsequent accounted results. The same holds true when studying social return by use of GDP or other aggregated indicators for economic activity.

25. Technological opportunities vary between industries. Some mature industries develop slowly or not at all, and the relative positions of incumbent firms are stable. The expectation is that profitability is relatively equally distributed, and that firms survive over time without being challenged by entrants. In other industries there is rapid technological development and existing competitors and entrant firms constantly challenge incumbent firms. In such situations we expect much investment in innovation, but not necessarily

high profitability as a result. The high innovation rate may be a prerequisite to be in the industry and survive. Survival over time becomes the relevant performance indicator whereas profit rate hardly give additional information.

26. Innovation surveys try to capture information about outputs of innovation activities directly. The preferred measure is share of sales consisting of new or changed products, new to the firm or the market, sometimes with similar indicators for process innovations. Such measures are valuable as indicators for intermediate results but have nothing to say about the ultimate goal of business firms.

27. Innovations take varying time to develop and profit, and their expected life cycles also vary accordingly. The real value of an innovation is not only the observed flow of resources stemming from it, but the expected future returns, which affects the current value of the innovating firm. Over time additional inputs are added and the original idea or innovation is gradually, or radically, transformed. This makes it very demanding to track the actual outcome of innovations; what are the relevant time lags, what are the relevant inputs explaining the observed outputs? As taken up above also the rate of depreciation of knowledge needs to be taken into account but is hard to assess.

28. A last issue to mention concerns firms operating internationally, which makes it hard to identify all relevant inputs and outputs. Statistical data usually refer to national territories whereas knowledge and other inputs to the innovation process can be produced and utilized globally. This issue also relates closely to technology transfer and spillovers both within an industry and across industries (Damijan and Knell, 2005).

Concluding remarks

29. Innovation surveys based on the Oslo Manual guidelines have added significantly to our understanding of innovation processes. Still there remain considerable challenges that expose some of the shortcomings in how innovation, knowledge and learning are conceptualized and understood, as well as how core concepts are operationalized and measured. To advance our knowledge we think that much can still be achieved by improving on modeling and analysis of existing data. We also see short-term opportunities in supplementing innovation data from existing sources. Over time we think that improvements in the innovation survey methodology itself can and must be undertaken, in particular to solve problems with timing of inputs and outputs. Lastly the renewed empirical insights should be brought forward into improving formal theorizing about knowledge and learning to better guide and make comparable analytical work.

30. Recent studies like the CDM model and follow up work have demonstrated that existing data can produce reasonable results, but that results are very sensitive to model specifications. It is therefore extremely important to properly take the econometric problems into account, and to develop and use a standardized setup if comparative work is undertaken. The CDM model also addresses selectivity issues and differences in initial states of the enterprises. The estimates are, however, hampered by lack of available unrelated instrumental variables for the selection part of the estimations. The model also has an identification problem that is caused by endogeneity. Another difficulty is that the model is linear, whereas the distribution of firms according to most innovation variables is highly skewed. This is probably possible to overcome by alternative specifications of the model. The same applies partly to the lack of instrumental variables that can be linked in from supplementary sources. Lastly, the internal timing problem of inputs and

(intermediate) outputs in the existing innovation surveys is a serious problem that can only be overcome by future changes in survey design or panel data. At present including supplementary time series information on performance is the only solution we see.

31. The supplementary data we think should be utilized include time series accounting data at the firm level. Other census-based sources include patenting activity education or other information on human resources, information on entry and exit of firms and other types of information on firm demography. To approximate technological opportunities in different industries patenting, R&D investments and investments in machinery and equipment could be utilized. To better assess the relationship between innovation at the level of business firms and at the level of the economy, a combination with input-output tables could be utilized. An important type of information concerns market shares and competitive environment of the firms. This type of information is difficult to get but may be sought in for instance trade registers and business associations. Lastly, new surveys are being developed that collect information of relevance that should be coordinated with the innovation surveys. An example is the surveys on knowledge management carried out in Canada, France and Denmark.

32. The final point concerns improvements of the innovation surveys themselves. This is a difficult task, as consistent time series of observations are called for, requiring as little changes as possible to the survey design from time to time. We do think, however, that some changes can be implemented without changing the contents of the survey. Basically they concern building panels and having sufficient frequency of data collection to be able to track development over time of single firms. This would solve the internal timing problem where recorded outputs occur before or at the same time as inputs. Also one could think of adding information (possibly by replacing redundant information). Candidates to include are better output information on process innovation, and input information on innovation activities related to organizational and marketing innovations.

References

Arellano, M. (2003), *Panel Data Econometrics*, Oxford: Oxford University Press.

Arellano, M. (2005), Comments on Griffith, R. E. Huergo, J. Mairesse, and B. Peters (2005), memeo, Madrid.

Arrow, K. (1962), "The economic implications of learning by doing", *The review of Economic Studies* 29, 155-173.

Benavente, J. M. (2006), "The role of research and innovation in promoting productivity in Chile", *Economics of Innovation and New Technology* 15, 301-315.

Boulding, K.E. (1966), "The economics of knowledge and the knowledge of economics", *American Economic Review* 56, (Proceedings), 1-13.

Crépon, B., E. Duguet and J. Mairesse (1998), "Research and Development, Innovation and Productivity: An Econometric Analysis at the Firm Level", *Economics of Innovation and New Technology* 7, 115-158.

Criscuolo, C. and J. Haskel (2002), "Innovations and productivity growth in the UK", CeRiBa discussion paper.

Damijan J.P. and M. Knell (2005), "How important is trade and foreign ownership inclosing the technology gap? Evidence from Estonia and Slovenia", *Review of World Economics*, 141, 271-295.

Duguet, E. (2000), "Knowledge diffusion, technological innovation and TFP growth at the firm level: Evidence from French manufacturing", EUREQua 2000.105.

Griliches, Z. ed. (1984), *R&D, Patents and Productivity*, Chicago: University of Chicago Press.

Griliches, Z. (1998), *R&D and Productivity: The Econometric Evidence*, Chicago: University of Chicago Press.

Griliches, Z. (1979), "Issues in Assessing the contribution of R&D to productivity growth", *Bell Journal of economics*, 10, 92-116.

Griliches, Z. (1995), "R&D and Productivity: Econometric Results and Measurement Issues", in P. Stoneman (ed.), *Handbook of the Economics of Innovation and Technical Change*, Blackwell Handbooks in Economics,

Griffith, R. E. Huergo, J. Mairesse, and B. Peters (2005), Innovation and productivity across four European Countries, memeo.

Hall, B. and J. Mairesse (2006), "Empirical studies of innovation in the knowledge-driven economy', *Economics of Innovation and New Technology* 15, 289-299.

Janz, N., H. Lööf, and B. Peters, (2004). "Firm level Innovation and Productivity: Is there a Common Story?" *Problems and perspectives in management* 2, 184-204.

Jefferson, G., B. Huamao, G. Xioajing and Y. Xiaoyun (2002), "R and D Performance in Chinese Industry", *Economics of Innovation and New Technology* 15, 345-366.

Kleinknecht, A. and P. Mohnen (2002) (eds.), *Innovation and Firm Performance. Econometric Explorations of Survey Data*. Palgrave, Hampshire and New York.

Lööf, H. and A. Heshmati (2002), "Knowledge capital and performance heterogeneity: A firm-level innovation study", *International Journal of Production Economics* 76, 61-

85.

Lööf, H. and A. Heshmati (2006), “On the relationship between innovation and performance: A sensitivity analysis”, *Economics of Innovation and New Technology* 15, 317-344.

Lööf, H., A. Heshmati, R. Apslund and S.O. Nås (2002), “Innovation and performance in manufacturing firms: A comparison of the Nordic countries”, International Journal of Management Research, 5-36.

Mohnen, P. (2006): Using innovation survey micro data: Recent experience. OECD, DSTI/EAS/STP/NESTI(2006)8.

Mohnen, P., J. Mairesse, and M. Dagenais (2006), “Innovativity: A comparison across seven European countries”, *Economics of Innovation and New Technology* 15, 391-413.

Mohnen, P. and P. Therrien (2003), “Comaring the innovation performance of manufacturing firms in Canada and in selected European countries: An econometric analysis”, in F. Gault, (ed.), *Understanding Innovation in Canadian Industry*, Montreal: McGill-Queens University Press, 313-339.

OECD, (2005), *Oslo Manual: Guidelines for Collecting and Interpreting Innovation Data*, Paris: OECD.

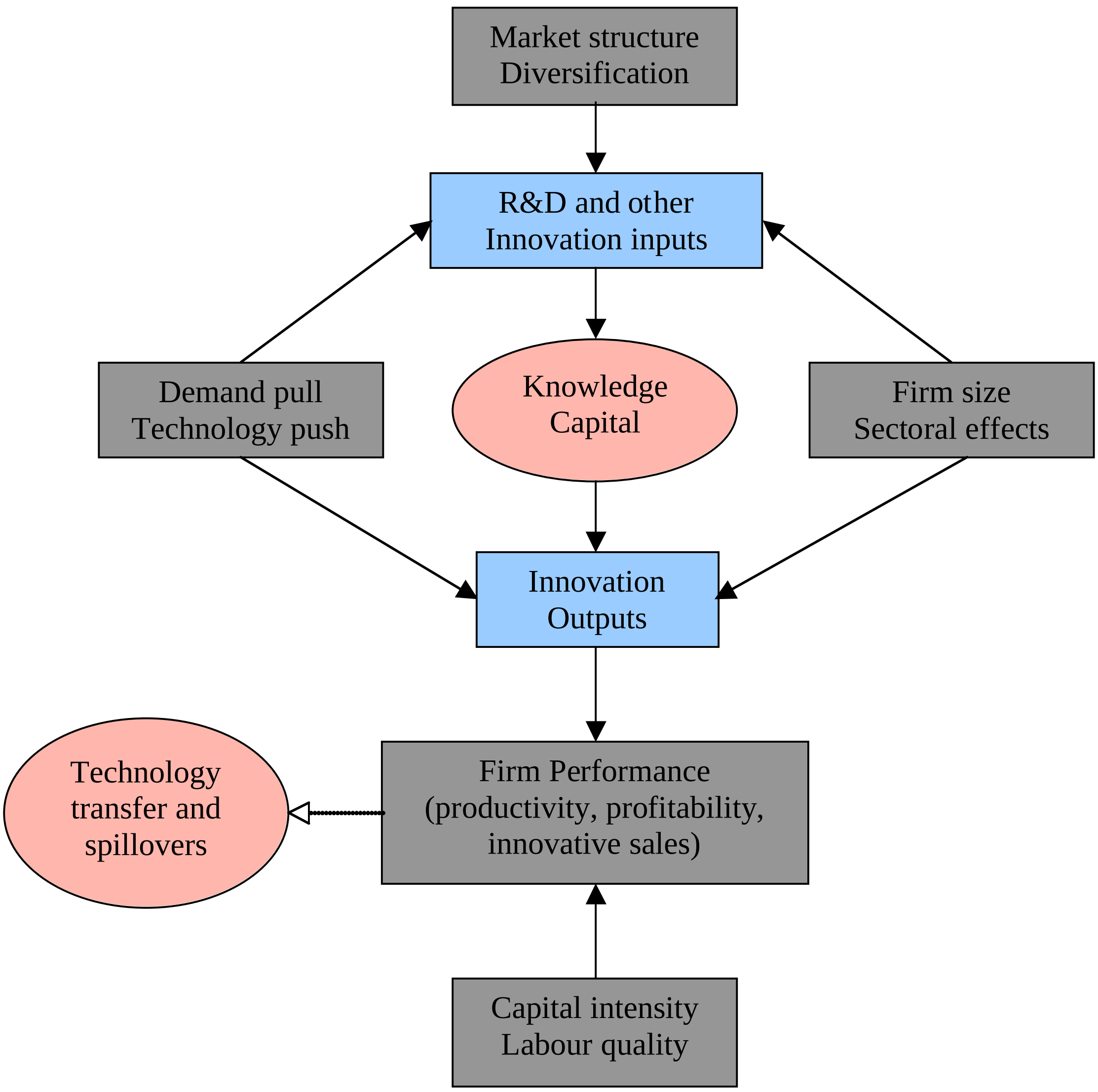
Pakes, A. and Z. Griliches (1984), “Patents and R&D at the firm level: A first look” in Griliches (1984).

Peters, B. 2005. Persistence of innovation: Stylised facts and panel data evidence, Discussion paper 05-81, ZEW.

Raymond, W. P. Mohnen, F. Palm, S. Schim van der Loeff (2006), Persistence of innovation in Dutch manufacturing: Is it spurious?, Cirano Scientific Series 2006s-04.

van Leeuwen, G. and L. Klomp (2006), “On the contribution of innovation to multi- factor productivity growth”, *Economics of Innovation and New Technology* 15, 367-390.

Figure 1: Innovation and productivity



Source: Authors adaptation of figure 1 in Hall and Mairesse (2006)

Biographical notes:

Mark Knell joined NIFU STEP in January 2005, as a Senior Researcher. In total, he has more than fifteen years of experience in teaching and research in the fields of technology and economic policy research.

Svein Olav Nås is senior researcher at NIFU STEP in Norway. Fields of work include indicator development and analysis within the broad field of innovation studies. Involved in preparation of the CIS surveys, revisions of the OM, and NESTI work.