11 LIST OF REFERENCES

Adams, J.D., 1990. Fundamental stocks of knowledge and productivity growth. Journal of political economy, Vol. 98, pp. 673-702.

Arrow K.J., 1962. The economic implications of learning by doing. Review of economic studies, Vol. 29, No. 3, pp. 155-173.

Aghion, J., Howitt, P., 1992. A model of Growth through Creative Destruction. *Econometrica*, Vol. 60, pp. 323-351.

Aghion, J., Howitt, P., 1998. *Endogenous Growth Theory*. Cambridge, Mass.: MIT Press

Anandakrishnan, M., Morita-Lou, H., 1984. *Indicators and Technology for Development*. In: Morita-Lou, Hiroko, 1984. *Science and Technology Indicators for Development*. London: Westview Press, pp. 25-33.

Archibugi D., Howells J., Michie, J., 1999. Innovation policy in a global economy. Cambridge University Press.

Archibugi D., Michie J., 1997. Technological Globalisation or National Systems of Innovation? *Futures*, Vol. 29, No. 2, pp.121-137.

Arthur W. B., 1994. *Increasing returns and path dependency in the economy*. The University of Michigan Press: Ann Arbor.

Balzat, M., 2002. *Theoretical basis and the theoretical treatment of National Innovation Systems*. University of Augsburg, Institute for Economics, Beitrag nr 232.

Barlas, Y., 1989. Multiple tests for validation of system dynamics type of simulation models, *European Journal of Operational Research*, Vol 42, No. 1, pp. 59-87.

Barlas, Y. 1990. An autocorrelation function-test for output validation, *Simulation*, Vol. 55, No. 1, pp. 7-16.

Barlas, Y., 1996. Formal aspects of model validity and validation in system dynamics. *System Dynamics Review*, Vol. 12, No. 3, pp.183-210.

Blanchard, B.S., Fabrycky W.J., 1998. *Systems Engineering and Analysis*, Prentice Hall, New Jersey.

Blalock, H.M., 1961. *Causal Inferences in Nonexperimental Research*. Toronto Library - University of North Carolina press.

Botha, A., 1997. *The Development of a system dynamics model to simulate the dynamic behaviour of technology substitution*. Dissertation, (MBA), University of Pretoria.

Bozeman, B., 2000. Technology transfer and public policy: a review of research theory. *Research Policy*, Vol. 29, pp. 627-655.

Bozeman, B., Rogers, J.D., 2000. A churn model of scientific knowledge value: Internet researchers as a knowledge value collective. *Research Policy*, Vol.31, pp. 769-794.

Brush, G. G., 1988. *Volume 12: How to choose the proper sample size*. American Society for quality control, ASQC Quality Press. Milwaukee, Wisconsin.

Buys A. 2001, *Technological decolonisation of South Africa by backwards integration of the national system of innovation*. Pretoria: Department of Engineering and Technology Management, University of Pretoria.

Buys A., 2002. *Industrial development in South Africa by backwards integration of the National system of Innovation*. Pretoria: Department of Engineering and Technology Management, University of Pretoria.

Buys A., 2004. *Characterization of the South African National System of Innovation*. Pretoria: Department of Engineering and Technology Management, University of Pretoria.

Bifani, P., 1984. *Environment and Indicators of Technological Development: Methodological Considerations*. In: Morita-Lou, Hiroko, 1984. *Science and Technology Indicators for Development*. London: Westview Press, pp. 61 – 76.

Blankley, W., Kahn, M., 2005. *The history of research and experimental development measurement in South Africa and some current issues*. South African Journal of Science, Vol. 101, March/April 2005.

Bochet, J., 1984. *Intrinsic Indicators of Technological Development: Preliminary Outline of the Issues Involved*. In: Morita-Lou, Hiroko, 1984. *Science and Technology Indicators for Development*. London: Westview Press. p. 77-106.

Boshoff, N., Mouton, J., 2003. HRD review, Chapter nine, *Science Policy indicators*, pp. 209 – 233.

Carlsson, B., 1995. *Technological Systems and Economic Performance: The Case of Factory Automation*. Kluwer Academic Publishers, Boston, Dordrecht.

Carlsson, B., 1997. *Technological Systems and Industrial Dynamic*. Kluwer Academic Publishers, Boston, Dordrecht.

Carlsson, B., Stankiewicz, R., 1995. On the nature, function and composition of technological systems. In Carlsson, B., 1995. *Technological Systems and Economic*

Performance: The Case of Factory Automation. Kluwer Academic Publishers, Boston, Dordrecht.

Carlsson, B., Jacobsson, S., Holmen, M., Rickne, A., 2002. Innovation Systems: analytical and methodological issues. *Research Policy*, Vol.31, pp. 233-245.

CENIS, 2000. *Science in South Africa: History institutions and statistics*. University of Stellenbosch.

Christensen J.L., 1992. *The Role of Finance in National Systems of Innovation*. In: Lundvall B. A., 1992. *National Systems of Innovation*., Pinter: London, pp. 116-127.

Cohen, W.M., Levinthal D.A., 1990. Absorptive Capacity: A new perspective on learning and innovation, *Administrative Science Quarterly*, vol. 35, pp. 128-152.

Coyle R.G., 1996. *System Dynamics Modelling*. Chapman & Hall:London.

Coyle, G. 2000. Qualitative and quantitative modeling in system dynamics: some research questions. *System Dynamics Review*, Vol. 16, No. 3, pp. 225–244.

CPROST, 1997. The Use Of National Systems Of Innovation Models To Develop Indicators Of Innovation And Technological Capacity. Report 6, Simon Fraser University at Harbour Centre

Crush, J. et al., 2000. Losing out minds: Skills migration and the South African brain drain. *SAMP Migration Policy Series*, No. 18, pp.43-59.

DST, 2004. *High Level Key Results. South African National Survey of Research and Experimental Development.*Department of Science and Technology, South Africa.

DACST, 2002. *South African Science and Technology: Key facts and figures*, National Advisory Council on Innovation (NACI), submitted to Branch: Science and Technology, Department of Arts, Culture, Science and Technology (DACST).

DACST, July 2002. *South Africa's National Research and Development Strategy*. Government of the Republic of South Africa.

DACST, 4 September 1996. *White Paper on Science and Technology: Preparing for the 21st century*, submitted to Department Arts, Science, Culture and Technology.

Dangerfield, B., Roberts, C., 1996. An overview of strategy and Tactics in System Dynamics Optimization. *Journal of the Operational Research Society*,Vol. 47, pp. 405-423.

Department of Finance, 1996. *Growth Employment and Redistribution Strategy*. Pretoria

De Wet, 2001. *Emerging from the technology colony: A view from the South. Pretoria*. Department of Engineering and Technology Management, University of Pretoria

DTI, 2002. *Accelerated Growth and Development: The contribution of an Integrated Manufacturing Strategy*, Pretoria.

Du Toit, C., Koekemoer, R., Ground, M., 2004. *Estimating technical progress for South Africa*. Ninth annual conference on econometric modelling for Africa.

Edquist C., 1997. *Systems of Innovation: Technologies, Institutions and Organizations*. A Cassel Imprint, Wellington house, London.

Eksteen, M., 2001. *Growth and Innovation Report. National Science & Technology forum*. CSIR, Pretoria.

Fabian Y., 1979. *Conceptual and Methodological Concerns in the Development of International S&T Indicators: the OECD Science and Technology Indicators System*, Seminar on Research Needs and Applications for Indicators Based on the Scientific and Technical Literature, Philadelphia, Institute for Scientific Information.

Fedderke, J., 2002. *Toward an Understanding of the Growth Absence: Reviewing the Evidence that can Account for the Poor Growth Performance of the South African Economy*. TIPS working paper

Fedderke, J., De Kadt, R.H.J., & Luiz, J., 2003, A Capstone Tertiary Educational system: Inefficiency, Duplication and Inequity in South Africa's Tertiary Education System, 1910-93, *Cambridge Journal of Economics*, Vol. 27, No. 3, pp. 377-400.

Foray D., *Generation and Distribution of Technological knowledge: Incentives, Norms, and Institutions*. In: Edquist, C., 1997, *Systems of innovation: technologies, institutions and organisations*. London: Pinter, pp. 64 -82

Forrester J. W., 1961. *Industrial dynamics*. John Wiley & Sons: New York.

Forrester, J. W. 1968. *Principles of Systems*. Cambridge MA: Productivity Press.

Forrester, J.W., 1973. *Confidence in Models of social behaviour – With emphasis on System Dynamics Models*. MIT System Dynamics Group Memo D-1967, Cambridge, MA 02139 (unpublished)

Forrester J.W., Senge, P., 1980. *Tests for building confidence in System Dynamics models*. In: Legasto A.A., Forrester J.W., Lyneis J.M., 1980. System Dynamics. Studies in the Management sciences, *System Dynamics*, Vol. 14, pp. 11-21.

Forrester J.W., Lyneis J.M., 1980. System Dynamics. Studies in the Management sciences. *System Dynamics*, Vol.14. pp. 209-228.

FRD, 1993. *South African Science and Technology indicators – 1993*. Directorate for Science and Technology Policy, Pretoria.

FRD, 1990. *South African Science and Technology indicators – 1990*. Directorate for Science and Technology Policy, Pretoria.

Freeman C., 1987. *Technology Policy and Economic Performance. Lessons from Japan*. Pinter: London.

Galli, R., Teubal M., 1997. *Paradigmatic shifts in national innovation systems*. In: Edquist, C., 1997, *Systems of innovation: technologies, institutions and organisations*. London: Pinter, pp. 343-370.

Gelsing, L., 1992. *Innovation and the Development of industrial networks*. In: Lundvall B. A., 1992. *National Systems of Innovation*. Pinter: London, pp. 116-127.

Godin, B., 2006. The Knowledge based economy: Conceptual framework or buzzword?. *The Journal of Technology Transfer***.** Vol. 31, No. 1, pp 17-30.

Griliches, Z., 2000. *R&D, Education and Productivity*. Cambridge, Mass.: Harvard University Press.

Grilliches, Z., 1979. Issues in Assessing the Contribution of Research and Development to Productivity Growth. *The Bell Journal of Economics*, Vol. 10, pp.92-116.

Grossman, G., Helpman, E., 1991a. Innovation and Growth in the Global Economy. Cambridge, Mass.: MIT Press.

Grossman, G., Helpman, E., 1991b. Quality Ladders in the Theory of Growth. Quarterly Journal of Economics, vol. 106 p. 557-586.

Grossman, G., Helpman, E., 1991c. *Quality Ladders in the Theory of Growth*. Rev. Economic Studies, vol. 58 p. 43-61.

Gunasekara, C., 2006. *Reframing the Role of Universities in the Development of Regional Innovation Systems*. The Journal of Technology Transfer. Vol. 31, No. 1, pp. 101 - 113

Guerrieri, P., Tylecote, A., 1997. *Inter-industry differences in technical change and national patterns of technological accumulation*. In: Edquist, C., 1997, *Systems of innovation: technologies, institutions and organisations*. London: Pinter., pp. 107-125.

Guilhon, B. 2001. *Technology and markets for knowledge: Knowledge creation, diffusion and exchange within a growing economy*. Boston/Dordrecht/London: Kluwer Academic Publisher

Hall, H., Jaffe, A.B., Trajtenberg, M., October 2001. *The NBER patent citation data file: Lessons, Insights and methodological issues*. Working Paper 8498. http://www.nber.org/papers/w848. National Bureau of Economic Research.

Hall, B., Van Reenen J., 2000. How effective are fiscal incentives for R&D? A review of the evidence. *Research Policy*, Vol. 29, p. 449-469.

Hamilton, J.D., 1994. *Time Series Analysis*. Princeton University Press, Princeton, New Jersey.

Hill, C., Roessner, J.D., 2000. *New directions in federal laboratory partnerships with industry*. Report to the Jet propulsion Laboratory.

Hilmola, O., Helo, P., Kekale, T., 2004. *Economic Dynamics of R&D: Analysis of Technology and Development*. System Dynamics Conference 2004

Hofstede, G.H., 1991. *Cultures and Organisations: Software of the mind*. McGraw-Hill, London.

Howells J., 1999, *Regional systems of innovation?*. In: Archibugi D., Howells J., Michie, J., 1999. *Innovation policy in a global economy*. Cambridge University Press, pp. 67-92

HSRC, 2004. *South African National Survey of Research and Experimental Development 2001/02*.

IDRC, 1993. *Towards a Science and Technology Policy for a Democratic South Africa*. Ottawa: International Development Resource Centre.

Jan, T-S., Jan C-G, 2000. Development of weapon systems in developing countries: a case study of long range strategies in Taiwan. *Journal of the Operational Research Society*, Vol. 51, pp. 1041-1050.

Janszen F. H. A., Degenaars G. H., January 1998. A dynamic analysis of the relations between the structure and the process of National Systems of Innovation using computer simulation; the case of the Dutch biotechnological sector. *Research Policy*, Vol. 27, No. 38.

Johnson B. *Institutional Learning*. In: Lundvall B. A., 1992. *National Systems of Innovation*. Pinter: London, pp. 23-45

Jones, C.I., 1995. R&D based models of economic Growth. *Journal of Political economy*, Vol. 103 No. 4.

Kaplan, D.E. 1999. *On the literature of the economics of technological change*. The South African Journal of Economics, Vol. 67, No. 4.

Kleinbaum, D.G., Kupper, L.L., et al, 1998. *Applied Regression Analysis and Multivariate Methods*. Duxbury Press.

Kleijnin J. P. C., 1995. Sensitivity analysis and optimization of system dynamics models: Regression analysis and statistical design of experiments, *System Dynamics Review*, Vol. 11, No. 4, pp. 275-288.

Kline S.J., 1985. "Innovation is not a Linear Process". *Research Management*, Vol. 28, pp. 36-45.

Kline S.J. and Rosenberg N., 1986. *An Overview of Innovation in ihe Positive Sum Strategy: Harnessing Technology for Economic Growth*, The National Academy Press, Washington D.C.

Kahn, M., 2004. *Then and now: an innovation system in transition*. Human Science Research Council, Cape Town.

Kaplan, D., 1995. *The sate of Science in South Africa: New priorities, new policies*. Development Policy Research Unit Working Paper, University of Cape Town

Karlsson, C., Flensburh, P., Horte, S., 2004. *Knowledge spillovers and knowledge Management*. Edward Elgar, Cheltenham, UK.

Koekemoer, R., Du Toit, C., Ground, M., 2004. *Estimating Technological Progress for South Africa*. Ninth annual conference for econometric modelling for Africa.

Lall, S. 2003. *Indicators of the relative importance of IPRs in developing countries. Research Policy*. Working Papers qehwps85, Queen Elizabeth House, University of Oxford

Lane, D.C., 1998. *Can we have confidence in generic structures?*. Journal of the Operational Research Society, Vol. 49, pp. 936-947.

Legasto A.A., Forrester J.W., Lyneis J.M., 1980. System Dynamics. Studies in the Management sciences: *System Dynamics*, Vol.14.

Levinthal D., Myatt J., Winter 1994. Co-evolution of capabilities and industry: The evolution of mutual fund processing. *Strategic Management Journal*, Vol. 15, pp. 45-62.

Luna-Reyes F. L., Andersen D. L., Winter 2003. Collecting and analyzing qualitative data for system dynamics: methods and models. *System Dynamics Review*, Vol. 19, No. 4, p. 271.

Lundvall, B.A., Andersen, E.S., Dalum, B., 2002. National systems of production, innovation and competence building. *Research Policy*, Vol. 3 pp. 1213–1231.

Lundvall B. A., 1992. *National Systems of Innovation*. Pinter: London.

Mani, S., 2001. *Government and Innovation Policy: An analysis of the South African experience since 1994.* The United Nations University, UNU/INTECH Discussion Papers, ISSN 1564-8370.

Mansfield, E., 1986. The R&D Tax credit an Other Technology Policy Issues. *The American Economic Review*, Vol 76, No. 2, pp. 190-194.

Marais, H.C., 1999. *Perspectives on Science Policy in South Africa*. Network publishers, Menlo Park. South Africa.

Matthew R. McGrail, Claire M. Rickard, Rebecca Jones, 2006. Publish or perish: a systematic review of interventions to increase academic publication rates. *Higher Education Research and Development*, Vol. 1, No. 25, pp. 19 – 35.

McKelvey, M., 1991. *How do national systems of innovation differ?: a critical analysis of Porter, Freeman, Lundvall and Nelson.* In: Hodgson, G.M., E. Screpanti, 1991, *Rethinking economics. Markets, technology, and economic evolution*. London: Edward Elgar, pp. 117-137.

Meyer, M., 2002. Tracing knowledge flows in innovation systems. Jointly published by Akademiai Kiado, Budapest and Kluwer Academic Publishers, Dordrecht. *Scientometrics* Vol. 54, No. 2, pp. 193 – 212.

Meadows D. L., 1974, Behrens W. W. et al., *Dynamics of growth in a finite world*. Wright-Allen Press: Cambridge, Massachusetts.

Milling, P.M., 2002. Understanding and managing innovation processes. *System Dynamics Review*, Vol. 18, no 1, p. 73;

Moles R., O'Regan B., January 2003. Modelling policies and decisions: A case study in mineral extraction. Information & Management, Vol. 40, No. 3, pp. 147-157.

Morecroft J. D. W., 1998. System dynamics and micro worlds for policymakers. *European Journal of Operational Research*, Vol. 1, No. 59, pp. 9-27.

Morita-Lou, Hiroko, 1984. *Science and Technology Indicators for Development*. London: Westview Press.

Mouton J., Boshoff, S.C., 2001. *South African Science in Transition*. CENIS: Stellenbosch

Mowery D. C., Rosenberg N., 1989. *Technology and the pursuit of Economic Growth*. Cambridge University Press.

NACI, 2004. A profile of Postgraduate Higher Education and Academic Research Community in South Africa. National Advisory Council on Innovation (NACI).

NACI, 2002. *South African Science and Technology: Key facts and figures*, National Advisory Council on Innovation (NACI).

Nasierowski, W., Arcelus, F.J., 1999. Interrelationships among elements of national innovation systems: A statistical evaluation. *European Journal of Operational Research,* Vol. 119, pp. 235-253.

NCHE, 1996. *A Framework for Transformation*, Department of Education, Pretoria.

Nelson, R., R. N. 1993. *Technical innovation and national systems*. In: Nelson R.R., 1993. *National Innovation Systems: A comparative Analysis*. Oxford University Press: Oxford.

Nelson R.R., 1993. *National Innovation Systems: A comparative Analysis*. Oxford University Press: Oxford.

Nelson, R., 1995. Recent evolutionary theorizing about economic change. *Journal of economic literature*, Vol. 33, pp. 48-90.

Nelson, R.R. and Winter S.G., 1982. *An evolutionary theory of economic change,* Cambridge M.A: The Belknap press of Harvard university Press

Nicolis G., Prigogine L., November 2004. *The University. Facing up its European responsibilities*. The fourth European Framework program and research on complex systems.

Niosi J., February 2002. National systems of innovations are "x-efficient" (and x-effective). *Research Policy*, Vol. 31, No. 2, pp. 291-302.

North D.C. June 1994. Economic Performance Through Time. The American Economic Review, pp. 359-368.

OECD, 1992. *Technology and the Economy: The key relationships*, OECD, Paris.

OECD, 1994 b. *Using patent data as science and technology indicators, patent manual*. OECD, Paris.

OECD, 1996. *The OECD Jobs Strategy – Technology, Productivity and Job Creation*, Volume 1, OECD, Paris.

OECD, 1997. *Manual on the measurement of human resources devoted to S&T*, “Canberra Manual”, OECD, Paris.

OECD, 2002 a. *Frascati Manual. Proposed Standard practice for surveys on research and experimental development*, OECD, Paris.

OECD, 2002 b. *The Measurement of Scientific and Technological Activities: Proposed guidelines for collecting and interpreting technological innovation data, Oslo manual*. OECD, Paris.

OECD, 2002 c. *STI Outlook 2000, Science, Technology and Innovation: Recent Policy Development in South Africa*. OECD, Paris.

OECD, 2002 d. *Dynamising National Innovation Systems*. OECD, Paris.

Oerlemans L. A. G., Pretorius M. W. et al., December 2003. *South African Innovation Survey 2001*. University of Pretoria, Eindhoven University of Technology, report no. 120.

Oliva R., Dec 2003. Model calibration as a testing strategy for system dynamics models. *European Journal of Operational Research*, Vol. 151, pp. 552-568.

Page C., Meyer D., 2000. *Applied research design for business and management*. McGraw-Hill Company: Australia, NSW.

Park, H., Oh, S., Kim, S., 2004. *Leverage Strategy to National R&D Investment in Korea: A System Dynamics Approach.* System Dynamics Conference 2004

Porter M., 1990. *Competitive advantage of nations*. Free Press: New York.

Pouris, 1995. *Science and Technology policy in South Africa: A system in transition*. In: Van der Berg, H., Prinsloo, R., Pienaar, D., 1995. *Directory of human sciences research organizations and professional associations in South Africa*., Human Sciences Research Council, South Africa.

Pouris A., February 1992. Perspectives on the Institutional Framework for Science and Technology Policy-Making, *South African Journal of Science*, Vol 88.

Pouris, A., 2003. Towards a South African R&D tax incentives scheme: fiscal policies and social benefits. *South African Journal of Science*, Vol. 99, No.5, pp.195-1999.

Pouris, A., 2005. Technological performance judged by American patents awarded to South
African inventors. *South African Journal of Science*, Vol. 101, pp.23-26.

Pouris, A., 2006. Assessing scientific strengths of academic institutions: the example of the University of Pretoria. *South African Journal of Science*, Vol. 102, pp.23-26.

Prinsloo, C., January 1997. *An analysis of the links between academic research, competitiveness and the National System of Innovation of South Africa*, ITI working paper, no.97-1.

Repenning, N.P., 1999. *A Simulation-Based Approach to Understand the Dynamics of Innovation Implementation*. Sloan school of Management, MIT.

Roberts, R, June 1998. Managing innovation: The pursuit of competitive advantage and the design of innovation intense environments. *Research Policy*, Vol. 27, No. 2, pp.159-175.

Roberts, E.B. 1978. *Managerial Applications of system dynamics*. MIT Press/Wright – Allen series.

Richardson, G.P., 1999. Reflections for the future of system dynamics. *Journal of the Operational Research Society*, Vol. 50, pp. 440-449.

Romer, P.M., 1990. Endogenous Technological Change. *Journal of Technological Economy*. Vol. 98, pp. 71-102.

Rosenberg, N. 1976. *Perspectives on Technology*. Cambridge: Cambridge University Press.

RSA, 1998. Act 23 of 1998, The National Research Foundation Act, South Africa

Rubinstein M.F., Firstenberg I.R., 1995. *Patterns of Problem Solving*. Prentice Hall, Englewood, New Jersey.

SAS Institute Inc., 1995. *SAS/ETS software: time series forecasting system: version 6*. Cary, NC.

Saviotti, P. P., 1997. *Innovation systems and evolutionary theories*. In: Edquist, C., 1997, *Systems of innovation: technologies, institutions and organisations*. London: Pinter, pp. 180-222.

Smith, K. 1998. *Science, Technology and Innovation indicators – A guide for policy makers*. IDEA Report 5.

Smits P., den Hertog, P., van der Schaft, E., May 1992. *Initiatives to Strengthen the Interface Between EC&RD and Society*. TNO Centre for Technology and Policy Studies.

STATSSA, September 2003. *Labour force Survey–March 2003*. Statistics South Africa: Pretoria.

Sterman John, D., 2000. *Business Dynamics: Systems Thinking and Modeling for a Complex World*. Irwin McGraw-Hill

Sundbo, J., 1998. *The theory of innovation: entrepreneurs, technology and strategy*. Edward Elgar Publishers

UNESCO, 1998. *World Science Report 1998*, France.

UNDP, 2003. South Africa Human Development Report. The Challenge of Sustainable Development in South Africa: Unlocking People's Creativity. Oxford University Press: Oxford.

UP, 2005. University of Pretoria Bureau for Institutional Research and planning.

Van der Berg, H., Prinsloo, R., Pienaar, D., 1995. Directory of human sciences research organizations and professional associations in South Africa., Human Sciences Research Council, South Africa.

Verbeek, A., Debackere, K., Luwel, M., Van Looy, B., Andries, P., Van Hulle, M., Deleus, F., 2000. Linking science to technology: Using bibliographic references in patents to build linkage schemes. *Scientometrics*, Vol. 54, No. 3, pp.399–420

Verspagen, B., 2005. *The impact of Academic Knowledge on Macroeconomic productivity Growth. An exploratory study*. Eindhoven Centre of Innovation Studies. A study commissioned by the Adviesraad voor Wetenschaps- en Technologiebeleid (AWT)

Weil, B.H., Bergan, T.A., Roberts, E,B., *The Dynamics of R&D strategy*. In: Roberts, E.B. 1978. *Managerial Applications of system dynamics*. MIT Press/Wright –Allen series.

Werker, C., Brenner, T., 2004. *Empirical Calibration of Simulation Models*. ECIS, Eindhoven, Netherlands, Working Paper 04.13.

Wolstenholme EF, Coyle RG. 1983. The development of system dynamics as a rigorous procedure for system description. *Journal of the Operational Research Society*, Vol. 34, pp.569–581.

Wolstenholme E.F., 1999. Qualitative vs Quantitative modeling: The evolving balance. *Journal of the Operational Research Society*, Vol. 50, pp. 422-428.

Yaffe, R., MacGee, M., 2000. *Introduction to time series analysis and forecasting with applications in SAS and SPSS*. Academic Press, San Diego, California

Zhang, S., Qu, R., 2004. *Research on the System Dynamics Model for the Development and Operation of the Sustainable Regional Innovation System*. Proceedings of Globelics conference 2004, Beijing.

12 APPENDIX A

12.1 South Africa's Patenting at the USPTO

The database used for data analysis on the patenting statistics of South Africa at the United States Patent Office was obtained from the National Bureau of Economic Research (NBER) in Cambridge, Massachusetts in the Unite States of America. This is available for free download at the following location: http://www.nber.org/patents/

The main data set extends from January 1, 1963 through December 30, 1999 (37 years), and includes (Hall et al, 2001):

- All the ***utility*** patents granted during that period, totalling 2,923,922 patents
- The citations file, includes all citations made by patents granted in 1975-1999, totalling 16,522,438 citations.
- Data on inventors and assignees.

The following table includes fields in the database, which are relevant to the information extracted in this study. For a more detailed account of what is included in the database, see NBER (2001).

Table 12-1: Field included in the database relevant to the analysis for this study

Field Name	Field Used
Patent number	
Grant year	
Grant date6	
Application year (starting in 1967)	Used
Country of first inventor	Used
State of first inventor (if U. S.)	
Assignee identifier, if the patent was assigned (starting in 1969)	
Assignee type (i.e., individual, corporate, or government; foreign or domestic)	Used
Technological category	
Technological sub-category	
Number of citations made	Used
Number of citations received	Used
Percent of citations made by this patent to patents granted since 19637	
Measure of "generality"	

Each patent document includes the date when the inventor filed for the patent (the ***application*** date), and the date when the patent was granted. The data contains the grant ***date*** and the grant ***year*** of all patents in the file (i.e., of all utility patents granted since 1963) and the application ***year*** for patents granted since 1967. The grant date depends upon the review process at the Patent Office, which takes on average about 2 years with a significant variance (NBER, 2001).

For the analysis, the actual timing of the patented inventions is closer to the application date than to the (subsequent) grant date. This is so because inventors have a strong incentive to apply for a patent as soon as possible following the completion of the innovation.

The following table lists the Patents filed where the first inventor was a South African. For the purpose of this discussion, from this point forward this subset of data within the database is referred to as "SA Patents" .The "References" column lists the summation of all the citations made from the patent to prior art.

Table 12-2: South African Patent outputs at the USPTO

Application Year	World patent output	SA patents registered at USPTO	References
1980			
1981	63910	71	426
1982	65009	78	502
1983	61563	74	424
1984	67071	90	504
1985	71442	89	622
1986	75088	89	692
1987	81458	130	908
1988	90134	95	654
1989	96077	124	839
1990	99254	98	733
1991	100016	88	619
1992	103307	96	765
1993	106848	143	1272
1994	120380	138	1204
1995	137661	98	922
1996	131450	107	928
1997	114881	81	739
1998	33780	25	217
1999	1560	2	25
2000	-	-	-
2001	-	-	-

Hall et al (NBER ,2001) warns against ***truncation*** problem: as the time series move closer to the last date in the data set, patent data timed according to the application date will increasingly suffer from missing observations consisting of patents filed in recent years that have not yet been granted. This issue is dealt with by only making use of the time series up to 1996.

Type of assignees

The USPTO classifies patents according to the type of assignees (owners of the patent rights), into the following seven categories (the figures are the percentages of each of these categories in our data):

Table 12-3: USPTO Assignee type categories

Assignee #	Assignee type
1	Unassigned
2	US non-government organization (mostly corporations)
3	Non-US non-government organization (mostly corporations)
4	US individuals
5	Non-US individuals
6	US Federal Government
7	Non US Government

"Unassigned" patents are those for which the inventors have not yet granted the rights to the invention to a legal entity such as a corporation, university or government agency, or to other individuals. The original inventors thus still owned these patents at the time of patenting, and they may or may have not transferred their patent rights at a later time (the NBER do not have data on transfers done after the grant date).

The SA Patent dataset was now further analysed, the analysis yields the following distribution of patents to the assignee types. The year is the application year – the date earliest to the actual invention date.

Table 12-4: South African Patent counts at the USPTO

	Category number							
Year	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**Total**
1977	43	1	26	0	4	0	0	74
1978	39	4	40	0	4	0	0	87
1979	42	2	44	0	6	0	0	94
1980	35	6	42	0	3	0	0	86
1981	28	1	37	0	4	0	1	71
1982	25	2	49	0	1	0	1	78
1983	28	3	42	0	1	0	0	74
1984	38	1	46	0	5	0	0	90
1985	28	3	56	0	2	0	0	89
1986	34	1	47	0	7	0	0	89
1987	66	2	57	0	4	0	1	130
1988	43	2	46	1	3	0	0	95
1989	48	4	65	0	5	0	2	124
1990	38	5	50	0	3	0	2	98
1991	38	3	44	0	2	0	1	88
1992	36	5	53	0	2	0	0	96
1993	62	3	74	0	4	0	0	143
1994	53	8	74	0	3	0	0	138
1995	39	5	47	0	7	0	0	98
1996	32	5	61	1	8	0	0	107
1997	29	11	62	0	1	0	0	103
1998	31	11	64	0	3	0	0	109
1999	30	4	73	1	5	0	0	113

It can be noticed from the above table that the two groups with the most patents are the “unassigned” group as well as the “Non-US non-government organization (mostly corporations)” group.

Certain assumptions are made regarding the assignee classifications and the implications on categorising it to the three sectors in the model developed in this study (Higher Education sector, Public sector and Private sector)

- All patents in category 2 and 3 are patents originating from companies in the “Private sector”
- An assumption is made that the unassigned entities will never be assigned and will remain in the names of the inventors who originally submitted the application for patents.

These assignee codes are then assigned to one of three categories.
- No Sector: Unassigned (1) and Individuals (4,5)
- Private sector: Non governmental organisations (2, 3)
- Governmental organisations (7) – Public sector

This yields the following result:

Table 12-5: Patent count analysis of South African Patents at the USPTO

Patents	No sector	Private sector	Public sector	Total
1977	47	27	0	74
1978	43	44	0	87
1979	48	46	0	94
1980	38	48	0	86
1981	32	38	1	71
1982	26	51	1	78
1983	29	45	0	74
1984	43	47	0	90
1985	30	59	0	89
1986	41	48	0	89
1987	70	59	1	130
1988	47	48	0	95
1989	53	69	2	124
1990	41	55	2	98
1991	40	47	1	88
1992	38	58	0	96
1993	66	77	0	143
1994	56	82	0	138
1995	46	52	0	98
1996	41	66	0	107
1997	30	73	0	103
1998	34	75	0	109
1999	36	77	0	113

12.1.1 Absorption of knowledge from the USPTO database

Table 12-6: South African Patent reference counts at the USPTO

	Category number							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**Total**
1977	256	30	144	0	18	0	0	448
1978	272	30	243	0	26	0	0	571
1979	302	12	294	0	45	0	0	653
1980	226	30	233	0	13	0	0	502
1981	201	5	178	0	31	0	11	426
1982	180	15	299	0	5	0	3	502
1983	172	22	223	0	7	0	0	424
1984	225	10	240	0	29	0	0	504
1985	199	39	368	0	16	0	0	622
1986	279	3	377	0	33	0	0	692
1987	494	21	358	0	32	0	3	908
1988	284	11	333	11	15	0	0	654
1989	331	37	451	0	19	0	1	839
1990	265	49	388	0	18	0	13	733
1991	284	4	312	0	19	0	0	619
1992	324	42	388	0	11	0	0	765
1993	546	24	632	0	70	0	0	1272
1994	446	165	575	0	18	0	0	1204
1995	351	79	418	0	74	0	0	922
1996	336	44	469	4	75	0	0	928

1997	159	120	460	0	0	0	0	739
1998	45	15	157	0	0	0	0	217

Table 12-7: Patent reference count analysis of South African Patents at the USPTO

Patents	No sector	Private sector	Public sector	Total
1977	274	174	0	448
1978	298	273	0	571
1979	347	306	0	653
1980	239	263	0	502
1981	232	183	11	426
1982	185	314	3	502
1983	179	245	0	424
1984	254	250	0	504
1985	215	407	0	622
1986	312	380	0	692
1987	526	379	3	908
1988	310	344	0	654
1989	350	488	1	839
1990	283	437	13	733
1991	303	316	0	619
1992	335	430	0	765
1993	616	656	0	1272
1994	464	740	0	1204
1995	425	497	0	922
1996	415	513	0	928
1997	159	580	0	739
1998	45	172	0	217
1999	36	77	0	113

12.1.2 South Africa's Patenting at the South African Patent Office

The South African Patent office works differently from the United States Patent Office. Patents are not examined in order to be granted by the South African Patent Office.

This by implication has the effect that the patents therefore is not necessarily novel, in relation to existing art. It does not necessary indicates a degree of inventiveness as is necessary to patent at the European or United States Patent Office. Since patenting at the USPTO is extremely expensive and only a very small number of South African patents are granted on a yearly basis, the possibility of making use of the South African patent office data was examined.

Finding reliable data on the South African patent database, specifically to find data on patents granted to South Africans proved to be a very frustrating and almost impossible task. Unlike the NBER database that exists for the USPTO, no such database exists at present for the South African Patent Office.

On acceptance of a patent by the patent office, the patent description is published in the South African Patent Journal. The patent is granted on the publication date of the appropriate issue of the Patent Journal. The Patent journal is published every month.

A file system with the patent journals scanned in (viewable only in pdf format) was obtained from Hahn & Hahn Patent attorneys. The patent journals (published monthly) for the period 1986 to 2004 were examined. All patent granted to a South African entity (priority country must be South Africa).

The ***priority country*** is the country where the patent is filed the first time. It is therefore assumed that the great majority of South Africans patenting will first patent in South Africa after which they might or might not patent their ideas in other countries. It is acknowledged that this approach has it's weaknesses, but still is the most simplified and the only feasible option to be able to search for South African patents in the South African Patent Journal.

This approach however also has it stronger points since by looking at patent published in the Patent Journal, only granted patents are taken into account. Although this is not examined, the process of self-elimination by patentees is used as a filter mechanism.

Table 12-8: Patent data gathered from the South African Patent Office journal

	company	individual	university	government	Total
1985	402	357	2	20	781
1988	410	471	1	16	898
1990	391	438	2	30	861
1993	364	374	3	27	768
1995	407	316	7	34	764
1998	431	385	19	22	857
2000	320	296	15	18	649
2001	353	355	16	14	738
2004	331	359	15	18	723

The following trends can be seen from the data gathered from the South African patent journal.

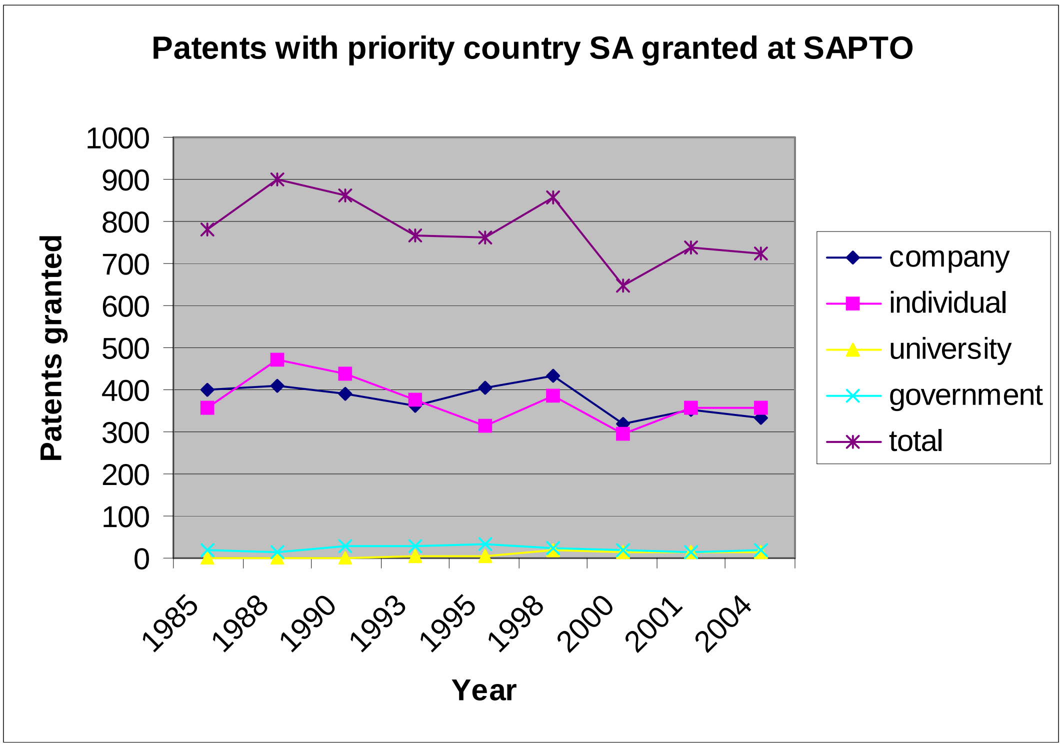

Figure 12-1 Patents granted at the South African Patent Office

12.2 South Africa's Publication data: Data gathering and Analysis

12.2.1 The Database of scientific papers originating from South Africa

A Database of all scientific papers in the ISI Web of science written by at least one author with a South African address was constructed from data. The database is constructed from text files downloaded from the ISI web of science database (Science Citation Index, Social Sciences Citation Index, Arts and Humanities Citation Index). These text files were then imported into an excel spreadsheet. The scientific publication was included in the database if the author's address or the Reprint address is a South African address.
The data was gathered from access to the ISI Web of science via the following organisations:

- Years 1987-2004 from Eindhoven University of Technology, Netherlands
- Years 1981-1986 from the Radboud University, Netherlands

The most important fields included in the data downloaded are listed in the following table.

Table 12-9: Fields in Database of South African Scientific Output

Name of Field	Description
AU	Author names
TI	Title
SO	Journal Name
C1	Author address: Addresses of authors writing scientific Publication
RP	Reprint address: Address where permission must be asked if the article is to be reprinted
CR	List of References made in the Journal
NR	Number of References made in the Journal
TC	Times Cited
PY	Publication Year
SC	Scientific Field

Field important for the analysis of the data are the

- RP and CI (Reprint address and the Author's address): According to these addresses the paper is categorised into either the Higher Education sector, Public sector and the Private sector. In the case when the addresses referring to a specific scientific publication has two sectors, each sector is awarded a 0.5 publication.
- NR: The number of references to other scientific output from the scientific publication
- TC: Times the paper has been cited by another author publishing scientific output

12.2.2 Analysis of the scientific publication data in the database

An analysis is conducted on the database to establish the distribution of publication rates per year for the three sectors modelled in the model building study. In other words the sampling exercise is executed to find answers to the following questions:

Q1: From the years 1981 to 2001 what was the amount of papers published in ISI journals by people from the Higher Education sector, Public sector and Private sector respectively?

Q2: From the years 1981 to 2001 what was the amount of references made in scientific publications created in the Higher Education sector, Public sector and Private sector respectively?

Since the amount of papers included in the database is too large to analyse the information by examining every single record, a sampling strategy is followed. Since much variation in the address names exist, the safest option of obtaining and categorising the journals into one of the sectors, was through inspection of the addresses.

The error margin cannot be controlled for in these circumstances. Whereas by following the strategy of doing a sample on the data, we can get a pretty good idea of the inaccuracy introduced though the whole process.

Methodology

A stratified random sample was taken from the database. In the Excel database, every 20th paper was extracted to a new "Sample database". This led to the following number of papers being extracted to a second excel spreadsheet. The column with the "References in papers" is a summed total of the references made in the papers that have been sampled.

Table 12-10: Sample taken for analysis of the distribution of Publications

Year	Total # Papers	Total # References made in papers	Stratified Random sample size
1981	3075	39113	153
1982	3518	41868	175
1983	3581	43009	179
1984	3461	46950	173
1985	3968	56568	198
1986	4694	66769	234
1987	4758	66235	239
1988	4632	65301	231
1989	4183	62072	209
1990	3949	66850	197
1991	4134	73512	206
1992	4000	72629	200
1993	4195	73382	209
1994	4291	84125	214
1995	4503	85292	224
1996	4479	96584	224
1997	4398	100120	220
1998	4498	103269	225
1999	4755	112023	237
2000	4461	109612	223
2001	4691	122392	234
Total	**88224**	**1587675**	**4404**

The selected entries in the sample database was then categorised into the three sectors through the classification system as discussed in the Frascati manual:

- Higher Education sector (Universities and Technikons, Academic Hospitals etc.)
- Public sector (Science Councils, National Facilities)
- Private sector (Companies)

Table 12-11: Results from the Sample

	# HES papers	# Pub papers	# Bus papers		# HES references	# Pub references	# Bus references	
1981	113	32	15		1371	385	150	
1982	126	44	16		1509	539	200	

1983	128	49	10		1426	616	18	
1984	137	36	7		1684	528	40	
1985	149	49	10		2040	740	77	
1986	178	56	15		3031	798	71	
1987	189	64	6		2359	571	41	
1988	192	54	7		2596	740	23	
1989	164	59	4		2557	908	19	
1990	155	51	8		2562	704	89	
1991	166	50	11		2878	910	46	
1992	168	42	8		3564	468	120	
1993	176	42	8		3286	572	90	
1994	174	47	8		3133	412	174	
1995	191	42	10		3175	585	114	
1996	177	52	14		4129	1301	126	
1997	179	44	18		3785	678	144	
1998	172	52	13		3854	1068	237	
1999	188	43	21		4857	1006	309	
2000	180	45	16		4781	1008	273	
2001	202	35	11		5004	804	86	

The Margin of error is computed through the following computation:

By taking a stratified sample, the effective sample size is increased by a factor between $n/0.8$ to $n/0.9$ if n denotes the sample size.

Since the sample size is 4404, and taking the design effect computation to be $n/0.85$ the effective sample size (N) therefore is 5181. To compute the margin of error the following equation is used (Page and Meyer, 2000):

Margin of error = $2*\sqrt{\frac{p(p-1)}{N}}$

Die margin of error is therefore computed to make use of the most conservative value for $p = 0.5$. This results in a Margin of 0.0139 or approximately 1%.

From the sample analysis, the percentage of papers created in the three sectors is computed. The same is done for the percentage of references made in the papers.

Table 12-12: Distribution of scientific paper output and reference counts

	HES % papers	Pub % papers	Bus % papers		HES % references	Pub % references	Bus % references	
1981	0.72	0.19	0.09		0.72	0.20	0.08	
1982	0.68	0.23	0.09		0.67	0.24	0.09	
1983	0.69	0.26	0.05		0.69	0.30	0.01	
1984	0.77	0.19	0.04		0.75	0.23	0.02	
1985	0.73	0.23	0.04		0.71	0.26	0.03	
1986	0.73	0.21	0.05		0.78	0.20	0.02	
1987	0.75	0.23	0.02		0.79	0.19	0.01	
1988	0.78	0.19	0.03		0.77	0.22	0.01	
1989	0.74	0.24	0.02		0.73	0.26	0.01	
1990	0.74	0.22	0.04		0.76	0.21	0.03	
1991	0.75	0.20	0.05		0.75	0.24	0.01	
1992	0.79	0.17	0.04		0.86	0.11	0.03	
1993	0.80	0.17	0.03		0.83	0.14	0.02	
1994	0.78	0.19	0.03		0.84	0.11	0.05	

1995	0.81	0.15	0.04		0.82	0.15	0.03	
1996	0.74	0.21	0.05		0.74	0.23	0.02	
1997	0.76	0.17	0.08		0.82	0.15	0.03	
1998	0.74	0.21	0.05		0.75	0.21	0.05	
1999	0.77	0.16	0.08		0.79	0.16	0.05	
2000	0.77	0.17	0.06		0.79	0.17	0.05	
2001	0.84	0.12	0.04		0.85	0.14	0.01	

The next step now is to apply these findings to the total population to get an estimation of the total papers produced in the different sectors by year from 1981 to 2001. The percentages computed and displayed in Table 12-12: is used to find an estimate of how many papers were created in a sector for the whole population.

Table 12-13: Scientific paper publication counts and reference counts

	HES papers	Pub papers	Bus papers		HES references	Pub references	Bus references	
1981	2214	584	277		28161	7823	3129	
1982	2392	809	317		28052	10048	3768	
1983	2471	931	179		29676	12903	430	
1984	2665	658	138		35213	10799	939	
1985	2897	913	159		40163	14708	1697	
1986	3427	986	235		52080	13354	1335	
1987	3569	1094	95		52326	12585	662	
1988	3613	880	139		50282	14366	653	
1989	3095	1004	84		45313	16139	621	
1990	2922	869	158		50806	14039	2006	
1991	3101	827	207		55134	17643	735	
1992	3160	680	160		62461	7989	2179	
1993	3356	713	126		60907	10273	1468	
1994	3347	815	129		70665	9254	4206	
1995	3647	675	180		69939	12794	2559	
1996	3314	941	224		71472	22214	1932	
1997	3342	748	352		82098	15018	3004	
1998	3329	945	225		77452	21686	5163	
1999	3661	761	380		88498	17924	5601	
2000	3435	758	268		86593	18634	5481	
2001	3940	563	188		104033	17135	1224	

It is clear from the data that the Business sector makes a much smaller contribution in terms of scientific paper output than the HES or Public sector. For this reason the production of paper in the Business sector is not included in the model.

12.2.3 The Depreciation of knowledge

Adams (1990) developed a production function model to measure the impact of 'fundamental stocks of knowledge' productivity growth at the sectoral level. In this model he makes use of stocks of publication data. Adams also made use of an accumulated stock of knowledge in his model.

$$N_t = N_{t-1}(1-\delta) + P_t \qquad \textbf{12-1}$$

where N stands for the stock of knowledge in period t (or t-1), δ is a depreciation factor (estimated to be equal to 0.13), and P is the number of papers published in year t.

In a study conducted by Cabellero and Jaffe, the found that the rate of ideas' obsolescence has increased from 3% in the 1900's to up to 12% in 1990. This percentage is also close to the estimated value form Adams' (1990) study.

The depreciation rate is also used for the rates of depreciation of stocks of knowledge in the model developed in this study. As citation data is available for the South African papers, it is used to check this estimation for the purpose of this model.

The citation pattern of the Papers generated in South Africa is investigated. This analysis is done from the database of papers generated in South Africa. It is assumed that the citation pattern can be used for the papers generated in all the sectors. i.e. it is assumed that the difference in citation patterns between sectors is negligible.

Table 12-14: Citations received by South African Scientific journals.

	# Papers	**# citations received**	**avg # citations received per paper**
1981	3075	25229	8.20
1982	3518	29365	8.35
1983	3581	27855	7.78
1984	3461	28404	8.21
1985	3968	31722	7.99
1986	4694	35210	7.50
1987	4758	34459	7.24
1988	4632	35926	7.76
1989	4183	31896	7.63
1990	3949	36425	9.22
1991	4134	36418	8.81
1992	4000	33066	8.27
1993	4195	33580	8.00
1994	4291	33835	7.89
1995	4503	33971	7.54
1996	4479	31217	6.97
1997	4398	29101	6.62
1998	4498	29776	6.62
1999	4755	27818	5.85
2000	4461	23627	5.30
2001	4691	20726	4.42
2002	5068	14893	2.94
2003	4990	9553	1.91

When the data is presented as a curve the following can be seen.

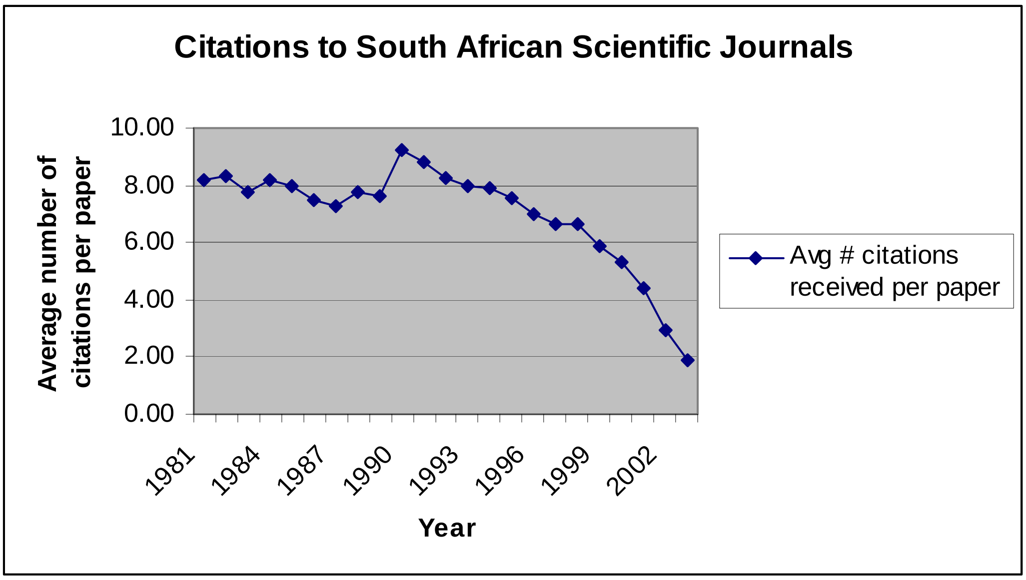

Figure 12-2 Citations to South African scientific papers and journals

It seems that the citations curve reaches an average maximum at around 1994 to 1992. From 2005 this is about 10 to 13 years. It can therefore be concluded that the estimation from Adams' model is a realistic assumption for the system.

From this reasoning the paper citation graph for South African scientific papers is therefore used to estimate an average period scientific knowledge remains relevant to the scientific community. The time knowledge remains relevant is approximated to be about 10 years (after which the citation curve seems to flatten).

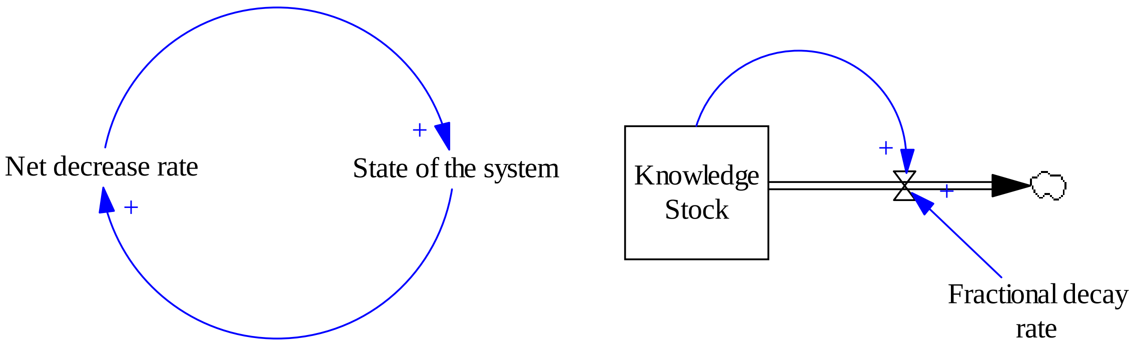

Figure 12-3 Dynamic structure of the decay of knowledge

The decay rate of knowledge is modelled through a first order material delay.

Outflow = Knowledge Stock/AT, where AT is the average delay time.

13 APPENDIX B

13.1 Inputs to R&D

The Frascati manual is devoted to measuring R&D inputs. This includes formal R&D in formal R&D units as well as informal or occasional R&D in other units. For statistical purposes the following R&D inputs are measured (OECD, 2002:20):

- Full-time-equivalent or person years spent on R&D
- Expenditures for R&D performed within a statistical unit or sector in the economy (Intramural expenditures).
    - Current costs: labour costs and other current costs.
    - Capital expenditures: annual gross expenditures on gross fixed assets.
- Extramural expenditures, which cover payment of R&D performed outside the statistical unit sector.

R&D is an activity involving significant transfers of resources among units, organisations and sector (especially between government and other performers). The aim of the Frascati manual is to establish specifications for R&D input data and therefore to establish specifications for the collection of the data.

In order to facilitate the collection of data, the description of flows of R&D funds and the interpretation of R&D data, the data gathered from reporting units (units from where data is collected) are grouped into sectors of the economy (OECD, 2002:53). The sectors the aggregates for R&D data is grouped and includes the following:

- Business enterprise sector
- Government
- Private non-profit
- Higher education
- Abroad

International aggregate expenditure comparisons are done on the gross domestic expenditure on R&D (GERD) performed on national territory in a given year. This includes R&D financed from abroad but excludes R&D funds performed abroad.

13.1.1 Types of research

In the description of the South African R&D system, reference will be made to different types of R&D performed within different sectors. These concepts are therefore shortly defined as in the Frascati manual (OECD, 2002 a):

***Basic Research*:** Basic research is experimental or theoretical work undertaken primarily to acquire new knowledge of the underlying foundations of phenomena and observable facts, without any particular application or use in view.

Pure Basic Research is carried out for the advancement of knowledge without seeking long-term economic or social benefits or making any effort to apply the results to practical problems or to transfer the results to sectors responsible for their application.\

Oriented Basic Research (Strategic Research) is carried out with the expectation that it will produce a broad base of knowledge likely to form the basis of the solution to recognised or expected, current or future problems or possibilities.

Applied Research: Applied research is an original investigation undertaken in order to acquire new knowledge. It is however, directed primarily towards a specific practical aim or objective.

Experimental Development: Experimental development is systematic work, drawing on knowledge gained from research and practical experience that is directed to new processes, systems and services; or to improving substantially those already produced or installed.

13.1.2 Measurement of R&D output

The reason interest in R&D depends more on the social and economic effects through the acquisition of new knowledge than on the activity itself. R&D output indicators are far more complicated to determine, define and produce.

The Output of R&D or science and technology (S&T) in general can be measured in several ways (OECD, 2002 a):

- Innovation surveys or existing data sources are methods of measuring the effects of the innovation process of which R&D plays an important role.
- Existing data sources such as bibliometrics, patent data, and the analysis of trade data in terms of "technology intensity" of products or industries concerned.

There has been only one Oslo-type innovation survey carried out by the University of Pretoria in co-operation with the Eindhoven University of Technology in the Netherlands (Oerlemans L. A. G., Pretorius M. W. et al, 2003)

Since time series data is important for the successful completion of the study, R&D outputs will be measured through existing data sources. The use of these sources is however not without problems. The following sections deals with limitations of the use of these indicators.

13.1.3 South African Frascati style R&D surveys

Blankley and Kahn (2005) published a paper on the history of South Africa's Frascati style surveys. The following sector discusses the methodological issues and measurement issues that exist in the time series data of the SA R&D survey data as well as the measurement of R&D in general.

Definition has expanded

Blankley and Kahn (2005) published a paper on the history of South Africa's Frascati style surveys. R&D is not as well defined and it is not easy to establish its boundaries as one would like to have it. There is a global debate about what qualifies as R&D;

this debate shifts as global changes occur regarding economic activity. Changes have occurred in the definition of R&D over the past 30 years. Where it was mainly hard technologies that were measured about 30 years ago, the measurement The definition has been expanded in some instances from only including 30 years ago only hard technologies now also includes services and software development.

Who does the reporting?

Measurement of the R&D is also influenced by who does the reporting. Researchers, managers and financial people are unlikely to report the same estimates of R&D spending – they all have different approaches. This may also influence the accuracy with which the survey is conducted.

South Africa's R&D surveys and changes in methodology

South Africa has been conducting Frascati-style R&D surveys since shortly after the first manual was released in 1963 (Boshoff *et al*, 2003). The first R&D survey based on OECD guidelines was conducted in 1966 and over the next 25 years up to 1993/4, South Africa conducted 18 regular official surveys. Social sciences R&D expenditure have only been included in the survey since 1977/78

The CSIR in partnership with the HSRC started to conducted survey fieldwork. The survey has been conducted under the auspices of different bodies until 1991.

- The CSIR conducted the R&D survey for Natural sciences
- The HSRC was responsible for the gathering of data for social sciences from 1977/8.

In general the 1991/2 survey was more thorough than the surveys from prior years. 57% increase in R&D expenditure from 1989/90 to 1991/2 in comparison with an increase of 34% between 1987/8 to 1989/90. This is largely so because of the companies surveyed. The register with the businesses surveyed in the Business sector was revised and populated with 44% more enterprises than the 1989 survey. Defence organisations were also included in the survey which was not included in the 1989 survey. The defence R&D expenditure made up roughly 15% of the R&D budget in 1991.

The CSIR in partnership with the HSRC started to conducted survey fieldwork. The survey has been conducted under the auspices of different bodies until 1991. The survey was put open on tender from after 1991/2. After 1991 it was of lower priority to the state resulting in the 1993 and 1997 tenders were being awarded to private consultancies. Kahn (2004) comments on the inconsistency in the time series data after 1991 as follows: "Unfortunately instability after 1991 led to the conduct of the Survey migrating across a number of agencies with consequent inconsistency of methodology, gaps in the time series; and loss of institutional memory and capacity. However the Surveys are the only series of R&D data and therefore must be used. In addition to the Frascati data one has higher education, grant maker and bibliometric databases that can be used to corroborate evidence."

The survey was carried out on biannual basis until 1993/4. No survey was carried out in the 1995/6 (possibly because of the National Research and Technology Audit) or

1999/2000 cycles, leaving gaps in the time series data (HSRC, 2003: 211; Boshoff, 2003). The 1993/4 and 1997/8 R&D surveys used a different methodology from the R&D surveys conducted in the past. The survey was handed over to a private consultant for the 1993/4 and 1997/8 survey years. In the 1993/4 R&D surveys a similar approach of previous years were followed. Yet more companies were added to the register of companies surveyed. Pouris (2006) states that the 1993/4 and 1997/8 methodology is comparable to the methodology used in the previous years when considering the data gathered for the Private and Public sector. The main methodological differences in the surveys exist for the 1997/8 survey with regards to the Higher Education Sector.

Up to 1997 SAPSE data was used for the calculation of HR data in the HES. Higher Education institutions were also surveyed regarding the time spent on R&D and R&D expenditure in these institutions. In the 1997 survey another approach was followed. The SAPSE data was utilised but the percentage time spent by researchers on R&D was estimated by making use of research coefficients. Universities were arranged in high, medium and low intensity groups. Research coefficients were used to estimate the research time spent in these institutions as well as the R&D expenditure in these institutions. In some instances the Non-profit and Public sector is combined as the Services sector. This however does not have a big impact since the non-profit sector makes a very small contribution to R&D expenditure.

Pouris states the following regarding the 1993/4 and 1997/8 surveys: "During the surveys I supervised dip-stick surveys were undertaken of the remaining enterprises in order to identify new-comers. Each time approximately 1000 enterprises were approached. Businesses are included according to various criteria, namely:
1) on the basis that they were included in previous surveys,
2) if they received R&D funds from government programmes,
3) they are identified by respondents as contractors of research
4) have been identified by journals, popular press etc as undertaking R&D. "

Mouton (2001:44) reports a suspicion that the 1997/8 R&D survey might underestimate R&D spending in South Africa.

No survey was carried out in the 1995/6 (possibly because of the National Research and Technology Audit) or 1999/2000 cycles, leaving gaps in the time series data (HSRC, 2003: 211). Another inconsistency with the methodology followed in the 1997/8 R&D Survey is that it does not discriminate between the Non-profit and Public sector.

For the years 2001/2 R&D survey, the responsibility of executing the survey was handed to the HSRC. The methodology followed in these surveys is comparable to that followed in the 1991 survey. The Higher Education sector was again fully surveyed as with the pre-1997 surveys.

Again the register with businesses has been updated with more businesses known to conduct R&D. In the 2001/2 survey a number of 139 Business BERD questionnaires were received back. The register was increased and 2003/4 a number of 339 non-nil response BERD questionnaires were received.

It can however be argued that although the increase in businesses surveyed no doubt leads to a potentially more accurate survey, the top 20 business sector performers accounted for 79% of business expenditure. Therefore we can conclude that a change in the size of the sample will add to accuracy but that the general level of R&D expenditure should not be affected too much.

We can therefore conclude from this discussion that there were some methodological changes in the surveys. Concentration however aids in the accuracy of the different surveys.

There are definite methodological inconsistencies in the survey method for the Higher Education Sector over the time period in consideration. For this reason the Higher Education Sector model does not make use of the Survey data, but makes use of HEMIS data in terms of Academic and Research staff employed in the system.

13.1.4 The HEMIS data base

A telephonic interview was conducted with Jean Skene the director of HEMIS at the Department of Education on 14 March 2006. In this interview it was formally confirmed that the comparison of the data in the HEMIS database between different years is appropriate.

The data is gathered from the Higher Education Institutions with the categories of the human resources identified as by the Department of Education. The definitions of the fields of data gathered from the Higher Education institutions remained consistent over the time period in consideration. No changes in definitions of the HR component have been made in the time series data.

From 1986 to 1998 Public Higher Education Institutions were required to submit data in the format of aggregated tables for headcounts, graduates and full time equivalent students and staff as specified by the Department of Education. The submission of data was stream lined in 1999: From 1999 the Public Higher Education Institutions were required to submit unit record databases for students and staff. These institutional databases are then loaded into a National database from which the Department of Education generates the aggregated tables.

From this can therefore be concluded that the use of time series data in the HEMIS database is therefore appropriate for the purpose of this study.

13.2 The Higher Education Sector

13.2.1 R&D Expenditure in the HES

R&D Survey data is gathered from the R&D Surveys (1977 to 2003). The following table reflects figures for R&D spending in the Higher Education sector. This is referred to in the main body of the thesis.

Table 13-1: Sector Source funding (Financiers) of the HES

Year	HES (R)	Public sector (R)	Private sector (R)
1977	30126000	8196000	1909000
1979	47409000	10222000	2903000
1981	72671136	15263000	5408000
1983	120989000	17489000	10403000
1985	239731000	40818000	22346000
1987	261135000	43874000	31151000
1989	406693000	56570000	51742000
1991	552457000	68914000	65882000
1993	336708000	46462000	29574000
1995	N/A	N/A	N/A
1997	406000000	33000000	57000000
1999	N/A	N/A	N/A
2001	581560000	1187075000	380075000
2003	346132000	848,554,000	478734000

Table 13-2: R&D expenditure in the Higher Education sector

Year	R&D investment (R)	Expenditure HR	% Expenditure HR	Expenditure on Capital	% Expenditure on Capital
1977	40944000	20481000	50.02%	2311000	5.64%
1979	62109000	33306000	53.63%		0.00%
1981	94424210	34053000	36.06%	5618240	5.95%
1983	151352000	53124000	35.10%	5985000	3.95%
1985	306534000	173232000	56.51%	13496000	4.40%
1987	339194000	175463000	51.73%	14826000	4.37%
1989	517566000	258324000	49.91%	35879000	6.93%
1991	690439000	341904000	49.52%	30062000	4.35%
1993	415648000	230435000	55.44%	15669000	3.77%
1995	N/A	N/A	N/A	N/A	N/A
1997	496000000	253100000	51.03%	N/A	N/A
1999	N/A	N/A	N/A	N/A	N/A
2001	1896156000	1127710000	59.47%	115953000	6.12%
2003	2071351000	925255000	44.67%	162380000	7.84%
Average			49.42%	Average	4.85%
Standard Deviation			7.51%	St Dev	2.07%

Table 13-3: R&D expenditure on Human resources in the HES

Year	Total HR Spending (R'000)	Researcher (R'000)	%	Technicians (R'000)	%	Support (R'000)	%
1977	20480	17721	86.53%	2232	10.90%	527	2.57%
1979	33306	28014	84.11%	4445	13.35%	847	2.54%
1981	34052	25814	75.81%	6903	20.27%	1335	3.92%
1983	53124	43357	81.61%	7715	14.52%	2052	3.86%
1985	173232	146712	84.69%	17995	10.39%	8525	4.92%
1987	175463	159281	90.78%	10783	6.15%	5399	3.08%
1989	258324	225510	87.30%	25230	9.77%	7584	2.94%
1991	341904	323377	94.58%	13179	3.85%	5348	1.56%
1993	230435	220381	95.64%	5734	2.49%	4320	1.87%
1995	N/A	N/A	N/A	N/A	N/A	N/A	N/A

1997	253100	240445	95.00%	7593	3.00%	5062	2.00%
1999	N/A	N/A	N/A	N/A	N/A	N/A	N/A
2001	1127710	1071325	95.00%	33831	3.00%	22554	2.00%
Average			88.28%		8.88%		2.84%
Standard Deviation			6.52%		5.76%		1.04%

Table 13-4: R&D spending on type of research

Year	**R&D investment (R'000)**	**Basic Research (R'000)**	**%**	**Applied Research (R'000)**	**%**	**Experimental Development (R'000)**	**%**
1979	62109	35598	57.32%	22289	35.89%	4757	7.66%
1981	94424	50381	53.36%	36807	38.98%	7325	7.76%
1983	151352	83124	54.92%	56200	37.13%	12027	7.95%
1985	306534	170941	55.77%	107237	34.98%	28357	9.25%
1987	339194	176415	52.01%	132168	38.97%	30612	9.02%
1989	517566	242332	46.82%	210725	40.71%	64508	12.46%
1991	690439	359788	52.11%	273757	39.65%	56895	8.24%
1993	415648	207319	49.88%	172351	41.47%	35978	8.66%
1995	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1997	496000	N/A	49.80%	N/A	38.00%	N/A	12.20%
1999	N/A	N/A	N/A	N/A	N/A	N/A	N/A
2001	1896156	862067	45.46%	706108	37.24%	234529	12.37%
2003	2071351	915971	44.22%	827,209	39.94%	328170	15.84%
Average	51.06%		38.45%		10.13%		
Standard Deviation	4.27%		2.01%		2.68%		

13.2.2 Human Resources in the HES

The Following table reflects figures for the FTE researchers employed in the Higher Education system. From the figures can be seen that by far the greatest share of FTE research personnel are FTE researchers.

Table 13-5: Human Resource data from the R&D Surveys

	Total HC	**HC Researchers**	**%**	**HC Technicians**	**%**	**HC Support**	**%**
1977	6425	5053	78.65%	926	14.41%	446	6.94%
1979	8181	6406	78.30%	1272	15.55%	503	6.15%
1981	6116	4044	66.12%	1456	23.81%	616	10.07%

1983	11465	8841	77.11%	1523	13.28%	1101	9.60%
1985	17889	13588	75.96%	2156	12.05%	2145	11.99%
1987	19943	15417	77.31%	1645	8.25%	2881	14.45%
1989	19682	13978	71.02%	2758	14.01%	2946	14.97%
1991	16514	14540	88.05%	962	5.83%	1012	6.13%
1993	10835	9916	91.52%	511	4.72%	408	3.77%
1995	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1997	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1999	N/A	N/A	N/A	N/A	N/A	N/A	N/A
2001	15767	12626	80.08%	827	5.25%	2314	14.68%
2003	19377	14055	72.53%	2594	13.39%	2728.5	14.08%
Average			77.88%		11.87%		10.26%
Standard Deviation			7.17%		5.62%		4.06%

Table 13-6: Full time equivalent researchers in the HES

Year	**Amount of R&D workers (FTE)**	**FTE Researchers**	**%**	**FTE Technicians**	**%**	**FTE Support Personnel**	**%**
1977	2555	1938	75.85%	447	17.50%	170	6.65%
1979	3216	2399	74.60%	623	19.37%	194	6.03%
1981	2253	1425	63.25%	627	27.83%	200	8.88%
1983	4128	3384	81.98%	532	12.89%	212	5.14%
1985	6810	5183	76.11%	928	13.63%	699	10.26%
1987	6610	5780	87.44%	473	7.16%	357	5.40%
1989	6353	5160	81.22%	837	13.17%	355	5.59%
1991	6533	5984	91.60%	289	4.42%	260	3.98%
1993	4450	4096	92.04%	234	5.26%	120	2.70%
1995	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1997	4693	N/A	N/A	N/A	N/A	N/A	N/A
1999	N/A	N/A	N/A	N/A	N/A	N/A	N/A
2001	4042	3425	84.74%	217	5.37%	401	9.92%
2003	4553.99	3373.78	74.08%	763.33	16.76%	416.88	9.15%
Average			80.26%		13.03%		6.70%
Standard Deviation			8.60%		7.22%		2.51%

Table 13-7: Data from HEMIS database for years 1986 to 2003[1]

	Academic and Research Personnel	**Professional Personnel**	**Total Personnel at Universities**
1986	9271	11232	29061
1987	9392	11368	30362
1988	9665	11697	31261
1989	N/A	N/A	N/A
1990	9615	11614	32618
1991	9971	12181	33855
1992	10211	12501	33819
1993	10357	12758	33864
1994	10268	12720	33745
1995	10489	12994	34181
1996	10567	13263	34368

[1] Data from the universities of Transkei, Northwest and Venda are not included as they are not available in the HEMIS database. Durban Westville is not included for 1990.

1997	10774	13606	34400
1998	10415	13108	32525
1999	N/A	N/A	N/A
2000	10390	13460	31226
2001	10010	13028	28482
2002	10552	13686	30048
2003	10641	13977	30211

Table 13-8: HEMIS data of Ageing if the researchers

	< 25	25 - 34	35 - 44	45 – 54	55 +
1986	138	2502	3284	2307	1039
1987	116	2364	3387	2447	1078
1988	104	2452	3346	2666	1097
1989	117	2393	3340	2650	1084
1990	129	2383	3494	2831	1171
1991	97	2255	3438	2941	1240
1992	97	2197	3477	3077	1363
1993	96	2109	3593	3219	1340
1994	88	2071	3558	3212	1339
1995	122	2220	3708	3333	1522
1996	140	2236	3667	3356	1618
1997	141	2265	3793	3380	1670
1998	86	2138	3486	3292	1658
1999	NA	NA	NA	NA	NA
2000	174	2392	3500	3525	1524
2001	177	2120	3036	3084	1373
2002	195	2367	3415	3563	1680
2003	151	2329	3404	3631	1747

13.2.3 Students in the Higher Education system

The following table documents data in Student enrolment in the South African Higher Education System (Universities only) gathered from two main sources namely:

- "1990 SA science and technology indicators" for years 1980 - 1988 (FRD, 1990).
- HEMIS database for years 1986 to 2003 (HEMIS, 2005)

The Model input column reflects the values used as model input by integration of the two sources.

Figure 13-1 Students in the Higher Education sector

	Student Numbers (HEMIS, 2005)	Student Numbers (FRD, 1990)	Model input
1980		152346	144000
1981		154833	154000
1982		158834	164000
1983		173116	174000
1984		185261	184000
1985		211756	198000
1986	211593	233625	211593
1987	223720	247694	223720
1988	242067	267608	242067
1989	257355		257355

1990	270399		270399
1991	282779		282779
1992	310384		310384
1993	318517		318517
1994	338470		338470
1995	361371		361371
1996	379825		379825
1997	412795		412795
1998	421316		421316
1999	431478		431478
2000	439810		439810
2001	426684		426684
2002	440204		440203.9
2003	435567		435567.4

The following is a graphical representation of the data presented in the table above, including the approximation of an integration of the two datasets, which is used as an input to the model.

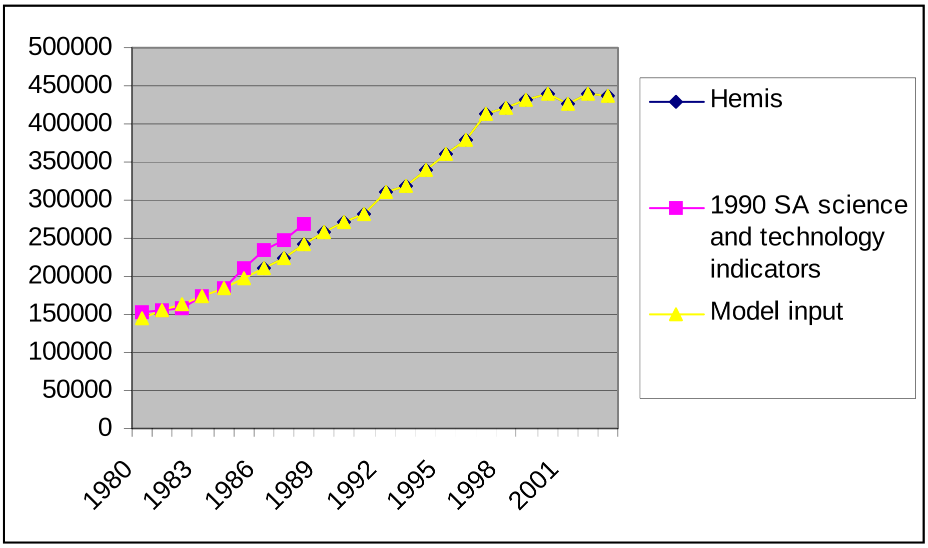

Figure 13-2 Student enrolment in South Africa's HES

13.2.4 The Student-Staff relationship

Figure 13-3 Student-to-staff relationship at HAU[2] and HDU[3] in South Africa

	HAU	HDU
1986	26.15	21.49
1987	27.33	24.15
1988	27.69	30.24
1989	28.15	34.49
1990	28.62	38.74
1991	30.46	35.38

[2] HAU – Historically Advantaged Universities
[3] HDU – Historically Disadvantaged Universities

1992	31.22	34.10
1993	32.10	32.57
1994	33.59	34.59
1995	34.78	34.59
1996	35.06	31.49
1997	35.42	27.11
1998	36.70	27.90
1999	38.36	25.41
2000	40.02	22.92
2001	43.58	31.00
2002	45.59	28.91
2003	46.99	31.66

Since roughly 90% of all R&D output is created in HAU, this analysis will only focus on the HAU. For this reason we use the student-to-staff ratio of these universities.

13.2.5 % time spent n R&D

The data from the Survey is used to find the percentage of the recorded HC spent on R&D to find the FTE. This figure is then used in the model and a regression is performed in order to model the effect the student staff relationship has on the percentage time Academic and Research staff have left to perform R&D duties.

Table 13-9: Student-to-staff ratio and the % time spent on R&D

	Professional Personnel	Total HC	Amount of R&D workers (FTE)	%	HC Researchers	FTE Researchers	
1977		6425	2555	39.77%	5053	1938	38.35%
1979		8181	3216	39.31%	6406	2399	37.45%
1981		6116	2253	36.84%	4044	1425	35.24%
1983		11465	4128	36.01%	8841	3384	38.28%
1985	11232	17889	6810	38.07%	13588	5183	38.14%
1987	11368	19943	6610	33.14%	15417	5780	37.49%
1989		19682	6353	32.28%	13978	5160	36.92%
1991	12181	16514	6533	39.56%	14540	5984	41.16%
1993	12758	10835	4450	41.07%	9916	4096	41.31%
1995	13425	13425	4571.5	34.05%	11017	N/A	N/A
1997	14128	14128	4693	33.22%	10655	N/A	N/A
1999		14128	4367.5	30.91%	10665	N/A	N/A
2001	13774	15767	4042	25.64%	12626	3425	27.13%
2003	14697	19377	4553.99	23.50%	14055	3374	24.01%

As far as possible the Survey data is used for the computation of the percentage time spent on R&D. There is evidence that the definitions used in the surveys have changed, it is however still the best source of time-series data of the past 20 years.

As no data values are available for the 1995 to 1999 surveys, these values are extrapolated. By now incorporating these values, the following time series data is used for the analysis.

Table 13-10: Constructed time series data for % time spent on R&D activities.

Year	HC Researchers	FTE Researchers	Percentage time spent
1985	13588	5183	38.14%
1987	15417	5780	37.49%
1989	13978	5160	36.92%
1991	14540	5984	41.16%
1993	12758	4450	34.88%
1995	13425	4571.5	34.05%
1997	14128	4693	33.22%
1999	13951	4367.5	31.31%
2001	12626	3425	27.13%
2003	14055	3374	24.01%

13.3 R&D data: Public sector

13.3.1 R&D expenditure

Table 13-11: R&D funding according to source

	Total expenditure in Sector	Source funding from HES	Funding Sourced from Private sector
1977	114371000	236000	7868000
1979	144293000	137000	7939000
1981	243617580	10763	6401610
1983	246780000	215203	7718000
1985	347357000	168000	12112000
1987	482567000	232000	24497000
1989	578008000	322000	49047000
1991	755018000	0	77147000
1993	810618000	169000	109220000
1995	N/A	N/A	N/A
1997	1591000000	3000000	222000000
1999	N/A	N/A	N/A
2001	1497564000	0	241860000
2003	2210860000	2716000	258426000

Table 13-12: R&D Expenditure in the Public sector

Year	R&D investment (R)	Expenditure HR	% Expenditure HR	Expenditure on Capital	% Expenditure on Capital
1977	114371000	62625000	54.76%	N/A	N/A
1979	144293000	50458000	34.97%	N/A	N/A
1981	243617580	86322000	35.43%	27976397	11.48%
1983	246780000	93906000	38.05%	27889000	11.30%
1985	347357000	130664000	37.62%	38792000	11.17%
1987	482567000	188457000	39.05%	43395000	8.99%
1989	578008000	253622000	43.88%	46470000	8.04%
1991	755018000	334622000	44.32%	37042000	4.91%
1993	810618000	408281000	50.37%	36393000	4.49%
1995	N/A	N/A	N/A	N/A	N/A
1997	1591000000	731300000	45.96%	65950000	4.15%
1999	N/A	N/A	N/A	N/A	N/A
2001	1497564000	657678000	43.92%	79783000	5.33%
2003	2210860000	1088712000	49.24%	183881000	8.32%
		Average	43.13%		7.82%
		Standard Deviation	6.28%		2.94%

Table 13-13: Expenditure on Human Resources by type of resources

	R ('000)	Researchers	%	Technicians	%	Support	
1977	62625	25112	40.10%	11653	18.61%	25860	41.29%
1979	50458	23416	46.41%	15847	31.41%	11195	22.19%
1981	86322	52726	61.08%	27564	31.93%	6032	6.99%
1983	93906	62134	66.17%	24493	26.08%	7279	7.75%
1985	130664	85907	65.75%	33154	25.37%	11603	8.88%

1987	188457	117649	62.43%	49324	26.17%	21484	11.40%
1989	253620	144231	56.87%	70935	27.97%	38454	15.16%
1991	334622	184977	55.28%	89964	26.89%	59681	17.84%
1993	408280	216432	53.01%	111757	27.37%	80091	19.62%
1995	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1997	731300	N/A	N/A	N/A	N/A	N/A	N/A
1999	N/A	N/A	N/A	N/A	N/A	N/A	N/A
2001	657678	N/A	N/A	N/A	N/A	N/A	N/A
2003	1088712	N/A	N/A	N/A	N/A	N/A	N/A
		Average	56.34%		26.87%		16.79%
		Standard Deviation	8.80%		3.86%		10.67%

Table 13-14: Investment in Research by type

	Total	Basic	%	Applied	%	Development	%
1979	144293	16674	11.56%	85867	59.51%	41752	28.94%
1981	243617.6	28796	11.82%	115703	47.49%	69493	28.53%
1983	234779	39637	16.88%	133451	56.84%	61689	26.28%
1985	347357	50164	14.44%	188689	54.32%	108505	31.24%
1987	482567	107057	22.18%	252188	52.26%	123322	25.56%
1989	578008	80323	13.90%	367204	63.53%	130481	22.57%
1991	755018	98081	12.99%	384481	50.92%	269966	35.76%
1993	810618	82414	10.17%	464640	57.32%	263564	32.51%
1995	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1997	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1999	N/A	N/A	N/A	N/A	N/A	N/A	N/A
2001	1497564	432260	28.86%	704864	47.07%	360109	24.05%
2003	2210860	694769	31.43%	1036447	46.88%	479644	21.69%
		Average	17.42%		53.61%		27.71%
		Standard Deviation	7.52%		5.70%		4.54%

13.3.2 Human Resources

Table 13-15: Recorded HC Research personnel in Frascati R&D Surveys

	total HC	HC Researchers	%	HC Tech	%	HD Support	%
1977	10202	2709	26.55%	2008	19.68%	5485	53.76%
1979	9268	2596	28.01%	2784	30.04%	3888	41.95%
1981	7355	3703	50.35%	2294	31.19%	1358	18.46%
1983	5764	3029	52.55%	1763	30.59%	972	16.86%
1985	7306	3739	51.18%	2049	28.05%	1518	20.78%
1987	8990	5114	56.89%	2030	22.58%	1846	20.53%
1989	8854	3564	40.25%	2791	31.52%	2499	28.22%
1991	8419	3116	37.01%	2129	25.29%	3174	37.70%
1993	8854	3113	35.16%	2392	27.02%	3348	37.81%
1995	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1997	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1999	N/A	N/A	N/A	N/A	N/A	N/A	N/A
2001	8012	2774	34.62%	1762	21.99%	3476	43.38%
2003	8805	3343	37.97%	1934	21.96%	3528	40.07%
		Average	**41.26%**		**26.79%**		**31.95%**
		Standard Deviation	**10.77%**		**4.24%**		**12.70%**

Table 13-16: Recorded FTE Research personnel in Frascati R&D Surveys

	Total FTE	Researchers	%	Technical personnel	%	Support personnel	%
1977	8512	2222	26.10%	1721	20.22%	4569	53.68%
1979	7678	2095	27.29%	2195	28.59%	3388	44.13%
1981	5563	2601	46.76%	2015	36.22%	947	17.02%
1983	4848	2457	50.68%	1564	32.26%	827	17.06%
1985	5216	2510	48.12%	1692	32.44%	1014	19.44%
1987	6374	3173	49.78%	1896	29.75%	1305	20.47%
1989	6426	2547	39.64%	2209	34.38%	1670	25.99%
1991	6654	2419	36.35%	1810	27.20%	2425	36.44%
1993	7060	2303	32.62%	1923	27.24%	2834	40.14%
1995	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1997	6400	2295	35.86%	1749	27.33%	2356	36.81%
1999	N/A	N/A	N/A	N/A	N/A	N/A	N/A
2001	5171	2134	41.27%	1195	23.11%	1842	35.62%
2003	5389	1900	35.25%	1304	24.19%	2186	40.56%
		Average	**39.14%**		**28.58%**		**32.28%**
		Standard Deviation	**8.40%**		**4.73%**		**12.01%**

For 2003, only the Science Council data is included – not available for other organisations in the Government sector

Table 13-17: Human Resources breakdown analysis on the Public sector

	Total HC	HC Researchers	Total FTE	FTE Researchers	% All staff	% Researchers
1977	10202	2709	8512	2222	83.43%	82.02%
1979	9268	2596	7678	2095	82.84%	80.70%
1981	7355	3703	5563	2601	75.64%	70.24%
1983	5764	3029	4848	2457	84.11%	81.12%
1985	7306	3739	5216	2510	71.39%	67.13%
1987	8990	5114	6374	3173	70.90%	62.05%
1989	8854	3564	6426	2547	72.58%	71.46%
1991	8419	3116	6654	2419	79.04%	77.63%
1993	8854	3113	7060	2303	79.74%	73.98%
1995	N/A	N/A	N/A	N/A	N/A	N/A
1997	N/A	N/A	6400	2295	N/A	N/A
1999	N/A	N/A	N/A	N/A	N/A	N/A
2001	8012	2774	5171	2134	64.54%	76.93%
2003	8805	3343	5389.41	2503	61.21%	74.87%
					75.04%	74.38%
					7.68%	6.26%

Table 13-18: Human Resources time spent on R&D analysis

	total HC	Total FTE	% time spent	HC Researchers	FTE Researchers	% time spent
1977	10202	8512	83.43%	2709	2222	82.02%
1979	9268	7678	82.84%	2596	2095	80.70%
1981	7355	5563	75.64%	3703	2601	70.24%
1983	5764	4848	84.11%	3029	2457	81.12%
1985	7306	5216	71.39%	3739	2510	67.13%
1987	8990	6374	70.90%	5114	3173	62.05%

1989	8854	6426	72.58%	3564	2547	71.46%
1991	8419	6654	79.04%	3116	2419	77.63%
1993	8854	7060	79.74%	3113	2303	73.98%
1995	N/A	N/A	N/A	N/A	N/A	N/A
1997	N/A	6400	N/A	N/A	2295	N/A
1999	N/A	N/A	N/A	N/A	N/A	N/A
2001	8012	5171	64.54%	2774	2134	76.93%
2003	8805	5389	61.21%	3343	2503	74.87%
		Average	75.04%			74.38%
		Standard Deviation	7.68%			6.26%

13.4 R&D Data: Private sector

13.4.1 R&D investment

Table 13-19: R&D expenditure by sources of funding in the Private sector

	Total expenditure in Sector	Source funding from HES	% funding from HES	Funding Sourced from Pub	% funding from Public sector
1977	68141000	0	0.00%	3443000	5.05%
1979	100594000	0	0.00%	2625000	2.61%
1981	185180000	0	0.00%	2989340	1.61%
1983	378550100	55813	0.01%	1846236	0.49%
1985	413462000	61000	0.01%	3185000	0.77%
1987	495836000	74000	0.01%	3333000	0.67%
1989	656951000	0	0.00%	13257000	2.02%
1991	1297602000	709000	0.05%	133766000	10.31%
1993	1336227000	828000	0.06%	60861000	4.55%
1995	N/A	N/A	N/A	N/A	N/A
1997	2216000000	1000000	0.05%	186000000	8.39%
1999	N/A	N/A	N/A	N/A	N/A
2001	4023576000	0	0.00%	392614000	9.76%
2003	5591325000	5133000	0.09%	354504000	6.34%
		Average	0.02%		4.38%
		Standard Deviation	0.03%		3.60%

Table 13-20: R&D Expenditure in the Private sector

	Total expenditure in Sector	Salary HR	% HR	Spending on Capital	% Capital
1977	68141000	36628000	53.75%	9287000	13.63%
1979	100594000	43916000	43.66%	N/A	N/A
1981	185180000	91956000	49.66%	33250000	17.96%
1983	378550100	165828000	43.81%	43816000	11.57%
1985	413462000	212679000	51.44%	32751000	7.92%
1987	495836000	258214000	52.08%	60277000	12.16%
1989	656951000	302719000	46.08%	70283000	10.70%
1991	1297602000	703578000	54.22%	171137000	13.19%
1993	1336227000	682289000	51.06%	135991000	10.18%
1995	N/A	N/A	N/A	N/A	N/A
1997	2216000000	1294600000	58.42%	627000000	2.83%
1999	N/A	N/A	N/A	N/A	N/A
2001	4023576000	1718373000	42.71%	782323000	19.44%
2003	5591325000	2488458000	44.51%	775849000	13.88%
		Average	**49.28%**		**14.35%**
		Standard Deviation	**5.06%**		**9.88%**

Table 13-21: R&D expenditure by type of R&D in the Private sector

	Total	Basic	%	Applied	%	Development	%
1979	100594	4948	4.92%	35668	35.46%	59978	59.62%
1981	185178	3164	1.71%	44650	24.11%	137364	74.18%
1983	378550	9437	2.49%	94560	24.98%	274553	72.53%
1985	413462	28053	6.78%	135383	32.74%	250026	60.47%
1987	495837	34664	6.99%	224048	45.19%	237125	47.82%

1989	656950	32645	4.97%	266940	40.63%	357365	54.40%
1991	1297622	68633	5.29%	514449	39.65%	714540	55.07%
1993	1336227	22617	1.69%	406738	30.44%	906872	67.87%
1995	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1997	2215900	225200	10.16%	668800	30.18%	1322000	59.66%
1999	N/A	N/A	N/A	N/A	N/A	N/A	N/A
2001	3736605	663819	17.77%	1397968	37.41%	1674818	44.82%
2003	5591325	759345	13.58%	1883082	33.68%	2948898	52.74%
Average			6.62%		34.43%		58.95%
Standard Deviation			5.22%		6.72%		9.98%

13.4.2 Human Resources

Table 13-22: Human Resources Headcount employed in the Business sector

	total HC	HC Researchers	%	HC Tech	%	HD Support	%
1977	6569	1790	27.25%	1742	26.52%	3037	46.23%
1979	6091	2180	35.79%	1685	27.66%	2226	36.55%
1981	7185	2403	33.44%	2098	29.20%	2694	37.49%
1983	8834	2676	30.29%	3203	36.26%	2955	33.45%
1985	9565	2744	28.69%	4040	42.24%	2781	29.07%
1987	9828	3000	30.53%	4005	40.75%	2823	28.72%
1989	7446	2396	32.18%	1960	26.32%	3090	41.50%
1991	11791	4688	39.76%	3444	29.21%	3659	31.03%
1993	9768	5157	52.79%	2585	26.46%	2026	20.74%
1995	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1997	8111	4113	50.71%	2208	27.22%	1790	22.07%
1999	N/A	N/A	N/A	N/A	N/A	N/A	N/A
2001	8284	3753	45.30%	2488	30.03%	2043	24.66%
2003	11608	5058	43.57%	3430	29.55%	3120	26.88%
Average			37.53%		30.95%		31.53%
Standard Deviation			8.75%		5.61%		7.80%

Table 13-23: FTE Human Resources employed in the Business sector

	Total FTE	Researchers	%	Technical personnel	%	Support personnel	%
1977	4237	1375	32.45%	1342	31.67%	1520	35.87%
1979	4088	1380	33.76%	1207	29.53%	1501	36.72%
1981	5494	1937	35.26%	1738	31.63%	1819	33.11%
1983	6771	1990	29.39%	2646	39.08%	2135	31.53%
1985	7196	2130	29.60%	3328	46.25%	1738	24.15%
1987	7257	2372	32.69%	3132	43.16%	1753	24.16%
1989	5008	2001	39.96%	1431	28.57%	1576	31.47%
1991	8481	3396	40.04%	2785	32.84%	2300	27.12%
1993	7649	4341	56.75%	1869	24.43%	1439	18.81%
1995	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1997	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1999	N/A	N/A	N/A	N/A	N/A	N/A	N/A
2001	6209	2952	47.54%	1916	30.86%	1341	21.60%
2003	9132	4153	45.48%	2605	28.53%	2374	26.00%
Average			38.45%		33.32%		28.23%
Standard Deviation			8.57%		6.69%		5.90%

Table 13-24: Research staff data in the Business sector

	HC (all)	FTE (All)		HC (Researchers)	FTE (Researchers)	
1977	6569	4237	64.50%	1790	1375	76.82%
1979	6091	4088	67.12%	2180	1380	63.30%
1981	7185	5494	76.46%	2403	1937	80.61%
1983	8834	6771	76.65%	2676	1990	74.36%
1985	9565	7196	75.23%	2744	2130	77.62%
1987	9828	7257	73.84%	3000	2372	79.07%
1989	7446	5008	67.26%	2396	2001	83.51%
1991	11791	8481	71.93%	4688	3396	72.44%
1993	9768	7649	78.31%	5157	4341	84.18%
1995	N/A	N/A	N/A	N/A	N/A	N/A
1997	8111	N/A	N/A	4113	N/A	N/A
1999	N/A	N/A	N/A	N/A	N/A	N/A
2001	8284	6209	74.95%	3753	2952	78.66%
2003	11608	9132	78.67%	5058	4153	82.11%
		Average	73.17%			77.52%
		Standard Deviation	4.86%			5.94%

13.5 The time value of money.

Inflation causes a currency's actual value to depreciate. In order to be able to make a meaningful comparison over the years regarding R&D investment and expenditure, it makes sense to look at it in terms of a constant Rand value. The consumer price index (StatsSA, 2005) was used to find the factor each year has to be multiplied with to find the 2001 Rand value. If one wants to express amount y from year Y in terms of Rand value in year x the following formula is used:

$$Factor = \frac{Index_{yearX}}{Index_{yearY}} \qquad \textbf{13-1}$$

Table 13-25: Time value of money computed from consumer price index (StatsSA, 2005)

Year	Index	Factor for 2001 Rand
1977	7.6	13.90789
1978	8.4	12.58333
1979	9.5	11.12632
1980	10.8	9.787037
1981	12.5	8.456
1982	14.3	7.391608
1983	16.1	6.565217
1984	17.9	5.905028
1985	20.8	5.081731
1986	24.7	4.279352
1987	28.7	3.682927
1988	32.4	3.262346
1989	37.1	2.849057
1990	42.4	2.492925
1991	49.0	2.157143
1992	55.7	1.897666
1993	61.2	1.727124
1994	66.6	1.587087
1995	72.4	1.459945

1996	77.7	1.36036
1997	84.4	1.25237
1998	90.2	1.17184
1999	94.9	1.113804
2000	100.0	1.057
2001	105.7	1

14 APPENDIX C

14.1 Absorption of Knowledge (HES)

The rate at which the system is able to produce new knowledge output is computed through the contribution made form different stocks in the system. The following expression is formulated for the R&D knowledge absorption rate in the system:

- $R_{Absorptionr}$: Absorption rate of knowledge in the system
- $S_{R\&Doutput} * S_{FTE}$: RD output stock interacting with the presence of full time equivalent people who can draw on the stocks of knowledge in system
- S_{World} / S_{HC}: Available external knowledge stock per Headcount personnel employed in the system

A multiplicative model is developed for the absorption rate per full time person working in the system:

$$\frac{R_{Absorptionr}}{R^*_{Absorption}} = f * \left(\frac{S_{R\&Doutput}}{S^*_{R\&Doutpu}} * \frac{S_{FTE}}{S^*_{FTE}}\right)^d * \left(\frac{S_{World}}{S^*_{World}} / \frac{S_{HC}}{S^*_{HC}}\right)^e \qquad \textbf{14-1}$$

This expression is linearised by taking the log-linear form:

$$\ln\left(\frac{R_{Absorptionr}}{R^*_{Absorption}}\right) = f + d * \ln\left(\frac{S_{R\&Doutput}}{S^*_{R\&Doutpu}} * \frac{S_{FTE}}{S^*_{FTE}}\right) + e * \ln\left(\frac{S_{World}}{S^*_{World}} / \frac{S_{HC}}{S^*_{HC}}\right)$$

14-2

This is the expression used to perform the regression for estimating the parameters d, e and f.

The section describes the variables included in the model to estimate the rate of knowledge absorption in the system. The following SAS program was used.

Table 14-1: SAS code for stationarity tests in variables AbsorbedR, rdfte and wsperhc

```
goptions reset=all cback=white colors=(black) lfactor=2
border;
title1 'Trend Plot';
proc gplotb data = HES.hesloglin;
plot (AbsorbedR rdfte wsperhc)*year;
plot AbsorbedR *(rdfte wsperhc);
run;

* test for stationarity of the 3 series using arima procedure
*;
proc arima data=hes.loglin;
identify var= AbsorbedR stationarity=(phillips=(0,1));
identify var=rdfte stationarity=(phillips=(0,1));
```

```
identify var=wsperhc stationarity=(phillips=(0,1));
                        run;
```

The following sections document and explain the output obtained from the SAS program.

14.1.1 Absorption rate of knowledge in the system

The following is the time plot output from the SAS program for the absorption rate per full time equivalent researchers in the system.

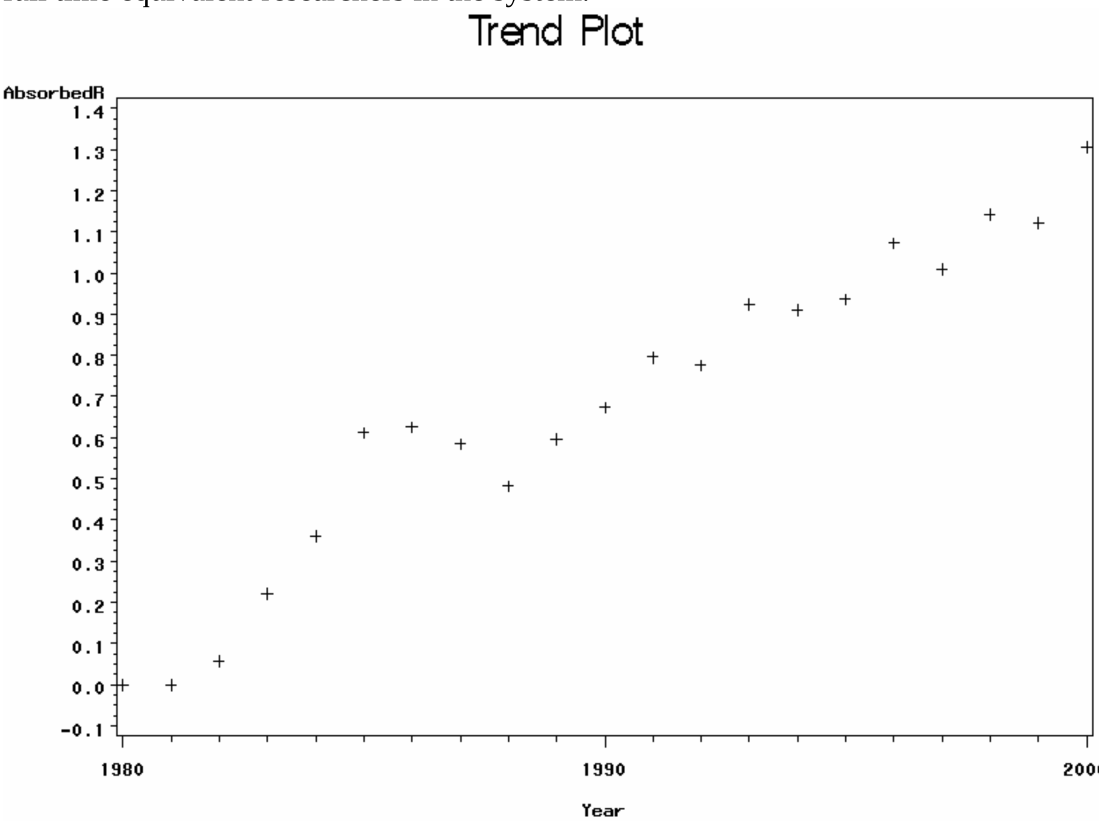

Figure 14-1 Time plot of the absorption rate in the Higher Education system

From Figure 14-1 can be seen that the time plot shows an upward trend line. We therefore make use of the “Trend” specification in the stationarity test output results.

Table 14-2: Phillips-Perron test output for variable “AbsorbedR”

```
                         The ARIMA Procedure

                     Name of Variable = AbsorbedR

                  Mean of Working Series    0.707911
                  Standard Deviation        0.393853
                  Number of Observations          22

                    Phillips-Perron Unit Root Tests

     Type           Lags         Rho    Pr < Rho        Tau    Pr < Tau

     Zero Mean         0      1.3497      0.9411       2.23      0.9912
                       1      1.3478      0.9409       2.21      0.9909
```

Single Mean	0	-0.8467	0.8869	-0.74	0.8157
	1	-0.7612	0.8944	-0.71	0.8224
Trend	0	-10.2431	0.3187	-2.46	0.3408
	1	-10.4180	0.3065	-2.48	0.3337

From the Phillips Perron test output obtained from SAS we read the following values for the probability statistics.

`Pr < Tau = 0.3408` for $\ell = 0$ en

`Pr < Tau = 0.3337` for $\ell = 1$.

Both *p*-values are greater than 0.05. We therefore cannot reject H_0: $d = 1$. We have to conclude that `AbsorbedR` has a unit root and is non-stationary.

14.1.2 R&D Knowledge Stock and FTE researchers

The following is the time plot output from the SAS program for the interaction of Full time equivalent personnel in the system with the R&D Knowledge stock.

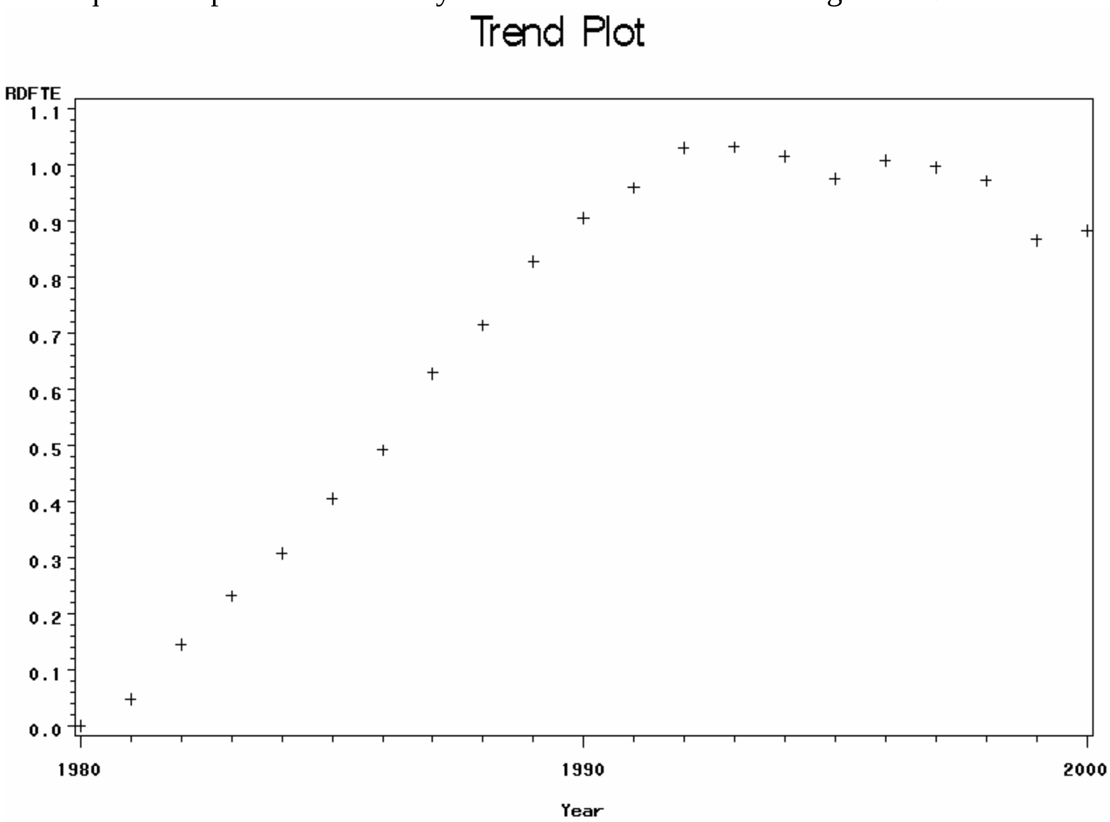

Figure 14-2 Time plot - FTE researcher interacting with R&D knowledge

From Figure 14-2 can be seen that the time plot shows an upward trend line. We therefore make use of the "Trend" specification in the stationarity test output results.

Table 14-3: Phillips-Perron test output for variable "`RDFTE`"

```
The ARIMA Procedure
```

```
                         Name of Variable = RDFTE

                    Mean of Working Series     0.701877
                    Standard Deviation         0.347422
                    Number of Observations           22

                    Phillips-Perron Unit Root Tests

   Type            Lags          Rho     Pr < Rho        Tau     Pr < Tau

   Zero Mean          0       0.7167       0.8424       1.65       0.9717
                      1       0.6539       0.8288       1.15       0.9302
   Single Mean        0      -2.0640       0.7546      -3.08       0.0438
                      1      -2.1431       0.7448      -2.72       0.0876
   Trend              0      -0.7058       0.9868      -0.49       0.9750
                      1      -0.9342       0.9839      -0.59       0.9684
```

From the Phillips Perron test output obtained from SAS we read the following values for the probability statistics.

`Pr < Tau = 0.9750` for $\ell = 0$ en

`Pr < Tau = 0.9684` for $\ell = 1$.

Both *p*-values are greater than 0.05. We therefore cannot reject H_0: $d = 1$. We have to conclude that `RDFTE` has a unit root and is non-stationary.

14.1.3 The World knowledge stock

The following is the time plot output from the SAS program for the World Knowledge Stock per Headcount person employed in the Higher Education system.

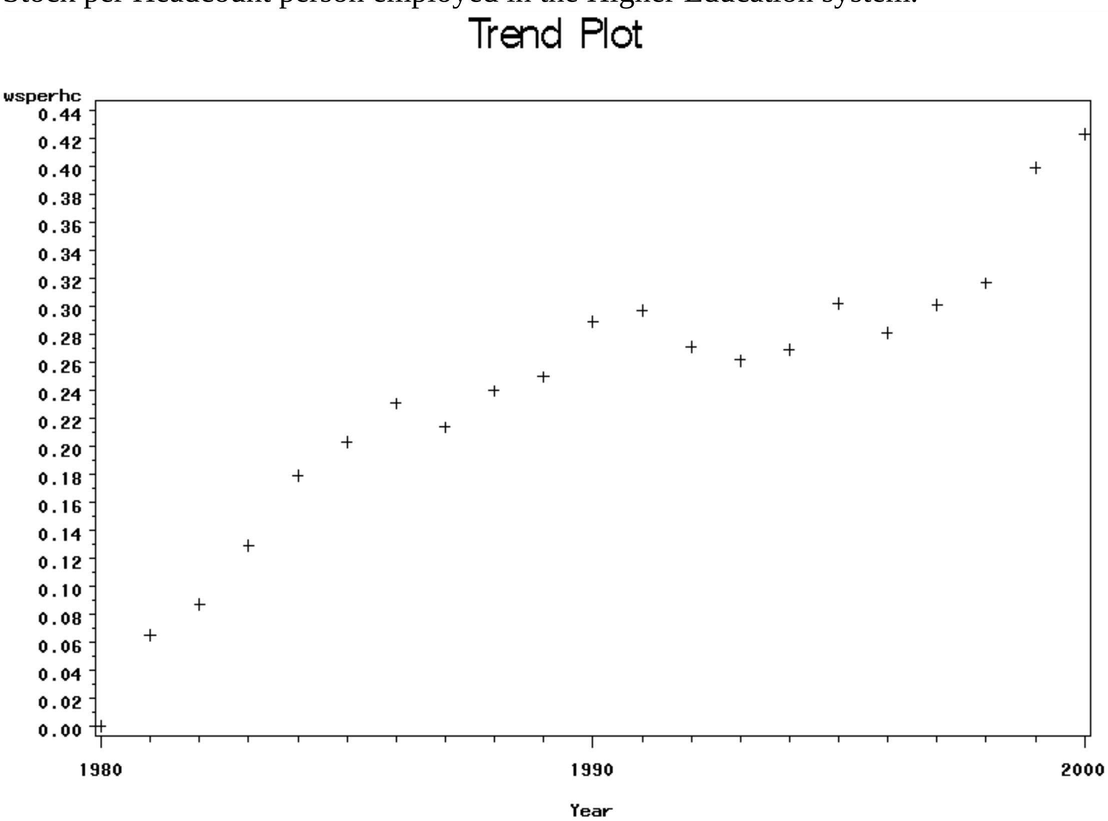

Figure 14-3 Time plot of the World stock of knowledge per HC researcher

From Figure 14-3 can be seen that the time plot shows an upward trend line. We therefore make use of the “Trend” specification in the stationarity test output results.

Table 14-4: Phillips-Perron test output for variable “wsperhc”

```
                         The ARIMA Procedure

                      Name of Variable = wsperhc

                  Mean of Working Series     0.247072
                  Standard Deviation         0.105268
                  Number of Observations           22

                    Phillips-Perron Unit Root Tests

 Type             Lags          Rho     Pr < Rho        Tau     Pr < Tau

 Zero Mean          0       1.1493       0.9175       2.10       0.9884
                    1       1.1216       0.9137       1.87       0.9815
 Single Mean        0      -2.3311       0.7210      -2.00       0.2838
                    1      -2.4006       0.7122      -1.96       0.3021
 Trend              0      -7.5941       0.5431      -2.70       0.2457
                    1      -8.1564       0.4906      -2.70       0.2456
```

From the Phillips Perron test output obtained from SAS we read the following values for the probability statistics.

`Pr < Tau = 0.2457` for $\ell = 0$ en

`Pr < Tau = 0.2456` for $\ell = 1$.

Both p-values are greater than 0.05. We therefore cannot reject H_0: $d = 1$. We have to conclude that `wsperhc` has a unit root and is non-stationary.

14.1.4 Colinearity tests

First however we should make sure that the variables are not collinear. The following is the test results obtained from SAS for the Colinearity test.

Table 14-5: Colinearity diagnostics for the model variables

```
                         Collinearity Diagnostics

                                                Condition
                    Number      Eigenvalue          Index

                         1         2.86046        1.00000
                         2         0.11563        4.97380
                         3         0.02392       10.93637

                         Collinearity Diagnostics

                        ---------Proportion of Variation---------
            Number        Intercept           RDFTE          wsperhc

                 1          0.01733         0.00605          0.00480
                 2          0.87833         0.11364          0.02912
                 3          0.10434         0.88032          0.96608
```

Larger values suggest potential near colinearity. Belsley, Kuh and Welsch (2000) recommend interpreting the Condition index greater or equal than 30 to reflect moderate to severe colinearity, worthy of further investigation. Since all the Condition indexes from the regression model are much smaller than 30, the conclusion can be made that colinearity is not a problem in this case.

Proportion of Variation

The variance proportion indicates for each predictor the proportion of total variance of its estimated regression coefficients associated with a particular principal component. The variance proportions suggest colinearity problems if more than one predictor has high variance proportions of at least 0.5, such a component suggest a problem. One should definitely be concerned when two or more ladings greater than 0.9 appear on a component with a large condition index (>30). This also does not seem to be a problem since the condition indexes are all smaller than 30.

14.1.5 Model estimation - Absorption rate (HES)

As all three variables are non-stationary, we should now fit a model and then test for co-integration in the residual to prove that the modelled relationship is non-spurious.

Table 14-6: SAS code for the model estimation procedure

```
proc autoreg data= HES.hesloglin;
model absorbedR = rdFTE wsperhc
/ method= ml nlag=1 dwprob;
output out=b r=residual;
run;
```

Table 14-7: SAS output for the model estimation of Absorptive capacity in the HES

```
                        The AUTOREG Procedure

                     Dependent Variable    AbsorbedR

                     Ordinary Least Squares Estimates

     SSE                     0.27533496    DFE                        19
     MSE                        0.01449    Root MSE              0.12038
     SBC                     -24.671383    AIC                 -27.94451
     Regress R-Square            0.9193    Total R-Square         0.9193
     Durbin-Watson               1.0282    Pr < DW                0.0014
     Pr > DW                     0.9986
NOTE: Pr<DW is the p-value for testing positive autocorrelation, and
      Pr>DW is the p-value for testing negative autocorrelation.

                          Phillips-Ouliaris
                         Cointegration Test

                     Lags           Rho           Tau

                        1       -11.8653       -2.9582

                                   Standard                  Approx
```

Variable	DF	Estimate	Error	t Value	Pr > |t|
Intercept	1	-0.1717	0.0655	-2.62	0.0167
RDFTE	1	0.2663	0.1444	1.84	0.0809
wsperhc	1	2.8038	0.4767	5.88	<.0001

Q and LM Tests for ARCH Disturbances

Order	Q	Pr > Q	LM	Pr > LM
1	0.1107	0.7393	0.0064	0.9360
2	1.0439	0.5934	0.6325	0.7289
3	2.1377	0.5443	1.4751	0.6880
4	5.2369	0.2638	3.5521	0.4700
5	6.1404	0.2928	3.5763	0.6119
6	6.3389	0.3863	3.8539	0.6964
7	8.5847	0.2839	4.9011	0.6720
8	13.1218	0.1077	11.4069	0.1797
9	13.8131	0.1291	13.0194	0.1617
10	14.6261	0.1463	13.0194	0.2226
11	14.7742	0.1931	13.0199	0.2920
12	15.2048	0.2304	13.0824	0.3631

Maximum Likelihood Estimates

SSE	0.18500614	DFE	18
MSE	0.01028	Root MSE	0.10138
SBC	-29.491802	AIC	-33.855972
Regress R-Square	0.6812	Total R-Square	0.9458
Durbin-Watson	2.0372	Pr < DW	0.4148
Pr > DW	0.5852		

NOTE: Pr<DW is the p-value for testing positive autocorrelation, and Pr>DW is the p-value for testing negative autocorrelation.

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|
Intercept	1	-0.0443	0.1432	-0.31	0.7606
RDFTE	1	0.6165	0.2642	2.33	0.0314
wsperhc	1	1.4308	0.6667	2.15	0.0457
AR1	1	-0.7526	0.1944	-3.87	0.0011

Autoregressive parameters assumed given.

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|
Intercept	1	-0.0443	0.1431	-0.31	0.7606
RDFTE	1	0.6165	0.2303	2.68	0.0154
wsperhc	1	1.4308	0.6188	2.31	0.0328

From the model estimation output obtained we can make the following conclusion:

The test for autocorrelation use is the Durban Watson test statistic. The `Durbin Watson` test statistic is `2.0372 with (Pr < DW = 0.4148) > 0.05 and (Pr < DW = 0.5852) < 0.95`. This indicates that we can we therefore can conclude that the model does not have autocorrelation.

Due to the small sample size and the limited number of data points available, the heteroscedasticity test (Q and LM test for ARCH disturbances) is only interpreted up to 2 time lags. The probability for arch disturbances in the model for lags 1 and 2 are larger than 0.05. We can therefore conclude that the modelled relationship does not suffer from heteroscedasticity.

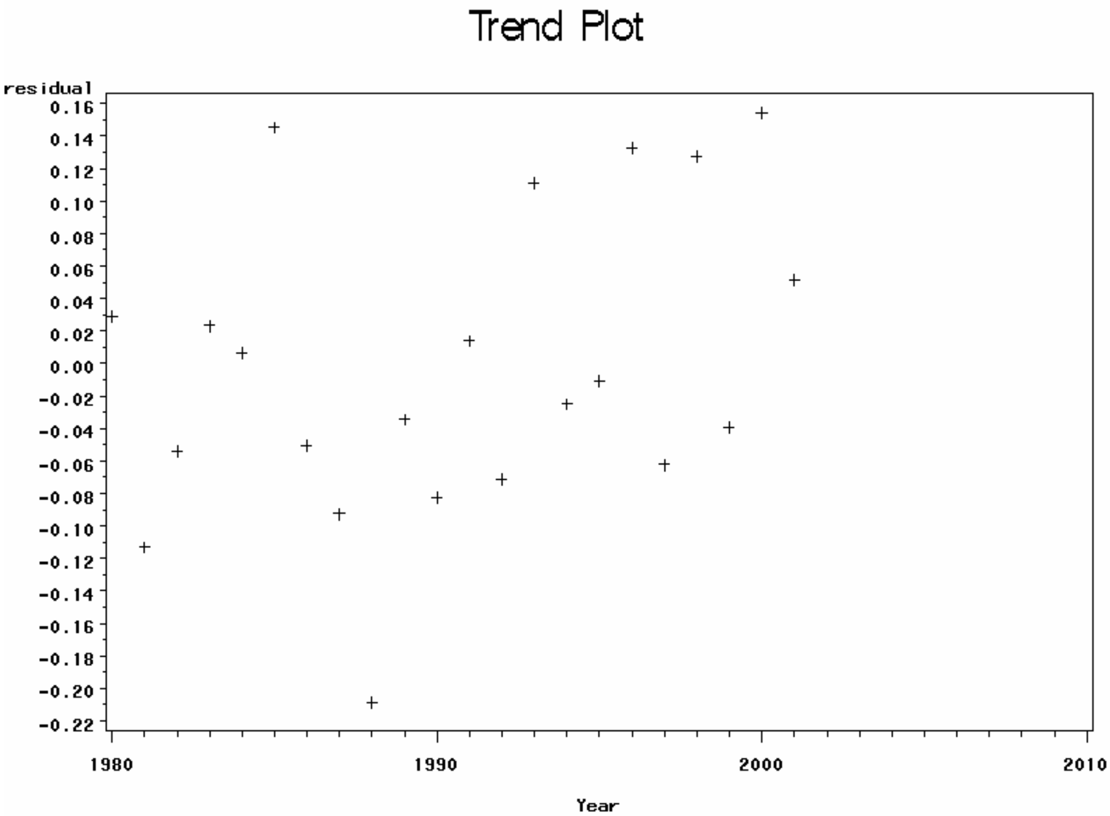

Figure 14-4: Time plot for the residual of the HES knowledge absorption

From Table 14-4 can be seen that the time plot seems to be scattered around 0. From the results we van also read the `Mean of Working Series -0.00219`. We therefore make use of the “Zero mean” specification in the stationarity test output results.

Table 14-8: Phillips Perron tests output for the residual

```
                        The ARIMA Procedure

                     Name of Variable = residual

                  Mean of Working Series    -0.00219
                  Standard Deviation        0.091676
                  Number of Observations          22

                    Phillips-Perron Unit Root Tests

 Type            Lags          Rho     Pr < Rho         Tau     Pr < Tau

 Zero Mean          0     -21.4679       0.0001       -4.55       0.0001
                    1     -21.7305       <.0001       -4.55       0.0001
 Single Mean        0     -21.5116       0.0008       -4.44       0.0023
                    1     -21.7978       0.0007       -4.44       0.0023
 Trend              0     -24.5472       0.0018       -5.17       0.0024
                    1     -25.2333       0.0012       -5.15       0.0026
```

Since an intercept is included in the model fitted, an intercept is included. For (n-1) = 2, the values are obtained from the Critical values for the Phillips Z Statistic or the Dickey Fuller t Statistic when applied to Residuals from Spurious Cointegration Regression (See **Error! Reference source not found.**). The critical value for the 1% level is -4.31.

From the Phillips Perron test output obtained from SAS we read the following values for the probability statistics.

`Tau = -4.55` for $\ell = 0$ en

`Tau = -4.55` for $\ell = 1$.

This means that we can therefore reject the null hypothesis of unit root since the `Tau` values are smaller than the critical value. The residues can be deemed stationary and the variables are cointegrated. We can therefore conclude that the regression is not spurious.

14.2 Creation of new knowledge (HES)

The rate at which the system is able to produce new knowledge output is computed through the contribution made form different stocks in the system. The following expression is formulated for the R&D output productivity per FTE researcher working in the system:

- R_{Paper} / S_{FTE} : R&D output rate per FTE researcher person on the system
- $S_{Experience} / S_{HC}$: Average Experience Stock of the people in the system.
- $S_{Absorbed} / S_{HC}$: Average Absorbed knowledge per person in the system.

A multiplicative model is developed for the development rate of papers per full time person working in the system:

$$\frac{R_{Paper}}{R^*_{Paper}} / \frac{S_{FTE}}{S^*_{FTE}} = c * (\frac{S_{Experience}}{S^*_{Experience}} / \frac{S_{HC}}{S^*_{HC}})^a * (\frac{S_{Absorbed}}{S^*_{Absorbed}} / \frac{S_{HC}}{S^*_{HC}})^b \quad \textbf{14-3}$$

This expression is linearised by taking the log-linear form:

$$\ln(\frac{R_{Paper}}{R^*_{Paper}} / \frac{S_{FTE}}{S^*_{FTE}}) = \ln(c) + a * \ln(\frac{S_{Experience}}{S^*_{Experience}} / \frac{S_{HC}}{S^*_{HC}}) + b * \ln(\frac{S_{Absorbed}}{S^*_{Absorbed}} / \frac{S_{HC}}{S^*_{HC}}) \quad \textbf{14-4}$$

This is then the expression used to perform the regression for estimating the parameters a, b and c.

The section describes the variables included in the model to estimate the rate of knowledge creation in the system. The following SAS program was used:

Table 14-9: SAS code for stationarity tests in variables prperfte, expperhc and absperhc

```
goptions reset=all cback=white colors=(black) lfactor=2
border;
title1 'Trend Plot';
proc gplot hes.rdloglin;
plot (prperfte expperhc absperhc)*year;
plot prperfte*(expperhc absperhc);
run;

* test for stationarity of the 3 series using arima procedure
*;
proc arima hes.rdloglin;
identify var=prperfte stationarity=(phillips=(0,1));
identify var=expperhc stationarity=(phillips=(0,1));
identify var=absperhc stationarity=(phillips=(0,1));
run;
```

The following sections document and explain the output obtained from the SAS program.

14.2.1 R&D output produced per FTE researcher

The following is the time plot output from the SAS program for the R&D output (papers) created per full time equivalent researcher in the system.

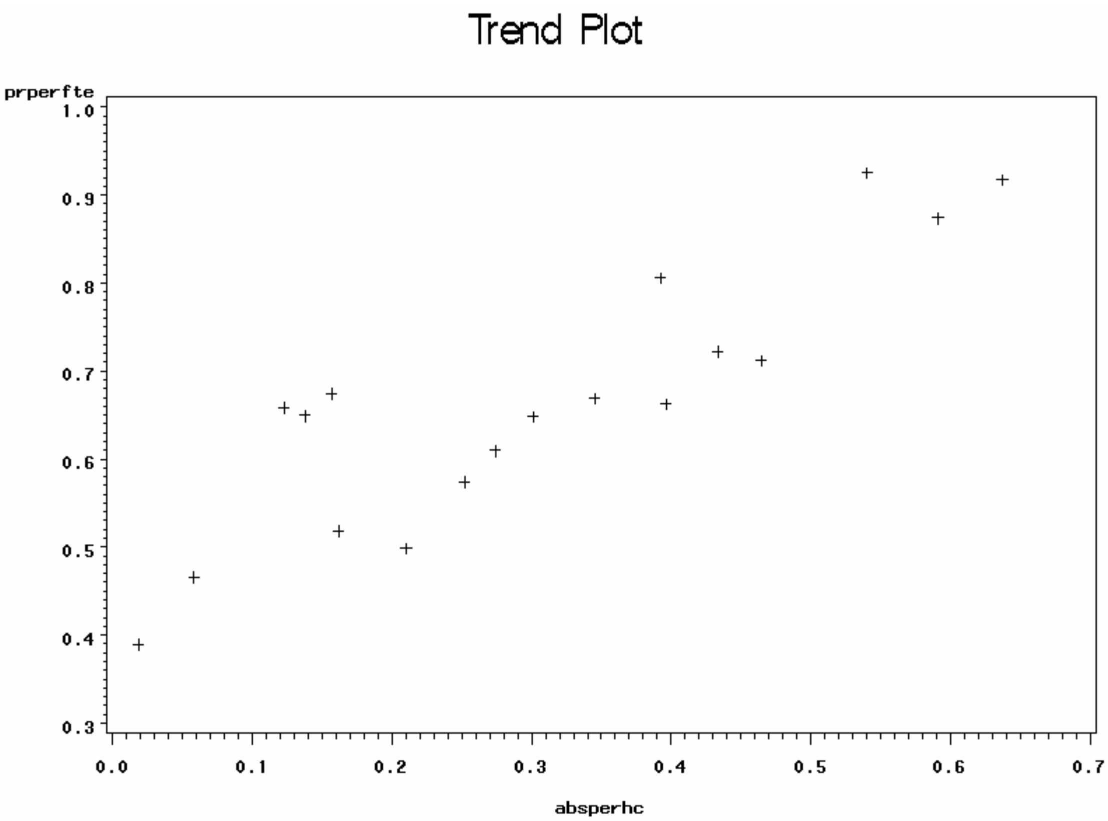

Figure 14-5 Time plot of the Knowledge creation rate per FTE

From Figure 14-5 can be seen that the time plot shows an upward trend line. We therefore make use of the "Trend" specification in the stationarity test output results.

Table 14-10: SAS output for Phillips Perron test for variable "prperfte"

```
                         The ARIMA Procedure

                    Name of Variable = prperfte

                 Mean of Working Series     0.669785
                 Standard Deviation         0.163311
                 Number of Observations           19

                   Phillips-Perron Unit Root Tests

 Type            Lags         Rho    Pr < Rho        Tau    Pr < Tau

 Zero Mean          0      0.6424      0.8229       0.99      0.9068
                    1      0.6918      0.8338       1.23      0.9372
 Single Mean        0     -4.7059      0.4187      -1.83      0.3549
                    1     -4.1852      0.4788      -1.78      0.3767
 Trend              0    -12.4121      0.1682      -3.03      0.1527
                    1    -12.2779      0.1749      -3.02      0.1543
```

From the Phillips Perron test output obtained from SAS we read the following values for the probability statistics.

`Pr < Tau = 0.1527` for $\ell = 0$ en

`Pr < Tau = 0.1543` for $\ell = 1$.

Both *p*-values are greater than 0.05. We therefore cannot reject H_0: $d = 1$. We have to conclude that `prperfte` has a unit root and is non-stationary.

14.2.2 Absorbed Stock per Headcount

The following is the time plot output from the SAS program for the absorbed knowledge stock in the system.

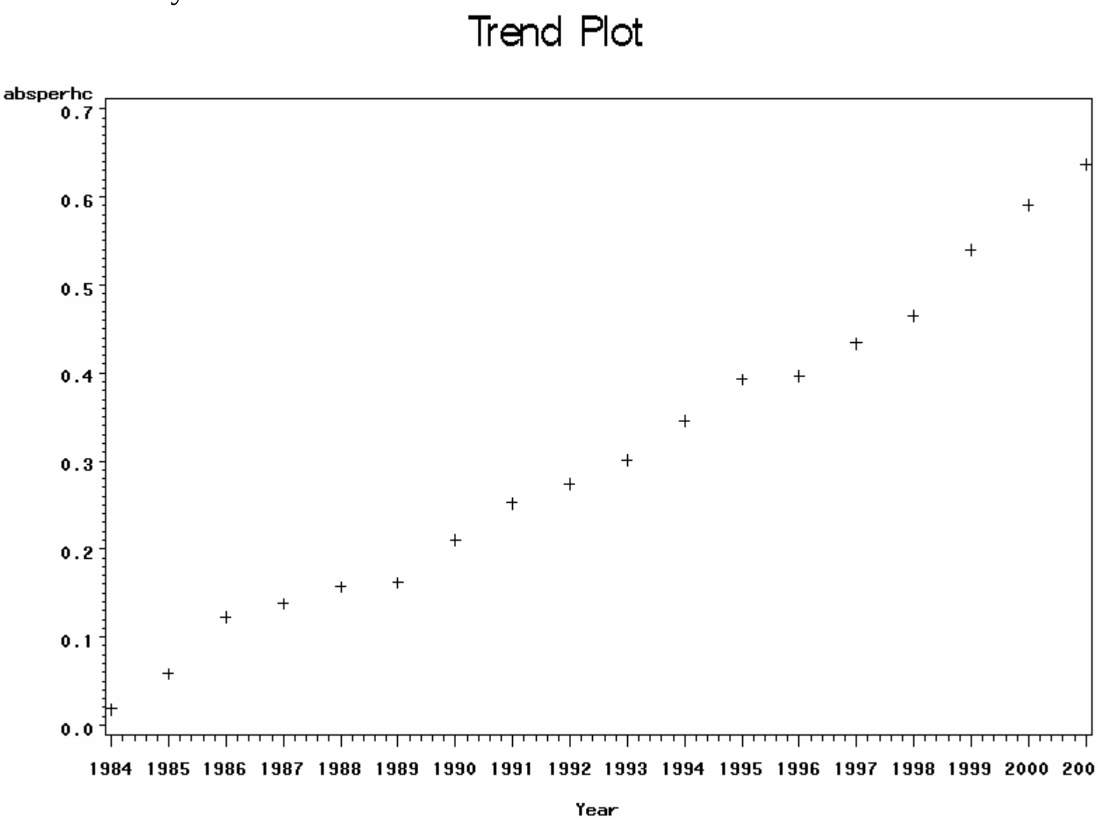

Figure 14-6 Time plot of the Absorbed knowledge stock per Headcount personnel

From Figure 14-6 can be seen that the time plot shows an upward trend line. We therefore make use of the “Trend” specification in the stationarity test output results.

Table 14-11: SAS output for Phillips Perron test for variable “absperhc”

```
                          The ARIMA Procedure

                       Name of Variable = absperhc

                    Mean of Working Series    0.549018
                    Standard Deviation        0.236942
                    Number of Observations          19

                      Phillips-Perron Unit Root Tests

     Type           Lags         Rho    Pr < Rho         Tau    Pr < Tau

     Zero Mean         0      1.2929      0.9324        5.49      0.9999
                       1      1.2825      0.9312        4.60      0.9999
     Single Mean       0     -0.3848      0.9218       -0.91      0.7625
                       1     -0.3868      0.9217       -0.90      0.7645
     Trend             0     -8.8861      0.4103       -2.71      0.2440
                       1     -9.1435      0.3877       -2.72      0.2410
```

From the Phillips Perron test output obtained from SAS we read the following values for the probability statistics.

Pr < Tau = 0.244 for $\ell = 0$ en

Pr < Tau = 0.241 for $\ell = 1$.

Both p-values are greater than 0.05. We therefore cannot reject H_0: $d = 1$. We have to conclude that absperhc has a unit root and is non-stationary.

14.2.3 Experience Stock per Headcount

The following is the time plot output from the SAS program for the Experience stock per Headcount in the system.

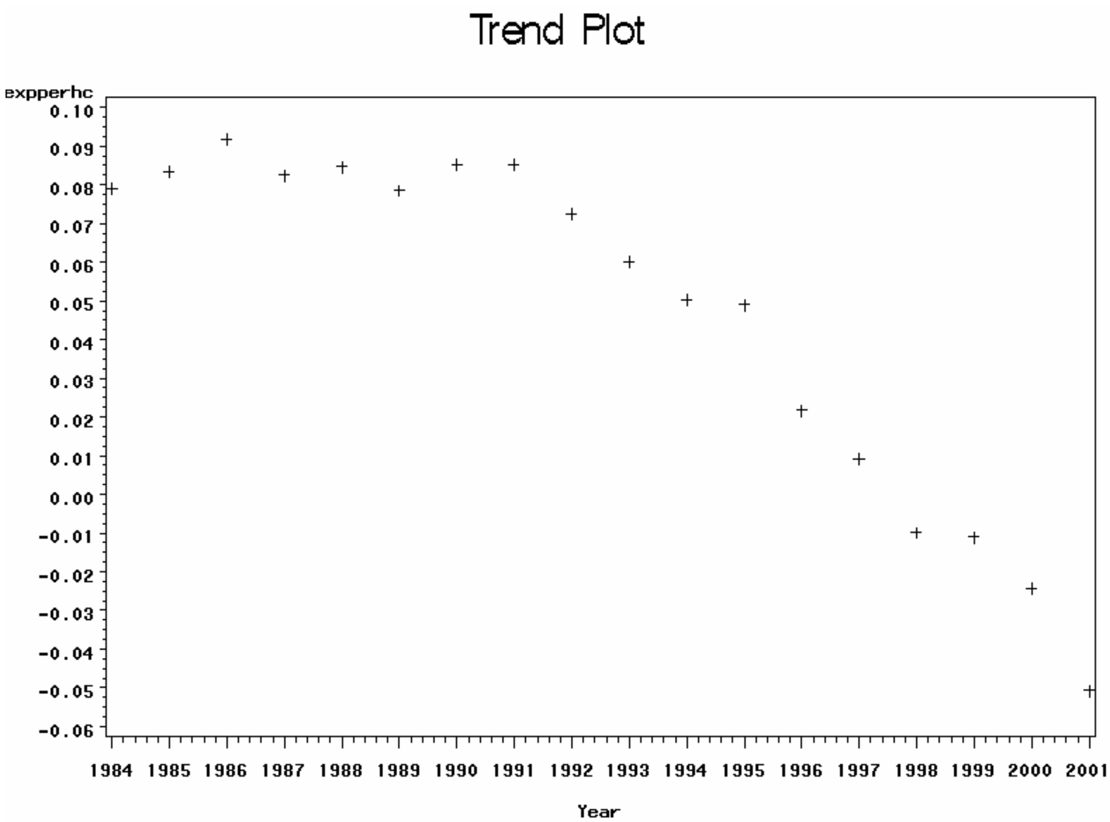

Figure 14-7 Time plot for the "Exptotal" variable in the system

From Figure 14-7 can be seen that the time plot shows a downward trend line. We therefore make use of the "Trend" specification in the stationarity test output results.

Table 14-12: SAS output for Phillips Perron test for variable "Expperhc"

The ARIMA Procedure

Name of Variable = expperhc

Mean of Working Series	0.195865
Standard Deviation	0.026819
Number of Observations	19

Phillips-Perron Unit Root Tests

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau
Zero Mean	0	-0.0565	0.6542	-0.15	0.6181
	1	-0.0591	0.6536	-0.15	0.6167
Single Mean	0	-4.5748	0.4334	-1.58	0.4732
	1	-5.1400	0.3726	-1.66	0.4324
Trend	0	-5.4205	0.7428	-2.37	0.3785
	1	-5.0065	0.7806	-2.42	0.3588

From the Phillips Perron test output obtained from SAS we read the following values for

the probability statistics.
`Pr < Tau = 0.3785` for ℓ = 0 en
`Pr < Tau = 0.3588` for ℓ = 1.

Both *p*-values are greater than 0.05. We therefore cannot reject H_0: $d = 1$. We have to conclude that `expperhc` has a unit root and therefore is non-stationary.

14.2.4 Colinearity tests

First however we should make sure that the variables are not collinear. The following is the test results obtained from SAS for the Colinearity test.

Table 14-13: Colinearity diagnostics for the model variables

```
                     Collinearity Diagnostics

                                            Condition
               Number      Eigenvalue           Index

                    1         2.28975         1.00000
                    2         0.70033         1.80818
                    3         0.00992        15.19053

                     Collinearity Diagnostics

                 ---------Proportion of Variation---------
       Number       Intercept         absperhc         expperhc

            1         0.00332          0.00460          0.00696
            2      0.00007497          0.01905          0.05398
            3         0.99660          0.97634          0.93906
```

Larger values suggest potential near colinearity. Belsley, Kuh and Welsch (2000) recommend interpreting the Condition index greater or equal than 30 to reflect moderate to severe colinearity, worthy of further investigation. Since all the Condition indexes from the regression model is much smaller than 30, the conclusion can be made that colinearity is not a problem in this case.

Proportion of Variation
The variance proportion indicates for each predictor the proportion of total variance of its estimated regression coefficients associated with a particular principal component. The variance proportions suggest colinearity problems if more than one predictor has a high variance proportions of at least 0.5 for such a components suggest a problem. One should definitely be concerned when two or more ladings greater than 0.9 appear on a component with a large condition index (>30). This also does not seem to be a problem since the condition indexes all have small values.

14.2.5 Model estimation the rate of Paper Development in the HES

As all three variables are non-stationary, we should now fit a model and then test for cointegration in the residual

Table 14-14: SAS code for the model estimation procedure

```
proc autoreg data= HES.hesloglin;
model prperfte = absperhc Expperhc
/ method= ml dwprob nlag = 1;
output out=b r=residual;
run;

* consider residual *;
proc gplot data=b;
plot residual*year;
run;

* test for cointegration using arima procedure *;
proc arima data=b;
identify var=residual
stationarity=(phillips=(0,1));
run;
```

Table 14-15: SAS output for the model estimation of Absorptive capacity in the HES

The AUTOREG Procedure

Dependent Variable prperfte

Ordinary Least Squares Estimates

SSE	0.12108175	DFE	19
MSE	0.00637	Root MSE	0.07983
SBC	-42.744877	AIC	-46.018004
Regress R-Square	0.9020	Total R-Square	0.9020
Durbin-Watson	0.9014	Pr < DW	0.0003
Pr > DW	0.9997		

NOTE: Pr<DW is the p-value for testing positive autocorrelation, and
Pr>DW is the p-value for testing negative autocorrelation.

Phillips-Ouliaris
Cointegration Test

Lags	Rho	Tau
1	-10.1658	-2.5256

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|
Intercept	1	0.0673	0.0578	1.16	0.2586
absperhc	1	0.6672	0.0683	9.77	<.0001
expperhc	1	1.1926	0.3490	3.42	0.0029

Q and LM Tests for ARCH Disturbances

Order	Q	Pr > Q	LM	Pr > LM
1	2.6944	0.1007	2.3836	0.1226
2	3.2668	0.1953	2.5181	0.2839
3	3.2901	0.3490	2.5182	0.4720
4	3.6991	0.4483	3.3194	0.5059
5	5.3127	0.3789	4.9323	0.4242
6	8.1806	0.2252	5.0488	0.5376
7	12.5365	0.0842	5.6996	0.5752
8	13.8929	0.0846	5.7748	0.6724
9	14.4766	0.1064	6.0550	0.7344
10	16.1185	0.0963	6.8146	0.7428
11	16.2700	0.1314	7.3181	0.7728
12	16.4044	0.1734	7.3242	0.8355

```
                    Maximum Likelihood Estimates

  SSE                  0.08224994    DFE                       18
  MSE                     0.00457    Root MSE             0.06760
  SBC                  -47.714313    AIC               -52.078483
  Regress R-Square         0.8178    Total R-Square        0.9334
  Durbin-Watson            1.8688    Pr < DW               0.2427
  Pr > DW                  0.7573
NOTE: Pr<DW is the p-value for testing positive autocorrelation, and
Pr>DW is the p-value for testing negative autocorrelation.


                                       Standard                 Approx
Variable         DF     Estimate         Error    t Value     Pr > |t|

Intercept         1    -0.004500        0.0807      -0.06       0.9562
absperhc          1       0.6998        0.1220       5.73       <.0001
expperhc          1       1.5280        0.4756       3.21       0.0048
AR1               1      -0.6004        0.1938      -3.10       0.0062


               Autoregressive parameters assumed given.

                                    Standard                 Approx
Variable         DF     Estimate         Error    t Value     Pr > |t|

Intercept         1    -0.004500        0.0797      -0.06       0.9556
absperhc          1       0.6998        0.1186       5.90       <.0001
expperhc          1       1.5280        0.4756       3.21       0.0048
```

From the model estimation output obtained we can make the following conclusion:

The **R-Square 0.8178** statistic indicate that the model accounts for 81% of the variation of the percentage time spent by staff on R&D activities.

The test for autocorrelation use is the Durban Watson test statistic. The Durbin Watson test statistic is 1.8688 with (Pr < DW = 0.2427 > 0.05 and (Pr < DW = 0.7573) < 0.95. This indicates that we therefore can conclude that the autoregressive model does not have autocorrelation.

Due to the small sample size and the limited number of data points available, the heteroscedasticity test (Q and LM test for ARCH disturbances) is only interpreted up to 2 time lags. The probability for arch disturbances in the model for lags 1 and 2 are larger than 0.05. We can therefore conclude that the modelled relationship does not suffer from heteroscedasticity.

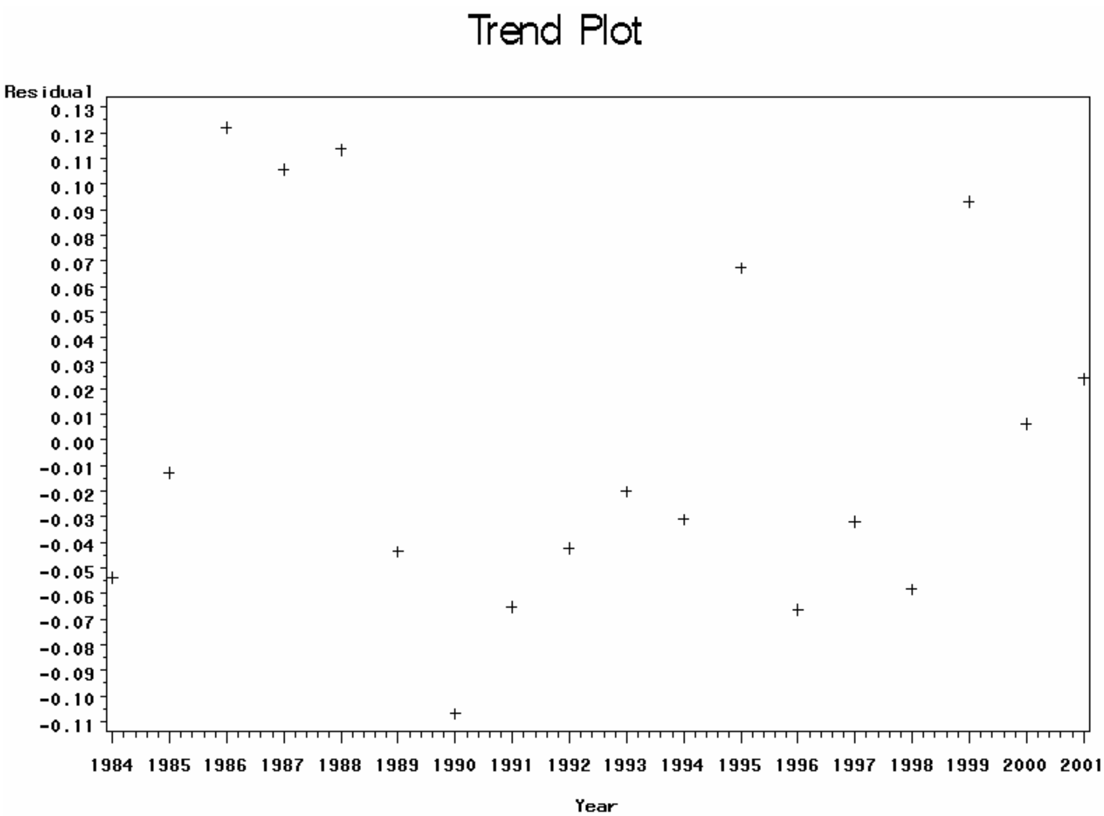

Figure 14-8 Time plot of the residual

From Figure 14-8 can be seen that the time plot seems to be scattered around 0. From the results we van also read the `Mean of Working Series is -0.0003`. We therefore make use of the "Zero mean" specification in the stationarity test output results.

Table 14-16: Phillips Perron test output for the residual

The ARIMA Procedure

Name of Variable = residual

Mean of Working Series	-0.00327
Standard Deviation	0.065914
Number of Observations	18

Phillips-Perron Unit Root Tests

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau
Zero Mean	0	-15.4783	0.0015	-3.66	0.0011
	1	-15.2356	0.0017	-3.66	0.0011
Single Mean	0	-15.6068	0.0079	-3.57	0.0187
	1	-15.3372	0.0089	-3.56	0.0190
Trend	0	-15.4989	0.0571	-3.42	0.0822
	1	-15.1530	0.0649	-3.40	0.0841

For (n-1) = 2, the values are obtained from the Critical values for the Phillips Z Statistic or the Dickey Fuller t Statistic when applied to Residuals from Spurious Cointegration Regression (See **Error! Reference source not found.**). The critical value for the 7.5% level is -3.58.

From the Phillips Perron test output obtained from SAS we read the following values for the probability statistics.

`Tau = -3.66` for ℓ = 0 en

`Tau = -3.66` for ℓ = 1.

This means that we can therefore reject the null hypothesis of unit root with a 10% significance level, since the `Tau` values are smaller than the critical value. The residues can be deemed stationary and the variables are cointegrated. We can therefore conclude that the regression is not spurious.

14.3 Student-to-Staff ratio and the % time spent on R&D model

The section describes the variables included in the model to estimate the percentage time staff has left as a function of the student-to-staff relationship. The following SAS program was used.

Table 14-17: SAS code for stationarity tests in variables "Percentage" and "studentstaff"

```
goptions reset=all cback=white colors=(black) lfactor=2
border;
title1 'Trend Plot';
proc gplotb data = hes.studstaff;
plot (Percentage studentstaff)*year;
plot Percentage*(studentstaff);
run;

proc arima data=hes.studstaff;
identify var=Percentage stationarity=(phillips=(0,1));
identify var=studentstaff stationarity=(phillips=(0,1));
run;
```

The following sections document and explain the output obtained from the SAS program.

14.3.1 Student to Staff ratio in the Higher Education system

The following is the time plot output from the SAS program for the absorption rate per full time equivalent researchers in the system.

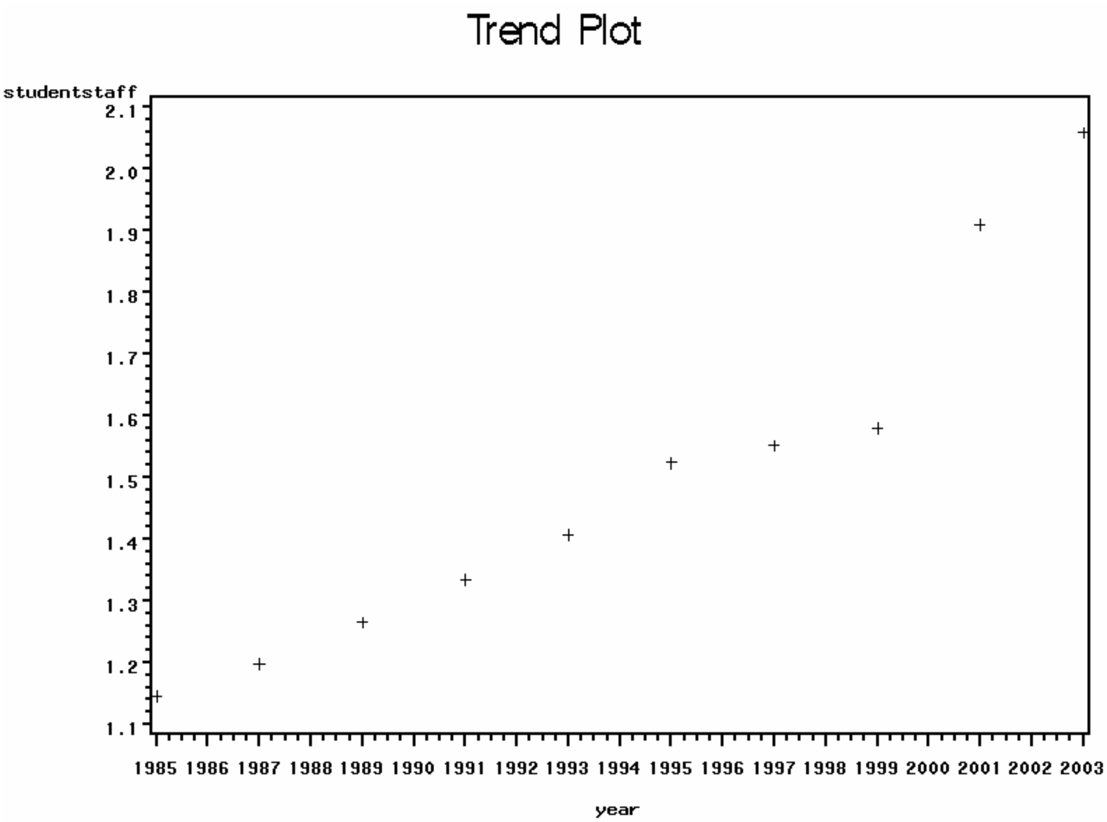

Figure 14-9 Time plot of the Student to staff ratio in the Higher Education system

From Figure 14-9 can be seen that the time plot shows an upward trend line. We therefore make use of the "Trend" specification in the stationarity test output results.

Table 14-18: Phillips-Perron test output for variable "`studentstaff`"

```
Name of Variable = studentstaff

Mean of Working Series    1.497557
Standard Deviation        0.282596
Number of Observations          10

Phillips-Perron Unit Root Tests
```

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau
Zero Mean	0	0.6552	0.8026	3.57	0.9984
	1	0.6573	0.8030	3.87	0.9989
Single Mean	0	1.4296	0.9825	1.15	0.9933
	1	1.6149	0.9854	1.59	0.9973
Trend	0	-3.8832	0.8411	-0.95	0.8959
	1	-3.9095	0.8390	-0.95	0.8949

From the Phillips Perron test output obtained from SAS we read the following values for the probability statistics.

`Pr < Tau = 0.8959` for $\ell = 0$ en

`Pr < Tau = 0.8949` for $\ell = 1$.

Both p-values are greater than 0.05. We therefore cannot reject H_0: $d = 1$. We have to conclude that `studentstaff` has a unit root and is non-stationary.

14.3.2 Percentage time spent on R&D

The following is the time plot output from the SAS program for the RD Knowledge stock with Full time equivalent personnel in the system.

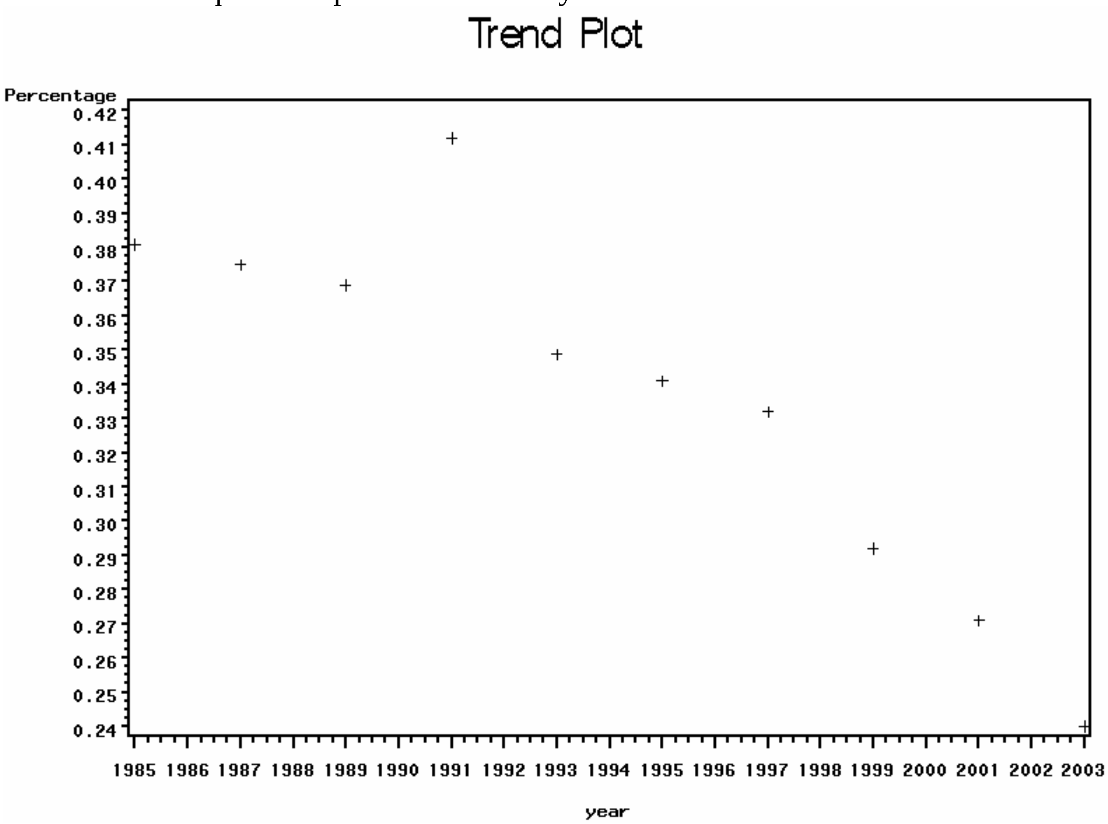

Figure 14-10 Time plot - FTE researcher interacting with R&D knowledge

From Figure 14-10 can be seen that the time plot shows an upward trend line. We therefore make use of the "Trend" specification in the stationarity test output results.

Table 14-19: Phillips-Perron test output for variable "`RDFTE`"

```
                         The ARIMA Procedure

                   Name of Variable = Percentage

                   Mean of Working Series     0.33618
                   Standard Deviation        0.050983
                   Number of Observations          10

                         Phillips-Perron Unit Root Tests

     Type             Lags          Rho     Pr < Rho         Tau     Pr < Tau
```

```
Zero Mean      0    -0.3958    0.5676    -1.57    0.1021
               1    -0.3875    0.5693    -1.82    0.0649
Single Mean    0     0.3426    0.9506     0.15    0.9501
               1     1.1173    0.9762     0.70    0.9830
Trend          0    -6.4884    0.5616    -1.89    0.5778
               1    -5.3570    0.6993    -1.76    0.6379
```

From the Phillips Perron test output obtained from SAS we read the following values for the probability statistics.

`Pr < Tau = 0.5778` for ℓ = 0 en

`Pr < Tau = 0.6379` for ℓ = 1.

Both *p*-values are greater than 0.05. We therefore cannot reject H_0: $d = 1$. We have to conclude that `Percentage` has a unit root and is non-stationary.

14.3.3 Model estimation – The time spent on R&D

As both variables are non-stationary, we should now fit a model and then test for cointegration in the residual.

Table 14-20: SAS code for the model estimation procedure

```
proc autoreg data= HES.hesloglin;
model absorbedR = rdFTE wsperhc
/ method= ml nlag=1 dwprob;
output out=b r=residual;
run;
```

Table 14-21: SAS output for the model estimation of Absorptive capacity in the HES

```
                          The AUTOREG Procedure

                    Dependent Variable     Percentage


                    Ordinary Least Squares Estimates

          SSE                  0.00378232    DFE                    8
          MSE                   0.0004728    Root MSE         0.02174
          SBC                  -45.816074    AIC           -46.421244
          Regress R-Square         0.8545    Total R-Square    0.8545
          Durbin-Watson            2.1775    Pr < DW           0.4613
          Pr > DW                  0.5387
NOTE: Pr<DW is the p-value for testing positive autocorrelation, and Pr>DW is the p-value
for testing negative autocorrelation.

                             Phillips-Ouliaris
                            Cointegration Test

                    Lags           Rho           Tau

                       1       -9.7900       -3.2686

                                             Standard                 Approx
        Variable        DF      Estimate        Error    t Value    Pr > |t|

        Intercept        1        0.5859       0.0371      15.80      <.0001
        studentstaff     1       -0.1668       0.0243      -6.85      0.0001


                     The REG Procedure
```

```
                          Model: MODEL1
                  Dependent Variable: Percentage

                       Analysis of Variance

                                     Sum of          Mean
Source                    DF        Squares        Square    F Value    Pr > F

Model                      1        0.02221       0.02221      46.98    0.0001
Error                      8        0.00378    0.00047279
Corrected Total            9        0.02599


             Root MSE              0.02174    R-Square     0.8545
             Dependent Mean        0.33618    Adj R-Sq     0.8363
             Coeff Var             6.46789


                       Parameter Estimates

                              Parameter       Standard
  Variable        DF          Estimate          Error    t Value    Pr > |t|

  Intercept        1           0.58592        0.03708      15.80      <.0001
  studentstaff     1          -0.16677        0.02433      -6.85      0.0001

                       The REG Procedure
                         Model: MODEL1
                  Dependent Variable: Percentage

                  Durbin-Watson D                  2.177
                  Number of Observations              10
                  1st Order Autocorrelation       -0.115
```

From the model estimation output obtained we can make the following conclusion:

The test for autocorrelation use is the Durban Watson test statistic. The `Durbin Watson` test statistic is `2.177 with` `(Pr < DW = 0.4613) > 0.05 and (Pr < DW = 0.5387) < 0.95.` This indicates that we can we therefore can conclude that the model does not have autocorrelation.

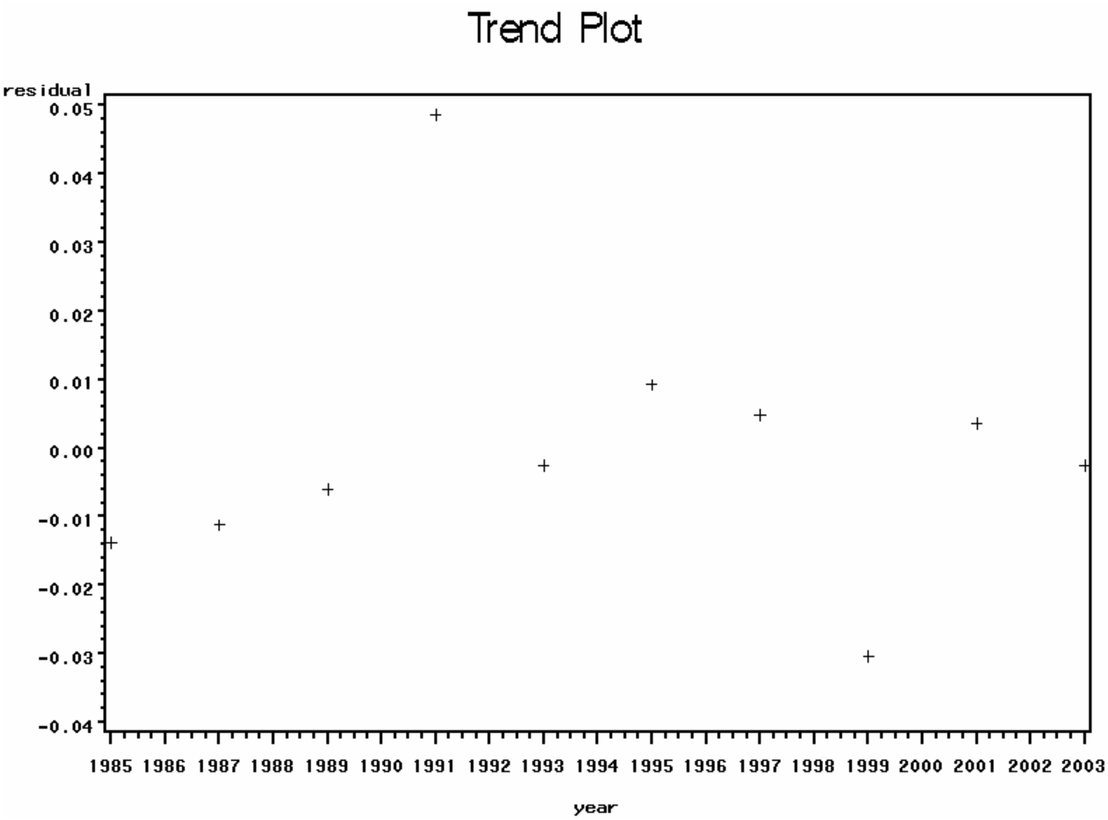

Figure 14-11 Time plot - residual of the HES knowledge absorption

From Figure 14-11 can be seen that the time plot seems to be scattered around 0. From the results we van also read the `Mean of Working Series 5.8E-17`. We therefore make use of the "Zero mean" specification in the stationarity test output results.

Table 14-22: Phillips Perron tests output for the residual

```
                    The ARIMA Procedure

                   Name of Variable = residual
                   Mean of Working Series      5.8E-17
                   Standard Deviation         0.019448
                   Number of Observations           10

                   Phillips-Perron Unit Root Tests
```

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau
Zero Mean	0	-10.0352	0.0080	-3.26	0.0044
	1	-9.7900	0.0092	-3.27	0.0044
Single Mean	0	-10.0445	0.0378	-3.06	0.0683
	1	-9.7461	0.0445	-3.07	0.0676
Trend	0	-10.0642	0.1852	-2.89	0.2062
	1	-9.4754	0.2329	-2.90	0.2032

Since an intercept is included in the model fitted, an intercept is included. For (n-1) = 1, the values are obtained from the Critical values for the Phillips Z Statistic or the Dickey Fuller t Statistic when applied to Residuals from Spurious Cointegration Regression (See table Table 14-22). The critical value for the 7.5% level is -3.20.

From the Phillips Perron test output obtained from SAS we read the following values for the probability statistics.

`Tau = -3.26` for $\ell = 0$ en

`Tau = -3.27` for $\ell = 1$.

This means that we can therefore reject the null hypothesis of unit root since the `Tau` values are smaller than the critical value. The residues can be deemed stationary and the variables are cointegrated. We can therefore conclude that the regression is not spurious.

15 APPENDIX D

15.1 Absorption of Knowledge (Pub)

The rate at which the system is able to produce new knowledge output is computed through the contribution made form different stocks in the system. The following expression is formulated for the R&D knowledge absorption rate in the system:

- $R_{Absorptionr}$: Absorption rate of knowledge in the system
- $S_{R\&Doutput}$: RD output stock in the system
- S_{FTE} : Stock of Full Time Equivalent people in the system
- S_{World} : Available external knowledge stock (Patents)
- S_{HC} : Headcount personnel employed in the system

A multiplicative model is developed for the absorption rate per Full Time Equivalent person working in the system:

$$\frac{R_{Absorptionr}}{R^*_{Absorption}} = f*(\frac{S_{R\&Doutput}}{S^*_{R\&Doutput}}*\frac{S_{FTE}}{S^*_{FTE}})^d*(\frac{S_{World}}{S^*_{World}}/\frac{S_{HC}}{S^*_{HC}})^e \qquad \text{15-1}$$

This expression is linearised by taking the log-linear form:

$$\ln(\frac{R_{Absorptionr}}{R^*_{Absorption}}) = f + d*\ln(\frac{S_{R\&Doutput}}{S^*_{R\&Doutput}}*\frac{S_{FTE}}{S^*_{FTE}}) + e*\ln(\frac{S_{World}}{S^*_{World}}/\frac{S_{HC}}{S^*_{HC}}) \qquad \text{15-2}$$

This is the expression used to perform the regression for estimating the parameters d, e and f. The regression is executed and the following estimates for the parameters are obtained:

The section describes the variables included in the model to estimate the rate of knowledge absorption in the system. The following SAS program was used.

Table 15-1: SAS code for stationarity tests

```
goptions reset=all cback=white colors=(black) lfactor=2
border;
title1 'Trend Plot';
proc gplot data=Pub.paploglinear;
plot (absorbedR RDftetype wsfte)*year;
plot absorbedR*(RDftetype wsfte);
run;

* test for stationarity of the 3 series using arima procedure
*;
proc arima data=Pub.paploglinear;
identify var=absorbedR  stationarity=(phillips=(0,1));
identify var=RDftetype stationarity=(phillips=(0,1));
identify var=wsfte stationarity=(phillips=(0,1));
run;
```

The following sections document and explain the output obtained from the SAS program.

15.1.1 Absorption rate of knowledge in the system

The following is the time plot output from the SAS program for the absorption rate per full time equivalent researchers in the system.

$$AbsorbedR = \ln(\frac{R_{Absorptionr}}{R^{*}_{Absorption}}) \qquad \textbf{15-3}$$

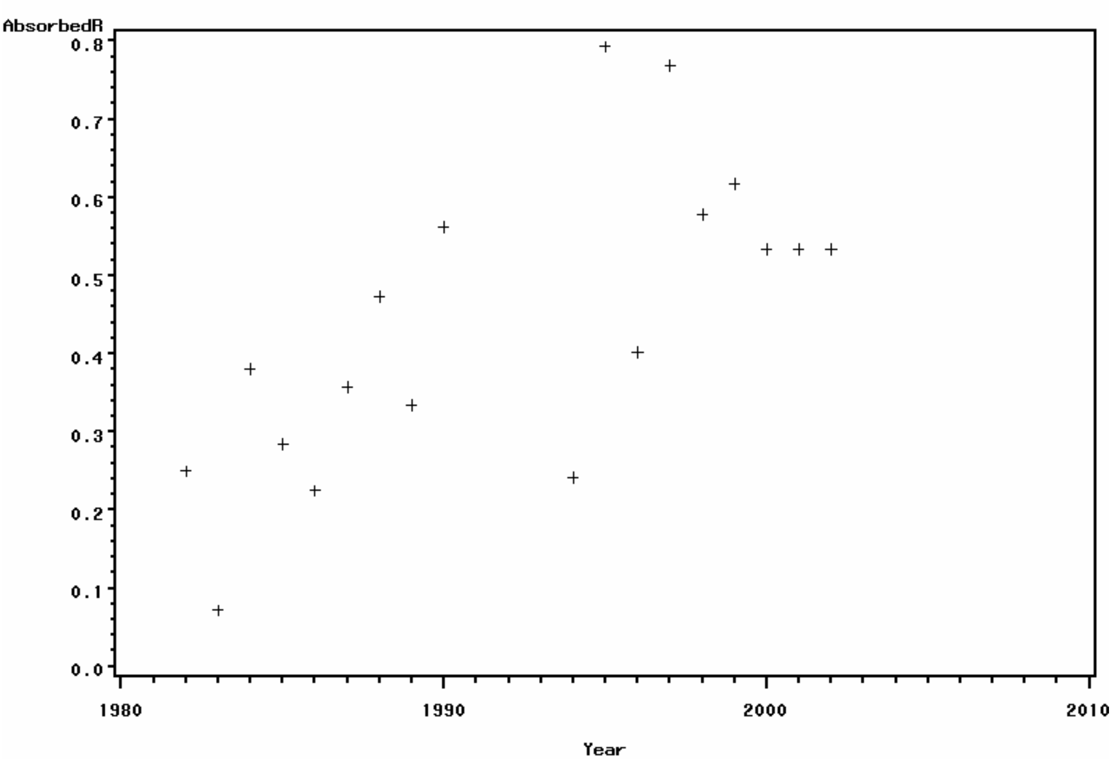

Figure 15-1 Time plot of the absorption rate in the Public sector

From Figure 15-1 can be seen that the time plot shows an upward trend. We therefore make use of the "Trend" specification in the stationarity test output results.

Table 15-2: Phillips-Perron test output for variable "AbsorbedR"

```
                    The ARIMA Procedure

                Name of Variable = AbsorbedR

              Mean of Working Series    0.286634
              Standard Deviation        0.274244
              Number of Observations          21

                    The ARIMA Procedure

               Phillips-Perron Unit Root Tests

Type           Lags        Rho    Pr < Rho       Tau    Pr < Tau

Zero Mean         0    -3.7431      0.1704     -1.30      0.1719
                  1    -2.4229      0.2734     -1.01      0.2690
Single Mean       0   -12.1070      0.0420     -2.78      0.0793
                  1   -11.0751      0.0603     -2.69      0.0922
Trend             0   -15.3530      0.0761     -3.27      0.1006
                  1   -15.0457      0.0841     -3.25      0.1036
```

From the Phillips Perron test output obtained from SAS we read the following values for the probability statistics.

`Pr < Tau = 0.1006` for $\ell = 0$ en

`Pr < Tau = 0.1036` for $\ell = 1$.

Both *p*-values are greater than 0.05. We therefore cannot reject H_0: $d = 1$. We have to conclude that `AbsorbedR` has a unit root and is non-stationary.

15.1.2 R&D Knowledge Stock and FTE researchers interaction

The following is the time plot output from the SAS program for the RD Knowledge stock with Full time equivalent personnel in the system.

$$RDfte = \ln(\frac{S_{R\&Doutput}}{S^*_{R\&Doutput}} * \frac{S_{FTE}}{S^*_{FTE}}) \qquad \textbf{15-4}$$

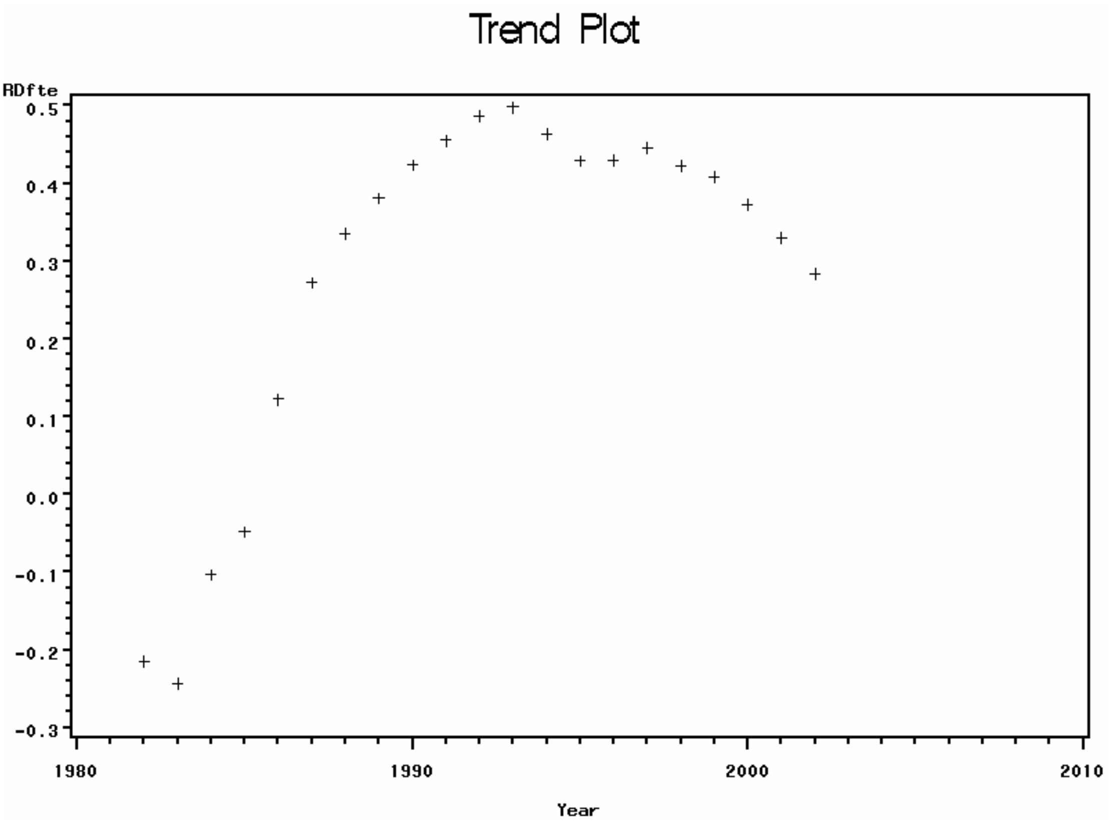

Figure 15-2 Time plot - FTE researcher interacting with R&D knowledge

From Figure 15-2 can be seen that the time plot shows a trend. We therefore make use of the "Trend" specification in the stationarity test output results.

Table 15-3: Phillips-Perron test output for variable "RDFTE"

Name of Variable = RDfte

Mean of Working Series	0.283461
Standard Deviation	0.229626
Number of Observations	21

Phillips-Perron Unit Root Tests

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau
Zero Mean	0	-0.2185	0.6177	-0.26	0.5806
	1	-0.4416	0.5695	-0.41	0.5217
Single Mean	0	-3.0877	0.6216	-2.94	0.0584
	1	-3.2346	0.6025	-2.71	0.0894
Trend	0	-0.4260	0.9895	-0.30	0.9842
	1	-0.4173	0.9896	-0.30	0.9844

From the Phillips Perron test output obtained from SAS we read the following values for the probability statistics.

Pr < Tau = 0.9842 for $\ell = 0$ en

Pr < Tau = 0.9844 for $\ell = 1$.

Both *p*-values are greater than 0.05. We therefore cannot reject H_0: $d = 1$. We have to conclude that RDFTE has a unit root and is non-stationary.

15.1.3 The external knowledge stock per headcount

The following is the time plot output from the SAS program for the World Knowledge Stock per R&D staff in the system.

$$Patwsperhc = \ln(\frac{S_{World}}{S^*_{World}} / \frac{S_{HC}}{S^*_{HC}})$$

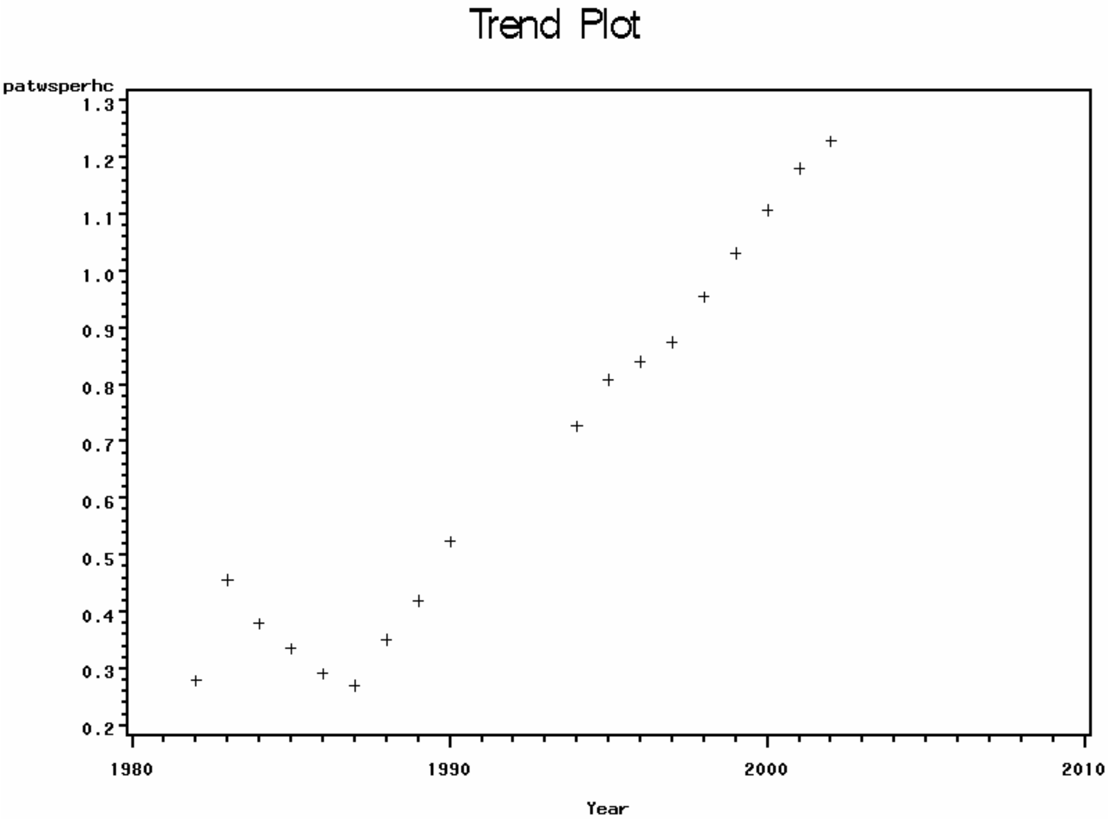

Figure 15-3 Time plot - World stock of knowledge per HC researcher

From Figure 15-3 can be seen that the time plot shows an upward trend line. We therefore make use of the "Trend" specification in the stationarity test output results.

Table 15-4: Phillips-Perron test output for variable "Patwsperhc"

Name of Variable = patwsperhc

Mean of Working Series 0.665825
Standard Deviation 0.30554
Number of Observations 21

Phillips-Perron Unit Root Tests

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau
Zero Mean	0	1.3564	0.9413	3.59	0.9996
	1	1.3514	0.9408	3.45	0.9994
Single Mean	0	0.6846	0.9750	0.73	0.9897
	1	0.6380	0.9736	0.64	0.9872
Trend	0	-4.5243	0.8265	-1.53	0.7840
	1	-4.7594	0.8075	-1.57	0.7697

From the Phillips Perron test output obtained from SAS we read the following values for the probability statistics.

Pr < Tau = 0.7840 for $\ell = 0$ en

Pr < Tau = 0.7697 for $\ell = 1$.

Both *p*-values are greater than 0.05. We therefore cannot reject H_0: d = 1. We have to conclude that `patwsperhc` has a unit root and is non-stationary.

15.1.4 Colinearity tests

First however we should make sure that the variables are not collinear. The following is the test results obtained from SAS for the Colinearity test.

Table 15-5: Colinearity diagnostics for the model variables

```
                    Collinearity Diagnostics

                                       Condition
     Number      Eigenvalue               Index

          1         2.68537             1.00000
          2         0.23340             3.39199
          3         0.08123             5.74964

                     ---------Proportion of Variation---------
     Number        Intercept          RDfte              patwsperhc
          1        0.02158            0.03438            0.01676
          2        0.23617            0.80117            0.02376
          3        0.74226            0.16444            0.95948
```

Larger values suggest potential near colinearity. Belsley, Kuh and Welsch (2000) recommend interpreting the Condition index greater or equal than 30 to reflect moderate to severe colinearity, worthy of further investigation. Since all the Condition indexes from the regression model are much smaller than 30, the conclusion can be made that colinearity is not a problem in this case.

Proportion of Variation

The variance proportion indicates for each predictor the proportion of total variance of its estimated regression coefficients associated with a particular principal component. The variance proportions suggest colinearity problems if more than one predictor has a high variance proportions of at least 0.5 for such a components suggest a problem. One should definitely be concerned when two or more ladings greater than 0.9 appear on a component with a large condition index (>30). This also does not seem to be a problem since the condition indexes are all smaller than 30.

15.1.5 Model estimation - Absorption rate

As all three variables are non-stationary, we should now fit a model and then test for cointegration in the residual.

Table 15-6: SAS code for the model estimation procedure

```
proc reg data = Pub.paploglinear ;
model  arperftecontract = RDftetype worldS
/tol vif collin;
output out=a r=residual;
run;
```

Table 15-7: SAS output for the model estimation of Absorptive capacity in the HES

```
                    The SAS System        13:11 Monday, January 23, 2006    7

                                The AUTOREG Procedure
```

```
                           Dependent Variable    AbsorbedR

                           Ordinary Least Squares Estimates

            SSE                     0.2953922    DFE                      15
            MSE                       0.01969    Root MSE            0.14033
            SBC                    -14.223913    AIC              -16.895028
            Regress R-Square           0.5317    Total R-Square       0.5317
            Durbin-Watson              3.3536    Pr < DW              0.9979
            Pr > DW                    0.0021
NOTE: Pr<DW is the p-value for testing positive autocorrelation, and Pr>DW is the p-value
for testing negative autocorrelation.

                         Q and LM Tests for ARCH Disturbances

              Order            Q    Pr > Q           LM    Pr > LM

                1         3.2076    0.0733       2.9992     0.0833
                2         3.7507    0.1533       2.9994     0.2232
                3         3.8340    0.2800       2.9994     0.3917
                4         7.5941    0.1076       6.0103     0.1984
                5        10.4175    0.0642       6.0293     0.3034
                6        13.7073    0.0331       6.3609     0.3840
                7        15.7191    0.0278       6.8765     0.4418
                8        15.7699    0.0458       8.3467     0.4004
                9        15.8418    0.0703      10.2484     0.3308
               10        15.8851    0.1030      13.4042     0.2019
               11        16.1797    0.1346      14.5944     0.2018
               12        16.2023    0.1821      15.1744     0.2320

                                                Standard                 Approx
          Variable          DF      Estimate       Error    t Value    Pr > |t|

          Intercept          1        0.2143      0.0750       2.86      0.0120
          RDfte              1        0.3880      0.1719       2.26      0.0393
          patwsperhc         1        0.1936      0.1211       1.60      0.1308

                            Estimates of Autocorrelations

    Lag    Covariance     Correlation    -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1

      0        0.0164        1.000000    |                    |********************|
      1       -0.0112       -0.685361    |      **************|                    |

                     The SAS System        13:11 Monday, January 23, 2006   8

                               The AUTOREG Procedure

                              Preliminary MSE     0.00870

                        Estimates of Autoregressive Parameters

                                                 Standard
                         Lag     Coefficient        Error    t Value

                           1        0.685361     0.194621       3.52

           Algorithm converged.

                             Maximum Likelihood Estimates

            SSE                    0.15202524    DFE                      14
            MSE                       0.01086    Root MSE            0.10421
            SBC                    -22.691358    AIC              -26.252845
            Regress R-Square           0.8553    Total R-Square       0.7590
            Durbin-Watson              1.9361    Pr < DW              0.2412
            Pr > DW                    0.7588
NOTE: Pr<DW is the p-value for testing positive autocorrelation, and Pr>DW is the p-value
for testing negative autocorrelation.

                                                Standard                 Approx
          Variable          DF      Estimate       Error    t Value    Pr > |t|

          Intercept          1        0.1991      0.0350       5.70      <.0001
          RDfte              1        0.3787      0.0819       4.62      0.0004
          patwsperhc         1        0.2187      0.0577       3.79      0.0020
          AR1                1        0.6712      0.1909       3.52      0.0034

                       Autoregressive parameters assumed given.

                                                Standard                 Approx
          Variable          DF      Estimate       Error    t Value    Pr > |t|

          Intercept          1        0.1991      0.0350       5.70      <.0001
          RDfte              1        0.3787      0.0819       4.62      0.0004
          patwsperhc         1        0.2187      0.0576       3.79      0.0020
```

From the model estimation output obtained we can make the following conclusion:

The test for autocorrelation use is the Durban Watson test statistic. The `Durbin Watson` test statistic is `1.9054 with (Pr < DW = 0.2205) > 0.05 and (Pr < DW = 0.7795) < 0.95`. This indicates that we can we therefore can conclude that the autoregressive model does not have autocorrelation.

Due to the small sample size and the limited number of data points available, the heteroscedasticity test is only interpreted up to 2 time lags. The probability for arch disturbances in the model for lags 1 and 2 are larger than 0.05. We can therefore conclude that the modelled relationship does not suffer from heteroscedasticity.

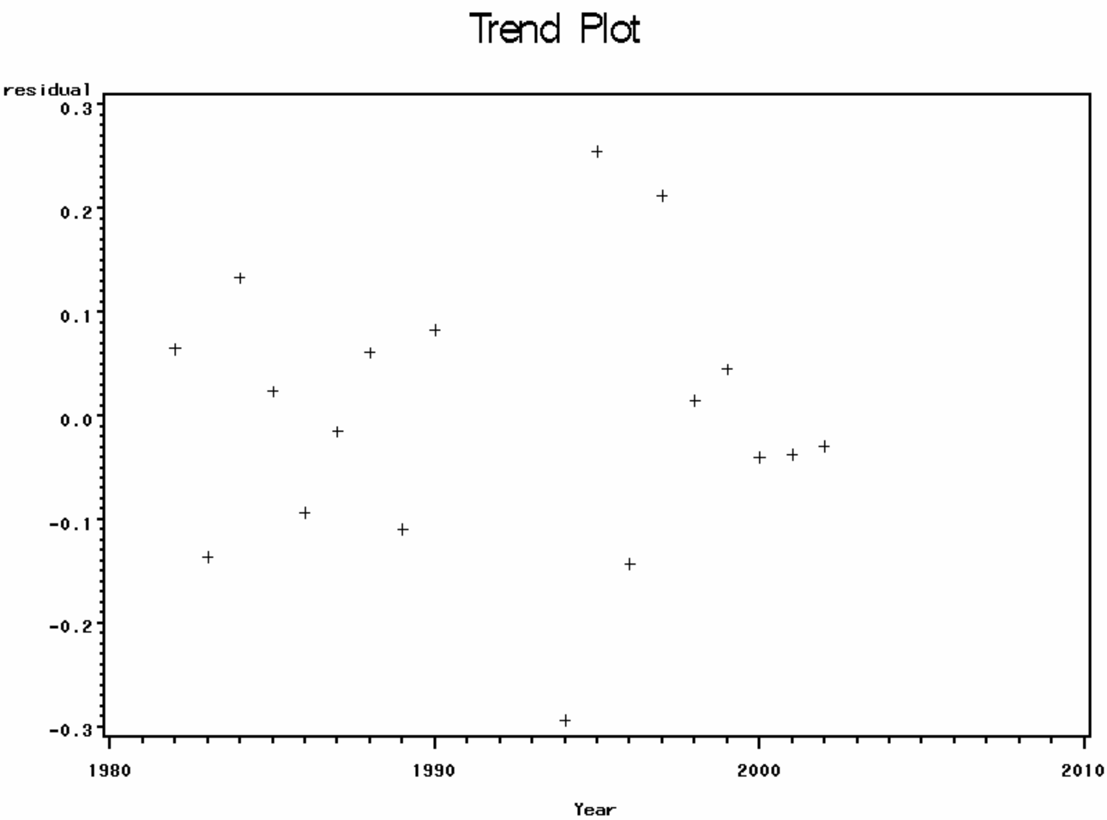

Figure 15-4 Time plot - residual of the HES knowledge absorption

From Figure 15-4 can be seen that the time plot seems to be scattered around 0. From the results we van also read the `Mean of Working Series -0.00167`. We therefore make use of the "Zero mean" specification in the stationarity test output results.

Table 15-8: Test for stationarity of the residual

```
                    The ARIMA Procedure

                Name of Variable = residual

            Mean of Working Series    3.46E-17
            Standard Deviation        0.128104
            Number of Observations          18

                 Phillips-Perron Unit Root Tests

Type             Lags   Rho         Pr < Rho    Tau        Pr < Tau

Zero Mean         0    -28.6839     <.0001      -9.39      <.0001
                  1    -29.0391     <.0001      -9.15      <.0001
Single Mean       0    -28.6795     <.0001      -9.10      <.0001
                  1    -29.0411     <.0001      -8.86      <.0001
Trend             0    -28.7210     <.0001      -8.85      0.0005
                  1    -29.0497     <.0001      -8.64      0.0005
```

Since an intercept is included in the model fitted, an intercept is included in the analysis. For (n-1) = 2, the values are obtained from the Critical values for the Phillips Z Statistic or the Dickey Fuller t Statistic when applied to Residuals from Spurious Cointegration Regression (See **Error! Reference source not found.**). The critical value for the 1% level is -4.31.

From the Phillips Perron test output obtained from SAS we read the following values for the probability statistics.

`Tau = -4.84` for $\ell = 0$ en

`Tau = -4.84` for $\ell = 1$.

This means that we can therefore reject the null hypothesis of unit root since the `Tau` values are smaller than the critical value. The residues can be deemed stationary and the variables are cointegrated. We can therefore conclude that the regression is not spurious.

15.2 Creation of new knowledge – Scientific papers (Public sector)

The rate at which the system is able to produce new knowledge output is computed through the contribution made form different stocks in the system. The following expression is formulated for the R&D output produced by human resources in the Public sector:

- R_{Paper} : Rate at which R&D output is generated in the system (Papers)
- S_{FTE} : Ratio of full time equivalent R&D staff in the system
- $S_{Absorbed}$: Absorbed knowledge stock in the system.
- $A_{Contract}$: The ration of research directed towards contract research
- $A_{Basi\&Applied}$: The ratio of research directed toward Basic and Applied research

A multiplicative model is developed for the development rate of papers per full time person working in the system:

$$\frac{R_{Paper}}{R^*_{Paper}} = d*\left(\frac{S_{Absorbed}}{S_{Absorbed}} * \frac{S_{FTE}}{S^*_{FTE}} * \frac{A_{Basic\&Applied}}{A^*_{Bsic\&Applied}}\right)^a \left(\frac{A_{State}}{A^*_{State}}\right)^b \left(\frac{S_{FTE}}{S^*_{FTE}}\right)^c \quad \textbf{15-5}$$

This expression is linearised by taking the log-linear form:

$$\ln(\frac{R_{Paper}}{R^*_{Paper}}) = \ln(d) + a * \ln(\frac{S_{Absorbed}}{S^*_{Absorbed}} * \frac{S_{FTE}}{S^*_{FTE}} * \frac{A_{Basic\&Applied}}{A^*_{Bsic\&Applied}}) + b * \ln(\frac{A_{State}}{A^*_{State}}) + \mathbf{c*} \ln(\frac{S_{FTE}}{S^*_{FTE}}) \quad \mathbf{15\text{-}6}$$

This is then the expression used to perform the regression for estimating the parameters a, b, c and d. The regression is executed and the following estimates for the parameters are obtained:

The section describes the variables included in the model to estimate the rate of knowledge creation in the system. The following SAS program was used.

Table 15-9: SAS program code for stationarity tests and trend plots

```
goptions reset=all cback=white colors=(black) lfactor=2 border;
title1 'Trend Plot';
proc gplot data=Pub.paploglinear;
plot (RDpapersr absftetype ftetot percstate)*year;

run;

* test for stationarity of the 3 series using arima procedure *;
proc arima data=Pub.paploglinear;
identify var=RDpapersr stationarity=(phillips=(0,1));
identify var=absftetype stationarity=(phillips=(0,1));
identify var=ftetot stationarity=(phillips=(0,1));
identify var=percstate stationarity=(phillips=(0,1));
run;
```

The following sections document and explain the output obtained from the SAS program.

15.2.1 R&D output produced

The following is the time plot output from the SAS program for the R&D output (papers) created in the system.

$$RDPapersR = \ln(\frac{R_{Paper}}{R^*_{Paper}}) \quad \mathbf{15\text{-}7}$$

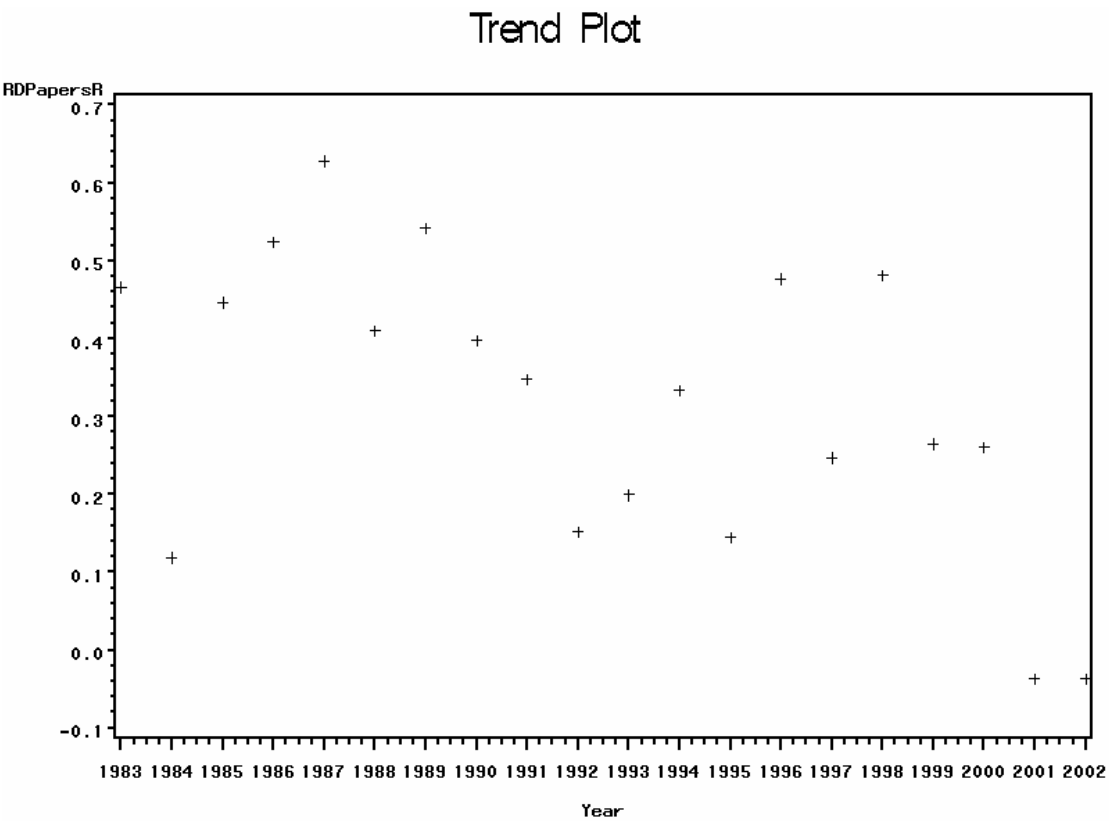

Figure 15-5 Time plot of the Knowledge creation rate per FTE

From Figure 15-5 can be seen that the time plot shows a downward trend. We therefore make use of the "Trend" specification in the stationarity test output results.

Table 15-10: SAS output for Phillips Perron test for variable "RDpapersr"

```
                         The ARIMA Procedure
                     Name of Variable = RDPapersR

                   Mean of Working Series    0.318457
                   Standard Deviation        0.182812
                   Number of Observations          20

                          The ARIMA Procedure

                     Phillips-Perron Unit Root Tests

 Type            Lags        Rho     Pr < Rho       Tau     Pr < Tau

Zero Mean         0      -3.5094      0.1848      -1.55      0.1098
                  1      -2.5793      0.2572      -1.44      0.1347
Single Mean       0     -11.6431      0.0475      -2.46      0.1399
                  1     -11.1338      0.0570      -2.41      0.1526
Trend             0     -16.0664      0.0545      -3.25      0.1040
                  1     -15.6509      0.0629      -3.23      0.1085
```

From the Phillips Perron test output obtained from SAS we read the following values for the probability statistics.

`Pr < Tau = 0.1040` for $\ell = 0$ en

`Pr < Tau = 0.1085` for $\ell = 1$.

Both *p*-values are greater than 0.05. We therefore cannot reject H_0: $d = 1$. We have to conclude that `RDpapersR` has a unit root and is non-stationary.

15.2.2 Absorbed Knowledge stock

The variable is the Absorbed knowledge stock and FTE R&D staff multiplied with the % time they are spending on Basic and Applied research.

$$Variable = absftetype = \ln\left(\frac{S_{Absorbed}}{S_{Absorbed}} * \frac{S_{FTE}}{S^{*}_{FTE}} * \frac{A_{Basic\&Applied}}{A^{*}_{Bsic\&Applied}}\right) \quad \textbf{15-8}$$

The following is the time plot output from the SAS program for the Absorbed Knowledge stock and the interaction with the FTE R&D staff focussing on Basic and Applied research.

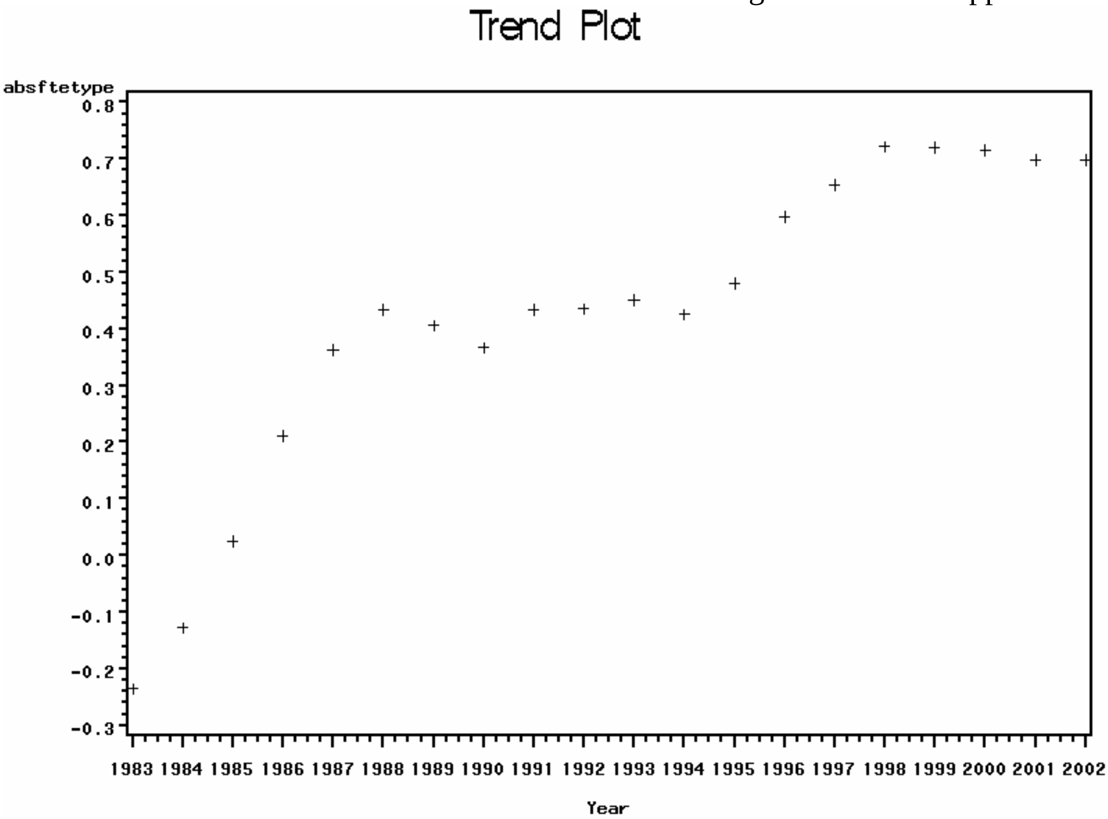

Figure 15-6 Time plot - Absorbed knowledge stock per HC

From Figure 15-6 can be seen that the time plot shows a trend line. We therefore make use of the "Trend" specification in the stationarity test output results.

Table 15-11: SAS output for Phillips Perron test for variable "Absftetype"

```
Name of Variable = absftetype

Mean of Working Series    0.424206
Standard Deviation        0.269425
Number of Observations          20

                    Phillips-Perron Unit Root Tests

Type           Lags        Rho    Pr < Rho       Tau    Pr < Tau

Zero Mean         0     0.6314      0.8221      0.85      0.8857
                  1     0.4476      0.7789      0.46      0.8034
Single Mean       0    -3.2008      0.6051     -3.81      0.0104
                  1    -3.3340      0.5878     -3.34      0.0274
Trend             0    -4.9511      0.7876     -2.38      0.3754
                  1    -5.6666      0.7222     -2.34      0.3968
```

From the Phillips Perron test output obtained from SAS we read the following values for the probability statistics.

`Pr < Tau = 0.7748` for $\ell = 0$ en

`Pr < Tau = 0.7764` for $\ell = 1$.

Both *p*-values are greater than 0.05. We therefore cannot reject H_0: $d = 1$. We have to conclude that `Absftetype` has a unit root and is non-stationary.

15.2.3 FTE total

The following is the time plot output from the SAS program for the Full time equivalent R&D staff in the system.

$$FTEtot = \ln(\frac{S_{FTE}}{S^{*}_{FTE}}) \qquad \textbf{15-9}$$

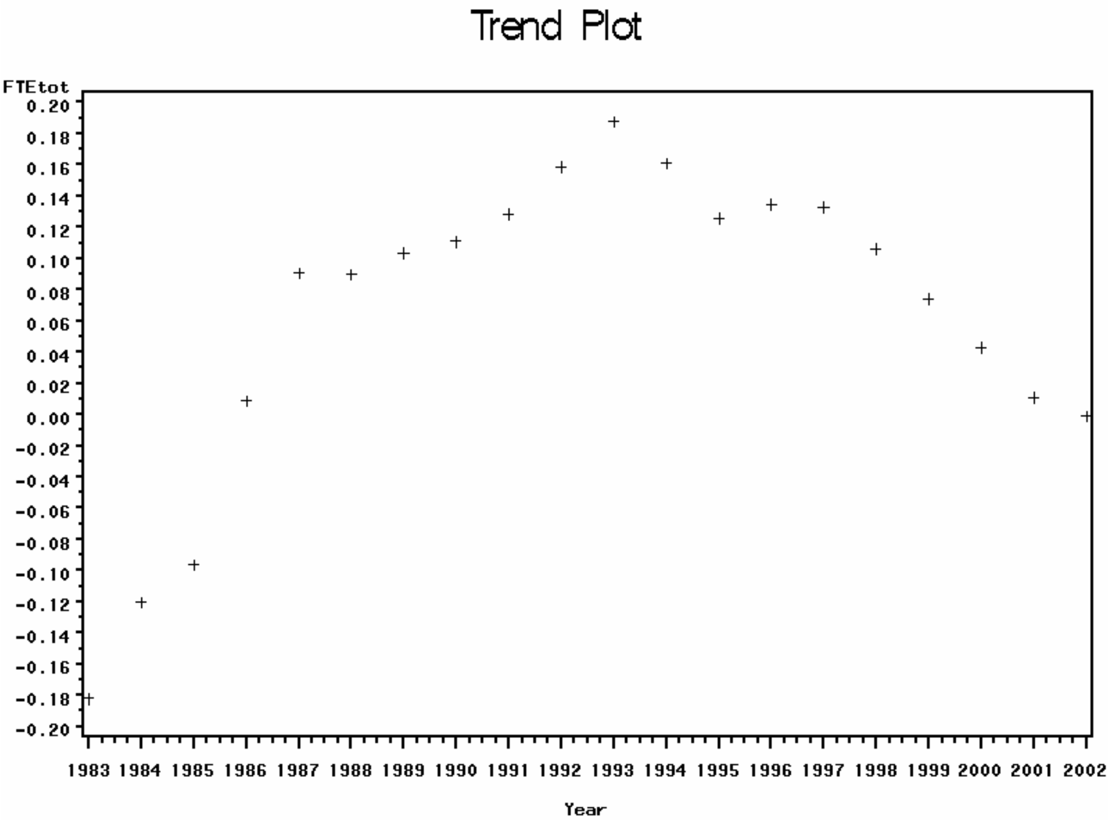

Figure 15-7 Time plot for the FTE variable in the system

From Figure 15-7 can be seen that the time plot shows a trend line. We therefore make use of the “Trend” specification in the stationarity test output results.

Table 15-12: SAS output for Phillips Perron test for variable “Ftetot”

Name of Variable = FTEtot

Mean of Working Series	0.063345
Standard Deviation	0.097222
Number of Observations	20

Phillips-Perron Unit Root Tests

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau
Zero Mean	0	-2.2394	0.2915	-1.59	0.1021
	1	-2.8032	0.2372	-1.58	0.1037
Single Mean	0	-4.5078	0.4436	-3.11	0.0426
	1	-4.9578	0.3943	-2.83	0.0728
Trend	0	-2.0418	0.9584	-1.55	0.7748
	1	-2.0614	0.9578	-1.54	0.7764

From the Phillips Perron test output obtained from SAS we read the following values for the probability statistics.

Pr < Tau = 0.7748 for $\ell = 0$ en

Pr < Tau = 0.7764 for $\ell = 1$.

Both *p*-values are greater than 0.05. We therefore cannot reject H_0: $d = 1$. We have to conclude that `FTEtot` has a unit root and therefore is non-stationary.

15.2.4 Percentage R&D funding from the State

The following is the time plot output from the SAS program for the percentage of total funding directed towards non-contract research.

$$Percstate = \ln(\frac{A_{State}}{A^*_{State}})$$

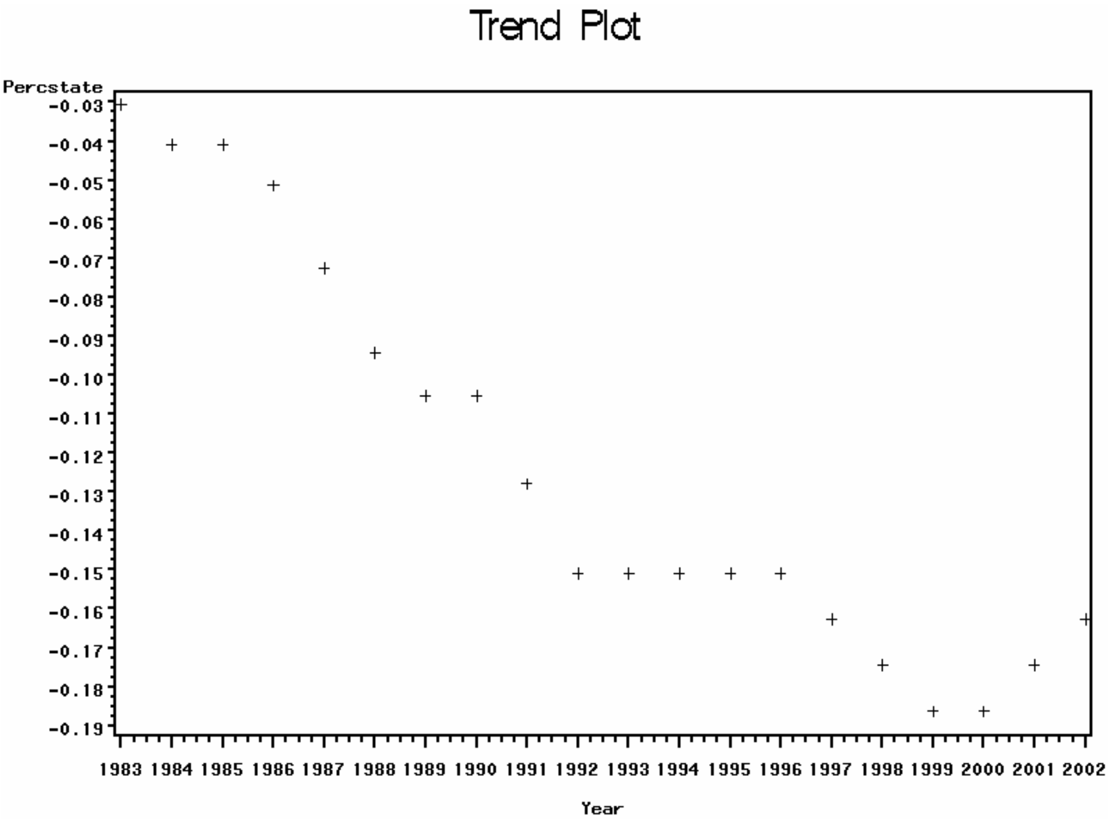

Figure 15-8 Time plot - "Percentage non-contract funding" variable

From Figure 15-8 can be seen that the time plot shows a downwards trend line. We therefore make use of the "Trend" specification in the stationarity test output results.

Table 15-13: SAS output for Phillips Perron test for variable "Percstate"

```
                    Name of Variable = Percstate

                  Mean of Working Series    -0.12347
                  Standard Deviation        0.050749
                  Number of Observations          20

                    Phillips-Perron Unit Root Tests

 Type              Lags         Rho      Pr < Rho        Tau     Pr < Tau

Zero Mean           0        0.9546       0.8874        1.91       0.9825
                    1        0.9006       0.8779        1.49       0.9610
Single Mean         0       -0.9254       0.8793       -1.20       0.6507
                    1       -0.9873       0.8735       -1.16       0.6694
Trend               0       -1.5686       0.9723       -0.59       0.9680
                    1       -2.3244       0.9512       -0.80       0.9483
```

From the Phillips Perron test output obtained from SAS we read the following values for the probability statistics.

`Pr < Tau = 0.9680` for $\ell = 0$ en

`Pr < Tau = 0.9483` for $\ell = 1$.

Both *p*-values are greater than 0.05. We therefore cannot reject H_0: $d = 1$. We have to conclude that `Percstate` has a unit root and therefore is non-stationary.

15.2.5 Colinearity tests

First however we should make sure that the variables are not collinear. The following is the test results obtained from SAS for the Colinearity test.

Table 15-14: Colinearity diagnostics for the model variables

```
                  Collinearity Diagnostics

                                        Condition
      Number        Eigenvalue          Index
        1             3.39433           1.00000
        2             0.48561           2.64383
        3             0.10669           5.64036
        4             0.01337          15.93298

              ---------------Proportion of Variation----------------
  Number    Intercept      absftetype       FTEtot          Percstate
  1         0.00675        0.00323          0.02198          0.00176
  2              0.05268        0.00000497       0.55275          0.00162
  3              0.35634        0.15964          0.42527          0.00892
  4              0.58423        0.83713          1.876643E-8      0.98771
```

Larger values suggest potential near colinearity. Belsley, Kuh and Welsch (2000) recommend interpreting the Condition index greater or equal than 30 to reflect moderate to severe colinearity, worthy of further investigation. Since all the Condition indexes from the regression model is much smaller than 30, the conclusion can be made that colinearity is not a problem in this case.

Proportion of Variation

The variance proportion indicates for each predictor the proportion of total variance of its estimated regression coefficients associated with a particular principal component. The variance proportions suggest colinearity problems if more than one predictor has a high variance proportions of at least 0.5 for such a components suggest a problem. One should definitely be concerned when two or more ladings greater than 0.9 appear on a component with a large condition index (>30). This also does not seem to be a problem since the condition indexes are all small values.

15.2.6 Model estimation the rate of Paper Development in the PubS

As all three variables are non-stationary, we should now fit a model and then test for cointegration in the residual.

Table 15-15: SAS code for the model estimation procedure

```
/*r2 = 58 all p's significant*/
proc reg data = Pub.paploglinear ;
model RDPapersR =  absftetype  ftetot percstate
/tol vif collin spec dw;
output out=b r=residual;
run;

proc reg data = Pub.paploglinear ;
model RDPapersR =  absftetype  ftetot percstate
/tol vif collin spec dw;
output out=b r=residual;
run;

proc gplot data=b;
plot residual*year;
run;

proc arima data=b;
identify var=residual
stationarity=(phillips=(0,1));
run;
```

Table 15-16: SAS output for the model estimation of Absorptive capacity in the HES

```
                          Dependent Variable: RDPapersR

                      Number of Observations Read          20
                      Number of Observations Used          20

                                Analysis of Variance

                                  Sum of           Mean
 Source                   DF     Squares         Square    F Value    Pr > F

 Model                     3     0.32021        0.10674       4.90    0.0133
 Error                    16     0.34820        0.02176
 Corrected Total          19     0.66840

                Root MSE              0.14752    R-Square     0.4791
                Dependent Mean        0.31846    Adj R-Sq     0.3814
                Coeff Var            46.32375

                                Parameter Estimates

                   Parameter     Standard                            Variance
Variable     DF     Estimate        Error    t Value    Pr > |t|    Tolerance     Inflation

Intercept     1     0.69529      0.10741      6.47      <.0001       .              0
absftetype    1     0.49421      0.33156      1.49      0.1555      0.13636        7.33352
FTEtot        1     0.87438      0.45615      1.92      0.0733      0.55327        1.80744
Percstate     1     5.19868      1.70890      3.04      0.0078      0.14467        6.91208

                              Test of First and Second
```

```
                          Moment Specification

                     DF     Chi-Square     Pr > ChiSq

                      9          11.97         0.2151

                Durbin-Watson D                1.959
                Number of Observations            20
                1st Order Autocorrelation     -0.117

                  Dependent Variable    RDPapersR

                  Ordinary Least Squares Estimates

   SSE                      0.34819932    DFE                         16
   MSE                         0.02176    Root MSE               0.14752
   SBC                      -12.273779    AIC                 -16.256708
   Regress R-Square             0.4791    Total R-Square          0.4791
   Durbin-Watson                1.9593    Pr < DW                 0.1850
   Pr > DW                      0.8150
NOTE: Pr<DW is the p-value for testing positive autocorrelation, and Pr>DW is the p-
value for testing negative autocorrelation.

                         Phillips-Ouliaris
                        Cointegration Test

                  Lags            Rho            Tau

                     1      -22.9280       -4.7653

                                             Standard                  Approx
 Variable            DF      Estimate         Error    t Value    Pr > |t|

 Intercept            1        0.6953        0.1074       6.47      <.0001
 absftetype           1        0.4942        0.3316       1.49      0.1555
 FTEtot               1        0.8744        0.4561       1.92      0.0733
 Percstate            1        5.1987        1.7089       3.04      0.0078
```

From the model estimation output obtained we can make the following conclusion:

The R-Square 0.4791 statistic indicate that the model accounts for 47.9% of the variation of the papers produced in the Public sector.

The test for autocorrelation use is the Durban Watson test statistic. The Durbin Watson test statistic is 1.9593 with (Pr < DW = 0.1850 > 0.05 and (Pr < DW = 0.8150) < 0.95. This indicates that we therefore can conclude that the autoregressive model does not have autocorrelation.

Chi-square tests for the first moment specification indicates that the model does not have heteroscedastic errors. The SPEC option performs a model specification test. The null hypothesis for this test maintains that the errors are homoscedastic, independent of the regressor and that several technical assumptions about the model specification are valid. With Pr = 0.2151 we fail to reject the null hypothesis. We can therefore conclude that no heteroscedasiticity is present in the model.

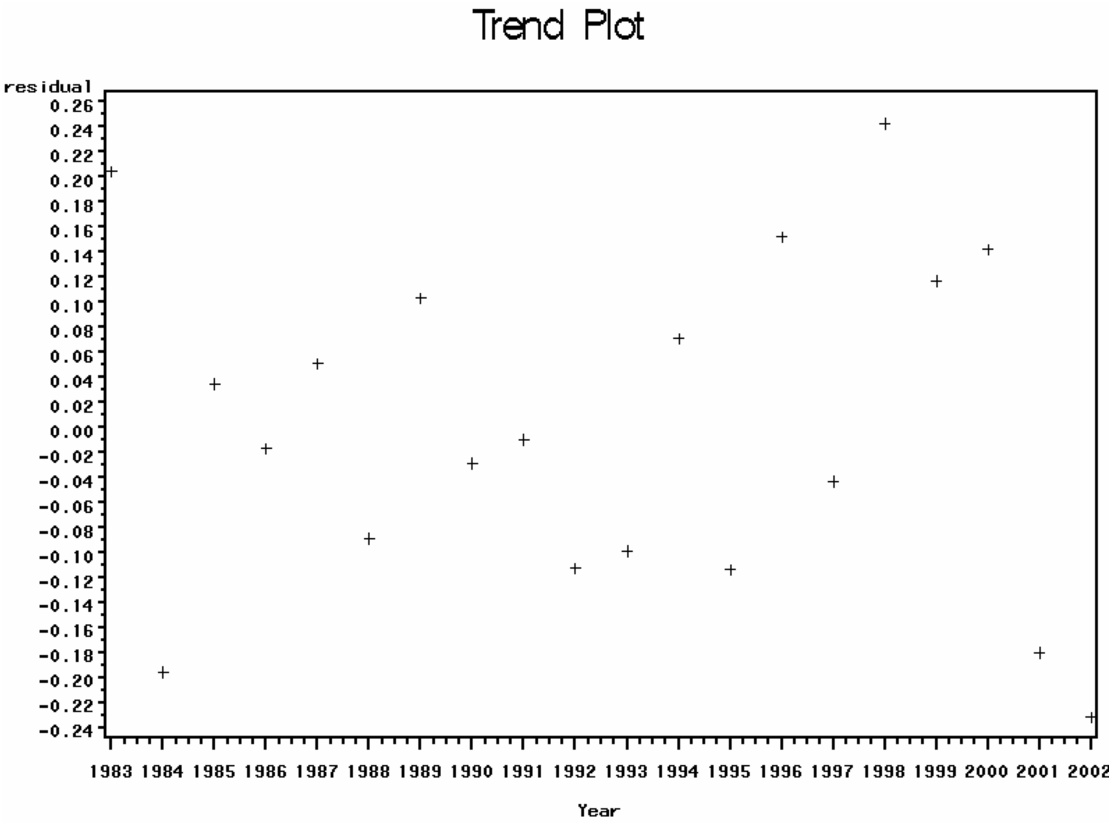

Figure 15-9 Time plot - "Percentage non-contract funding" variable

From **Error! Reference source not found.** can be seen that the time plot seems to be scattered around 0. From the results we van also read the `Mean of Working Series is 2.78E-17`. We therefore make use of the "Zero mean" specification in the stationarity test output results.

Table 15-17: SAS output for residual stationarity test

```
                          The ARIMA Procedure

                      Name of Variable = residual

                  Mean of Working Series    2.78E-17
                  Standard Deviation        0.131947
                  Number of Observations          20

                     Phillips-Perron Unit Root Tests

Type              Lags         Rho    Pr < Rho         Tau      Pr < Tau

Zero Mean           0     -21.6248      <.0001       -4.78       <.0001
                    1     -22.9280      <.0001       -4.77       <.0001
Single Mean         0     -21.4877      0.0005       -4.61       0.0020
                    1     -22.8079      0.0003       -4.60       0.0020
Trend               0     -21.6749      0.0054       -4.50       0.0107
                    1     -22.9344      0.0028       -4.51       0.0107
```

For (n-1) = 3, the values are obtained from the Critical values for the Phillips Z Statistic or the Dickey Fuller t Statistic when applied to Residuals from Spurious Cointegration Regression (See table Table 15-17). The critical value for the 5% level is -4.11.

From the Phillips Perron test output obtained from SAS we read the following values for the probability statistics.

`Tau = -4.78` for $\ell = 0$ en

`Tau = -4.77` for $\ell = 1$.

This means that we can therefore reject the null hypothesis of unit root with a 5% significance level, since the `Tau` values are smaller than the critical value. The residues can be deemed stationary and the variables are cointegrated. We can therefore conclude that the regression is not spurious.

15.3 Creation of new knowledge – Patents (Public sector)

The rate at which the system is able to produce new knowledge output is computed through the contribution made form different stocks in the system. The following expression is formulated for the R&D output productivity per FTE researcher working in the system:

- $R_{Patents}$: R&D output rate in the system (Patents)
- S_{FTE} : FTE researchers in the system
- A_{ExpDev} : Fraction of funding directed towards Experimental Development.
- A_{State} : The ratio of research expenditure funded by the state – assumed to be directed towards non-contract research.

A multiplicative model is developed for the development rate of papers per full time person working in the system:

$$\frac{R_{Patent}}{R^*_{Patent}} = b * (\frac{S_{FTE}}{S^*_{FTE}} * \frac{A_{ExpDev}}{A^*_{ExpDev}} * \frac{A_{Statet}}{A^*_{State}})^a \qquad \textbf{15-10}$$

This expression is linearised by taking the log-linear form:

$$\ln(\frac{R_{Patent}}{R^*_{Patent}}) = \ln(b) + a * \ln(\frac{S_{FTE}}{S^*_{FTE}} * \frac{A_{ExpDev}}{A^*_{ExpDev}} * \frac{A_{Statet}}{A^*_{State}}) \qquad \textbf{15-11}$$

This is then the expression used to perform the regression for estimating the parameters a and b. The regression is executed and the following estimates for the parameters are obtained:

The section describes the variables included in the model to estimate the rate of knowledge creation in the system. The following SAS program was used.

Table 15-18: SAS program code for stationarity tests and trend plots

```
goptions reset=all cback=white colors=(black) lfactor=2
border;
title1 'Trend Plot';
proc gplot data=Pub.patloglinear;
plot (RDout ftepattypestate )*year;
run;
```

```
* test for stationarity of the 3 series using arima procedure
*;
proc arima data=Pub.patloglinear;
identify var=RDout stationarity=(phillips=(0,1));
identify var=ftepattypestate  stationarity=(phillips=(0,1));
run;
```

The following sections document and explain the output obtained from the SAS program.

15.3.1 R&D patent output produced

The following is the time plot output from the SAS program for the R&D output (papers) created per full time equivalent researcher in the system.

$$Rdout = \ln(\frac{R_{Patent}}{R^{*}_{Patent}}) \qquad \textbf{15-12}$$

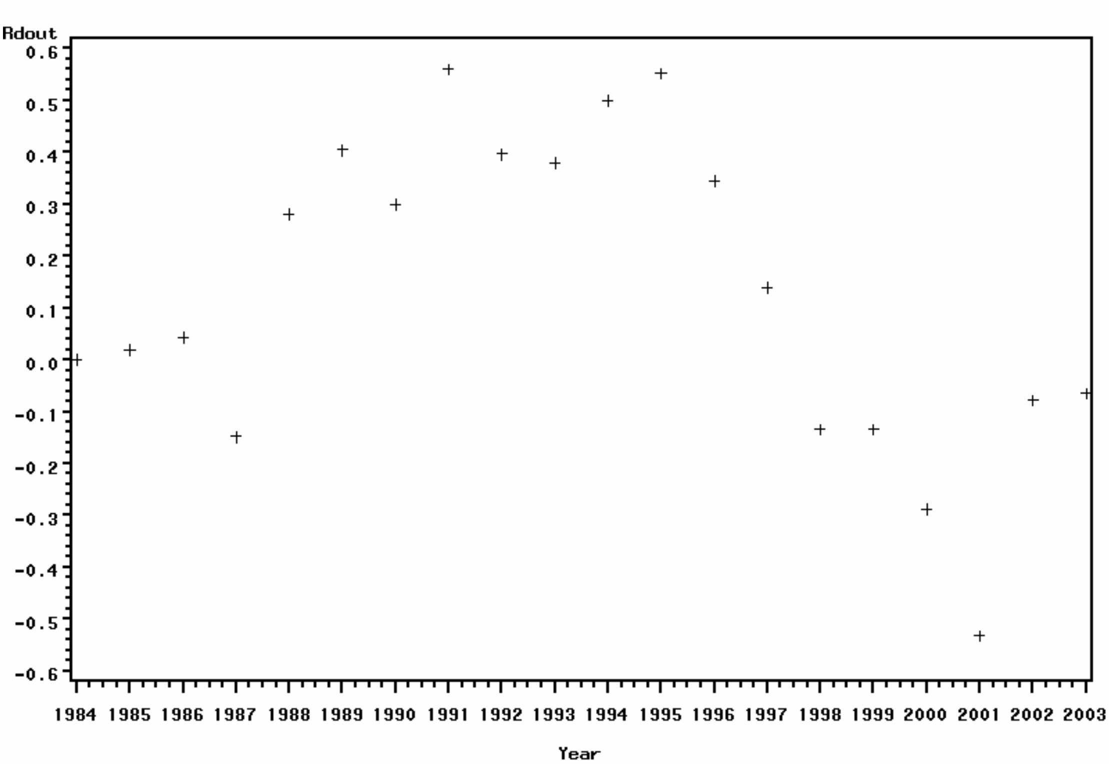

Figure 15-10 Time plot of the Knowledge creation rate per FTE

From Figure 15-7 can be seen that the time plot can be best be described though the "trend" specification. We therefore make use of the "Trend" specification in the stationarity test output results.

Table 15-19: SAS output for Phillips Perron test for variable "Rdout"

```
                    The ARIMA Procedure
                Name of Variable = Rdout
           Mean of Working Series      0.1272
```

```
                    Standard Deviation          0.29628
                    Number of Observations           20

                   Phillips-Perron Unit Root Tests

 Type             Lags         Rho     Pr < Rho        Tau      Pr < Tau

 Zero Mean          0      -3.6847      0.1738      -1.39       0.1475
                    1      -3.6518      0.1758      -1.38       0.1491
 Single Mean        0      -4.3549      0.4612      -1.46       0.5310
                    1      -4.3861      0.4576      -1.47       0.5284
 Trend              0      -5.5918      0.7293      -1.80       0.6669
                    1      -5.4777      0.7400      -1.78       0.6740
```

From the Phillips Perron test output obtained from SAS we read the following values for the probability statistics.

`Pr < Tau = 0.6669` for $\ell = 0$ en

`Pr < Tau = 0.6740` for $\ell = 1$.

Both *p*-values are greater than 0.05. We therefore cannot reject H_0: $d = 1$. We have to conclude that `RDout` has a unit root and is non-stationary.

15.3.2 Full time Staff (Experimental development research, non contract)

The following is the time plot output from the SAS program for the FTE people in the system doing Basic and Applied research. The variable is the FTE R&D staff multiplied with the % time they are spending on basic and applied research and the % of funding spent on non-contract related R&D activities.

$$ftepattypestate = \ln\left(\frac{S_{FTE}}{S^*_{FTE}} * \frac{A_{ExpDev}}{A^*_{ExpDev}} * \frac{A_{Statet}}{A^*_{State}}\right)$$

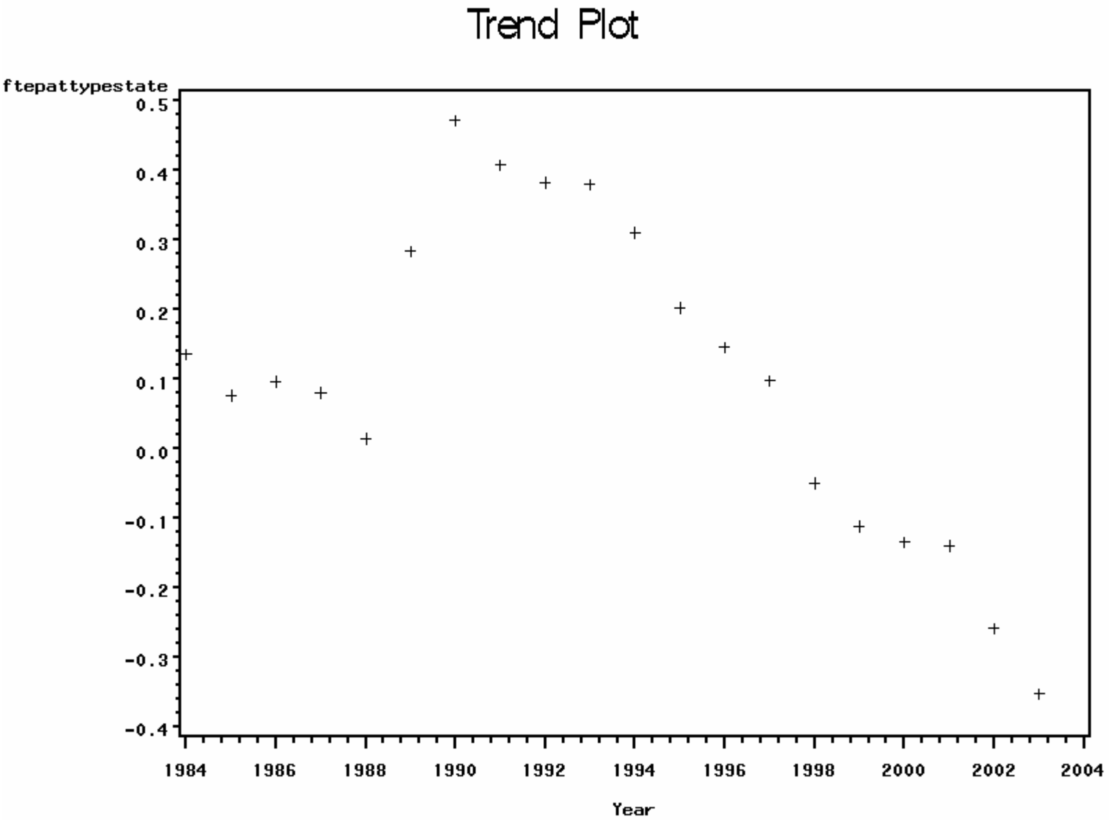

Figure 15-11 Time plot - Absorbed knowledge stock per HC personnel

From **Error! Reference source not found.** can be seen that the time plot shows an upward trend line. We therefore make use of the "Trend" specification in the stationarity test output results.

Table 15-20: SAS output for Phillips Perron test for variable "ftepattypestate"

The ARIMA Procedure

Name of Variable = ftepattypestate

Mean of Working Series	0.1021
Standard Deviation	0.224336
Number of Observations	20

Phillips-Perron Unit Root Tests

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau
Zero Mean	0	-0.7501	0.5096	-0.40	0.5242
	1	-1.3614	0.4062	-0.63	0.4303
Single Mean	0	0.4480	0.9666	0.20	0.9654
	1	-0.2294	0.9328	-0.11	0.9356
Trend	0	-1.9198	0.9619	-0.82	0.9450
	1	-2.3734	0.9476	-0.95	0.9287

From the Phillips Perron test output obtained from SAS we read the following values for the probability statistics.

Pr < Tau = 0.9450 for ℓ = 0 en

Pr < Tau = 0.9287 for ℓ = 1.

Both *p*-values are greater than 0.05. We therefore cannot reject H_0: d = 1. We have to conclude that ftepattypestate has a unit root and is non-stationary.

15.3.3 Colinearity tests

First however we should make sure that the variables are not collinear. The following is the test results obtained from SAS for the Colinearity test.

Table 15-21: Colinearity diagnostics for the model variables

Collinearity Diagnostics

			----Proportion of Variation--	
Number	Eigenvalue	Condition Index	Intercept	ftepattypestate
1	1.41424	1.00000	0.29288	0.29288
2	0.58576	1.55382	0.70712	0.70712

Larger values suggest potential near colinearity. Belsley, Kuh and Welsch (2000) recommend interpreting the Condition index greater or equal than 30 to reflect moderate to severe colinearity, worthy of further investigation. Since all the Condition indexes from the regression model is much smaller than 30, the conclusion can be made that colinearity is not a problem in this case.

Proportion of Variation

The variance proportion indicates for each predictor the proportion of total variance of its estimated regression coefficients associated with a particular principal component. The variance proportions suggest colinearity problems if more than one predictor has a high variance proportions of at least 0.5 for such a components suggest a problem. One should definitely be concerned when two or more ladings greater than 0.9 appear on a component with a large condition index (>30). This also does not seem to be a problem since the condition indexes are all small values.

15.3.4 Model estimation the rate of Patent Development in the PubS

As both variables entered in the model to be estimated are non-stationary, we should fit a model and then test for cointegration in the residual.

Table 15-22: SAS code for the model estimation procedure

```
proc reg data= Pub.patloglinear;
model Rdout = ftepattypestate
/ collin spec;
run;

/* r=0.60 p 0.00001  */
proc autoreg data= Pub.patloglinear;
model Rdout = ftepattypestate
/ dwprob method= ml archtest ;
output out=b r=residual;
run;
proc gplot data=b;
plot residual*year;
run;

proc arima data=b;
identify var=residual
stationarity=(phillips=(0,1));
run;
```

Table 15-23: SAS output for the model estimation of Absorptive capacity in the HES

```
                         The REG Procedure
                           Model: MODEL1
                     Dependent Variable: Rdout

              Number of Observations Read          20
              Number of Observations Used          20

                        Analysis of Variance

                                    Sum of          Mean
Source                    DF       Squares        Square    F Value    Pr > F

Model                      1       1.05369       1.05369      27.02    <.0001
Error                     18       0.70194       0.03900
Corrected Total           19       1.75563

             Root MSE              0.19748    R-Square     0.6002
             Dependent Mean        0.12720    Adj R-Sq     0.5780
             Coeff Var           155.24845

                        Parameter Estimates

                                Parameter       Standard
  Variable            DF         Estimate          Error    t Value    Pr > |t|
```

```
              Intercept             1         0.02274          0.04852         0.47        0.6450
              ftepattypestate       1         1.02316          0.19683         5.20        <.0001


                                         The REG Procedure
                                           Model: MODEL1
                                     Dependent Variable: Rdout

                                      Test of First and Second
                                        Moment Specification

                                     DF     Chi-Square     Pr > ChiSq

                                      2           2.11         0.3488

                                      The AUTOREG Procedure

                                    Dependent Variable    Rdout

                                  Ordinary Least Squares Estimates

                    SSE                     0.70194207    DFE                        18
                    MSE                        0.03900    Root MSE              0.19748
                    SBC                     -4.2437275    AIC                -6.2351921
                    Regress R-Square            0.6002    Total R-Square         0.6002
                    Durbin-Watson               1.5178    Pr < DW                0.0875
                    Pr > DW                     0.9125
  NOTE: Pr<DW is the p-value for testing positive autocorrelation, and Pr>DW is the p-
  value for testing negative autocorrelation.

                                Q and LM Tests for ARCH Disturbances

                         Order           Q     Pr > Q           LM     Pr > LM

                             1      0.0215     0.8834       0.0320      0.8580
                             2      0.1960     0.9067       0.0802      0.9607
                             3      0.9652     0.8097       0.5220      0.9140
                             4      2.5316     0.6390       2.3069      0.6795
                             5      2.5335     0.7714       2.4582      0.7828
                             6      4.7339     0.5784       5.5867      0.4710
                             7      5.0327     0.6560       5.7859      0.5650
                             8      5.3684     0.7176       7.9828      0.4352
                             9      6.8896     0.6486      10.8254      0.2879
                            10      8.3899     0.5908      11.1950      0.3425
                            11      8.4369     0.6737      12.4735      0.3291
                            12      9.1364     0.6912      12.4735      0.4084

                                                         Standard                  Approx
              Variable                DF     Estimate       Error     t Value     Pr > |t|

              Intercept                1       0.0227      0.0485        0.47       0.6450
              ftepattypestate          1       1.0232      0.1968        5.20       <.0001
```

From the model estimation output obtained we can make the following conclusion:

The **R-Square 0.6002** statistic indicate that the model accounts for 60% of the variation of the papers produced in the Public sector.

The test for autocorrelation use is the Durban Watson test statistic. The Durbin Watson test statistic is 1.9593 with (Pr < DW = 0.0875 > 0.05 and (Pr < DW = 0.9125) < 0.95. This indicates that we therefore can conclude that the autoregressive model does not have autocorrelation.

The Chi-square test for the first moment specification indicates that the model does not have heteroscedastic errors. The SPEC option performs a model specification test. The null hypothesis for this test maintains that the errors are homoscedastic, independent of the regressor and that several technical assumptions about the model specification are valid. With Pr = 0.3488 we fail to reject the null hypothesis. We can therefore conclude that no heteroscedasiticity is present in the model.

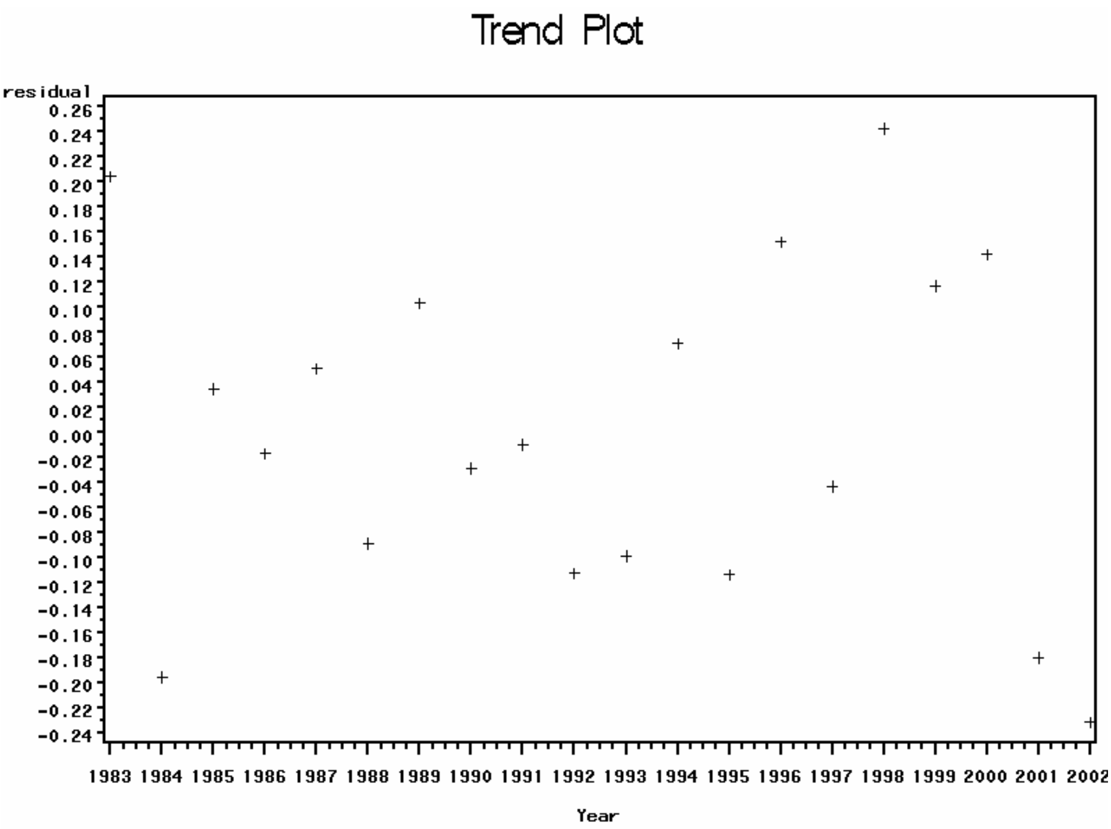

Figure 15-12 Trend plot for the residual

From Figure 15-12 can be seen that the time plot seems to be scattered around 0. From the results we van also read the `Mean of Working Series is -382E-19` in Table 15-24. We therefore make use of the "Zero mean" specification in the stationarity test output results.

Table 15-24: SAS output for residual stationarity test

```
                         The ARIMA Procedure

                    Name of Variable = residual

                  Mean of Working Series     -382E-19
                  Standard Deviation         0.187342
                  Number of Observations           20

                   Phillips-Perron Unit Root Tests

 Type           Lags    Rho         Pr < Rho     Tau      Pr < Tau

Zero Mean        0    -15.4024      0.0019      -3.37     0.0019
                 1    -15.7512      0.0016      -3.39     0.0019
Single Mean      0    -15.3087      0.0112      -3.25     0.0325
                 1    -15.6218      0.0098      -3.27     0.0313
Trend            0    -15.3592      0.0695      -3.17     0.1191
                 1    -15.7363      0.0611      -3.20     0.1144
```

For (n-1) = 1, the values are obtained from the Critical values for the Phillips Z Statistic or the Dickey Fuller t Statistic when applied to Residuals from Spurious Cointegration Regression (See table Table 15-24). The critical value for the 5% level is -3.37.

From the Phillips Perron test output obtained from SAS we read the following values for the probability statistics.

`Tau = -3.37` for $\ell = 0$ en

`Tau = -3.39` for $\ell = 1$.

This means that we can therefore reject the null hypothesis of unit root with a 5% significance level, since the `Tau` values are smaller than the critical value. The residues can be deemed stationary and the variables are cointegrated. We can therefore conclude that the regression is not spurious.

16 APPENDIX E

16.1 Absorption of Knowledge (Private sector)

The rate at which the system is able to produce new knowledge output is computed through the contribution made form different stocks in the system. The following expression is formulated for the R&D knowledge absorption rate in the system:

- $R_{Absorptionr}$: Absorption rate of knowledge in the system
- $S_{R\&Doutput} * S_{FTE}$: RD output stock interacting with the presence of people full time equivalent people who can draw on the stocks of knowledge person in system
- S_{World} / S_{HC} : Available external knowledge stock per Headcount personnel employed in the system

A multiplicative model is developed for the absorption rate per full time person working in the system:

$$\frac{R_{Absorptionr}}{R^*_{Absorption}} = f * \left(\frac{S_{R\&Doutput}}{S^*_{R\&Doutpu}} * \frac{S_{FTE}}{S^*_{FTE}}\right)^d * \left(\frac{S_{World}}{S^*_{World}} / \frac{S_{HC}}{S^*_{HC}}\right)^e \qquad \textbf{16-1}$$

This expression is linearised by taking the log-linear form:

$$\ln\left(\frac{R_{Absorptionr}}{R^*_{Absorption}}\right) = f + d * \ln\left(\frac{S_{R\&Doutput}}{S^*_{R\&Doutpu}} * \frac{S_{FTE}}{S^*_{FTE}}\right) + e * \ln\left(\frac{S_{World}}{S^*_{World}} / \frac{S_{HC}}{S^*_{HC}}\right)$$

16-2

This is the expression used to perform the regression for estimating the parameters d, e and f.

The section describes the variables included in the model to estimate the rate of knowledge absorption in the system. The following SAS program was used.

Table 16-1: SAS code for stationarity tests in variables AbsorbedR, RDfte and wsperhc

```
goptions reset=all cback=white colors=(black) lfactor=2
border;
title1 'Trend Plot';
proc gplotb data = priv.loglin;
plot (arperfte rdfte wsperhc)*year;
plot arperfte *(rdfte wsperhc);
run;

* test for stationarity of the 3 series using arima procedure
*;
proc arima data=priv.loglin;
identify var= arperfte stationarity=(phillips=(0,1));
identify var=rdfte stationarity=(phillips=(0,1));
identify var=wsperhc stationarity=(phillips=(0,1));
```

```
run;
```

The following sections document and explain the output obtained from the SAS program.

16.1.1 Absorption rate of knowledge in the system

The following is the time plot output from the SAS program for the absorption rate per full time equivalent researchers in the system.

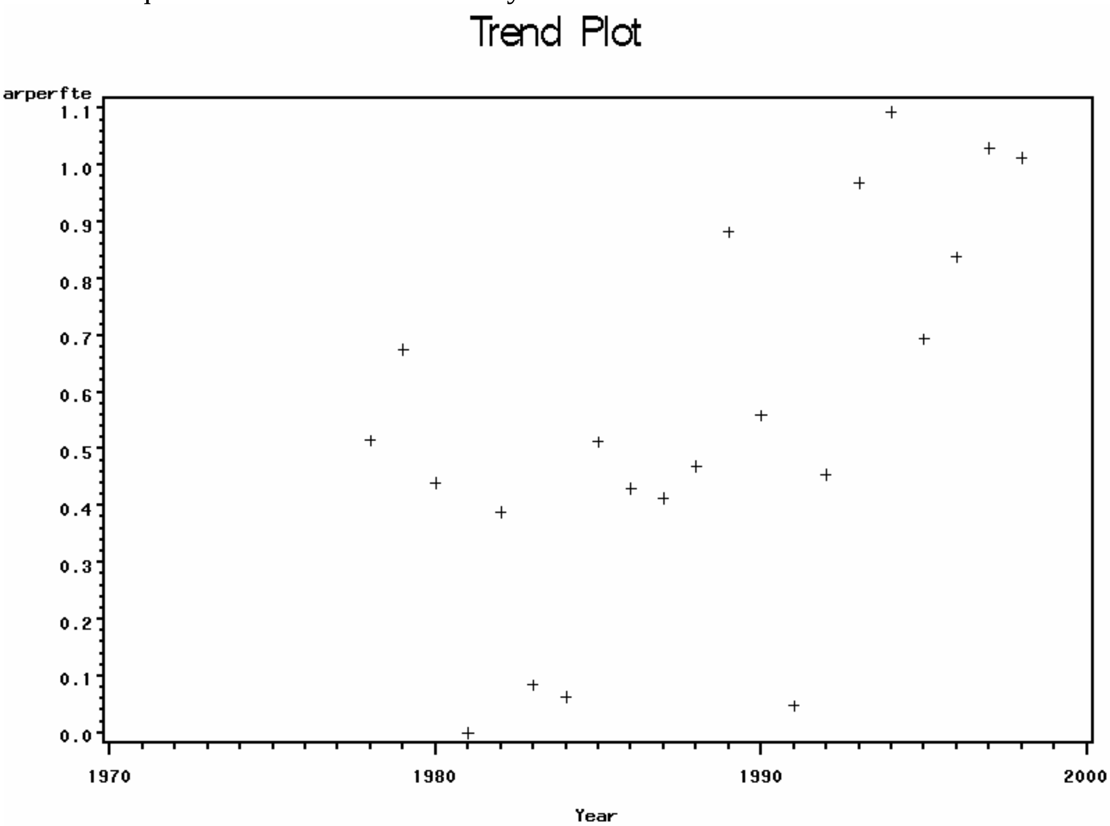

Figure 16-1 Time plot of the absorption rate in the Private sector

From **Error! Reference source not found.** can be seen that the time plot shows an upward trend. We therefore make use of the “Trend” specification in the stationarity test output results.

Table 16-2: Phillips-Perron test output for variable “`arperfte`”

```
                         The ARIMA Procedure

                      Name of Variable = arperfte

                    Mean of Working Series    0.503147
                    Standard Deviation        0.329908
                    Number of Observations          21

                         The ARIMA Procedure

                    Phillips-Perron Unit Root Tests

   Type            Lags         Rho    Pr < Rho        Tau    Pr < Tau

  Zero Mean          0      -1.7376      0.3537      -0.66     0.4172
```

	1	-1.7138	0.3569	-0.65	0.4202
Single Mean	0	-10.6238	0.0704	-2.53	0.1233
	1	-11.7998	0.0468	-2.64	0.1021
Trend	0	-15.7939	0.0658	-3.32	0.0921
	1	-17.2017	0.0401	-3.40	0.0800

From the Phillips Perron test output obtained from SAS we read the following values for the probability statistics.

`Pr < Tau = 0.0921` for $\ell = 0$ en

`Pr < Tau = 0.0800` for $\ell = 1$.

Both *p*-values are greater than 0.05. We therefore cannot reject H_0: $d = 1$. We have to conclude that `arperfte` has a unit root and is non-stationary.

16.1.2 R&D Knowledge Stock and FTE researchers interaction

The following is the time plot output from the SAS program for the RD Knowledge stock with Full time equivalent personnel in the system.

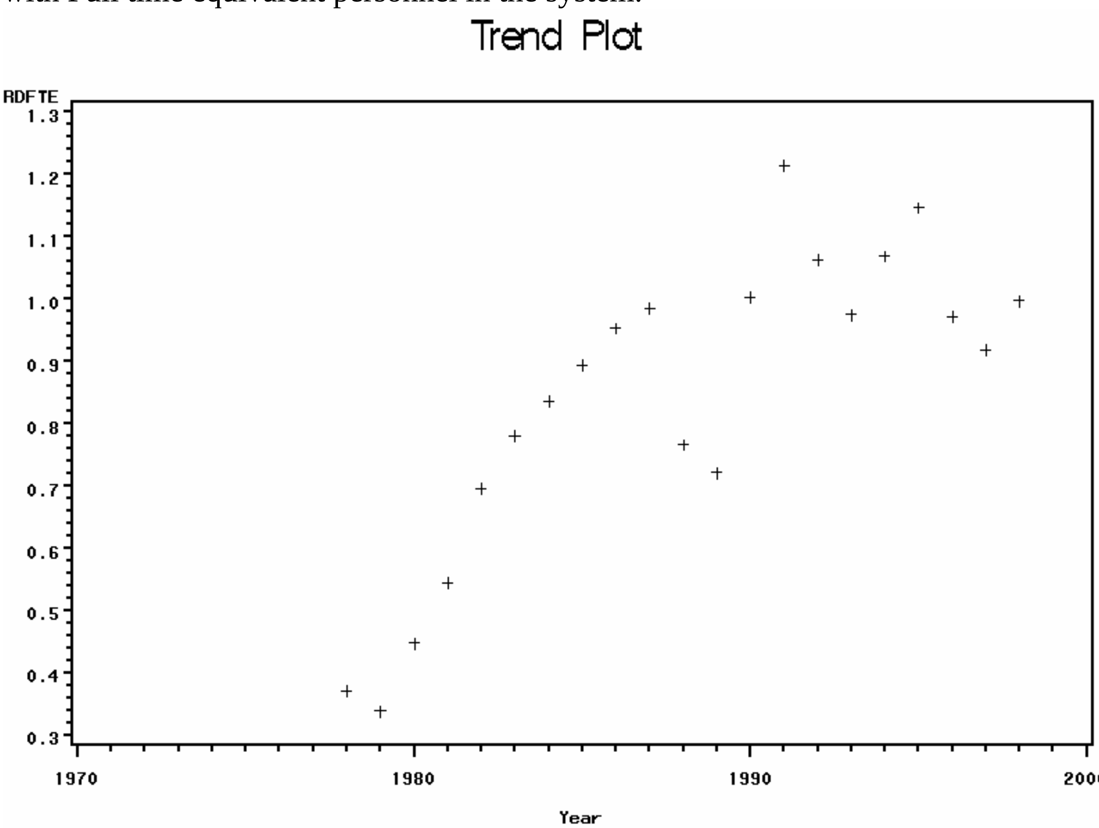

Figure 16-2 Time plot - FTE researcher interacting with R&D knowledge

From Figure 16-2 can be seen that the time plot shows an upward trend. We therefore make use of the "Trend" specification in the stationarity test output results.

Table 16-3: Phillips-Perron test output for variable "`RDFTE`"

```
                    The ARIMA Procedure

              Name of Variable = RDFTE

          Mean of Working Series    0.604104
          Standard Deviation         0.24815
          Number of Observations          21

                 Phillips-Perron Unit Root Tests

Type             Lags        Rho    Pr < Rho        Tau    Pr < Tau

Zero Mean           0     0.1144      0.6971       0.12      0.7093
                    1     0.0630      0.6844       0.06      0.6908
Single Mean         0    -5.6262      0.3308      -2.73      0.0860
                    1    -5.8495      0.3108      -2.69      0.0923
Trend               0    -8.7810      0.4304      -2.23      0.4476
                    1    -9.9789      0.3331      -2.36      0.3874
```

From the Phillips Perron test output obtained from SAS we read the following values for the probability statistics.

`Pr < Tau = 0.4476` for $\ell = 0$ en

`Pr < Tau = 0.3874` for $\ell = 1$.

Both *p*-values are greater than 0.05. We therefore cannot reject H_0: $d = 1$. We have to conclude that `RDFTE` has a unit root and is non-stationary.

16.1.3 The external knowledge stock

The following is the time plot output from the SAS program for the World Knowledge Stock per Headcount person employed in the Higher Education system.

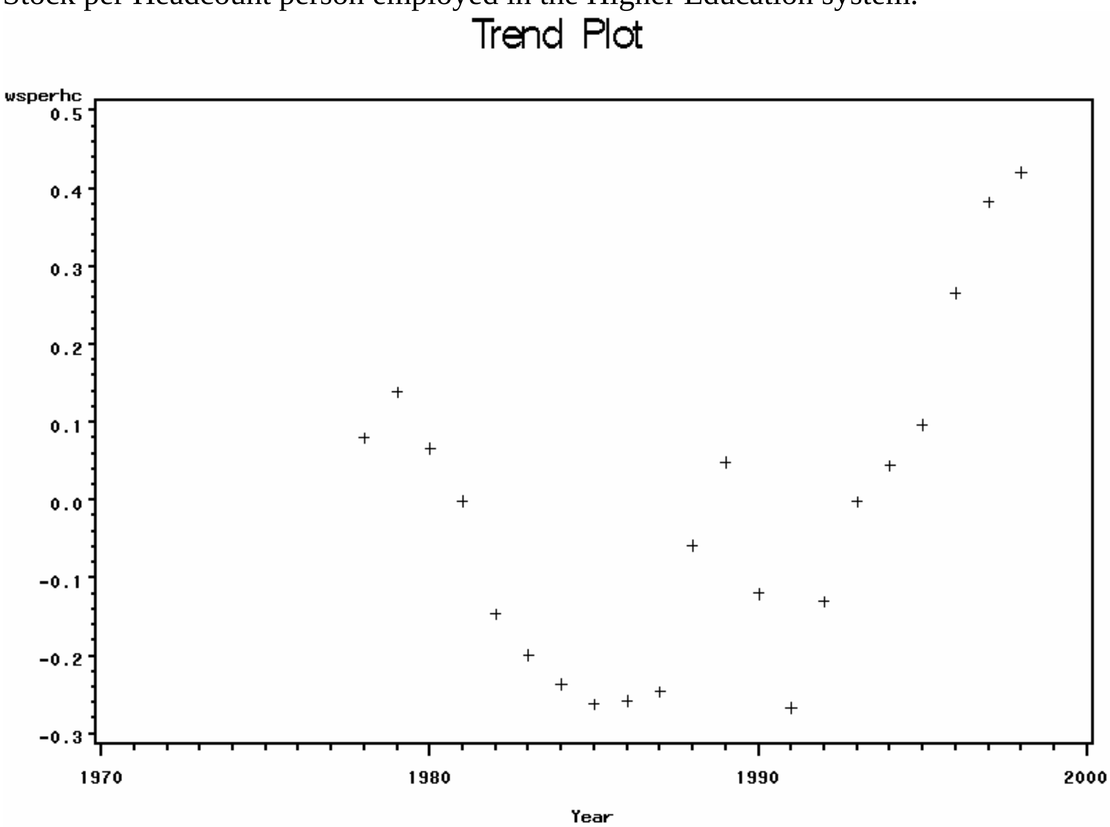

Figure 16-3 Time plot - World stock of knowledge per HC researcher

From **Error! Reference source not found.** can be seen that the time plot shows an upward trend. We therefore make use of the “Trend” specification in the stationarity test output results.

Table 16-4: Phillips-Perron test output for variable “WSperHC”

The ARIMA Procedure

Name of Variable = wsperhc

Mean of Working Series	-0.03798
Standard Deviation	0.175669
Number of Observations	21

Phillips-Perron Unit Root Tests

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau
Zero Mean	0	-1.4766	0.3903	-0.50	0.4856
	1	-3.2016	0.2061	-0.92	0.3043
Single Mean	0	-0.7238	0.8970	-0.22	0.9204
	1	-2.6069	0.6844	-0.71	0.8227
Trend	0	-1.2211	0.9791	-0.40	0.9797
	1	-2.3452	0.9505	-0.71	0.9583

From the Phillips Perron test output obtained from SAS we read the following values for the probability statistics.

`Pr < Tau = 0.9797` for $\ell = 0$ en

`Pr < Tau = 0.9583` for $\ell = 1$.

Both *p*-values are greater than 0.05. We therefore cannot reject H_0: $d = 1$. We have to conclude that `wsperhc` has a unit root and is non-stationary.

16.1.4 Colinearity tests

First however we should make sure that the variables are not collinear. The following is the test results obtained from SAS for the Colinearity test.

Table 16-5: Colinearity diagnostics for the model variables

Collinearity Diagnostics

Number	Eigenvalue	Condition Index
1	2.04882	1.00000
2	0.88054	1.52538
3	0.07064	5.38552

Number	---------Proportion of Variation--------- Intercept	RDFTE	wsperhc
1	0.02987	0.02973	0.04556
2	0.01250	0.00486	0.89138
3	0.95763	0.96540	0.06306

Larger values suggest potential near colinearity. Belsley, Kuh and Welsch (2000) recommend interpreting the Condition index greater or equal than 30 to reflect moderate to severe colinearity, worthy of further investigation. Since all the Condition indexes from the regression model are much smaller than 30, the conclusion can be made that colinearity is not a problem in this case.

Proportion of Variation

The variance proportion indicates for each predictor the proportion of total variance of its estimated regression coefficients associated with a particular principal component. The variance proportions suggest colinearity problems if more than one predictor has a high variance proportions of at least 0.5 for such a components suggest a problem. One should definitely be concerned when two or more ladings greater than 0.9 appear on a component with a large condition index (>30). This also does not seem to be a problem since the condition indexes are all smaller than 30.

16.1.5 Model estimation - Absorption rate

As all three variables are non-stationary, we should now fit a model and then test for cointegration in the residual.

Table 16-6: SAS code for the model estimation procedure

```
proc autoreg data=Priv.loglinear ;
model arperfte = RDFTe wsperhc
/ method=ml dwprob stationarity=(phillips=(1));
output out=abspriv r=residual;
run;
```

Table 16-7: SAS output for the model estimation of Absorptive capacity in the HES

```
                              The AUTOREG Procedure

                         Dependent Variable    arperfte


                        Ordinary Least Squares Estimates

               SSE                0.96423696    DFE                       18
               MSE                   0.05357    Root MSE             0.23145
               SBC                4.02923226    AIC               0.89566494
               Regress R-Square       0.5781    Total R-Square        0.5781
               Durbin-Watson          1.7635    Pr < DW               0.1533
               Pr > DW                0.8467
NOTE: Pr<DW is the p-value for testing positive autocorrelation, and Pr>DW is the p-value
for testing negative autocorrelation.

                      Q and LM Tests for ARCH Disturbances

                Order            Q    Pr > Q           LM    Pr > LM

                  1         0.2084    0.6481       0.1093     0.7409
                  2         1.3994    0.4967       1.4322     0.4887
                  3         3.5987    0.3082       2.3405     0.5048
                  4         7.0608    0.1327      10.3345     0.0352
                  5         7.7103    0.1729      10.4443     0.0636
                  6         8.4556    0.2066      11.4023     0.0767
                  7         9.8622    0.1965      11.4473     0.1203
                  8        10.0637    0.2606      13.2544     0.1034
                  9        10.2928    0.3273      14.3887     0.1092
                 10        10.4760    0.3998      17.8694     0.0572
                 11        11.6565    0.3900      18.3951     0.0729
                 12        12.1879    0.4307      19.1378     0.0853
```

```
                         Phillips-Ouliaris
                        Cointegration Test

                 Lags           Rho           Tau

                    1       -18.6489       -4.0045


                                            Standard                    Approx
  Variable          DF       Estimate          Error    t Value    Pr > |t|

  Intercept          1         0.1729         0.1351       1.28      0.2168
  RDFTE              1         0.6339         0.2117       2.99      0.0078
  wsperhc            1         1.3863         0.2991       4.63      0.0002

                       The AUTOREG Procedure

                  Dependent Variable    arperfte


                 Ordinary Least Squares Estimates

    SSE                    0.96423696    DFE                        18
    MSE                       0.05357    Root MSE              0.23145
    SBC                    4.02923226    AIC                0.89566494
    Regress R-Square           0.5781    Total R-Square         0.5781
    Durbin-Watson              1.7635    Pr < DW                0.1533
    Pr > DW                    0.8467
NOTE: Pr<DW is the p-value for testing positive autocorrelation, and Pr>DW is the p-value
for testing negative autocorrelation.

                         Phillips-Ouliaris
                        Cointegration Test

                 Lags           Rho           Tau

                    1       -18.6489       -4.0045


                                            Standard                    Approx
  Variable          DF       Estimate          Error    t Value    Pr > |t|

  Intercept          1         0.1729         0.1351       1.28      0.2168
  RDFTE              1         0.6339         0.2117       2.99      0.0078
  wsperhc            1         1.3863         0.2991       4.63      0.0002
```

From the model estimation output obtained we can make the following conclusion:

The test for autocorrelation use is the Durban Watson test statistic. The `Durbin Watson` test statistic is `1.7635 with (Pr < DW = 0.1533) > 0.05 and (Pr < DW = 0.8467) < 0.95`. This indicates that we can we therefore can conclude that the autoregressive model does not have autocorrelation.

The heteroscedasticity test (Q and LM test for ARCH disturbances) is only interpreted up to 2 time lags. The probability for arch disturbances in the model for lags 1 and 2 are larger than 0.05. We can therefore conclude that the modelled relationship does not suffer from heteroscedasticity.

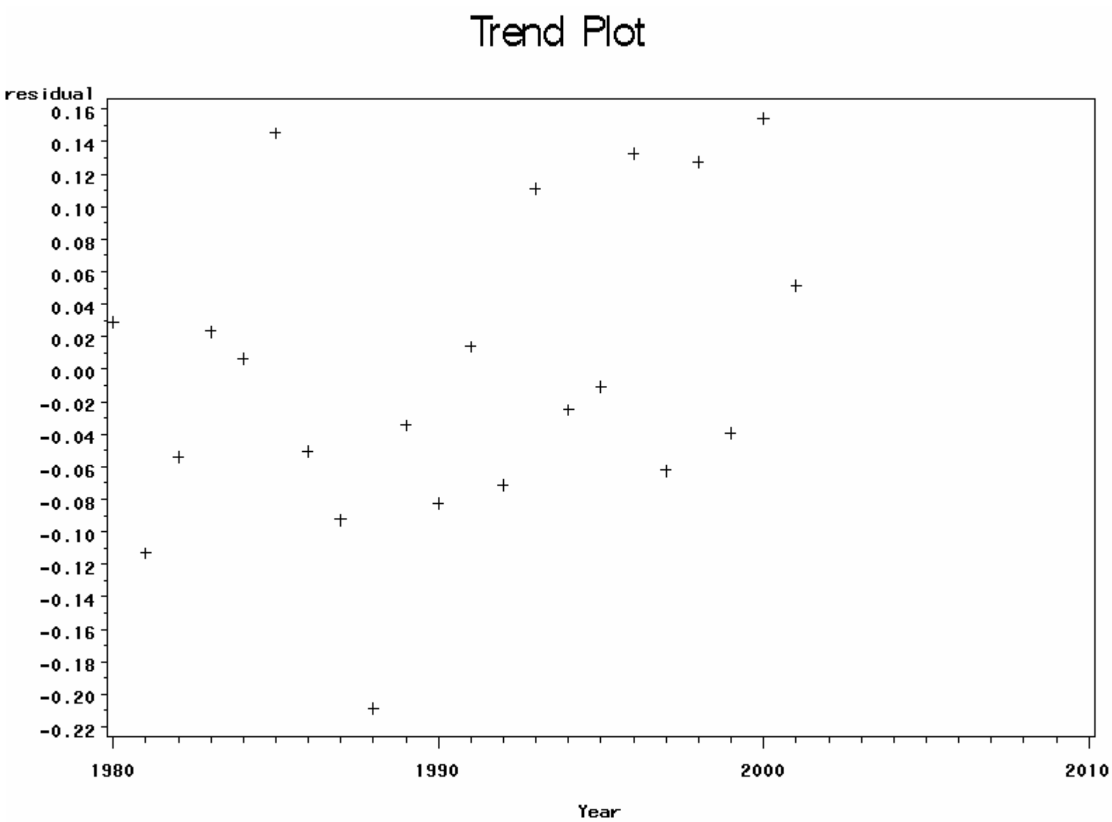

Figure 16-4 Time plot for the residual of the knowledge absorption model

From Figure 16-5 can be seen that the time plot seems to be scattered around 0. From the results we van also read the `Mean of Working Series -0.00219` in Table 16-8: . We therefore make use of the "Zero mean" specification in the stationarity test output results.

Table 16-8: Stationarity test for the residual

The ARIMA Procedure

Name of Variable = residual

Mean of Working Series	-889E-19
Standard Deviation	0.21428
Number of Observations	21

Phillips-Perron Unit Root Tests

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau
Zero Mean	0	-18.0392	0.0006	-3.99	0.0004
	1	-18.6489	0.0004	-4.00	0.0004
Single Mean	0	-18.0573	0.0037	-3.89	0.0084
	1	-18.6569	0.0028	-3.91	0.0081
Trend	0	-18.0173	0.0296	-3.75	0.0424
	1	-18.6237	0.0233	-3.77	0.0407

Since an intercept is included in the model fitted, an intercept is included. For (n-1) = 2, the values are obtained from the Critical values for the Phillips Z Statistic or the Dickey Fuller t Statistic when applied to Residuals from Spurious Cointegration Regression (See Table 16-8:). The critical value for the 5% level is -3.77.

From the Phillips Perron test output obtained from SAS we read the following values for the probability statistics.
`Tau = -3.99` for $\ell = 0$ en
`Tau = -4.00` for $\ell = 1$.

This means that we can therefore reject the null hypothesis of unit root since the `Tau` values are smaller than the critical value. The residues can be deemed stationary and the variables are cointegrated. We can therefore conclude that the regression is not spurious.

16.2 Creation of new knowledge (Private sector)

The rate at which the system is able to produce new knowledge output is computed through the contribution made form different stocks in the system. The following expression is formulated for the R&D output productivity per FTE researcher working in the system:

- R_{Paper} / S_{FTE} : R&D output rate per FTE researcher person on the system
- $S_{Experience} / S_{HC}$: Average Experience Stock of the people in the system.
- $S_{Absorbed} / S_{HC}$: Average Absorbed knowledge per person in the system.

A multiplicative model is developed for the development rate of papers per full time person working in the system:

$$\frac{R_{Paper}}{R^*_{Paper}} / \frac{S_{FTE}}{S^*_{FTE}} = c * \left(\frac{S_{Experience}}{S^*_{Experience}} / \frac{S_{HC}}{S^*_{HC}}\right)^a * \left(\frac{S_{Absorbed}}{S^*_{Absorbed}} / \frac{S_{HC}}{S^*_{HC}}\right)^b \quad \textbf{16-3}$$

This expression is linearised by taking the log-linear form:

$$\ln\left(\frac{R_{Paper}}{R^*_{Paper}} / \frac{S_{FTE}}{S^*_{FTE}}\right) = \ln(c) + a * \ln\left(\frac{S_{Experience}}{S^*_{Experience}} / \frac{S_{HC}}{S^*_{HC}}\right) + b * \ln\left(\frac{S_{Absorbed}}{S^*_{Absorbed}} / \frac{S_{HC}}{S^*_{HC}}\right) \quad \textbf{16-4}$$

This is then the expression used to perform the regression for estimating the parameters a, b and c.

The section describes the variables included in the model to estimate the rate of knowledge creation in the system. The following SAS program was used.

Table 16-9: SAS code for the stationarity tests procedure for “prperfte”, “ftetot”, “AbsS”

```
goptions reset=all cback=white colors=(black) lfactor=2
border;
title1 'Trend Plot';
proc gplot priv.loglin;
plot (prperfte ftetot AbsS)*year;
plot prperfte*(ftetot AbsS);
run;

* test for stationarity of the 3 series using arima procedure
*;
proc arima priv.loglin;
identify var=prperfte stationarity=(phillips=(0,1));
identify var=ftetot stationarity=(phillips=(0,1));
identify var=AbsS stationarity=(phillips=(0,1));
run;
```

The following sections document and explain the output obtained from the SAS program.

16.2.1 R&D output produced per FTE researcher

The following is the time plot output from the SAS program for the R&D output (papers) created per full time equivalent researcher in the system.

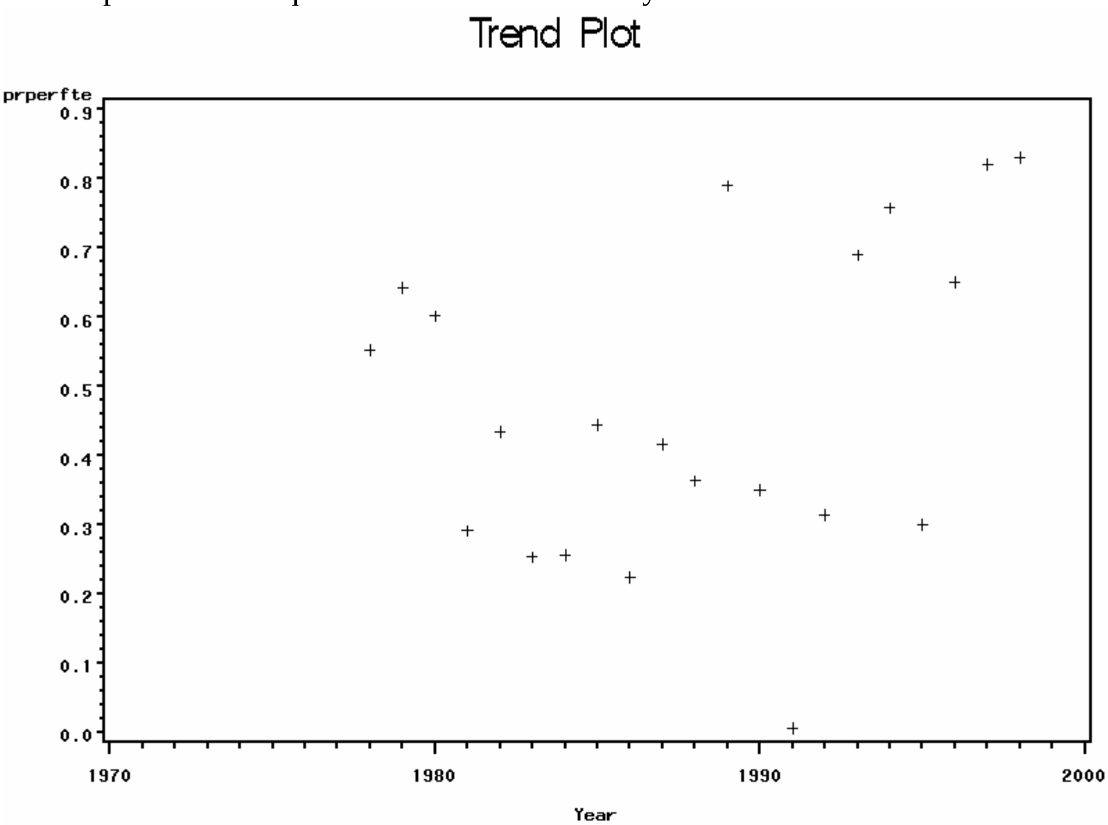

Figure 16-5 Time plot of the Knowledge creation rate per FTE

From Figure 16-5 can be seen that the time plot although scattered seems to follow an upward trend. We therefore make use of the "Trend" specification in the stationarity test output results.

Table 16-10: SAS output for Phillips Perron test for variable "prperfte"

```
                     The ARIMA Procedure

                   Name of Variable = prperfte

                Mean of Working Series    0.475973
                Standard Deviation         0.22445
                Number of Observations          21

                    Phillips-Perron Unit Root Tests

Type            Lags          Rho     Pr < Rho         Tau     Pr < Tau

Zero Mean          0      -1.9449       0.3271       -0.83       0.3434
                   1      -1.5327       0.3821       -0.71       0.3965
Single Mean        0     -13.5207       0.0249       -2.82       0.0725
                   1     -14.1240       0.0197       -2.87       0.0660
Trend              0     -14.7705       0.0918       -3.10       0.1331
                   1     -14.7723       0.0918       -3.10       0.1331
```

From the Phillips Perron test output obtained from SAS we read the following values for the probability statistics.

`Pr < Tau = 0.1331` for $\ell = 0$ en

`Pr < Tau = 0.1331` for $\ell = 1$.

Both *p*-values are greater than 0.05. We therefore cannot reject H_0: $d = 1$. We have to conclude that `rdperfte` has a unit root and is non-stationary.

16.2.2 Absorbed Stock

The following is the time plot output from the SAS program for the Absorbed knowledge stock in the system.

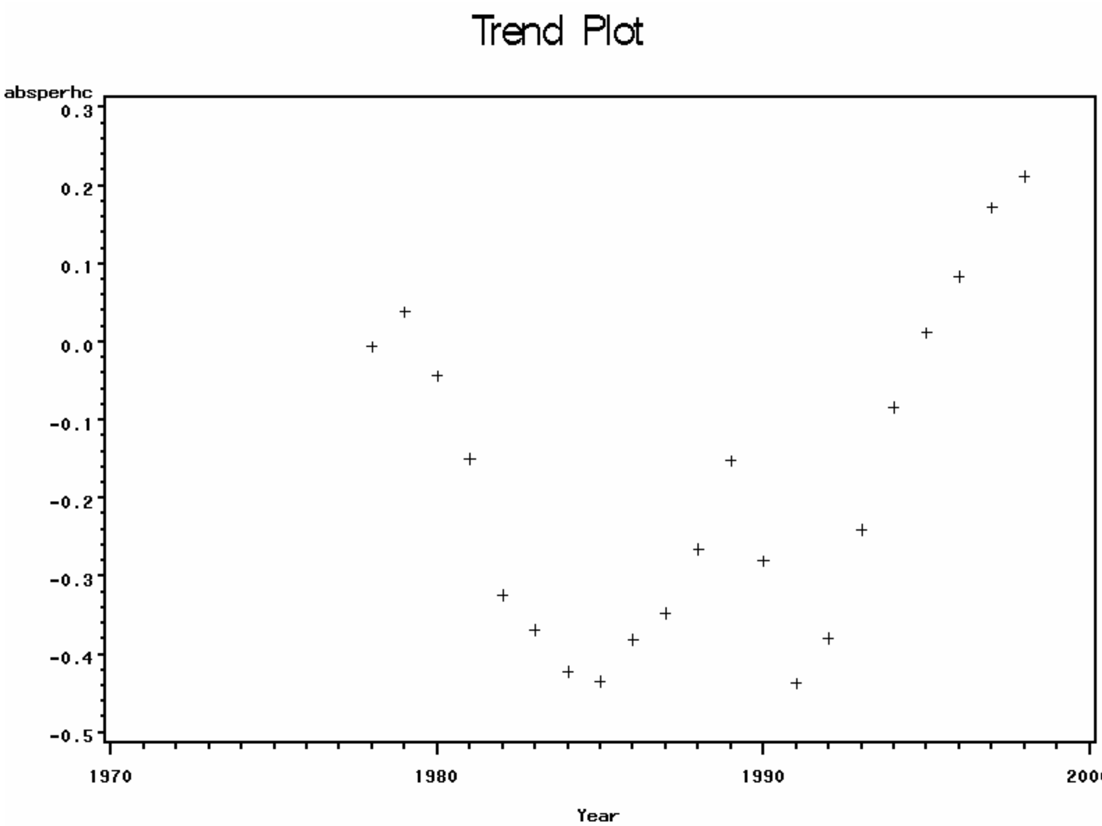

Figure 16-6 Time plot of the Absorbed knowledge stock per HC personnel

From **Error! Reference source not found.** can be seen that the time plot shows a trend line. We therefore make use of the "Trend" specification in the stationarity test output results.

Table 16-11: SAS output for Phillips Perron test for variable "Absperhc"

Name of Variable = absperhc

Mean of Working Series	-0.18065
Standard Deviation	0.203145
Number of Observations	21

Phillips-Perron Unit Root Tests

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau
Zero Mean	0	-0.9679	0.4709	-0.60	0.4439
	1	-1.6692	0.3630	-0.83	0.3422
Single Mean	0	-0.8287	0.8880	-0.34	0.9020
	1	-2.3245	0.7208	-0.79	0.7998
Trend	0	-1.9274	0.9634	-0.91	0.9346
	1	-2.5355	0.9435	-1.07	0.9102

From the Phillips Perron test output obtained from SAS we read the following values for

the probability statistics.
`Pr < Tau = 0.9346` for ℓ = 0 en
`Pr < Tau = 0.9102` for ℓ = 1.

Both *p*-values are greater than 0.05. We therefore cannot reject H_0: d = 1. We have to conclude that `absperhc` has a unit root and is non-stationary.

16.2.3 Experience per Headcount in the system

The following is the time plot output from the SAS program for the Full time personnel in the system.

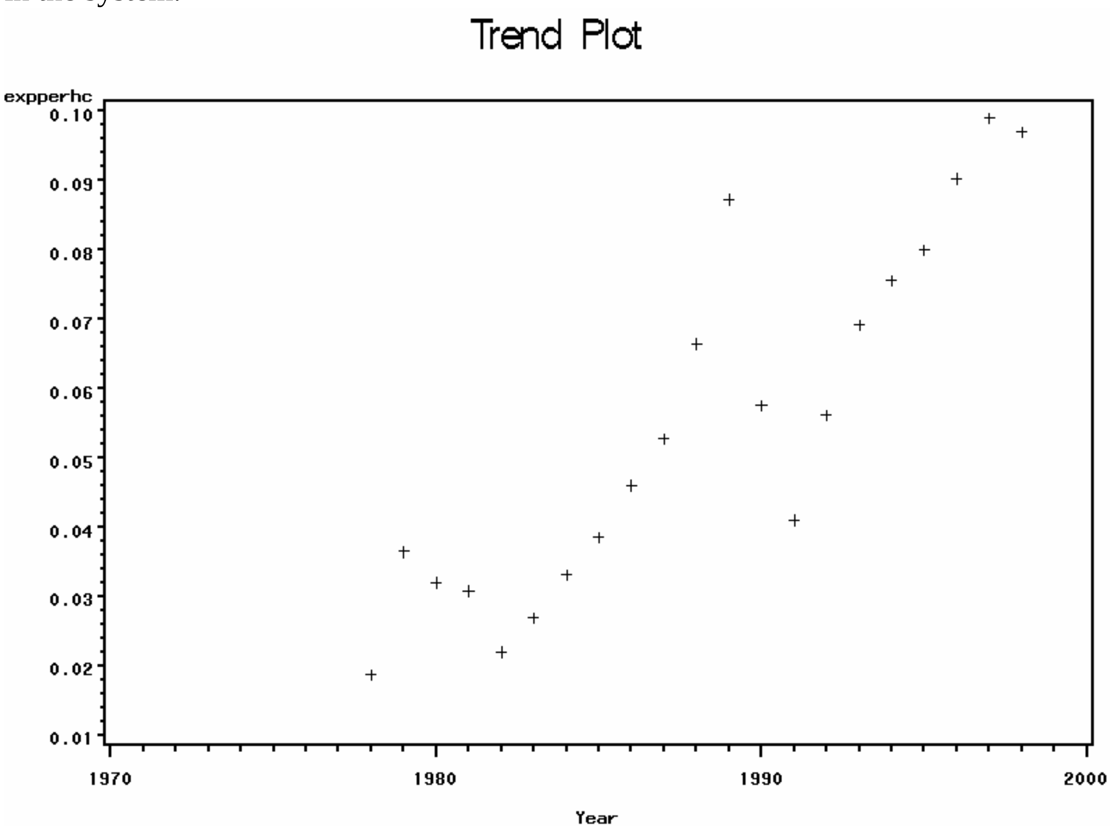

Figure 16-7 Time plot for the "Expperhc" variable in the system

From Figure 16-7 can be seen that the time plot shows an downward trend line. We therefore make use of the "Trend" specification in the stationarity test output results.

Table 16-12: SAS output for Phillips Perron test for variable "Expperhc"

```
                    Name of Variable = expperhc

                  Mean of Working Series      0.0551
                  Standard Deviation        0.024941
                  Number of Observations          21

                     Phillips-Perron Unit Root Tests

Type             Lags         Rho     Pr < Rho         Tau     Pr < Tau

Zero Mean           0      0.8955       0.8770        0.94       0.9011
                    1      0.8630       0.8710        0.87       0.8902
Single Mean         0     -2.0219       0.7589       -0.89       0.7699
                    1     -2.3803       0.7137       -0.98       0.7387
Trend               0    -10.7197       0.2812       -2.48       0.3342
                    1    -12.4559       0.1811       -2.64       0.2693
```

From the Phillips Perron test output obtained from SAS we read the following values for the probability statistics.
`Pr < Tau = 0.3342` for $\ell = 0$ en
`Pr < Tau = 0.2693` for $\ell = 1$.

Both *p*-values are greater than 0.05. We therefore cannot reject H_0: $d = 1$. We have to conclude that `Expperhc` has a unit root and therefore is non-stationary.

16.2.4 Colinearity tests

First however we should make sure that the variables are not collinear. The following is the test results obtained from SAS for the Colinearity test.

Figure 16-8 Colinearity diagnostics for the model variables

```
           Collinearity Diagnostics

                                Condition
Number      Eigenvalue          Index
   1         2.36140            1.00000
   2         0.59053            1.99969
   3         0.04807            7.00918

            ---------Proportion of Variation---------
 Number       Intercept        absperhc         Expperhc
    1         0.01425          0.04170          0.01752
    2         0.00357          0.46016          0.05796
    3         0.98218          0.49814          0.92452
```

Larger values suggest potential near colinearity. Belsley, Kuh and Welsch (2000) recommend interpreting the Condition index greater or equal than 30 to reflect moderate to severe colinearity, worthy of further investigation. Since all the Condition indexes from the regression model is much smaller than 30, the conclusion can be made that colinearity is not a problem in this case.

Proportion of Variation

The variance proportion indicates for each predictor the proportion of total variance of its estimated regression coefficients associated with a particular principal component. The variance proportions suggest colinearity problems if more than one predictor has a high variance proportions of at least 0.5 for such a components suggest a problem. One should definitely be concerned when two or more ladings greater than 0.9 appear on a component with a large condition index (>30). This also does not seem to be a problem since the condition indexes are all small values.

16.2.5 Model estimation the rate of Paper Development in the HES

As all three variables are non-stationary, we should now fit a model and then test for cointegration in the residual

Table 16-13: SAS code for the model estimation procedure

```
proc autoreg data=Priv.loglinear ;
model prperfte = absperhc expperhc
/ method=ml dwprob  stationarity=(phillips=(1));
output out=abspriv r=residual;
run;

proc reg data = Priv.loglinear ;
    model prperfte = absperhc expperhc
    /tol vif collin spec dw;
    output out=abspriv r=residual;
run;

* consider residual *;
proc gplot data= abspriv;
plot residual*year;
run;

* test for cointegration using arima procedure *;
proc arima data=abspriv;
identify var=residual
stationarity=(phillips=(0,1));
run;
```

Table 16-14: SAS output for the model estimation of patent output in the Private sector

```
                         The AUTOREG Procedure

                     Dependent Variable    prperfte

                    Ordinary Least Squares Estimates

 SSE                     0.44599109    DFE                       18
 MSE                        0.02478    Root MSE             0.15741
 SBC                     -12.162568    AIC               -15.296135
 Regress R-Square            0.5784    Total R-Square        0.5784
 Durbin-Watson               2.3068    Pr < DW               0.5945
 Pr > DW                     0.4055
NOTE: Pr<DW is the p-value for testing positive autocorrelation, and Pr>DW is the p-
value for testing negative autocorrelation.
```

Q and LM Tests for ARCH Disturbances

Order	Q	Pr > Q	LM	Pr > LM
1	0.2084	0.6481	0.1093	0.7409
2	1.3994	0.4967	1.4322	0.4887
3	3.5987	0.3082	2.3405	0.5048
4	7.0608	0.1327	10.3345	0.0352
5	7.7103	0.1729	10.4443	0.0636
6	8.4556	0.2066	11.4023	0.0767
7	9.8622	0.1965	11.4473	0.1203
8	10.0637	0.2606	13.2544	0.1034
9	10.2928	0.3273	14.3887	0.1092
10	10.4760	0.3998	17.8694	0.0572
11	11.6565	0.3900	18.3951	0.0729
12	12.1879	0.4307	19.1378	0.0853

Phillips-Ouliaris
Cointegration Test

Lags	Rho	Tau
1	-22.3665	-5.1370

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|
Intercept	1	0.4779	0.1193	4.01	0.0008
absperhc	1	0.6672	0.2013	3.31	0.0039
expperhc	1	2.1522	1.6392	1.31	0.2057

Dependent Variable: prperfte

Number of Observations Read	21
Number of Observations Used	21

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	0.61195	0.30597	12.35	0.0004
Error	18	0.44599	0.02478		
Corrected Total	20	1.05794			

Root MSE	0.15741	R-Square	0.5784
Dependent Mean	0.47597	Adj R-Sq	0.5316
Coeff Var	33.07080		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Tolerance	Variance Inflation
Intercept	1	0.47791	0.11925	4.01	0.0008	.	0
absperhc	1	0.66715	0.20126	3.31	0.0039	0.70586	1.41671
expperhc	1	2.15218	1.63923	1.31	0.2057	0.70586	1.41671

Test of First and Second
Moment Specification

DF	Chi-Square	Pr > ChiSq
5	4.30	0.5076

From the model estimation output obtained we can make the following conclusion:

The **R-Square 0.5784** statistic indicate that the model accounts for 57% of the variation of the percentage time spent by staff on R&D activities.

The test for autocorrelation use is the Durban Watson test statistic. The `Durbin Watson` test statistic is `2.3068 with` `(Pr < DW = 0.6945 > 0.05 and (Pr < DW = 0.4055) < 0.95.` This indicates that we therefore can conclude that the autoregressive model does not have autocorrelation.

The heteroscedasticity test (Q and LM test for ARCH disturbances) is only interpreted up to 2 time lags. The probability for arch disturbances in the model for lags 1 and 2 are larger than 0.05. We can therefore conclude that the modelled relationship does not suffer from heteroscedasticity.

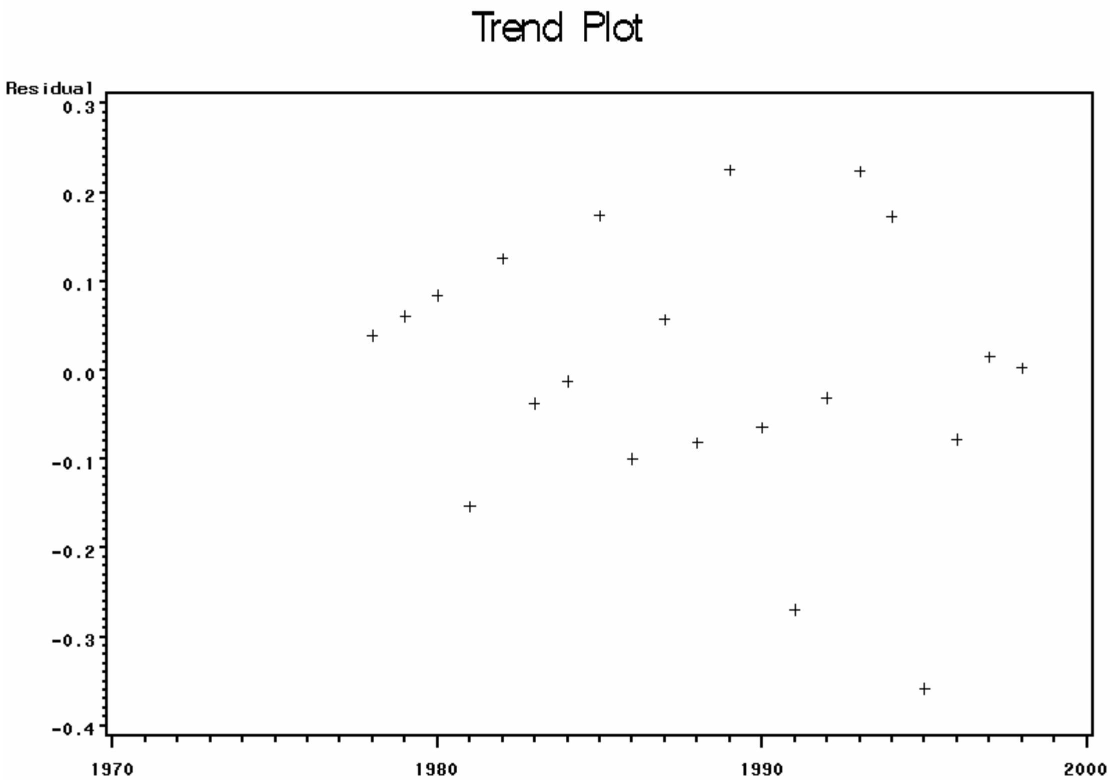

Figure 16-9 Time plot of the residual

From Figure 16-9 can be seen that the time plot seems to be scattered around 0. From the results we van also read the `Mean of Working Series is 8.02E-17` in **Error! Reference source not found.**. We therefore make use of the “Zero mean” specification in the stationarity test output results.

Table 16-15: Stationarity tests output for the residual

The ARIMA Procedure

Name of Variable = residual

Mean of Working Series	8.02E-17
Standard Deviation	0.145732
Number of Observations	21

Phillips-Perron Unit Root Tests

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau
Zero Mean	0	-23.1026	<.0001	-5.11	<.0001
	1	-22.3665	<.0001	-5.14	<.0001
Single Mean	0	-23.1028	0.0003	-4.97	0.0008
	1	-22.3618	0.0004	-5.00	0.0008
Trend	0	-23.7885	0.0022	-4.98	0.0039
	1	-22.8581	0.0035	-5.02	0.0036

For (n-1) = 2, the values are obtained from the Critical values for the Phillips Z Statistic or the Dickey Fuller t Statistic when applied to Residuals from Spurious Cointegration Regression (See table **Error! Reference source not found.**). The critical value for the 1% level is -4.31.

From the Phillips Perron test output obtained from SAS we read the following values for the probability statistics.

`Tau = -5.11` for $\ell = 0$ en

`Tau = -5.14` for $\ell = 1$.

This means that we can therefore reject the null hypothesis of unit root with a 1% significance level, since the `Tau` values are smaller than the critical value. The residues can be deemed stationary and the variables are cointegrated. We can therefore conclude that the regression is not spurious.

17 STATISTICAL TABLES

Table 17-1: Critical values for the Phillips Z statistic

TABLE B.9

Critical Values for the Phillips Z_t Statistic or the Dickey-Fuller t Statistic When Applied to Residuals from Spurious Cointegrating Regression

Number of right-hand variables in regression, excluding trend or constant	*Sample size*	*Probability that $(\hat{\rho} - 1)/\hat{\sigma}_{\hat{\rho}}$ is less than entry*						
$(n - 1)$	(T)	*0.010*	*0.025*	*0.050*	*0.075*	*0.100*	*0.125*	*0.150*
		Case 1						
1	500	−3.39	−3.05	−2.76	−2.58	−2.45	−2.35	−2.26
2	500	−3.84	−3.55	−3.27	−3.11	−2.99	−2.88	−2.79
3	500	−4.30	−3.99	−3.74	−3.57	−3.44	−3.35	−3.26
4	500	−4.67	−4.38	−4.13	−3.95	−3.81	−3.71	−3.61
5	500	−4.99	−4.67	−4.40	−4.25	−4.14	−4.04	−3.94
		Case 2						
1	500	−3.96	−3.64	−3.37	−3.20	−3.07	−2.96	−2.86
2	500	−4.31	−4.02	−3.77	−3.58	−3.45	−3.35	−3.26
3	500	−4.73	−4.37	−4.11	−3.96	−3.83	−3.73	−3.65
4	500	−5.07	−4.71	−4.45	−4.29	−4.16	−4.05	−3.96
5	500	−5.28	−4.98	−4.71	−4.56	−4.43	−4.33	−4.24
		Case 3						
1	500	−3.98	−3.68	−3.42	—	−3.13	—	—
2	500	−4.36	−4.07	−3.80	−3.65	−3.52	−3.42	−3.33
3	500	−4.65	−4.39	−4.16	−3.98	−3.84	−3.74	−3.66
4	500	−5.04	−4.77	−4.49	−4.32	−4.20	−4.08	−4.00
5	500	−5.36	−5.02	−4.74	−4.58	−4.46	−4.36	−4.28

The probability shown at the head of the column is the area in the left-hand tail.

Source: P. C. B. Phillips and S. Ouliaris, "Asymptotic Properties of Residual Based Tests for Cointegration," *Econometrica* 58 (1990), p. 190. Also Wayne A. Fuller, *Introduction to Statistical Time Series*, Wiley, New York, 1976, p. 373.

18 APPENDIX F

Please find the bitmap version of the model for the HES on the CD provided.

CD/HES/HES bitmap - SD model

19 APPENDIX G

Please find the bitmap version of the model for the Public sector on the CD provided.

CD/HES/Pub bitmap - SD model

20 APPENDIX H

Please find the bitmap version of the model for the Private sector on the CD provided.

CD/HES/Priv bitmap - SD model

19 APPENDIX I: SENSITVITY ANALYSIS - DELPHI STUDY

A sensitivity analysis is conducted of the Delphi responses obtained from the survey conducted in the research project. This analysis is done in order to determine if there is any reason for concern that there could exist some level of sub-aggregation within the individual group members of the three sectors surveyed (HES, Public Sector and Private Sector). If such a sub aggregation should exist, it could affect the reliability of the responses.

Sensitivity analysis of the responses is conducted by analysing responses aggregated in groups from the three sectors surveyed. The main purpose of doing this is to investigate if group aggregations do exist and if so to which extend. It must however also be kept in mind that dividing the already relatively small sample into the three sector groupings leave us with very small samples (5 responses for the three sectors – although 1 from the HES dropped out in the last round).

A simple analysis is done: The means and medians of the different groups are compared to investigate the possibility of a sub-aggregation within the groups members of the three sectors surveyed. If a significant difference in the mean or median is found the Standard deviation is investigated – if the sub-groups also have a high standard deviation, it can be concluded that there exists an overall low level of consensus. However if the overall level of consensus is low, but a high level of consensus is achieved in the sub-groups, it could indicate a sub-aggregation for that response – potentially skewing the results.

19.1 The Higher Education Sector: Delphi sensitivity analysis

In order to save space in the box plots as well as in the tables, the questions in the survey is reference to the question number. The following table serves as a reference for the coding of the issue categories.

Table 19-1: Numbering of survey questions

Issue Category
Q1. Lack of funding for R&D in the HES
Q2. Lack of multidisciplinary research projects
Q3. Poor linkages pose a threat to future capacity and the relevance of R&D performed in the system
Q4. Inability to retain and rejuvenate human resource stock in the system
Q5. The deterioration of quality of human resources working in R&D in the sector
Q6. Inadequate funding of equipment
Q7. The lack of female and black researchers for R&D to reach representative work force
Q8. Difficulty of successful R&D policy alignment with national priorities
Q9. Weak IP protection policies in HES

The following box plot is a visual representation of the responses received from respondents from all three sectors.

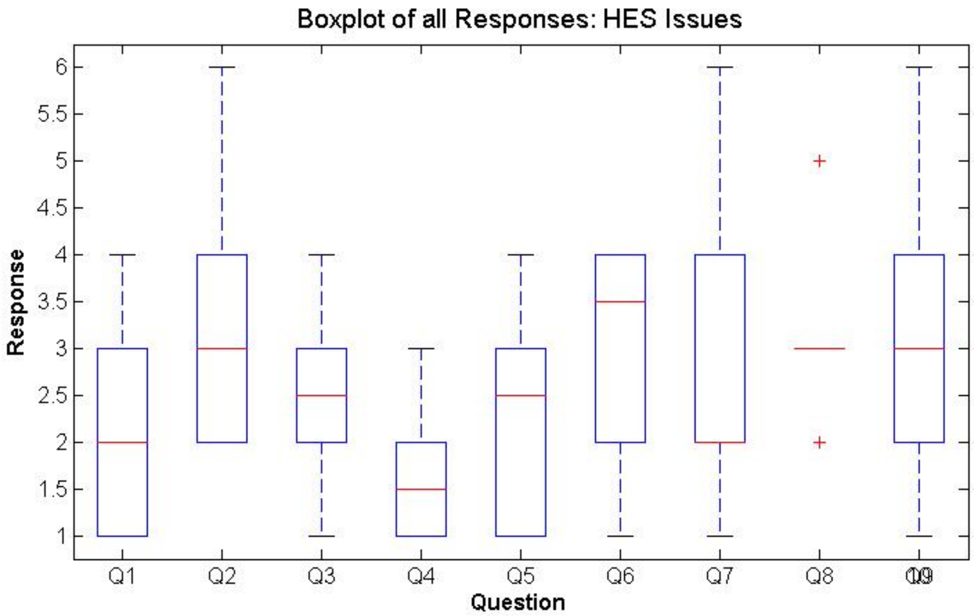

Figure 19-1: Box plot of Responses from all respondents: Issues in the HES

The following tables summarise the analysis done on the set of responses from all respondents.

Table 19-2: Summary of responses from all respondents: Issues in HES

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9
Mean	2.29	3.29	2.43	1.64	2.43	3.14	2.71	3.29	3.36
Median	2.00	3.00	2.50	1.50	2.50	3.50	2.00	3.00	3.00
St. Dev	1.14	1.59	0.85	0.74	1.16	1.03	1.54	0.99	1.39
Upper Quartile	4.00	6.00	4.00	3.00	4.00	4.00	6.00	5.00	6.00
Lower Quartile	1.25	2.00	2.00	1.00	1.25	2.25	2.00	3.00	2.25
Maximum	4	6	4	3	4	4	6	5	6
Minimum	1	2	1	1	1	1	1	2	1

Table 19-3: Summary of responses from HES respondents: Issues in HES

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9
Mean	1.3	3.8	2.8	1.5	1.8	2.3	3.0	3.3	4.3
Median	1.0	3.0	2.5	1.5	1.5	2.5	3.5	3.0	4.5
St. Dev	0.5	1.5	1.0	0.6	1.0	1.0	1.4	1.3	1.0
Upper Quartile	2.0	6.0	4.0	2.0	3.0	3.0	4.0	5.0	5.0
Lower Quartile	1.0	3.0	2.0	1.0	1.0	1.8	2.5	2.8	3.8
Maximum	2.0	6.0	4.0	2.0	3.0	3.0	4.0	5.0	5.0

Minimum	1.0	3.0	2.0	1.0	1.0	1.0	1.0	2.0	3.0

Table 19-4: Summary of responses from Public Sector respondents: Issues in HES

	Q1	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**	**Q7**	**Q8**	**Q9**
Mean	3	4	3	2	3	3	2	4	3
Median	2	4	3	2	3	4	2	3	3
St. Dev	1.00	1.71	0.50	0.58	0.82	0.96	0.00	1.00	0.58
Upper Quartile	4	6	3	2	4	4	2	5	3
Lower Quartile	2	2.75	2.75	1	2.75	2.75	2	3	2
Maximum	4	6	3	2	4	4	2	5	3
Minimum	2	2	2	1	2	2	2	3	2

Table 19-5: Summary of responses from Public Sector respondents: Issues in HES

	Q1	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**	**Q7**	**Q8**	**Q9**
Mean	2.8	2.7	2.0	1.8	2.5	3.7	3.0	3.2	3.3
Median	3.0	2.0	2.0	1.5	2.5	4.0	2.5	3.0	3.5
St. Dev	1.2	1.6	0.9	1.0	1.4	0.8	2.1	1.0	1.8
Upper Quartile	4.0	6.0	3.0	3.0	4.0	4.0	6.0	5.0	6.0
Lower Quartile	2.3	2.0	1.3	1.0	1.3	4.0	1.3	3.0	2.3
Maximum	4	6	3	3	4	4	6	5	6
Minimum	1	2	1	1	1	2	1	2	1

The following table summarises a comparison of means between the three sectors surveyed. The mean is rounded as to give a better sense which main category each group opinion fall into.

Table 19-6: Comparison of means (HES issues)

	Q1	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**	**Q7**	**Q8**	**Q9**
Mean	2	3	2	2	2	3	3	3	3
HES	1	4	3	2	2	2	3	3	4
PUB	3	4	3	2	3	3	2	4	3
Private	3	3	2	2	3	4	3	3	3

The following figure is a graphical representation of the data in the table:

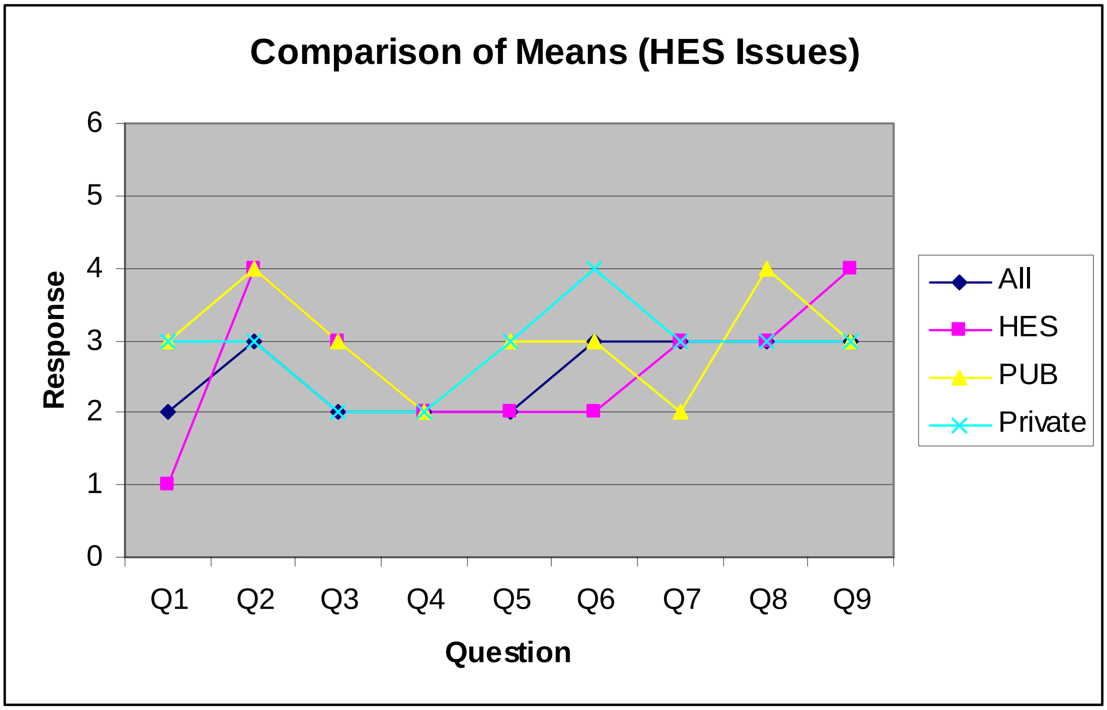

Figure 19-2: Graphical representation of the Comparison of Means (HES Issues)

The following table summarises the comparison of medians between the three sectors surveyed.

Table 19-7: Comparison of Medians (HES Issues)

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9
Median	2	3	2	2	3	4	2	3	3
HES	1	3	3	2	2	3	4	3.	5
PUB	2	4	3	2	3	4	2	3	3
Private	3	3	2	2	3	4	3	3	3

The following figure is a graphical representation of the data in the table:

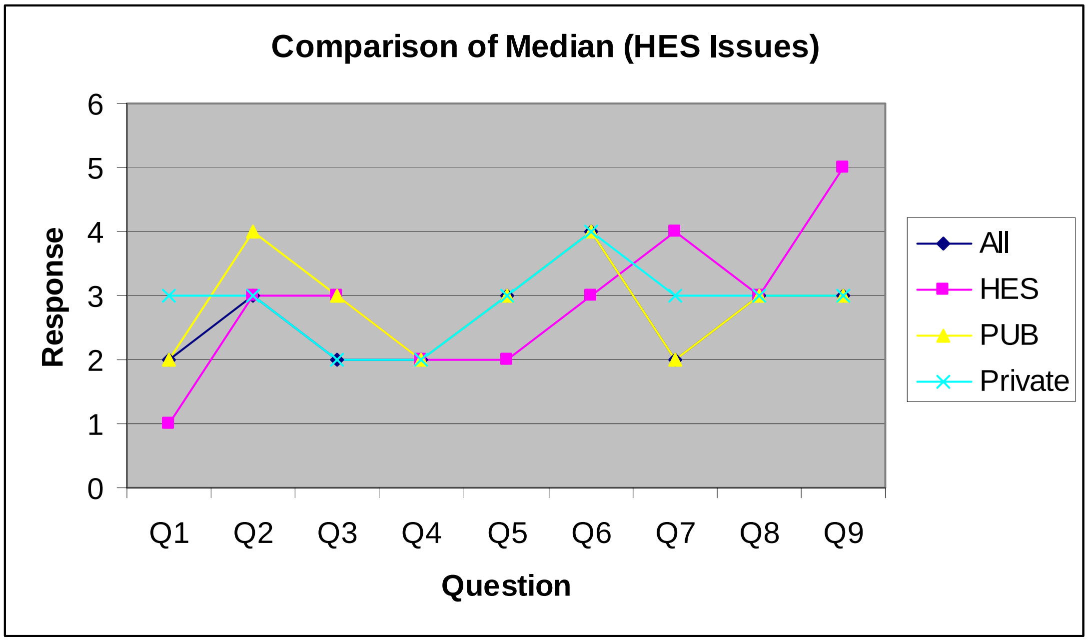

Figure 19-3: Graphical representation of the Comparison of Medians (HES Issues)

No consistent trend or major differences can be seen from the mean and the median of the three groups. It can therefore be concluded that no obvious sub-aggregation that could skew the overall result exist in the groups surveyed.

Although not explicitly included in the analysis, the box plots for each one of the sub aggregations are also provided.

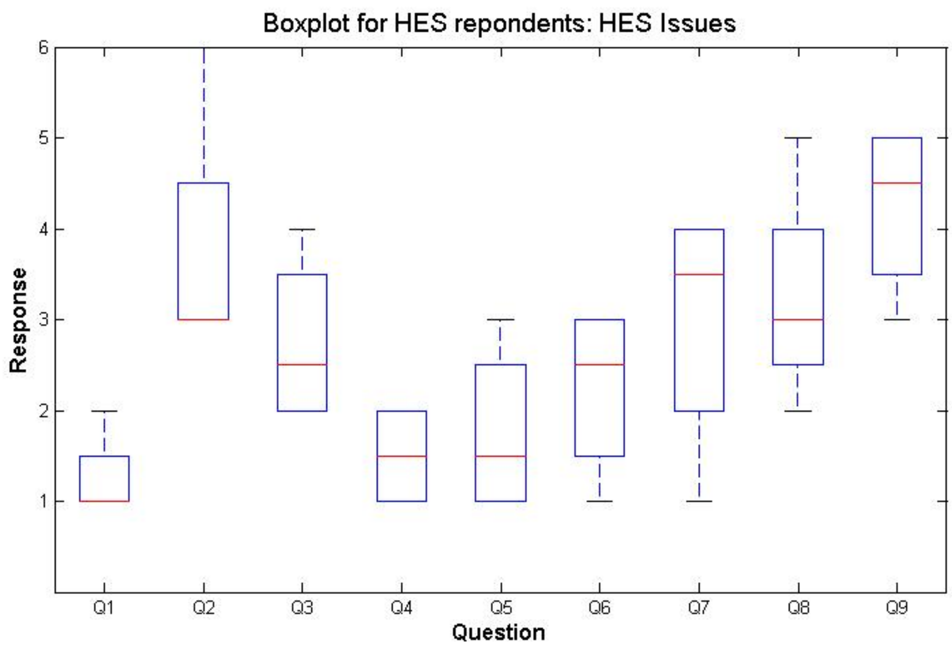

Figure 19-4: Box plot for HES respondents: HES issues

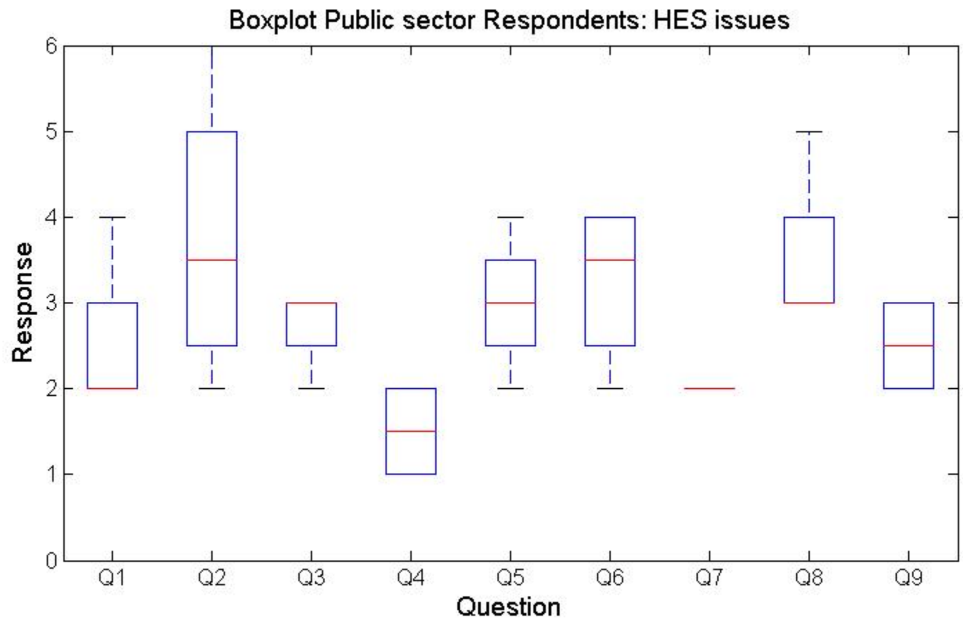

Figure 19-5: Box plot for Public sector respondents: HES issues

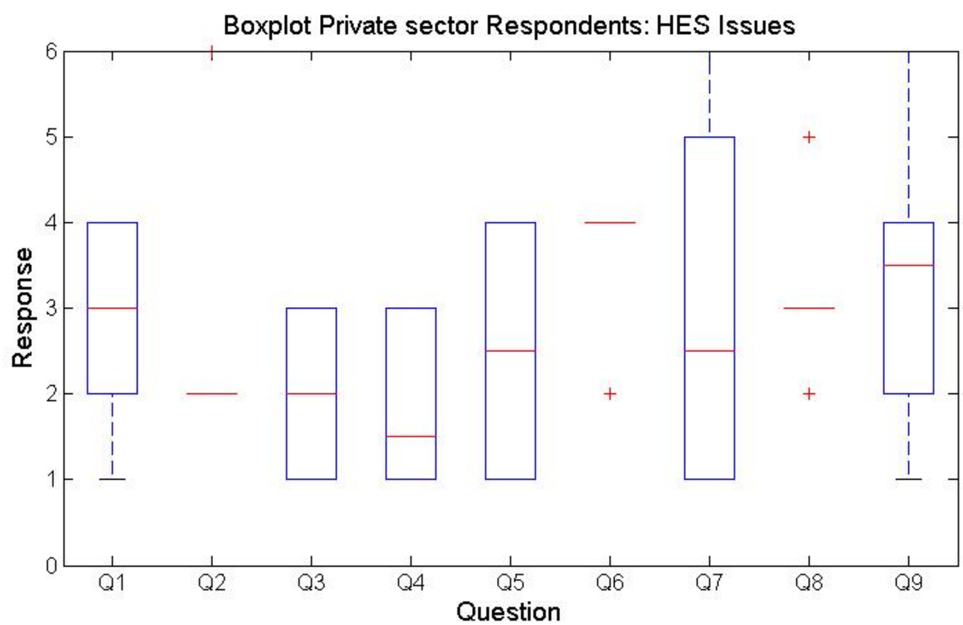

Figure 19-6: Box plot for Private sector respondents: HES issues

19.2 The Public Sector: Delphi sensitivity analysis

In order to save space in the box plots as well as in the tables, the questions in the survey is reference to the question number. The following table serves as a reference for the coding of the issue categories.

Table 19-8: Numbering of survey questions for Public sector

Issue Category
Q1. Inability to retain and rejuvenate the researchers stock in the system
Q2. Lack of government funding to the public sector to develop R&D and technology platforms
Q3. Deterioration of quality of human resources working in R&D
Q4. Current BEE policies having a negative effect on quality and R&D capacity
Q5. A lack of direction and leadership in science policy

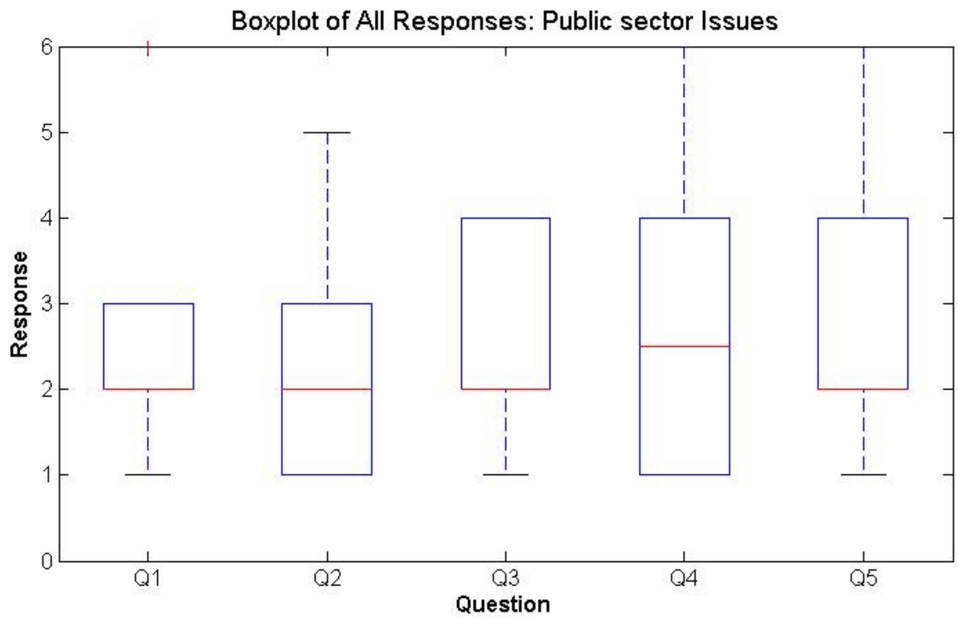

Figure 19-7: Box plot of Responses from all respondents: Issues in the Public sector

The following table summarises the analysis done on the set of responses from all respondents:

Table 19-9: Summary of responses from All respondents: Issues in Public Sector

	Q1	Q2	Q3	Q4	Q5
Mean	2.43	2.21	2.43	2.86	2.79
Median	2.00	2.00	2.00	2.50	2.00
St. Dev	1.16	1.31	1.16	1.70	1.48
Upper Quartile	6	5	4	6	6
Lower Quartile	2	1	2	1.25	2
Maximum	6	5	4	6	6
Minimum	1	1	1	1	1

Table 19-10: Summary of responses from HES respondents: Issues in Public Sector

	Q1	Q2	Q3	Q4	Q5
Mean	3.25	2.25	2	1.75	2
Median	2.5	2.5	2	1.5	2
St. Dev	1.89	0.96	0.82	0.96	0.82

Upper Quartile	6	3	3	3	3
Lower Quartile	2	1.75	1.75	1	1.75
Maximum	6	3	3	3	3
Minimum	2	1	1	1	1

Table 19-11: Summary of responses from Public Sector respondents: Issues in Public Sector

	Q1	**Q2**	**Q3**	**Q4**	**Q5**
Mean	1.75	1.50	3.00	3.50	4.25
Median	2.00	1.00	3.00	3.50	4.50
St. Dev	0.50	1.00	1.15	1.29	1.71
Upper Quartile	2	3	4	5	6
Lower Quartile	1.75	1	2	2.75	3.5
Maximum	2	3	4	5	6
Minimum	1	1	2	2	2

Table 19-12: Summary of responses from Private Sector respondents: Issues in Public Sector

	Q1	Q2	Q3	Q4	Q5
Mean	2.43	2.21	2.43	2.86	2.79
Median	2.00	2.00	2.00	2.50	2.00
Stdev	1.16	1.31	1.16	1.70	1.48
Upper Quartile	6	5	4	6	6
Lower Quartile	2	1	2	1.25	2
Maximum	6	5	4	6	6
Minimum	1	1	1	1	1

The following table summarises a comparison of means between the three sectors surveyed. The mean is rounded as to give a better sense which main category each group opinion fall into.

Table 19-13: Comparison of means (Public Sector issues)

	Q1	**Q2**	**Q3**	**Q4**	**Q5**
Mean	2	2	2	3	3
HES	3	2	2	2	2
PUB	2	2	3	4	4
Private	2	2	2	3	3

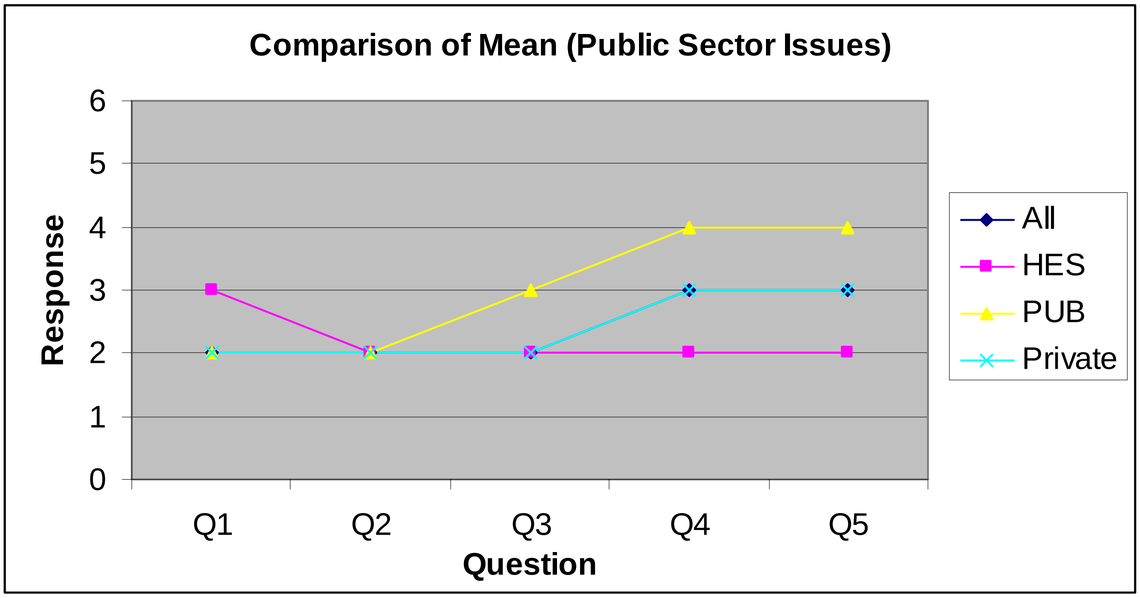

Figure 19-8: Graphical representation of the Comparison of Means (Pub. Sector Issues)

The following table summarises the comparison of medians between the three sectors surveyed.

Table 19-14: Comparison of Medians (HES Issues)

	Q1	Q2	Q3	Q4	Q5
Mean	2	2	2	2.5	2
HES	2.5	2.5	2	1.5	2
PUB	2	1	3	3.5	4.5
Private	2	2	2	2.5	2

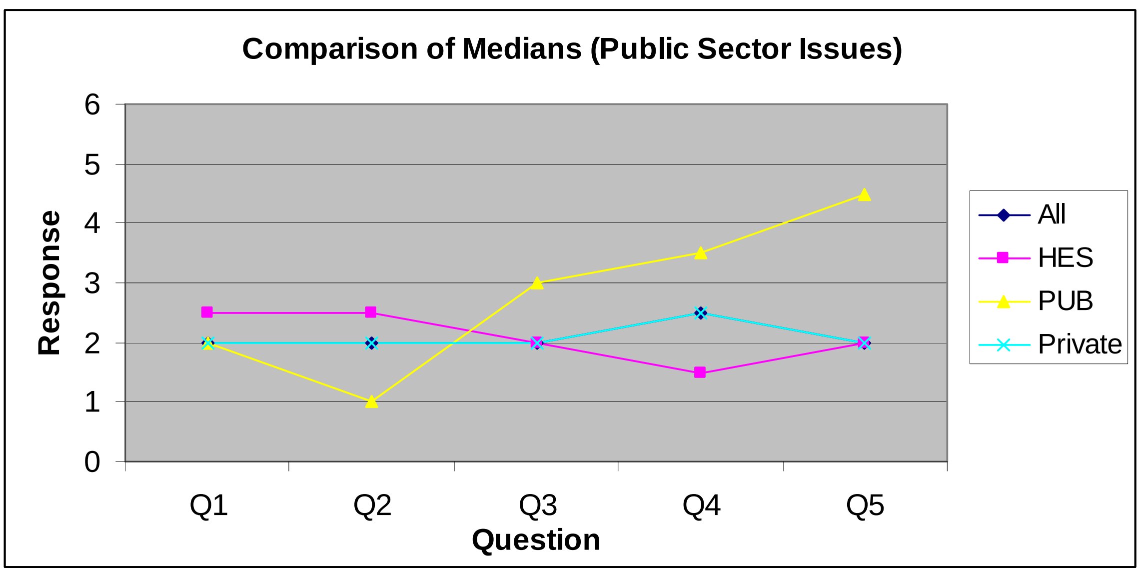

Figure 19-9: Graphical representation of the Comparison of Medians (Pub. Sector Issues)

Q4 and Q5 show larger differences between the three sectors in terms of the Median value. From this we have to investigate the possibility that homogeneity in the responses from the three sectors could be skewing the results.

Table 19-15: Summary of St.DEv (Private sector issues)

	Q1	Q2	Q3	Q4	Q5
Overall	1.16	1.31	1.16	1.70	1.48
HES	1.89	0.96	0.82	0.96	0.82
PUB	0.50	1.00	1.15	1.29	1.71
Private	1.16	1.31	1.16	1.70	1.48

We however find that the resulting high standard deviation for the whole response set is also present within the response sets from the three sectors for these three questions. The high Standard Deviation for Q4, Q5 indicates a low overall level of consensus within the groups. We can therefore conclude that the high overall standard deviation is not a low level of consensus due to homogeneity on responses from the three sectors. We can therefore conclude that for the purpose of this study we can be satisfied that the responses are heterogeneous for this sector.

Although not explicitly included in the analysis, the box plots for each one of the sub aggregations are also provided.

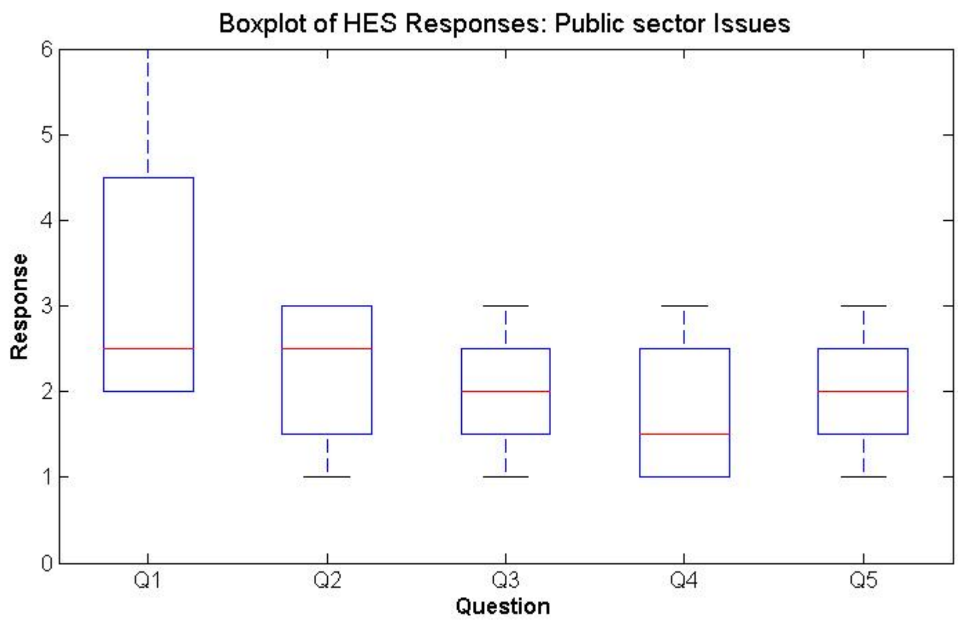

Figure 19-10: Box plot of Responses from HES respondents: Issues in the Public sector

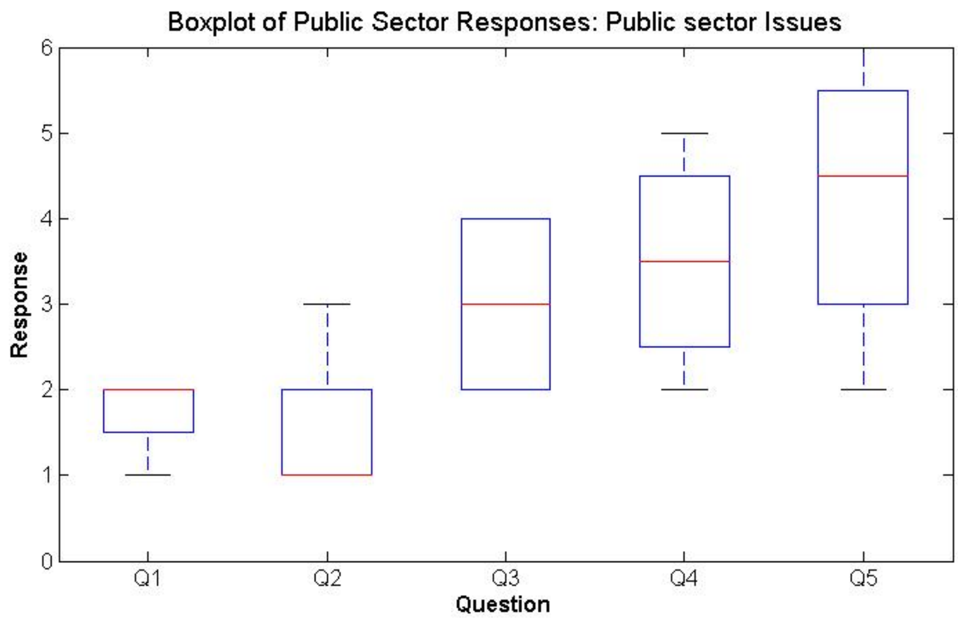

Figure 19-11: Box plot of Responses from Public sector respondents: Issues in the Public sector

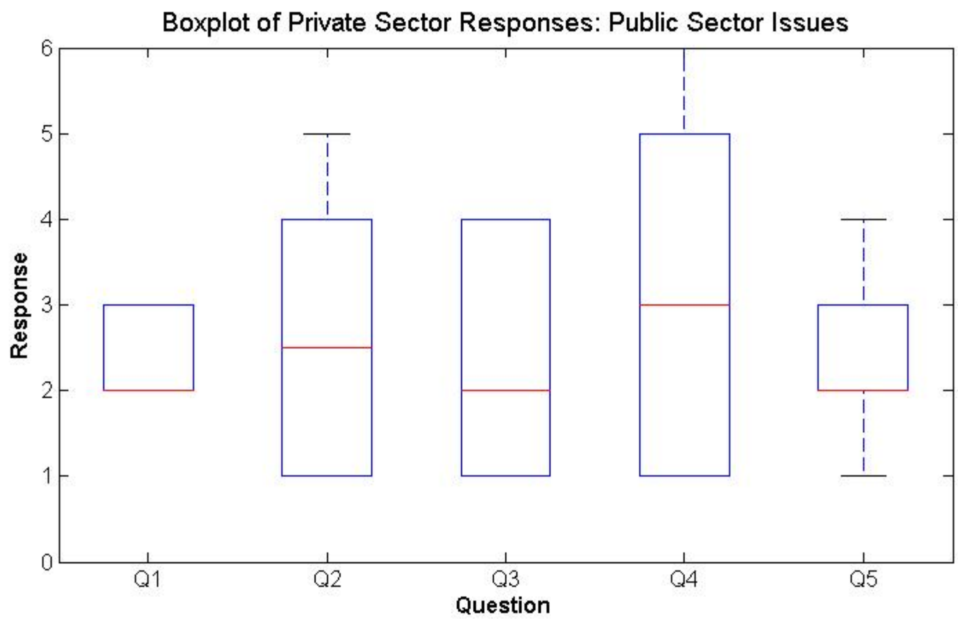

Figure 19-12: Box plot of Responses from Private sector respondents: Issues in the Public sector

19.3 The Private Sector: Delphi sensitivity analysis

In order to save space in the box plots as well as in the tables, the questions in the survey is reference to the question number. The following table serves as a reference for the coding of the issue categories.

Table 19-16: Numbering of survey questions for Public sector

Issue Category
Q1. Lack of research culture in South Africa
Q2. Lack of fiscal incentives from government to foster R&D culture in companies
Q3. Lack of funding of R&D
Q4. Inability to retain and rejuvenate the researchers stock in the system
Q5. Deterioration of quality (skill level) of human resources working in R&D
Q6. Current BEE policies will have a negative effect on South Africa's future R&D capacity
Q7. Poor linkages
Q8. Lack of direction and leadership in science policy
Q9. Restrictive communication infrastructure

The following box plot is a visual representation of the responses received from respondents from all three sectors.

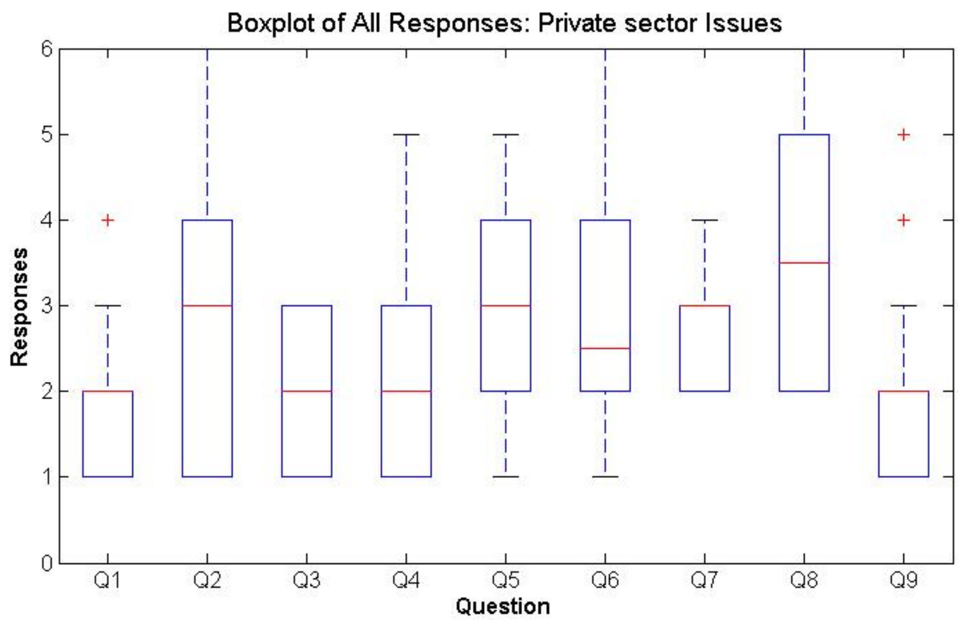

Figure 19-13: Box plot of Responses from all respondents: Issues in the Private sector

The following tables summarise the analysis done on the set of responses from all respondents.

Table 19-17: Summary of responses from all respondents: Issues in Private Sector

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9
Mean	2	3	2	2	3	3	3	4	2
Median	2	3	2	2	3	3	3	4	2
St. Dev	0.86	1.64	0.83	1.20	1.30	1.59	0.63	1.49	1.21
Upper Quartile	4	6	3	5	5	6	4	6	5

Lower Quartile	1	1	1.25	1.25	2	2	2	2.25	1
Maximum	4	6	3	5	5	6	4	6	5
Minimum	1	1	1	1	1	1	2	2	1

Table 19-18: Summary of responses from HES respondents: Issues in Private Sector

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9
Mean	2	1	2	2	2	2	3	3	2
Median	2	1	2	2	2	2	3	3	2
St. Dev	0.58	0.50	0.96	0.50	0.82	0.58	0.82	0.58	1.41
Upper Quartile	2	2	3	2	3	2	4	3	4
Lower Quartile	1	1	1	1.75	1.75	1	2.75	2	1
Maximum	2	2	3	2	3	2	4	3	4
Minimum	1	1	1	1	1	1	2	2	1

Table 19-19: Summary of responses from Public sector respondents: Issues in Private Sector

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9
Mean	2	5	3	2	4	4	2	5	2
Median	2	5	3	2	4	4	2	5	2
St. Dev	0	1.3	0.6	1	0.6	1.3	0.5	1.3	0
Upper Quartile	2	6	3	3	4	5	3	6	2
Lower Quartile	2	3.75	2	1	3	2.75	2	3.75	2
Maximum	2	6	3	3	4	5	3	6	2
Minimum	2	3	2	1	3	2	2	3	2

Table 19-20: Summary of responses from Private sector respondents: Issues in Private Sector

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9
Mean	2.0	2.5	2.0	3.0	3.3	3.5	2.7	4.0	2.2
Median	1.5	3.0	2.0	3.0	3.5	4.0	3.0	4.5	1.5
St. Dev	1.3	1.2	0.9	1.4	1.6	1.8	0.5	1.7	1.6
Upper Quartile	4.0	4.0	3.0	5.0	5.0	6.0	3.0	6.0	5.0
Lower Quartile	1.0	1.5	1.3	2.3	2.3	2.5	2.3	2.5	1.0
Maximum	4	4	3	5	5	6	3	6	5
Minimum	1	1	1	1	1	1	2	2	1

The following table summarises a comparison of means between the three sectors surveyed. The mean is rounded as to give a better sense which main category each group opinion fall into.

Table 19-21: Comparison of means (Private sector issues)

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9
Mean	2	3	2	2	3	3	3	4	2
HES	2	1	2	2	2	2	3	3	2
PUB	2	5	3	2	4	4	2	5	2
Private	2	3	2	3	3	4	3	4	2

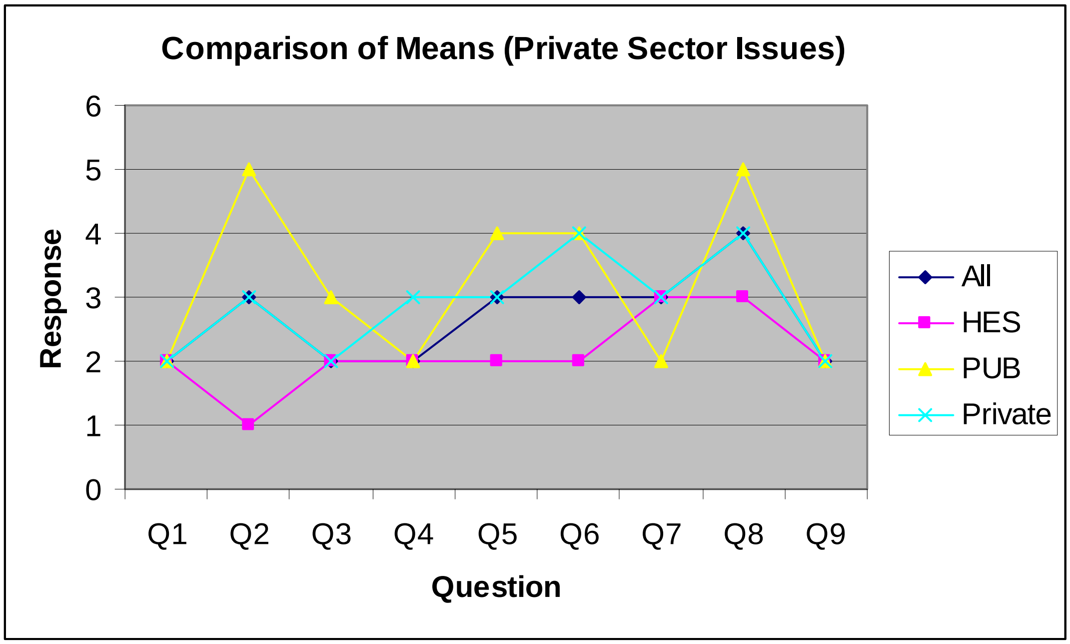

Figure 19-14: Graphical representation of the Comparison of Means (Private Sector Issues)

The following table summarises the comparison of medians between the three sectors surveyed.

Table 19-22: Comparison of Medians (Private sector Issues)

	Q1	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**	**Q7**	**Q8**	**Q9**
Median	2	3	2	2	3	3	3	4	2
HES	2	1	2	2	2	2	3	3	2
PUB	2	5	3	2	4	4	2	5	2
Private	1.5	3.0	2.0	3.0	3.5	4.0	3.0	4.5	1.5

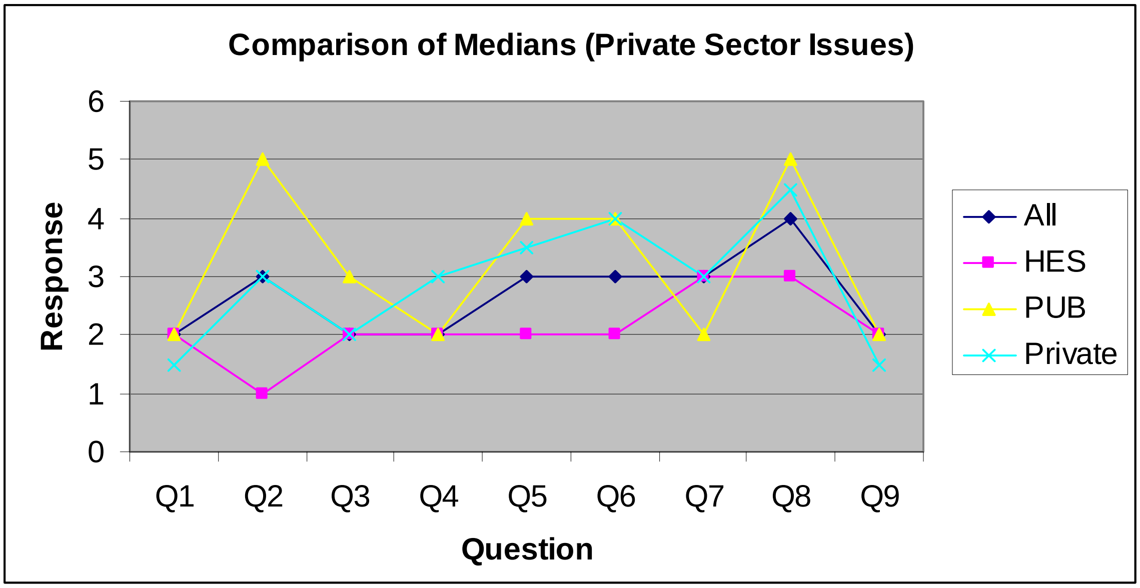

Figure 19-15: Graphical representation of the Comparison of Medians (Private Sector Issues)

For the Private sector analysis, the HES respondents in some cases seem to rank the issues as more serious than the other two groupings. However this trend was not evident in the analysis of the other two sectors.

Q2, Q5, and Q6 all show larger differences between the three sectors in terms of the Median value. From this we have to investigate the possibility that homogeneity in the responses from the three sectors could be skewing the results.

Table 19-23: Summary of St.DEv (Private sector issues)

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9
Overall	0.86	1.64	0.83	1.20	1.30	1.59	0.63	1.49	1.21
HES	0.58	0.50	0.96	0.50	0.82	0.58	0.82	0.58	1.41
PUB	0	1.3	0.6	1	0.6	1.3	0.5	1.3	0
Private	1.3	1.2	0.9	1.4	1.6	1.8	0.5	1.7	1.6

We however find that the resulting high standard deviation for the whole response set is also present within the response sets from the three sectors for these three questions. The high Standard Deviation for Q2, Q5 and Q6 indicates a low overall level of consensus within the groups. We can therefore conclude that the high overall standard deviation is not a low level of consensus due to homogeneity on responses from the three sectors. We can therefore conclude that for the purpose of this study we can be satisfied that the responses are heterogeneous for this sector.

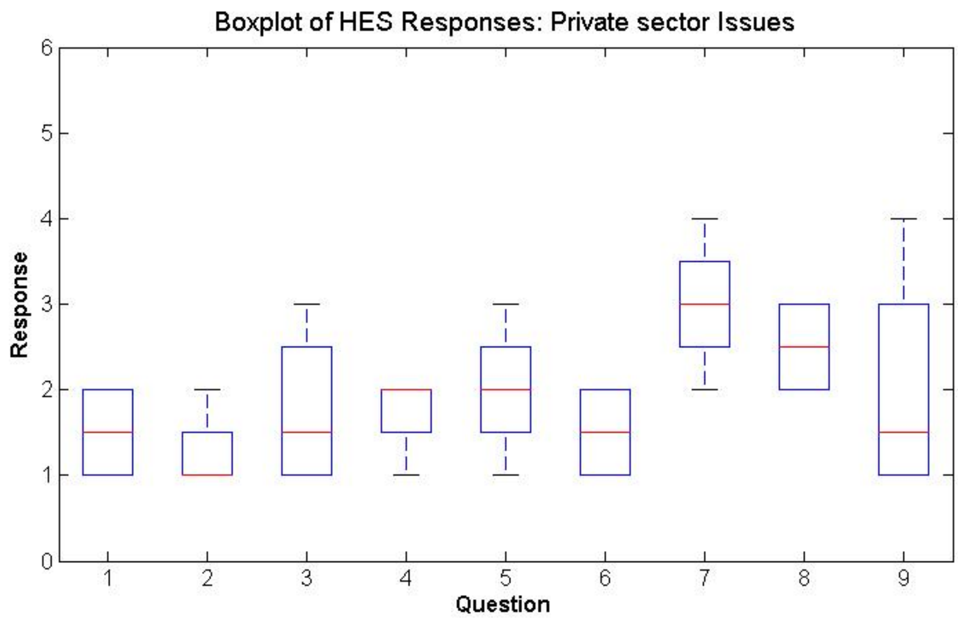

Figure 19-16: Box plot of Responses from HES respondents: Issues in the Private sector

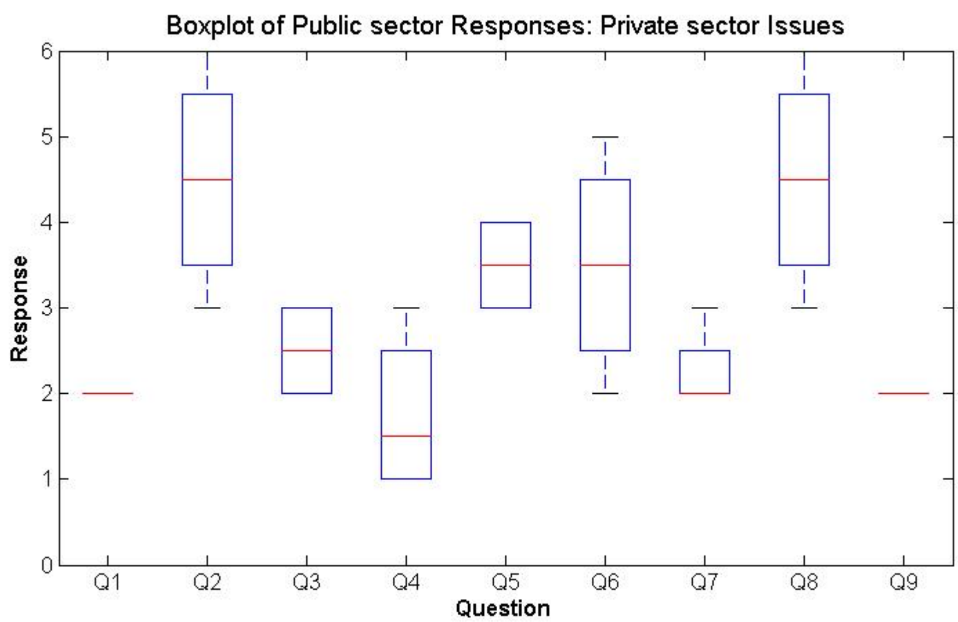

Figure 19-17: Box plot of Responses from Public sector respondents: Issues in the Private sector

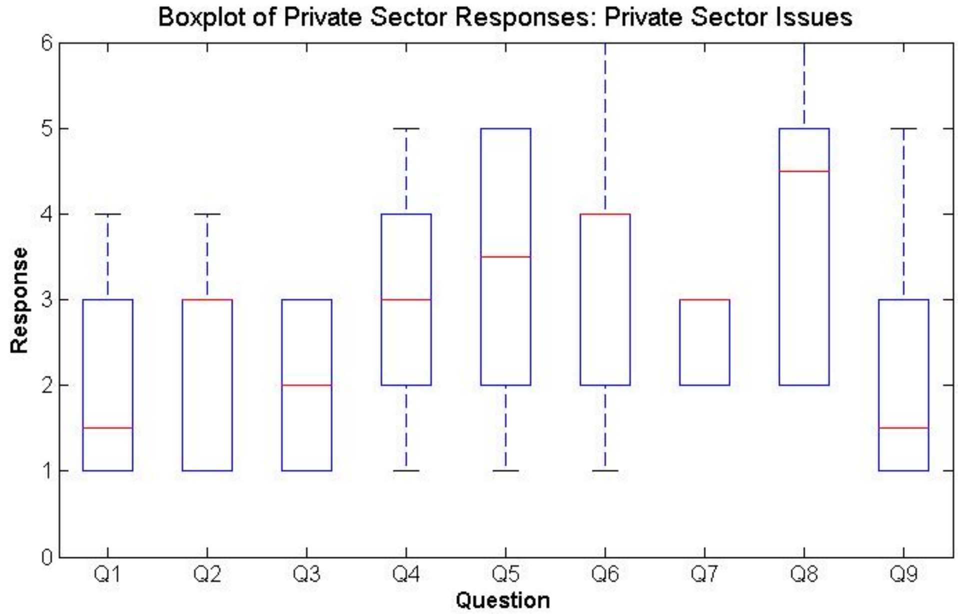

Figure 19-18: Box plot of Responses from Private sector respondents: Issues in the Private sector