Four Essays on Empirical Economics

Dissertation

zur Erlangung des Grades
Doktor der Wirtschaftswissenschaften (Dr. rer. pol.)
am Fachbereich Wirtschaftswissenschaften
der Universität Konstanz

vorgelegt von:

Michael Rauber
Hauptstr. 7a
88090 Immenstaad

Konstanz, den 5. August 2007

Tag der mündlichen Prüfung: 12. Dezember 2007

1. Referent: Prof. Dr. H. W. Ursprung, Universität Konstanz

2. Referent: Prof. Dr. W. Pohlmeier, Universität Konstanz

Table of Contents

Summary 1
Chapter 1: Foreign Aid and Developing Countries' Creditworthiness 4
1.1 Introduction 5
1.2. A simple model of aid and default risk 6
1.3 Aid and country creditworthiness: An empirical exploration 10
1.3.1 Data 10
1.3.1.1 Country creditworthiness 10
1.3.1.2 Aid 11
1.3.1.3 Control variables 12
1.3.1.4 Lagged dependent variable 13
1.3.2 Estimation 13
1.3.2.1 Specification 13
1.3.2.2 GMM estimation 14
1.3.2.3 Results 15
1.3.3 Robustness checks 16
1.4 Summary and conclusions 18
1.5 Tables and Figures 20
1.6 Data appendix 26
References 29
Chapter 2: Research Productivity of German Academic Economists 33
2.1 Evaluation of Researchers: A Life Cycle Analysis of German Academic Economists 34
2.1.1 Introduction 35
2.1.2 The data set 38
2.1.3 Describing the landscape of German academic research in economics 40
2.1.4 A micro-econometric investigation of life cycle productivities 43
2.1.5 Persistence of research productivity 46
2.1.6 Some new rankings for German economics departments 48

2.1.7 Tables and Figures 52
References 57
2.2 Life Cycle and Cohort Productivity in Economic Research: The Continental European Experience as exemplified by the case of Germany 59
2.2.1 Introduction 60
2.2.2 The Data 63
2.2.2.1 The sample 63
2.2.2.2 The dependent variable: Individual annual research productivity 64
2.2.2.3 The explanatory variables 65
2.2.3 Results 67
2.2.3.1 Identifying life cycles in research productivity 67
2.2.3.2 Quantile regressions and cohort-specific life cycles 70
2.2.3.3 Ability-specific life cycles 72
2.2.4 An exercise in deconstruction: Quality, quantity and co-authorship 73
2.2.5 Conclusions 76
2.2.6 Tables and Figures 79
2.2.7 Appendix 86
References 87
Chapter 3: Noncognitive Skills and Success in Life: The Importance of Motivation and Self-Regulation 90
3.1 Introduction 91
3.2 The Data 92
3.3 Noncognitive skills 93
3.4 Cognitive skills 97
3.5 Results 98
3.5.1 School performance 98
3.5.2 Professional career 101
3.5.2.1 Wage and Income 101
3.5.2.2 Occupational Prestige 104
3.5.3 Happiness 105
3.6 Selection 106
3.7 Conclusion 107
3.8 Tables and Figures 109
Appendix 115

References 117
References 120
Zusammenfassung 129
Erklärung 133
Abgrenzung 134

List of Tables

1.1: Benchmark specification: Alternative estimation methods 20
1.2: Different types of aid 21
1.3: Netting out interest payments and debt forgiveness grants 22
1.4: Different country groups 23
1.5: Structural breaks and nonlinear effects 24
1.6: Descriptive statistics 28
1.7: Correlations 28
2.1.1: Estimates 54
2.1.2: Transition probabilities: year 6 –year 12 55
2.1.3: Transition probabilities of cohort B 55
2.1.4: Department Rankings 56
2.2.1: Pooled estimates 79
2.2.2: Quantile estimates 80
2.2.3: Semiparametric Fixed Effects estimation 80
2.2.4: Ability groups 81
2.2.5: Density Deconstruction 82
2.2.6: Descriptive statistics 86
3.1: Descriptive statistics 109
3.2: Correlations between variables 109
3.3: School performance 110
3.4: Professional career: Wage 111
3.5: Professional career: Occupational Prestige 112
3.6: Happiness 113
3.7: Selection effects 114

List of Figures

1.1: The effect of raising A_1 on the equilibrium value of K_2 25
2.1.1: Percentiles 1969-1998 52
2.1.2: Percentiles for cohort 1969-1974 52
2.1.3: 80%- Percentile Lines for all cohorts 52
2.1.4: Probability of being active 53
2.1.5: Conditional productivity by rank 53
2.1.6: Tobit estimates by cohort 53
2.2.1: Pooled specification 83
2.2.2: Censored Quantile regressions 83
2.2.3: Semiparametric Fixed Effects estimation 84
2.2.4: Historical time and cohort effects 84
2.2.5: Ability specific regressions 84
2.2.6: Density Deconstruction 85
3.1: Attribution schemes 94
3.2: Distribution of net hourly wage 116
3.3: Distribution of IST 116
3.4: Distribution of Occupational Prestige 116

Summary

This dissertation is a collection of four research articles written during my doctoral studies at the University of Konstanz from April 2004 to July 2007. It has originated from the fundamental insight that empirical analysis is indispensable to our understanding of economic relationships. All of the papers have an empirical orientation in common and the research questions addressed are analysed using various econometric methods to gain a deeper understanding of these interactions. This chapter serves as an introduction and briefly summarizes the following chapters and central results.

Chapter 1 is based on the working paper *Foreign aid and developing countries' creditworthiness,* which is jointly written with Philipp Harms (RWTH Aachen) and descends from my diploma thesis. We explore whether foreign aid affects developing countries' creditworthiness, as proxied by the Institutional Investor's measure of country credit risk. We start our analysis by developing a simple model of international borrowing and lending to analyze how aid affects agents' demand for foreign credit and the likelihood that they will repay their debt. In this framework, a transfer in a given period lowers the net benefits of future default and therefore raises creditworthiness as perceived by international investors. We then test this hypothesis, using a data set that covers a large number of developing countries in the 1980s and 1990s. We find that larger aid inflows result in an improvement of the recipient country's Institutional Investor rating. However, different types of aid differ in their effects: while grants and technical assistance significantly raise a country's creditworthiness, we do not find a significant effect if we focus on the loan component of total aid flows.

Chapter 2 descends from two articles on the research productivity of German academic economists, both written together with Heinrich Ursprung (University of Konstanz). In the first paper *Evaluation of researchers: A life cycle analysis of German academic economists* we start off with the observation that all meaningful evaluations compare certain features of

a person with the features observed in a group of peers. For researchers the most important peer characteristic is age. Age features two distinct dimensions that are relevant in the evaluation context: vintage and career age. Both of these dimensions are liable to have a strong impact on research productivity because research production heavily relies on human capital that is determined, on the one hand, by the initial endowment (i.e. by ability and initial training) and, on the other hand, by experience and obsolescence of knowledge.

We therefore argue that any meaningful bibliometric evaluation of researchers needs to take into account that research productivity follows distinct life cycles. Using an encompassing data set portraying the research behavior of German academic economists, we confirm that research productivity crucially depends on career age and vintage. Based on the identified effects, we develop a simple formula that shows how a researcher's performance compares to that of his or her peers. This kind of information may serve as an input for performance-related remuneration and track-record based allocation of research grants. We then go on to investigate the persistence of individual productivity. The persistence issue is of special importance in the academic labor market because of the irrevocable nature of tenure. Our results indicate that reducing the review period from the German standard of eight years to the American standard of six years does not appear to come at an inordinate loss of information. However, such a reduction might make an academic career more appealing because of the shorter waiting period. Finally, we show how life cycle considerations can be used in evaluations of university departments in order to render the resulting rankings insensitive to the age structure of the evaluated faculties.

The second part of Chapter 2 derives from the paper *Life Cycle and Cohort Productivity in Economic Research: The Continental European Experience as exemplified by the case of Germany*, also written jointly with Heinrich Ursprung. One of the objectives of this article is to analyze how the traditional continental European science system has responded, on the one hand, to structural changes that provide more incentives for high research productivity, and, on the other hand, to the increased competition stemming from the fact that the European science system has become more and more exposed to the global standards set by the Anglo-Saxon system. We examine this transformation process by investigating the research productivity of German academic economists over their life cycles. Thereby we find that the pattern of research productivity over the life cycle is co-determined by economic incentives and by sociological factors. The influence of the economic incentives is reflected in the hump-shape of the identified life cycles, the sociological factors show up

in the marked cohort effects. As compared to the life cycles of their American peers, the life cycles of German economists turn out to be flatter and the level of research productivity appears to depend much more on cohort specific factors. These findings support our hypothesis that the academic environment in Germany has changed dramatically over the period of our investigation. A further result of our analysis that deserves special attention is that life cycles in research productivity are highly heterogeneous, especially with respect to ability and gender. Studies that attempt to identify the research behavior of the "representative" economist therefore miss a large part of the story. Finally, we find that not only overall productivity, but also research quality follows distinct life cycles.

Chapter 3 is based upon the paper *Noncognitive skills and success in life: The importance of motivation and self-regulation.* The recent econometric literature focuses on the importance of so called noncognitive skills on labor market outcomes and success. Whereas it is by now well known that cognitive skills are important for success in life, research with respect to noncognitive skills is still in its infancy. In this article I analyze the effect of motivation, self-regulation and social skills formed during early childhood on performance in school, achievement in professional life and happiness at the age of 43. For the analysis, a longitudinal panel study of 3240 high-school students from North Rhine-Westphalia was used. Having data on both cognitive and noncognitive skills at the time of adolescence and early adulthood, the panel structure of the dataset allows the explicit taking of timing and causality into account.

The analysis confirms recent insights that noncognitive skills formed during childhood have a significant and lasting impact on success in all three areas. With respect to performance in school, a high degree of self-regulation and motivation improves grades, but the effects differ between subjects and gender.

The analysis of labour market success indicates that wages, income and occupational prestige increase with the level of self-regulation. However, the strength of the effect differs between men and women. Finally, individuals with internal and unstable attribution styles seem to be more successful in the labour market than others. These results are in accordance with several theories from motivational psychology and indicate important possibilities for parents to contribute towards the long term development and success of their children.

Chapter 1

Foreign Aid and Developing Countries' Creditworthiness

1.1 Introduction

The goal of this chapter is to explore whether aid affects developing countries' creditworthiness, as reflected by the *Institutional Investor*'s evaluation of credit risk. Our interest in this question is driven by the observation that credit ratings play an important role for countries' ability to borrow abroad: as various studies document, a lower rating – interpreted as a greater likelihood that borrowers will default on their debt – raises the yield that has to be offered to compensate lenders for higher credit risk (Cantor and Packer, 1996; Larrain et al., 1997; Eichengreen and Mody, 1998; Cunningham et al., 2001; Ciocchini et al., 2003). Moreover, a negative assessment by rating agencies may induce creditors to require higher collateral, which implicitly raises the costs of borrowing. Finally, legal constraints in several industrialized countries prevent potential lenders from investing in countries whose rating is below a critical threshold (Haque et al., 1996).[1] We start our analysis by developing a simple model of international borrowing and lending to analyze how aid affects agents' demand for foreign credit and the likelihood that they will repay their debt. In this framework, a transfer in a given period lowers the net benefits of future default and therefore *raises* creditworthiness as perceived by international investors. The empirical results that we present in the second part of the paper provide support for this hypothesis: using a set of annual data for a large number of developing countries in the 1980s and 1990s, we find that larger aid inflows result in an improvement of the recipient country's *Institutional Investor* rating. However, different types of aid differ in their effects: while grants and technical assistance significantly raise a country's creditworthiness, we do not find a significant effect if we focus on the loan component of total aid flows.

Our study fits into the – by now quite voluminous – empirical literature that analyzes the impact of foreign aid on growth, investment and capital flows.[2] More specifically, it is related to the recent work on IMF program effectiveness – in particular the study of Mody and Saravia (2003), who investigate the impact of IMF-supported programs on emerging market bond spreads.[3] There are two features that distinguish our study from these

[1] While we are taking the consequences of credit ratings as given, we are not trying to assess the success of rating agencies in accurately *predicting* default and currency crises. For a study that accomplishes this task see Reinhart (2002).

[2] See Hansen and Tarp (2000, 2001), Easterly (2003), Roodman (2004) as well as Harms and Lutz (2004) for recent surveys on the aid-growth literature, and Harms and Lutz (2006) for a study of the relationship between aid and private foreign investment.

[3] Recent analyses of the growth effects of IMF programs are provided by Przeworski and Vreeland (2000) and by Barro and Lee (2004). Ramcharan (2003) offers a survey of this literature. Bird and Rowlands (2002) consider the effects of IMF programs on international capital flows.

contributions: first, by considering a much broader set of aid types and aid sources, we are able to assess whether aid has an impact on creditworthiness even if it is not subject to IMF conditionality. Second, since few developing countries are completely neglected by foreign donors, we do not run into the sample selection problems that are prominent in the literature on program effectiveness (see Przeworski and Vreeland, 2000).

The specification of our empirical model is influenced by earlier studies on the determinants of country ratings (Lee, 1993; Haque et al., 1996 and 1998; Reinhart et al., 2003) and by the literature that analyzes emerging market bond spreads (see Cunningham et al., 2001, for a recent survey). However none of the investigations in this field considers the role of foreign aid. While it might be argued that aid only matters indirectly – by influencing the current stock of foreign debt or foreign reserves – we find that aid flows have explanatory power even if we simultaneously include debt and reserve levels.

The rest of this chapter is structured as follows: Section 1.2 presents a simple model of international borrowing and endogenous default risk. Section 1.3 describes our data set, empirical strategy, and results. Section 1.4 summarizes and concludes. Information on data definitions, sources and summary statistics are given in the data appendix.

1.2. A simple model of aid and default risk

Our goal is to develop a simple model which highlights how aid flows affect agents' consumption and investment decisions and thus the likelihood that they will repay their debt in the future.[4] We consider a small open economy that is populated by a continuum of identical agents whose total mass we normalize to one. The representative agent lives for two periods and maximizes

$$\mathrm{E}[U] = u(C_1) + \beta \mathrm{E}[C_2]. \tag{1}$$

In (1), C_t is consumption in period t, β is the agent's subjective discount factor, E is the expectations operator, and u is a continuous function with $u' > 0$ and $u'' < 0$.

The agent's first-period consumption is subject to the constraints

$$C_1 = Y_1 + A_1 + D_2 - I_1 - R_1 D_1, \tag{2}$$

[4] Our theoretical framework is related to Asiedu and Villamil (2002) who explicitly consider the effect of aid on the sustainability of international lending. In their model it is the fear of losing access to *future* aid flows that prevents countries from defaulting on their debt. By contrast, we focus on the *instantaneous* effect of aid on creditworthiness.

$$K_2 = I_1. \tag{3}$$

In (2), Y_1 is (exogenous) first-period income, A_1 is a grant ("foreign aid") received from abroad in period 1, $R_1 D_1$ is a payment on (exogenous) initial foreign debt (interest and principal), D_2 represents *new* debt accumulated during period 1, and I_1 denotes first-period investment, which determines the average capital stock in period 2. We assume that the agent does not default in period 1, and to save on notation we define $Y_1^{net} \equiv Y_1 - R_1 D_1$. We also assume that the sum of net income and aid, $Y_1^{net} + A_1$, is low enough to guarantee that D_2 is strictly positive. Accordingly, the representative agent is a "borrower" on international capital markets. In the second period, the borrower uses the capital stock to produce $Y_2 = \phi G(K_2)$, where $\phi > 0$, $G' > 0$, $G'' < 0$, $\lim_{K_2 \to 0} G'(K_2) = \infty$, and $\lim_{K_2 \to \infty} G'(K_2) = 0$.

Before consumption takes place at the end of period 2, the agent decides whether to pay back his debt or not. We exclude the possibility of a partial default. Hence he repays his *entire* debt or nothing at all. Due to the risk of default, international investors charge an interest rate $\rho \in (r, \infty)$ whose implicit premium above the risk-free international interest rate r depends on the (endogenous) likelihood of repayment. If the borrower defaults on his debt, he faces a punishment Π, which can be expressed as a pure loss in income, i.e. the income of defaulting borrowers is reduced without raising the income of lenders. We assume that the punishment has the following form:[5]

$$\Pi = s(1+\rho)\gamma K_2. \tag{4}$$

In (4), $s \in [0, \infty)$ is a random variable with distribution function $F(s)$ which is realized after production in period 2, and γ is a strictly positive constant. The assumption that Π is unknown in period 1 is meant to reflect the fact that the response of creditors to a default depends on a host of random political and economic factors, which cannot be perfectly anticipated. Moreover, we argue that richer economies are more vulnerable to debtor retaliation – e.g. because of their deeper integration with the world economy –, and we therefore make Π dependent on K_2. Finally, our assumption that the punishment in case of

[5]This specification is inspired by Eaton et al. (1986) as well as Aizenman (1989a, 1989b). It gets some empirical support from Rose (2005) who demonstrates that, in the past, defaulting countries suffered substantial declines in their international trade.

default is proportional to the gross interest rate (including the risk premium) guarantees the existence and uniqueness of an equilibrium.[6]

It follows from (4) that the borrower strictly prefers to default on his debt in the second period if $Y_2 - (1+\rho)D_2 < Y_2 - s(1+\rho)\gamma K_2$. Hence, default takes place if $s < D_2/\gamma K_2$: a high level of debt relative to the onus of punishment makes it unattractive to honor one's payment obligations. Using this result and defining $\omega_2 \equiv D_2/\gamma K_2$, we can rewrite the borrower's expected utility as

$$\mathrm{E}[U] = u(Y_1^{net} + A_1 + D_2 - K_2) + \beta\int_0^{\omega_2} [\phi G(K_2) - s(1+\rho)\gamma K_2]dF(s)$$
$$+\beta\int_{\omega_2}^{\infty} [\phi G(K_2) - (1+\rho)D_2]dF(s). \tag{5}$$

When choosing the optimal values of D_2 and K_2 in period 1, the individual borrower takes into account how his decision affects the likelihood of future default. Straightforward maximization of (5) with respect to D_2 and K_2 therefore yields the following first-order conditions:

$$u'(Y_1^{net} + A_1 + D_2 - K_2) = \beta(1+\rho)[1 - F(\omega_2)] \tag{6}$$

$$u'(Y_1^{net} + A_1 + D_2 - K_2) = \beta[\phi G'(K_2) - (1+\rho)\gamma\int_0^{\omega_2} sdF(s)]. \tag{7}$$

The LHS in (6) reflects the marginal utility of additional debt in period 1, while the RHS gives the marginal cost of borrowing, adjusted for the likelihood of future default, which is $F(\omega_2)$. Expression (7) equates the marginal disutility of saving in period 1 with the expected marginal return on investment, which is lower than the marginal product of capital since the agent anticipates the possibility of default and the associated costs.

To close the model, we consider the supply side of the international capital market. We assume that loans are provided by risk-neutral foreign investors who are aware of domestic agents' incentives to repay their debt, and who are willing to supply credit as long as the yield compensates them for the possibility of default:

$$(1+\rho)[1 - F(\omega_2)] = 1 + r, \tag{8}$$

[6]Without this assumption, the incentive to deny repayment would depend on the risk premium which, in turn, depends on the likelihood of default. As a consequence, there might be a unique equilibrium, multiple equilibria or no equilibrium at all.

where r represents the risk-free interest rate offered by international capital markets. Combining (6) and (8) yields

$$\omega_2 = \frac{\Omega - A_1}{\gamma K_2} + \frac{1}{\gamma}, \tag{9}$$

where $\Omega \equiv u'^{-1}(\beta'(1+r)) - Y_1^{net}$. Using (6) – (9) we get

$$(1+r) = \phi G'(K_2) - \gamma[1+\rho(K_2, A_1)]\int_0^{\omega_2(K_2,A_1)} sdF(s), \tag{10}$$

where it follows from (8) and (9) that ω_2 and ρ are decreasing in K_2 and A_1. The RHS in (10) can be written as a function $\Lambda(K_2, A_1)$. Due to the properties of the production function and our assumption that ρ is finite, we have $\lim_{K_2 \to 0} \Lambda = \infty$. Moreover, we know that $\lim_{K_2 \to \infty} \Lambda = 0$. We can thus draw Λ as a function of K_2 as in Figure 1.1. Optimal first-period investment is given by the point of intersection of this curve with a horizontal line at $(1+r)$.[7] For a given level of K_2, an increase in A_1 raises Λ, shifting the curve upward and resulting in a higher equilibrium value of K_2. It follows from (9) that this reduces the likelihood of default.

The economic intuition behind these results is straightforward: for a given second-period capital stock, raising aid in period 1 reduces the amount agents wish to borrow relative to their investment. By lowering ω_2, this reduces the likelihood of future default. Since default is associated with costs that are proportional to the capital stock, a higher likelihood of repayment raises the expected return on investment, thus increasing the optimal level of the second-period capital stock. Note that higher aid may (but need not) result in *both* higher investment *and* in higher second-period debt. If A_1 raises the equilibrium value of D_2 while reducing the likelihood of default, aid acts as a catalyst for private capital flows while improving recipient countries' creditworthiness.

Our model has been designed to highlight a particular channel through which aid affects creditworthiness – namely, by lowering the expected net benefits of future default. We are aware that we have neglected several important aspects: first, while we have focused on the impact of aid on countries' *willingness to pay*, a default may also be triggered by a

[7]Note that we cannot be sure that Λ is monotonically decreasing, as drawn in Figure 1.1, and there may be multiple values of K_2 satisfying (10). However, since expected lifetime utility of borrowers is increasing in K_2, we will focus on the equilibrium that entails the highest second-period capital stock.

low *ability to pay*: due to exogenous shocks, countries may fail to honor their foreign debt even if the costs of default outweigh the benefits. We could have accounted for this aspect by assuming that second-period income is random, thus allowing for the possibility that available resources fail to cover repayment obligations. Without spelling out this extension, we believe that it would not change our key result: aid would still raise creditworthiness, both by reducing future debt and by expanding future production possibilities. Moreover, we have not considered the potential role of aid as a *signal* to foreign investors: on the one hand, aid may raise creditworthiness by indicating that a countries' economic policies are approved by international donors. On the other hand, large aid flows may be a sign of financial trouble and may thus be associated with lower credit ratings. While these effects are beyond the scope of our model, they should be taken into account when we interpret the empirical findings presented in the following section.

1.3 Aid and country creditworthiness: An empirical exploration

1.3.1 Data

1.3.1.1 Country creditworthiness

Our aim is to test whether foreign aid actually has a positive effect on countries' creditworthiness, as measured by the country credit ratings published in the *Institutional Investor* (in what follows, we will use the abbreviation IICCR).[8] The use of the IICCR allows us to consider a much broader set of countries than related studies on the determinants of emerging market spreads. As documented by Gelos et al. (2003), many low-income countries do not have access to international bond markets, but it would be wrong to conclude that perceived creditworthiness is irrelevant in these cases: the likelihood of default may still affect the availability of bank loans, trade credit etc.

The IICCR ranks countries on a scale from 0 to 100, with a lower rating reflecting a higher likelihood that borrowers in this country will default on their debt. The ratings are "...based on information provided by senior economists and sovereign risk analysts at leading global banks and money management and securities firms" (Institutional Investor,

[8]While Haque et al. (1996) consider the indexes published by *Euromoney* and the *Economist Intelligence Unit* as alternative measures of creditworthiness, they observe that there is a "substantial degree of cross-sectional agreement among the ratings" (Haque et al. 1996:699). We therefore use the IICCR as a "representative" proxy for international lenders' assessment of default risk.

2002:170).[9] The scores have been published regularly since 1979, and the number of countries covered has increased from 96 in 1980 to 145 in 2000. When we started to assemble our data set, availability of the IICCR was a prerequisite for accepting a country in the sample.[10]

The IICCR is published every six months (in the March and September issues of the *Institutional Investor*), while most regressors are only available on an annual basis. We decided to transform the original time series into annual data by computing the (unweighted) average of the March and September scores. However, our results are not driven by this choice: although the IICCR of a given country may vary between March and September, the estimated coefficients and significance levels did not change by much when we used only March (or September) values instead of averages.

Finally, the fact that the IICCR is bounded from below and above suggests to transform the data. The transformation we chose follows Haque et al. (1996) as well as most of the other predecessor studies:

$$IICT_{it} = 100 \cdot ln\left(\frac{IICCR_{it}}{100 - IICCR_{it}}\right). \tag{11}$$

However, this logistic transformation does not drive our qualitative results, and our main conclusions still hold if we use the untransformed IICCR.

1.3.1.2 Aid

The aid variable used in our analysis is provided by the OECD's Development Assistance Committee (DAC) data base (OECD 2005), and is referred to as "official development assistance and net official aid" (henceforth ODA). It consists of grants and of loans with a grant element of at least 25 percent; deducted from this are repayments of loan principal.[11] We control for country size by dividing through total population, and nominal aid flows are transformed into constant international dollars by using the World Bank's (2005) PPP-conversion factors and the US GDP deflator. Since a log-linear specification turned out to

[9] As reported by Haque et al. (1996), the individual criteria used by banks to assess default risk are not specified. Hence, we have no information on whether observed aid flows directly enter the ratings.

[10] The other criteria were that a country was classified as a middle-income or low-income country in 2000, and that its population exceeded one million in the year 2000.

[11] Chang et al. (1998) have created an alternative measure – *effective development assistance* (EDA) – which only includes the grant component of concessionary loans. Unfortunately, the Chang et al. (1998) data are only available through 1995. In order to make use of a larger sample, we decided to stick to the original ODA series. However, since the evolution of EDA closely follows the time path of official development assistance, we do not expect this to be crucial for our results.

best fit the data, we decided to transform the original aid-per-capita data into natural logarithms.[12]

Later on we will replace ODA per capita by less aggregate variables, namely the loan component of aid, pure grants, and technical assistance. We will also differentiate between aid offered by multilateral donors and "bilateral" aid received from individual countries.

1.3.1.3 Control variables

Apart from establishing a positive effect of aid on $IICT$, our model suggests that countries' creditworthiness is affected by the following variables: current income (Y_1), the initial stock of debt (D_1), variables that influence future productivity (ϕ), and variables that reflect countries' vulnerability to creditor sanctions (γ). To account for the influence of Y_1 and D_1, we include the logarithm of countries' gross national income (*GNI*) and total external debt (*DEBT*), both in per-capita terms. Again, the original data are transformed into constant international dollars by using PPP-conversion factors and the US GDP-deflator. Based on our theoretical analysis, we expect *GNI* to raise *IICT*, while *DEBT* should have a negative effect.

The positive influence of growth prospects on creditworthiness that is suggested by our model is captured by the following three variables: the growth rate of real per-capita income (*GROWTH*), the inflation rate (*INFLATION*) as a proxy for macroeconomic stability, and a measure of "governance" (*GOV*), which reflects the absence of corruption, the quality of the bureaucracy, and the rule of law.[13] Our decision to control for the quality of governance is motivated by the recent literature on aid, growth, and capital flows, which emphasizes the role of countries' economic and institutional environment. It is also suggested by Ciocchini et al. (2003) who find that higher corruption raises countries' interest rate spreads. Note, however, that the fact that *Political Risk Services* started to publish its index in 1982 and introduced a new scaling for their governance variables in 1998 limits our sample to this time interval.

[12]The loss of data due to negative ODA values is negligible (10 observations). We also estimated an equation with the log of aid and the log of population as separate regressors. An F-test supported the restriction associated with using the log of per-capita aid.

[13]Each of these features is captured by an index that is published in the *International Country Risk Guide* and assembled in *Political Risk Services'* IRIS3 database. The measures range from 0 to 6, with a higher value reflecting a better business climate, and the composite measure we use is an unweighted average of the three indexes. Both the inflation rate and the governance index are transformed into natural logarithms.

To control for countries' vulnerability to creditor sanctions (γ) and the resulting incentive to refrain from default, we use a measure of countries' trade openness (*TRADE*), which is the sum of exports and imports divided by GDP. Finally, we include the ratio of reserves over imports (*RESERVES*) to account for the fact that a lot of default episodes were triggered by balance-of-payments crises (Kaminsky and Reinhart, 1999) and that countries' ability to defend their pegs depends on their stock of foreign reserves. While this regressor is not derived from our model, it figures prominently in related studies (Haque et al., 1996), and omitting it would raise the risk of producing biased estimates.

1.3.1.4 Lagged dependent variable

In addition to the variables mentioned above, we use the lagged value of IICT as a regressor. Such a dynamic specification is suggested by Haque et al. (1996, p.718) who find that "there is considerable persistence in the ratings, so that a country tends to retain its rating over time unless significant adverse or positive developments occur".

1.3.2. Estimation

1.3.2.1 Specification

The equation we estimate is

$$IICT_{it} = \delta IICT_{i(t-1)} + \beta a_{i(t-1)} + \sum_{k=1}^{K} \gamma_k x_{k,i(t-1)} + \alpha_i + \xi_t + \varepsilon_{it}, \qquad (12)$$

In (12), α_i is an unobserved ("fixed" or "individual") effect that may be arbitrarily correlated with the other regressors. ξ_t is a time dummy which is meant to capture variations in industrialized countries' interest rates, but also general changes in investor sentiment. The variable $a_{i(t-1)}$ is the logarithm of per-capita aid received by country i in period $t-1$, while $x_{k,i(t-1)}$ is the control variable k for country i in period $t-1$. Finally, ε_{it} is the usual error term. The t-statistics presented below are based on a robust covariance matrix that allows for heteroskedastic disturbances. The fixed effects capture all country-specific, but time-invariant features. Given the considerable heterogeneity of our sample with respect to countries' political institutions, cultural background, and geographical conditions, their inclusion is particularly important to reduce the extent of omitted variable

bias.[14] By using lagged values of the regressors we are trying to catch two birds with one stone: first, it is likely that the IICCR value for a given country in period t is formed on the basis of economic circumstances in period $t-1$, especially since 50 percent of the assessment is published in the month of March. Second, using lagged values is a simple strategy to reduce endogeneity bias.[15]

1.3.2.2 GMM estimation

It is well-known that estimating equation (12) by OLS leads to biased coefficients.[16] We therefore follow the procedure suggested by Arellano and Bond (1991): the first step is to eliminate the country-specific effects by taking differences on both sides of equation (12). This yields

$$\Delta IICT_{it} = \delta \Delta IICT_{i(t-1)} + \beta \Delta a_{i(t-1)} + \sum_{k=1}^{K} \gamma_k \Delta x_{k,i(t-1)} + \Delta \xi_t + \Delta \varepsilon_{it}, \qquad (13)$$

where $\Delta IICT_{it} \equiv IICT_{it} - IICT_{i(t-1)}$. The second step is to estimate (13) by GMM. Arrelano and Bond (1991) demonstrate that, by using lagged levels of both the endogenous variable and of the regressors as instruments, one arrives at a set of moment conditions which allow to efficiently estimate the model's parameters. These estimates are consistent if the error term ε_{it} is serially uncorrelated – an assumption that can be checked by testing the hypothesis of no second-order serial correlation in the first-differenced residuals (Arellano and Bond, 1991).

While the validity of the overidentifying restrictions used in GMM estimation can be assessed by considering Hansen's J-statistic (Hansen, 1982), past levels of the right-hand variables are weak instruments for current differences if the time series involved are highly persistent. The solution suggested by Arellano and Bover (1995) as well as Blundell and Bond (1998) is to add further moment conditions by simultaneously estimating equations (12) and (13). As stated above, the *IICT* series are likely to be very persistent. We therefore decided to use the "Systems-GMM" estimator of Blundell and Bond (1998). Most of our results are based on an efficient two-step GMM estimator that applies an endogenous

[14] An F-test that compares a pooled regression with the fixed-effects specification strongly supports our inclusion of country-specific dummies.
[15] Our results did not change by much when we experimented with other specifications, e.g. the September value of the IICCR and contemporaneous values of the regressors.
[16] The bias disappears in panels with infinitely long time series (Nickell, 1981). Bond (2002) and Wooldridge (2002) offer excellent surveys of the problems associated with dynamic panel data estimation and of the available approaches to arrive at consistent estimates.

weighting matrix to the moment conditions, and standard errors are computed by using Windmeijer's (2005) finite sample correction. Since exploiting all available lags as instruments results in a proliferation of moment conditions, and since the finite-sample bias of the GMM estimator is exacerbated if the number of instruments exceeds the number of countries (Judson and Owen, 1999), our main results are derived by using a restricted set of instruments – usually the first to fourth lags of the regressors. Moreover, we follow the strategy of Roodman (2004) who further reduces the size of instrument matrix by summing up individual moment conditions.[17]. While we will stick to this specification for most of the paper, we will also report the consequences of using alternative approaches.

1.3.2.3 Results

Column 1 of Table 1.1 presents the results of estimating (13). Most importantly, aid has a positive effect, and the coefficient is significant at the one-percent level.[18] Moreover, most control variables have the expected sign, although not all of them are significant. The p-value for the Arellano-Bond (m2) test statistic supports the hypothesis that the instruments used are exogenous and that we can confidently reject the hypothesis that the disturbances are serially correlated. Finally, the results confirm the observation of Haque et al. (1996) that credit ratings are very persistent. For countries that are stigmatized by a bad rating this implies that investors are slow at changing their assessment even when the country is hit by a positive aid shock. Combining the coefficients of aid and of the lagged dependent variable, one finds that, in the long run, a permanent one-percent increase of aid per capita raises creditworthiness by approximately 0.25 percent.

While the two-step estimator used in column (1) is superior to the one-step estimator in terms of efficiency, those efficiency gains may be rather modest, and the one-step alternative which uses an exogenous weighting matrix may be more reliable in small samples. In column (2) of Table 1.1 we therefore present results which demonstrate that using the one-step estimator delivers the same qualitative results as the two-step alternative. To make sure that our result is not due to a few influential observations, we proceeded by applying Hadi's (1994) procedure to identify multivariate outliers and excluded those

[17]This implies that, e.g., the two moment conditions $E(x_{i,t-2}\Delta\varepsilon_{i,t}) = 0$ and $E(x_{i,t-3}\Delta\varepsilon_{i,t}) = 0$, merge into $E(x_{i,t-2}\Delta\varepsilon_{i,t} + x_{i,t-3}\Delta\varepsilon_{i,t}) = 0$. Reducing the number of instruments obviously comes at the cost of lower efficiency, but in small samples it reduces the risk of overfitting the model. The "collapse" option is part of the *xtabond2* Stata routine written by David Roodman.

[18]Evaluated at the mean, the estimated coefficient of 5.94 implies a short-run elasticity of 0.04.

observations from the sample. The results in column (3) of Table 1.1 suggest that the effect of aid and of most other regressors becomes stronger if we omit outliers.

It is an open issue whether GMM estimation really improves upon the fixed-effects estimator in dynamic panel-data models with small samples.[19] An alternative approach is to estimate (12) by OLS and to apply the bias-correction suggested by Kiviet (1995) and Bruno (2005). As indicated by column (4) of Table 1.1, using this "corrected LSDV (LSDVC)" estimator does not change our main finding that aid raises creditworthiness.[20]

1.3.3 Robustness checks

In this subsection, we will report the results of replacing total aid per capita in equation (13) by different types of aid, of running this regression for various country groups and time periods, and of experimenting with non-linear specifications. Apart from testing the robustness of our previous findings, these variations provide important insights on the channels through which aid affects country creditworthiness.

Table 1.2 differentiates between various types of aid: column 1 considers only pure grants, while columns 2 and 3 consider technical assistance and loans, respectively. While grants and technical assistance have a stronger impact on creditworthiness than total aid, the coefficient for loans is much smaller and not significantly different from zero. This seems intuitive: both grants and technical assistance correspond to the type of transfer modelled in Section 2, with technical assistance being more likely to be used productively and to raise future income.[21] Conversely, loans which raise the future debt burden seem to be less suited to improve a country's standing vis-a-vis international capital markets, even in the short run.[22] Columns 4 and 5 of Table 1.2 show that bilateral aid has a much stronger impact on creditworthiness than multilateral aid. We conjecture that this difference reflects the fact that a large part of multilateral aid consists of loans while the dominant share of bilateral aid

[19]Judson and Owen (1999) demonstrate that the bias of the fixed-effects estimator depends on the length of the time series relative to the cross-sectional dimension.

[20]To compute these results, we used the *xtlsdvc* routine developed for Stata by Giovanni Bruno.

[21]Our result is also in line with Asiedu and Villamil (2002) whose model implies that the threat of withholding productivity-enhancing aid is especially powerful in preventing default. A more skeptical view is expressed by Roodman (2004:6) who notes that technical assistance "...funds not so much recipient governments as consultants." Note, however, that Roodman's statement does not exclude the possibility that the advise offered by those consultants raises productivity.

[22]For those countries where loan repayments exceed new disbursements, net loans per capita are negative and the logarithm is not defined. We checked whether the resulting reduction of the sample was driving our results by omitting those observations from the benchmark regression in Table 1. It turned out that, in this smaller sample, total aid still had a significantly positive effect on creditworthiness, which suggests that the result in column (3) of Table 2 is not driven by the modified sample.

comes in the form of pure grants (OECD, 2005). Hence, the findings in columns (4) and (5) of Table 1.2 seem to replicate the result that grants have a strong impact on creditworthiness while loans have none.

In some countries, net aid flows – i.e. new disbursements minus repayments of loans – are substantially reduced by interest payments on existing debt. To check the possibility that the impact of aid changes if interest payments are netted out, we subtracted those payments from both total aid and from the grant component of aid. The results are reported in columns (1) and (2) of Table 1.3. Apparently, our main finding is not affected by this modification: neither the coefficient of total aid nor its significance changes dramatically. The impact of grants shrinks, but it is still significantly positive and slightly bigger than the effect of total aid. Note, however, that the number of observations is reduced by almost 10 percent if we subtract interest payments. Since we want to preserve the biggest possible sample and since netting out interest payments does not affect our main findings, we decided to move on using net aid without the adjustment for interest payments. Columns (3) and (4) of Table 1.3 investigate the possibility that the effect of aid on creditworthiness merely reflects the consequences of one-time debt-writeoffs, e.g. in the wake of the Brady deals of the late eighties and early nineties. We therefore subtracted "debt forgiveness grants" as reported by the OECD (2005) from total aid and grants, respectively. The results demonstrate that this variation has almost no influence on the size of the estimated coefficients. This suggests that the impact of aid on creditworthiness goes beyond the effect of debt forgiveness.[23]

Table 1.4 considers various subsets of our original sample. We started by removing individual regions from our sample (columns 1 to 3):[24] If we remove the Latin American countries, the coefficient of aid drops somewhat, but it is still significantly different from zero. Omitting Sub-Saharan African or Asian countries (columns (4) and (5)) leads to similar results: while the coefficient of aid and of some control variables may change, our general finding that aid raises creditworthiness does not seem to be driven by any particular region.

A strict reading of our theoretical model suggests that it should only apply to those countries who have access to international capital markets. Data on market access are provided by Gelos et al. (2003), and column (4) of Table 1.4 shows the consequences of

[23]When we tested the importance of debt forgiveness for our results by removing countries involved in Brady deals, we found that this did not alter our findings.
[24]Removing regions instead of considering them in isolation helps to keep the subsample at a reasonable size.

omitting countries who never issued international bonds. Apparently, this does not alter our main findings.

Finally, we checked whether our results were driven by the rather volatile assessment of transition countries' creditworthiness in the early 1990s, and removed those countries from the sample. Column (5) of Table 1.4 shows that our key result is not affected by this sample modification. It is notable, however, that the coefficient of the lagged dependent variable rises substantially, which indicates that transition countries were more likely to see their creditworthiness reassessed than other countries.

Columns (1) and (2) of Table 1.5 report the results from running the regression for observations before and after 1990. While this break point is somewhat arbitrary, it is likely that aid disbursement criteria and thus the impact of aid changed after the end of the cold war. The numbers indicate that there are, indeed, substantial differences between the two decades: while the coefficients and t-statistics suggest a significantly positive effect in both periods, aid had a much stronger impact on credit ratings during the 1980s than during the 1990s. This may be due to the fact that transition countries' credit ratings were particularly volatile in the early 1990s, and that most of the outliers identified above fall into this period. In fact, if we omit transition countries and outliers the significance of aid during the 1990s improves substantially (see column (3) of Table 1.5). We also investigated the proposition brought forward by Hansen and Tarp (2000) (among others) that there are diminishing returns to aid, and used the squared value of aid as an additional regressor. The numbers in column (4) of Table 1.5 do not support this notion: the coefficient of aid squared is positive, but not significantly different from zero. Finally, we checked whether the effect of aid on creditworthiness depends on the institutional environment and therefore included an interactive term – the logarithm of aid per capita times our governance variable *GOV* – as an additional regressor. As column (5) of Table 1.5 demonstrates, the notion that "money matters – in a good policy environment" (World Bank, 1998:28) is not supported in our context: the coefficient of the interactive term is negative, but insignificant.

1.4 Summary and conclusions

When we started this investigation, we were curious whether aid could possibly raise developing countries' creditworthiness and thus act as a "catalyst" for private capital flows. In this respect, our results are both encouraging and disheartening: aid has a significantly positive effect on the *Institutional Investor*'s index of country credit risk, but the size of this

effect is rather modest. Moreover, credit ratings are extremely persistent, such that a temporary increase in aid flows is unlikely to improve the ratings of countries whose economic and institutional weaknesses taint their standing vis-a-vis international capital markets.

Our results also shed light on the channels through which aid may improve creditworthiness: technical cooperation and grants seem to be more effective than loans, suggesting that aid improves a country's reputation when it lowers future repayment obligations relative to future income and thus reduces the potential benefits from default. This conjecture is also supported by the observation that bilateral aid has a stronger impact on the *Institutional Investor*'s ratings than multilateral aid. On a more general level, our results thus emphasize the importance to disentangle the different components of aid when assessing the effect of aid on macroeconomic variables. While this chapter has limited its attention to the relationship between aid flows and creditworthiness, we are quite sure that this insight generalizes to other parts of the aid-effectiveness debate.

1.5 Tables and Figures

Table 1.1: Benchmark specification: Alternative estimation methods
(Dependent variable: Transformed index of country credit risk)

	(1) SYS-GMM two-step lags 1 to 4	(2) SYS-GMM one-step lags 1 to 4	(3) SYS-GMM two-step no outliers	(4) LSDVC
AID	5.939***	6.269***	7.278***	4.066***
	(3.77)	(2.89)	(3.20)	(4.26)
GNI	21.746**	11.564	20.552**	-0.250
	(2.29)	(1.33)	(2.01)	(0.09)
DEBT	-23.084***	-27.306***	-25.594***	-11.315***
	(2.75)	(3.36)	(4.03)	(6.66)
GROWTH	0.664***	0.695***	0.622***	0.706***
	(3.59)	(4.51)	(3.60)	(7.15)
INFLATION	-1.560	-2.358*	-1.527	-1.039*
	(1.52)	(1.96)	(1.55)	(1.86)
GOV	14.173**	9.335*	14.997**	3.560*
	(2.04)	(1.67)	(2.47)	(1.84)
TRADE	0.389**	0.393**	0.340*	0.219***
	(2.10)	(2.29)	(1.98)	(4.35)
RESERVES	0.281**	0.360***	0.323***	0.217***
	(2.41)	(2.76)	(2.65)	(6.74)
Lagged IICT	0.832***	0.918***	0.865***	0.914***
	(13.89)	(12.48)	(12.18)	(45.94)
Observations	837	837	814	837
Countries	70	70	68	70
Instruments	61	61	61	
J -statistic (p value)	0.30	0.43	0.25	
Arellano-Bond stat. (p value)	0.38	0.43	0.58	

Notes: In parentheses: Absolute values of t-statistics, based on a robust covariance-matrix. ***, **, *: significance levels of 1, 5, 10 percent. All regressions include time dummies. Column (1): Two-step Systems-GMM estimator applying Windmeijer's (2005) finite-sample correction to compute standard errors. Lags 1 to 4 of regressors used as instruments. Reduction of moment conditions by using the "collapse" option suggested by Roodman (2004). Column (2): One-step Systems-GMM estimator. Lags 1 to 4 of regressors used as instruments. Column (4): The corrected fixed effects (LSDVC) estimator suggested by Kiviet (1995) and Bruno (2005).

Table 1.2: Different types of aid
(Dependent variable: Transformed index of country credit risk)

	(1)	(2)	(3)	(4)	(5)
	Grants	Technical A.	Loans	Bilateral	Multilateral
AID	8.657***	11.314***	0.575	6.136***	3.108**
	(4.44)	(3.53)	(0.53)	(2.99)	(2.45)
GNI	13.095	12.529	30.986***	21.546**	24.227**
	(1.15)	(1.04)	(3.05)	(2.16)	(2.11)
DEBT	-20.507**	-18.757**	-38.691***	-27.123***	-27.654***
	(2.25)	(2.13)	(5.22)	(3.88)	(3.14)
GROWTH	0.735***	0.649***	0.530***	0.613***	0.638***
	(4.04)	(3.77)	(2.80)	(3.59)	(3.27)
INFLATION	-2.260**	-2.137**	-1.521	-1.589	-1.913
	(1.90)	(2.48)	(0.95)	(1.33)	(1.34)
GOV	13.025*	14.858**	5.586	12.583*	15.036*
	(1.78)	(2.00)	(0.57)	(1.68)	(1.91)
TRADE	0.298	0.288	0.637***	0.411**	0.364**
	(1.64)	(1.57)	(4.41)	(2.59)	(2.03)
RESERVES	0.240*	0.233*	0.222	0.253**	0.345**
	(1.85)	(1.96)	(1.31)	(2.18)	(2.29)
Lagged IICT	0.883***	0.810***	0.777***	0.827***	0.843***
	(14.58)	(15.58)	(9.36)	(12.76)	(14.07)
Observations	847	846	702	829	812
Countries	70	70	69	70	70
Instruments	61	61	61	61	61
J -statistic (p value)	0.29	0.61	0.34	0.27	0.23
Arellano-Bond (p value)	0.29	0.29	0.57	0.31	0.67

Notes: In parentheses: Absolute values of t-statistics, based on a robust covariance-matrix. ***, **, *: significance levels of 1, 5 and 10 percent. All regressions include time dummies. The estimator used is the Blundell and Bond (1998) two-step Systems-GMM estimator applying Windmeijer's (2005) finite-sample correction to compute standard errors. Lags 1 to 4 of regressors used as instruments. Reduction of moment conditions by using the "collapse" option suggested by Roodman (2004).

Table 1.3: Netting out interest payments and debt forgiveness grants
(Dependent variable: Transformed index of country credit risk)

	(1)	(2)	(3)	(4)
	Aid w/o int.paym.	Grants w/o int.paym.	Aid w/o debt forg.	Grants w/o debt forg.
AID	4.870**	4.911***	5.480***	8.445***
	(2.49)	(2.88)	(3.16)	(4.35)
GNI	16.320**	14.937**	23.847**	15.907
	(2.09)	(2.00)	(2.41)	(1.53)
DEBT	-19.399**	-19.679***	-26.417***	-21.841***
	(2.44)	(2.82)	(3.11)	(-2.73)
GROWTH	0.583***	0.583***	0.641***	0.680***
	(3.26)	(3.80)	(3.30)	(3.93)
INFLATION	-2.165*	-2.226*	-1.707	-2.227*
	(1.71)	(1.83)	(1.54)	(1.94)
GOV	11.094	10.876	14.401*	13.539*
	(1.56)	(1.42)	(1.81)	(1.86)
TRADE	0.277	0.336**	0.448***	0.334*
	(1.68)	(2.14)	(2.39)	(1.89)
RESERVES	0.249	0.329**	0.312**	0.242*
	(1.42)	(2.33)	(2.56)	(1.94)
Lagged IICT	0.909***	0.892***	0.840***	0.864***
	(12.16)	(14.60)	(14.08)	(14.47)
Observations	782	797	832	847
Countries	64	65	70	70
Instruments	61	61	61	61
J -statistic (p value)	0.19	0.35	0.19	0.32
Arellano-Bond (p value)	0.31	0.42	0.43	0.20

Notes: See Table 1.2.

Table 1.4: Different country groups
(Dependent variable: Transformed index of country credit risk)

	(1)	(2)	(3)	(4)	(5)
	no Lat.Am.	no Subs. Afr.	no Asia	Cap.M. access	No transition.
AID	3.688**	5.903***	4.681**	5.685***	6.491***
	(2.36)	(3.34)	(2.43)	(2.92)	(2.87)
GNI	14.026	5.169	16.766*	15.891*	14.53*
	(1.59)	(0.50)	(1.99)	(1.79)	(1.99)
DEBT	-12.908*	-22.395**	-25.676***	-32.821***	-27.995***
	(1.89)	(2.56)	(3.31)	(3.89)	(3.49)
GROWTH	0.505***	0.999***	0.5448***	0.848***	0.6310***
	(2.91)	(5.36)	(3.63)	(5.41)	(4.23)
INFLATION	-2.298**	-2.666**	-2.28*	-0.47	-2.014*
	(2.11)	(2.39)	(1.81)	(0.42)	(1.73)
GOV	18.604***	12.489**	1.100	10.328	7.415
	(2.66)	(2.09)	(0.16)	(1.61)	(1.23)
TRADE	0.076	0.232	0.389**	0.475***	0.397**
	(0.90)	(1.21)	(2.00)	(2.82)	(2.33)
RESERVES	0.244	0.388***	0.397***	0.338***	0.360**
	(2.12)	(3.45)	(2.64)	(3.12)	(2.43)
Lagged IICT	0.877***	0.908***	0.901***	0.844***	0.930***
	(14.16)	(13.29)	(11.61)	(11.52)	(12.00)
Observations	518	602	690	733	804
Countries	49	46	59	60	65
Instruments	61	61	61	61	61
J -statistic (p value)	0.65	0.65	0.45	0.44	0.58
Arellano-Bond (p value)	0.25	0.95	0.40	0.36	0.25

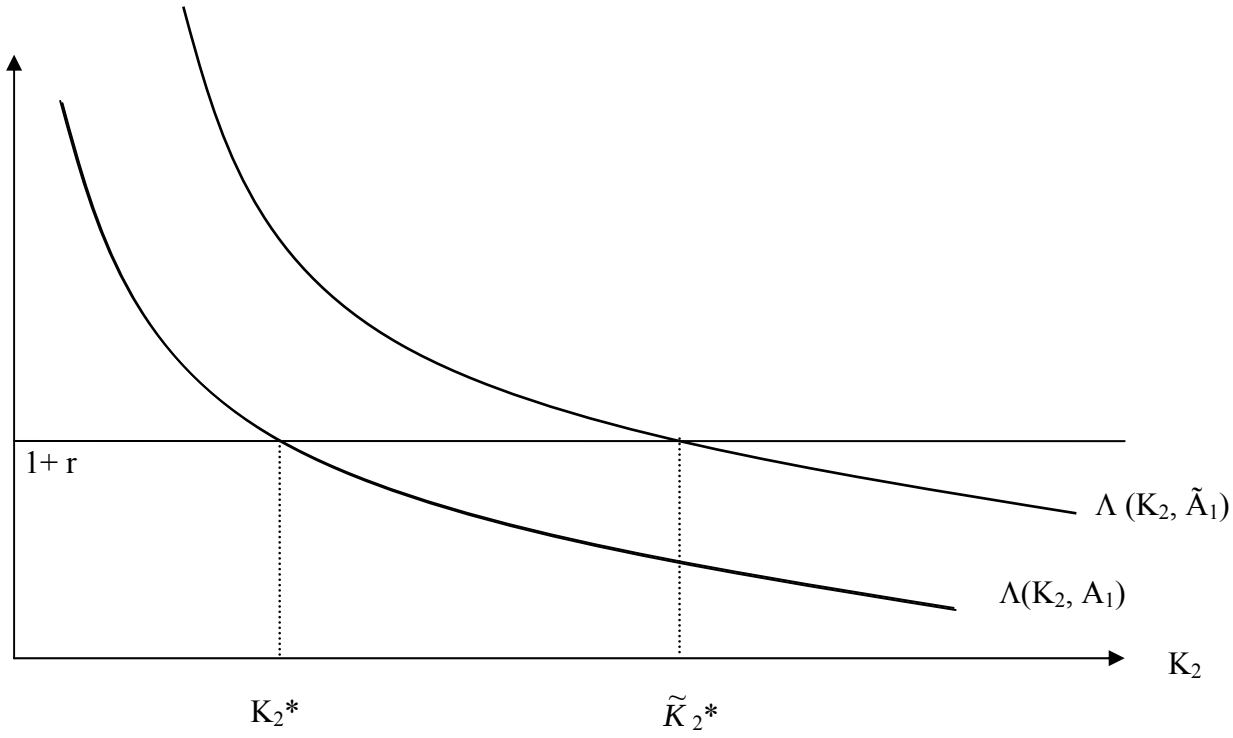
Notes: In parentheses: Absolute values of t-statistics, based on a robust covariance-matrix. ***, **, *: significance levels of 1, 5, 10 percent. All regressions include time dummies. The estimator used is the Blundell an Bond (1998) one-step Systems-GMM estimator. Lags 1 to 4 of regressors used as instruments. Reduction of moment conditions by using the "collapse" option suggested by Roodman (2004).

Table 1.5: Structural breaks and nonlinear effects
(Dependent variable: Transformed index of country credit risk)

	(1)	(2)	(3)	(4)	(5)
	Through 1990	After 1990 Full sample	After 1990 no outliers no trans. count.	AID sq.	Aid * GOV
AID	4.853**	1.760*	2.171***	6.194***	8.582**
	(2.25)	(1.97)	(2.87)	(5.53)	(2.48)
AID squared				0.068	
				(0.31)	
AID * GOV					-1.015
					(0.84)
GNI	23.315*	11.765*	13.111***	16.734*	19.842*
	(1.89)	(1.85)	(3.17)	(1.80)	(1.92)
DEBT	-26.153**	-6.167	-8.743**	-19,779**	-19.732**
	(2.48)	(0.88)	(2.11)	(2.31)	(2.46)
GROWTH	0.684**	0.673***	0.530***	0.751***	0.681***
	(2.34)	(3.22)	(3.95)	(4.07)	(3.49)
INFLATION	-0.834	-1.578	0.172	-1.868*	-1.545
	(0.50)	(1.45)	(0.21)	(2.08)	(1.58)
GOV	16.888	-0.766	0.201	14.825*	23.338**
	(1.26)	(0.17)	(0.00)	(2.07)	(2.37)
TRADE	0.746***	0.073	0.125*	0.299*	0.304**
	(3.12)	(0.86)	(1.67)	(1.86)	(2.00)
RESERVES	0.233	0.197***	0.303***	0.201*	0.214
	(1.46)	(2.65)	(4.14)	(1.99)	(1.64)
Lagged IICT	0.717***	0.803***	0.818***	0.837***	0.815***
	(9.76)	(13.30)	(19.26)	(13.41)	(12.00)
Observations	365	472	424	837	837
Countries	54	68	62	70	70
Instruments	53	63	62	66	66
J -statistic (p value)	0.37	0.38	0.71	0.21	0.27
A.-B. (p value)	0.27	0.56	0.90	0.32	0.30

Notes: See Table 1.2.

Figure 1.1: The effect of raising A_1 on the equilibrium value of K_2 $\left(\tilde{A}_1 > A_1\right)$



1.6 Data appendix

Definitions and sources

Institutional Investor Country Credit Rating (IICCR): Country Credit Ratings published in the Institutional Investor magazine every March and September since 1979. Source: Institutional Investor magazine, various issues.

Aid: Logarithm of official development assistance and net official aid per capita in constant international dollars. (To transform the flow of aid denoted in current US dollars into constant international dollars, we used the World Bank's PPP-conversion factors and the US GDP deflator.) Source: OECD (2005)

Grants: Logarithm of grants per capita. Grants are transfers in cash or in kind for which no legal debt is incurred by the recipient in constant international dollars. Source: OECD (2005).

Technical cooperation: Logarithm of technical cooperation per capita. Technical co-operation is the provision of know-how in the form of personnel, training, research and associated costs in constant international dollars. Source: OECD (2005).

Loans: Logarithm of loans per capita. Loans are transfers in cash or in kind for which the recipient incurs a legal debt in constant international dollars. Source: OECD (2005).

Bilateral Aid: Logarithm of bilateral aid per capita. Bilateral transactions are those undertaken by a donor country directly with an aid recipient (in constant international dollars. Source: OECD (2005).

Multilateral Aid: Logarithm of multilateral aid per capita. Total net aid flows minus bilateral aid in constant international dollars. Source: OECD (2005).

GNI: Logarithm of gross national income per capita in constant international dollars. Source: World Bank (2005).

Debt: Logarithm of total external debt per capita in constant international dollars. Source: World Bank (2005).

Growth: Annual percentage growth rate of gross domestic product per capita based on constant local currency. Source: World Bank (2005).

Inflation: Logarithm of the annual percentage inflation rate, as measured by the consumer price index. Source: World Bank (2005).

Governance: Logarithm of a governance indicator which is an unweighted average of three International Country Risk Guide (ICRG) indices, ranging from 0 to 6: Corruption in Government: Lower scores indicate "high government officials are likely to demand special payments" and that "illegal payments are generally expected throughout lower levels of government" in the form of "bribes connected with import and export licenses, exchange controls, tax assessment, police protection, or loans." Rule of Law: This variable "reflects the degree to which the citizens of a country are willing to accept the established institutions to make and implement laws and adjudicate disputes." Higher scores indicate: "sound political institutions, a strong court system, and provisions for an orderly succession of power." Lower scores indicate: "a tradition of depending on physical force or illegal means to settle claims." Upon changes in government new leaders "may be less likely to accept the obligations of the previous regime." Quality of the Bureaucracy: High scores indicate "an established mechanism for recruitment and training," "autonomy from political pressure," and "strength and expertise to govern without drastic changes in policy or interruptions in government services" when governments change. Source: Political Risk Services

Trade: Trade is the sum of exports and imports of goods and services measured as a share of gross domestic product. Source: World Bank (2005).

Reserves: Net international reserves (excludes gold) divided by imports of goods and services. Source: World Bank (2005).

Countries: Total sample

Algeria , Angola, Argentina, Bangladesh, Bolivia, Botswana, Brazil, Bulgaria, Burkina Faso, Cameroon, Chile, China, Colombia, Congo Rep., Costa Rica, Cote d'Ivoire, Democratic Republic of Congo, Dominican Republic, Ecuador, Egypt Arab Rep., El Salvador, Ethiopia, Gabon, Ghana, Guatemala, Haiti, Honduras, Hungary, India, Indonesia, Iran, Jamaica, Jordan, Kenya, Malawi, Malaysia, Mali, Mexico, Morocco, Mozambique, Nicaragua, Nigeria, Oman, Pakistan, Panama, Papua New Guinea, Paraguay, Peru, Philippines, Poland, Romania, Russian Federation, Senegal, Sierra Leone, South Africa, Sri Lanka, Sudan, Syrian Arab Republic, Tanzania, Thailand, Togo, Trinidad and Tobago, Tunisia, Turkey, Uganda, Uruguay, Venezuela RB, Vietnam, Zambia, Zimbabwe

Countries without access to international capital markets

Burkina Faso, Bolivia, Botswana, Georgia, Haiti, Mali, Malawi, Nicaragua, Nepal, Sierra Leone, Togo, Uganda.

Descriptive statistics and correlations

Table 1.6: Descriptive statistics

	Mean	**Std. Dev.**	**Min.**	**Max.**
IICR	27.84	14.33	4.05	72.3
Total aid p.c.	106.16	150.89	-54.57	1785.99
Grants p.c.	79.57	129.77	0.38	1946.98
Techn. Ass. p.c.	24.96	28.96	-13.74	308.24
Loans p.c.	26.60	48.94	-266.30	505.99
Bilat. aid p.c.	78.51	122.43	-37.69	1494.75
Multilat. aid p.c.	27.65	40.94	-75.71	386.04
GNI p.c.	3623.12	2484.42	213.38	19146.11
Debt p.c.	2429.45	2637.29	125.31	37116.06
GROWTH	0.86	4.89	-20.90	16.54
Inflation	153.24	1102.781	0.06	23773.13
Governance	2.77	0.92	0.67	5.33
Trade	57.83	27.54	12.35	192.11
Reserves	26.41	28.06	0.42	276.91

Notes: Summary statistics refer to the 847 observations used in the regression underlying column (1) of Table 2. While the aid-per-capita variables, Gross National Income (GNI), debt per capita, inflation and governance entered our regressions in logs, the summary statistics refer to the original data.

Table 1.7: Correlations

	AID	GNI	DEBT	GROWTH	INFL.	GOV	TRADE	RES.
AID	1.00							
GNI	-0.24	1.00						
DEBT	0.11	0.63	1.00					
GROWTH	-0.09	0.09	-0.09	1.00				
INFLATION	-0.07	0.01	0.15	-0.25	1.00			
GOV	-0.17	0.32	0.27	0.18	-0.05	1.00		
TRADE	0.33	0.12	0.39	0.03	-0.22	0.26	1.00	
RESERVES	-0.23	0.25	-0.02	0.16	0.00	0.22	-0.05	1.00

Notes: Summary statistics refer to the 847 observations used in the regression underlying column (1) of Table 2. All correlations refer to the variables as used in the regressions, i.e. the logarithm of aid per capita, the logarithm of GNI per capita etc.

References

Aizenman J. (1989a): Country Risk and Contingencies, *International Economic Journal,* 3, 81-102.

Aizenman J. (1989b): Country Risk, Incomplete Information and Taxes on International Borrowing, *Economic Journal,* 99, 147-161.

Arellano M. and S. Bond (1991): Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations, *Review of Economic Studies,* 58, 277-297.

Arellano M. and O. Bover (1995): Another Look at the Instrumental-Variable Estimation of Error-Components Models, *Journal of Econometrics*, 68, 29-52.

Asiedu E. and A.P. Villamil (2002): Imperfect enforcement, foreign investment, and foreign aid, *Macroeconomic Dynamics*, 6, 476-495.

Barro, R.J. and J-W. Lee (2004): IMF programs: Who is chosen and what are the effects?, *mimeo, Harvard University*.

Bird, G. and D. Rowlands (2002): Do IMF Programmes Have a Catalytic Effect on Other International Capital Flows?, *Oxford Development Studies,* 30, 229-249.

Blundell R. and S.R. Bond (1998): Initial conditions and moment restrictions in dynamic panel data models, *Journal of Econometrics,* 87, 115-143.

Bond S.R. (2002): Dynamic Panel Data Models: A Guide to Micro Data Methods and Practice, *CEMMAP Working Paper,* No. CWP09/02.

Bruno, G.S.F. (2005): Approximating the bias of the LSDV estimator for dynamic unbalanced panel data models, *Economics Letters,* 87, 361-66.

Cantor R. and F. Packer (1996): Determinants and Impact of Sovereign Credit Ratings, *FRBNY Economic Policy Review*, October 1996.

Chang C., E. Fernandez-Arias, and L. Serven (1998): Measuring Aid Flows: A New Approach, *World Bank, Development Research Group, Washington D.C.*

Ciocchini F. and E. Durbin and D. Ng (2003): Does Corruption Increase Emerging Market Bond Spreads?, *Journal of Economics and Business,* 55, 503-28.

Cunningham, A., L. Dixon, and S. Hayes (2001): Analysing yield spreads on emerging market sovereign bonds, *Bank of England Financial Stability Review,* December 2001, 175-186.

Easterly W. (2003): Can Foreign Aid Buy Growth?, *Journal of Economic Perspectives,* 17, 23-48.

Easterly W., R. Levine, and D. Roodman (2004): Comment: Aid, Policies, and Growth, *American Economic Review*, 94, 774-780.

Eaton J., M. Gersovitz, and J. Stiglitz (1986): The Pure Theory of Country Risk, *European Economic Review,* 30, 481-513.

Eichengreen B. and A. Mody (1998): What explains changing Spreads on Emerging Market Debt: Fundamentals or Market Sentiment?, *NBER Working Paper No. 6408*.

Gelos, R.G., R. Sahay, and G. Sandleris (2004): Sovereign Borrowing by Developing Countries: What Determines Market Access?, *IMF Working Paper No. 04/221*.

Hadi A.S.(1994): A Modification of a Method for the Detection of Outliers in Multivariate Samples, *Journal of the Royal Statistical Society, Series B,* 56, 393-396.

Hansen L. (1982): Large Sample Properties of Generalized Method of Moments Estimators, *Econometrica,* 50, 1029-1054.

Hansen H. and F. Tarp (2000): Aid effectiveness disputed, *Journal of International Development,* 12, 375-398.

Hansen H. and F. Tarp (2001): Aid and growth regressions, *Journal of Development Economics,* 64, 547-570.

Haque U. N., M. Kumar, N. Mark, and D. Mathieson (1996): The Economic Content of Indicators of Developing Country Creditworthiness, *IMF-Staff Papers,* 43, 688-724.

Haque U. N., N. Mark, and D. Mathieson (1998): The Relative Importance of Political and Economic Variables in Creditworthiness Ratings, *IMF Working Paper WP/98/46*.

Harms P. and M. Lutz (2006): Aid, Governance, and Private Foreign Investment: Some Puzzling Findings for the 1990s, *The Economic Journal*, 116, 773-790.

Harms P. and M. Lutz (2004): The Macroeconomic Effects of Aid, *University of St. Gallen Discussion Paper 2004-11*.

Judson R. and A. Owen (1999): Estimating Dynamic Panel Data Models: A Guide for Macroeconomists, *Economics Letters*, 65, 9-15.

Kaminsky G. and C. Reinhart (1999): The Twin Crises: The Causes of Banking and Balance-of-Payments Problems, *American Economic Review,* 89, 473-500.

Kiviet, J.F. (1995): On bias, inconsistency, and efficiency of various estimators in dynamic panel data models, *Journal of Econometrics,* 68, 53-78.

Larrain F., H. Reisen, and J. von Maltzan (1997): Emerging Market Risk and Sovereign Credit Ratings?, *OECD Development Centre, Technical Paper No. 124*.

Lee S. (1993): Are the credit ratings assigned by bankers based on the willingness of LDC borrowers to repay?, *Journal of Development Economics*, 40, 349-359.

Mody, A. and D. Saravia (2003): Catalyzing capital flows: Do IMF-supported programs work as commitment devices?, *IMF Working Paper 03/100*.

Nickell S. (1981): Biases in Dynamic Models with Fixed Effects, *Econometrica,* 49, 1417-26.

OECD (2005): CRI online database, *Development Assistance Committee*.

Political Risk Services: *International Country Risk Guide*, various issues.

Przeworski, A. and J.R. Vreeland (2000): The effect of IMF programs on economic growth, *Journal of Development Economics*, 62, 385-421.

Ramcharan, R. (2003): Assessing IMF program effectiveness, *IMF Research Bulletin,* 4(2).

Reinhart, C. (2002): Default, Currency Crises and Sovereign Credit Ratings, *NBER Working Paper No. 8738*.

Reinhart, C.M., K.S. Rogoff, and M.S. Savastano (2003): Debt Intolerance, *Brookings Papers on Economic Activity*, 2003(1), 1-74.

Roodman, D. (2004): The Anarchy of Numbers: Aid, Development, and Cross-country Empirics, *Center for Global Development Working Paper 32*.

Rose, A. (2005): One reason countries pay their debts: renegotiation and international trade, *Journal of Development Economics,* 77, 189-206.

Windmeijer, F. (2005): A finite-sample correction for the variance of linear efficient two-step GMM estimators, *Journal of Econometrics,* 126, 25-51.

Wooldridge J. (2002): Econometric Analysis of Cross Section and Panel Data, *Cambridge, Massachusetts: The MIT Press*.

World Bank (1998): Assessing aid: what works, what doesn't, and why, *A World Bank policy research report, Oxford University Press*.

World Bank (2005): World Development Indicators on CD-ROM, *Washington DC, The World Bank*.

Chapter 2

Research Productivity of German Academic Economists

Evaluation of Researchers:
A Life Cycle Analysis of German Academic Economists

2.1.1. Introduction

Evaluations compare certain features of a person with the features observed in a group of peers. A worthwhile evaluation needs to explicitly define the relevant comparison group and to make a case for the employed choice. In many cases, the contemporaries of the person to be evaluated represent the relevant peer group, the best example being the standard IQ test whose name even refers to the fact that intelligence is measured in relation to some denominator, which is, of course, the respective person's age. In sports, where evaluation almost represents the raison d'être, it is also quite common to compare contestants of the same age group, but other comparison groups, based, for example, on body weight or professional status, are also widely employed.

Research evaluations that are based on scientometric methods are still surrounded by a touch of controversy. Nevertheless, it is generally accepted that reasonable scientometric evaluations need to focus on narrowly defined disciplines; how the disciplines should be delineated is, of course, another matter. Many scientometric studies are, moreover, restricted to specific geographic regions and types of institutions. Apart from these public-domain characteristics, the relevant peer group is also described by personal characteristics, arguably the most important one being the researcher's age.

Age features two distinct dimensions that are relevant in the evaluation context: vintage and career age. Both of these dimensions are liable to have a strong impact on research productivity because research production heavily relies on human capital that is determined, on the one hand, by the initial endowment (i.e. by ability and initial training) and, on the other hand, by experience and obsolescence of knowledge. Since initial training (graduate education) is related to the *age cohort*, whereas experience and obsolescence of knowledge are related to *career age*, both of these age dimensions represent personal characteristics that are associated with generally recognized peer groups (class of 2005, assistant professors in their sixth year, etc.).

Precisely because life-cycle and vintage effects are liable to influence any researcher's productivity, research evaluations which are undertaken to implement incentive-compatible managerial reward or penalty schemes, need to take these age dimensions into account. In principle, this statement is not controversial. Tenure and promotion committees have always compared the track records of the applicants with precedents. Alternatively, they have judged whether the track records are compatible with an established policy or standard. These standards, however, have evolved over time by investigating research oeuvres of applicants who, by the very fact that they aspired to take a

certain career step, constitute a peer group defined by *career age*. Decisions with respect to performance-related pay have likewise been based on comparisons of track records. Since remuneration, unlike tenure and rank, does not represent a time-invariant prize, the applicant's age at the time of the application, i.e. his or her *cohort* or *vintage*, is always implicitly taken into account by the responsible authorities.

Even though of great importance for management decisions, studies dealing with the evaluation of *economic* research have hitherto rather neglected the age dimensions. This neglect applies especially to studies that evaluate entire groups of researchers, for example university departments or research institutes. An exception is the ranking study by Combes and Linnemer (2003). These authors, who rank 600 economic research institutions from 14 European countries, present, among others, one research productivity index that takes the respective researcher's career age into account. Even though the employed method of normalization with respect to career age is purely ad hoc, and the career age of the economists is estimated by rule of thumb, this study is groundbreaking because it spells out the demands that high-quality rankings should meet.

The available literature on life cycles in research productivity is oddly disconnected from the evaluation issue. The studies investigating life cycles are usually motivated by Gary Becker's human capital theory that predicts that investment in human capital decreases over the life cycle, thereby generating hump-shaped individual life cycles in labor productivity and earnings. Some scholars have extended the human capital approach to analyze the processes which are specific to research production. Others have used the standard human capital approach in order to guide their attempts to empirically identify the determinants of labor productivity; these scholars focus on research production mainly because measuring research productivity is, in many respects, easier than measuring labor productivity in other fields. The AER paper by Levin and Stephan (1991) followed both of these routes and was instrumental in kicking off the field that is now known as the *economics of science*.

Surprisingly few studies on research productivity were written by economists or investigate the economics profession. This has already been deplored by Paula Stephan in her (1996) JEL survey. Recent work on the economics profession include Kenny and Studley (1996), Oster and Hamermesh (1998) and Baser and Pema (2004) whose empirical results are compatible with a hump-shaped progression of individual research productivity over the life cycle as hypothesized by Becker's human capital theory. Goodwin and Sauer (1995), on the other hand, who do not clamp the life cycle in the Procrustes bed of a

quadratic specification, identify a bi-modal life cycle. Hutchinson and Zivney (1995) and Hartley et al. (2001) do not find any evidence supportive of the standard life cycle hypothesis at all.

Among the many considerable econometric problems that arise when estimating life-cycles in research productivity, the most challenging one arguably consists of separating career age and cohort effects, an endeavor that is confounded by the fact that publication behavior has changed over time. In order to estimate life cycle and cohort effects separately, an extensive panel data set comprising many cohorts is indispensable, otherwise the potentially considerable cohort-specific influences cannot be estimated, and the resulting estimates of the life cycle pattern will be biased.[25] It is conceivable that, because of these econometric problems, the empirical evidence with respect to cohort effects is somewhat elusive. Basar and Pema (2004) do not find any cohort effects at all, and Goodwin and Sauer (1995) report only marginally significant effects that are tainted since they may well reflect the fact that the members of the analyzed cohorts differ in age, implying that the older cohorts are composed of academic survivors and thus liable to have been more productive on the average.

The identification problem becomes even more challenging if one acknowledges that the publication behavior of economists has changed over time. Even if these changes have been relatively small, they may become significant in the course of a time period that allows estimating cohort effects. Since, however, career time, historical time and cohort affiliation depend on each other in a linear manner (career time = historical time – cohort "birth" year), only two out of the three effects can be estimated subject to some assumption about the development of the third one. This is the reason why all estimates of life cycle and cohort effects need to be interpreted with some caution.[26]

This chapter unfolds as follows. In the next section we present a new data set that describes the research behavior of German academic economists, and in section 2.1.3 we describe the heterogeneity of research production with respect to both age dimensions (career age and cohort affiliation). Our investigation of heterogeneity culminates in the presentation of a simple formula that translates any German economist's research oeuvre into a ranking vis-à-vis his or her peers. Section 2.1.4 describes the result of some life cycle

[25] Cohort-specific influences are, for example, the knowledge base incurred during graduate education, the rate of obsolescence, access to resources, opportunities provided by the socio-economic environment, and modes of behavior imprinted on the fledgling scientists. See Stephan (1996), pp. 1216-7.

[26] For a detailed exposition of the econometric methods that have been proposed to identify age, cohort, and period effects on individual research productivity, see Hall et al. (2005).

regressions. Since tenure represents the arguably most important special feature of the academic labor market, we analyze, in section 2.1.5, the persistence of individual research productivity in order to assess at what career stage promotion to a tenured position is justifiable. In section 2.1.6, we turn to evaluations of whole research units (German economics departments) and present some rankings that take the age dimension into account.

2.1.2 The data set

Most studies of research productivity over the life cycle employ a sample of scientists who are relatively active in research. The rationale for this approach is twofold. On the one hand, the behavior of choice researchers is better documented than that of less active ones. On the other hand, the standard econometric methods are better suited to process steady streams of activities than time series with many periods of inactivity. Since it is our intention to develop an evaluation scheme for all kinds of scientists, we did not follow this restricted approach and compiled a dataset that compromises, in principle, all academic economists currently working in Germany.

Since we use the EconLit data base we had to restrict ourselves to economists who received their doctoral degree at the earliest in 1969, the first year covered by EconLit. Considering that German academic economists receive their doctoral degrees when they are about 30 years old, this implies that the oldest economists in our data set were about 65 years old in 2004, the last year covered in our study. For these economists, we thus have complete life cycles. For the younger ones, the available life cycle becomes, of course, increasingly shorter. The shortest life cycles that we decided to consider have a length of six years which corresponds to a career age at which promising academic economists are granted tenure. We thus only consider scholars who received their doctoral degrees between 1969 and 1998 and who were employed at a German university in the year 2004 or have retired from such a position shortly before.

On the basis of these restrictions we have analyzed the publication records of more than 600 economists.[27] To be more precise, our data set is comprised of all EconLit-listed journal publications (up to the year 2004) authored or co-authored by the economists included in our sample. Evaluating only the set of journals referenced in EconLit excludes journals whose scope is not aligned with the current mainstream of economic research, new

[27] We thank Gerhard Foth , Robert Hofmeister and Christina Voigt for valuable research assistance.

economics journals, and journals that do not meet EconLit's quality standards. Whereas scope and timeliness are issues to be considered (scholars with peripheral or interdisciplinary specializations and scholars working on emerging fields may be underrated), exclusion because of insufficient quality does not appear to be an issue since the minimum quality standard set by EconLit is rather soft.

The quality standards set by the journals indexed in EconLit are of course quite diverse. Any study working with this data base therefore needs to capture quality differences in one way or another. If a reward scheme does not take these quality differences into account, the scientists would no longer attempt to produce research output of the highest possible quality but would rather shift their efforts towards producing results that are just about publishable in the journals with the softest quality standards. In other words: "Gresham's law of research evaluation" would see to it that mediocre research drives good research out of circulation.

A popular approach to controlling for journal quality is to use a subset of journals whose prime quality is uncontested. The ranking study by Kalaitzidakis et al. (2003), for example, followed this strategy. Restricting the journal set in this manner comes, however, at a significant cost. First of all, information especially about less accomplished scientists who do not publish in prime journals, is lost with the consequence that reward schemes based on such a set of journals would not provide any incentives for this class of employees. A second drawback of restricting the journal set is that this strategy would prohibit us from investigating changes in research *quality* over the life cycle. For these reasons we decided to work with the whole set of journals indexed in EconLit, and to explicitly control for journal quality.

The evaluation of journal quality represents a field of its own. From the plethora of weighting schemes we chose the "CLpn" scheme proposed by Combes and Linnemer (2003) because it is based on the journals' relative (subjectively perceived) reputation and (objectively measured) impact, and thus appears to provide a well-balanced rating over the whole quality range.[28] The CLpn-scheme converts each journal publication in standardized units of AER-page equivalents. The quality weight of the five top-tiered journals is normalized to unity. The sixteen second-tiered journals' imputed weight amounts to two thirds. Weights then decline in discrete steps (one half, one third, one sixth) down to the

[28] One disadvantage of this method is that journal quality is kept constant over the period of investigation that covers, after all, a time-span of 36 years.

minimum weight of one twelfth. Our variable that measures research productivity of researcher i on an annual basis (year T) is defined as follows:

$$CLpn_i(T) = \sum_k \frac{p_{k(i)} w_{k(i)}}{n_{k(i)}}, \quad (1)$$

where $p_{k(i)}$ and $n_{k(i)}$ denote the number of pages and the number of authors of researcher i's publications k, while $w_{k(i)}$ denotes the appropriate journal quality weight. The CLpn-index thus not only controls for quality but also for the number of authors and the length of the journal articles.[29]

In order to obtain comparable individual life cycles of research productivity, we merged the annual records of individual research productivity with the year in which the respective researcher obtained the doctoral degree, i.e. we align the individual life cycles by this reference year. Our data set also contains some coarse information about the included economists' field of specialization, and we also documented the researchers' gender. Only about 7.5% of our academic economists are women. 15% of the economists in our sample specialize in microeconomics, 26% in macroeconomics, 34% in public economics and 16% in econometrics. Economists who could not be assigned to any one of these fields were assigned to the field OTHER.

2.1.3 Describing the landscape of German academic research in economics

In order to obtain a first impression of the size and distribution of the oeuvres of German academic economists, we cumulate the annual research outputs defined in equation (1) from career year -5 until career year t, where 0 denotes the year in which the economists were granted their doctoral degrees:

$$R_i(t) = \sum_{T=-5}^{t} CLpn_i(T), \quad (2)$$

and then compute for all career ages t the borderline values of R for the following percentiles: 25%, 50%, 80%, and 90%. The resulting information is depicted in Figure 2.1.1.

Averaging over all economists in our sample we observe, first of all, that the oeuvre of the median researcher is quite modest. During his whole career the median German economist does not manage to produce more than 10 AER-equivalent pages. Assuming that all of this research has been published in journals belonging to the lowest quality tier, this

[29] We did not, however, take into account that the number of words per page differs across journals.

implies that the median economist publishes about 6 journal articles (20 pages each) during his research career, i.e. one article every six years. Second, Figure 2.1.1 reveals that the distribution of the individual research oeuvres is skewed to the right and exhibits a large variation. These characteristics do, of course, not come as a surprise. Rather, they constitute stylized facts that have transpired from many related studies.[30] More interesting is the fact that the percentile borderlines are not monotonous and exhibit a marked "overall" concavity. The violation of monotonicity of the stock variable R is not as puzzling as it might appear at first sight; it simply reflects cohort effects in our unbalanced panel. If research productivity increases dramatically across cohorts, the stock of the scientists at a young career age (measured across *all* cohorts) may well be larger than the stock of the scientists at an older career age (measured across only those cohorts who have reached this career age). The concavity of the percentile borderlines admits two interpretations: it may either reflect decreasing marginal productivity over the life cycle or it may again represent an artifact of cohort effects in our unbalanced panel.

In order to discriminate between the *decreasing marginal productivity* interpretation and the interpretation that presumes *cohort effects*, we analyzed the career-time oeuvres of different cohorts. For that purpose, we divided our sample of economists into five cohorts, each comprising six age groups. The oldest cohort comprises the age groups 1969-1974, and the youngest one the age groups 1993-1998. The members of the oldest cohort thus look back on a career of at least 30 years, while the members of the youngest one have had a career of at least six years. The percentile borderlines are now monotonous, indicating that vintage effects within the cohorts are relatively small. Figure 2.1.2 presents the percentile borderlines for the oldest cohort.[31]

Two interesting insights transpire. First, eyeballing of the cohort-specific percentile borderlines does not suggest any pronounced concavity. An S-shaped life cycle productivity pattern supporting the factors portrayed by the standard human capital model thus cannot be identified, at least not at the aggregate level. To shed some more light on this issue, we will, therefore, further investigate our economists' life cycles with the help of micro-econometric methods in Section 2.1.4. The second feature that emerges is more conclusive. The German economics profession is characterized by striking cohort effects in research productivity: the percentile borderlines become increasingly steeper for younger cohorts. The increase in

[30] The highly skewed nature of publication was first observed by Lotka in 1926 in a study on physics journals (cf. Stephan, 1996, p. 1203).
[31] The working paper version of this article also presents the evidence for the other cohorts.

cohort-specific research productivity is illustrated in Figure 2.1.3 in which the 80%-lines of the five cohorts are superimposed. This representation shows that it took an economist who tops 80% of his peers in the oldest cohort about 18 years to accumulate an oeuvre of 20 AER-equivalent pages, whereas a top-80% economist of the second cohort managed to do so in 12 years. This time span is reduced to 8 and 4.5 years for the two following cohorts, respectively, and the top-80% economist of the youngest cohort only needs 3.5 years to produce 20-AER equivalent pages.

From our data set we can extract information that is directly relevant for the evaluation of individual researchers. In particular, we can assign each economist a peer-specific performance rank at each point of career time. This kind of information is of prime importance for a university management that wants to pursue a rational performance-related remuneration policy. Information about the standing of individual researchers vis-à-vis their peers is, moreover, a prerequisite for department rankings that are insensitive to the age structure of the evaluated faculties. We will turn to this issue in section 2.1.6. Whole career profiles in terms of relative performance are, finally, of vital importance to assess the persistence of research performance. The crucial question in this context is whether it is possible to forecast a scientist's research performance from his track record, and if so, at which stage of a scientist's career such forecasts are sufficiently accurate to serve as a basis for management decisions such as granting tenure or awarding substantial research grants. The persistence issue will be dealt with in section 2.1.5. Here we will follow up the first issue and ask ourselves how the information about the *current* cohort-specific ranking of *individual* economists can be condensed in such a way that it can serve as a simple management information device.

To do so, we consider the standard situation faced by a university management or a research foundation that would like to assess an economist's relative research standing in the German academic profession. Usually, the evaluator has only access to this person's CV including publication list. With the help of the publication list it is easy enough to compute via equations (1) and (2) the accumulated research output R at the end of the year 2004. Dividing this output R by the *adjusted* career age τ (τ = 2010-Y, where Y denotes the year in which the evaluated economist received his or her doctoral degree) yields the average research productivity P.[32] How does the average research productivity P of an economist translate into a ranking vis-à-vis his or her peers? Since the relative research standing

[32] We let the productive time of a researcher start five years before the doctorate. Since the doctorate takes place in career year t=0, the adjusted career age τ = 2004-Y+6 = 2010-Y.

depends on the average research productivity as well as on the cohort age of the person to be evaluated, we are seeking a formula of the form S=f(P,Y), where S denotes the evaluated economist's relative research standing in percentiles.[33] Regressing S on Y and P yields the following formula:

$$S = 18.3 - \frac{9.2}{1000} \cdot Y + 0.55 \cdot \sqrt[3]{P} \ .^{34} \qquad (3)$$

For evaluation purposes, the negative residuals of our regression (overestimation) clearly present the relevant downward risk. Since the distribution of residuals resembles a normal distribution with a standard deviation of 0.077, the probability of overestimating a candidate by 10 percentiles is about 10%. This appears to be a risk well worth taking in a situation in which the alternative is to rely on peer evaluations and recommendations that are notoriously biased.

2.1.4 A micro-econometric investigation of life cycle productivities

The empirical evidence presented in the previous section suggests that life cycles in economic research productivity are rather flat. This evidence refers, however, to highly aggregated data. In order to do justice to the heterogeneity in our population of economists we exploited the micro-structure of our data set by regressing individual research productivity not only on career-time and cohort membership, but also on the field of specialization, on a gender dummy variable, and on a measure of ability. Following Goodwin and Sauer (1995), we ranked the researchers according to their cohort-specific average life-time productivity. We then defined quintile ranks within the distribution for each three-year cohort and assigned each researcher the appropriate *ability rank*.

Since about three quarters of our observations of the dependent variable (research productivity of economist i in year t) are zeroes, one cannot apply OLS. To accommodate this high degree of censoring we used the *hurdle model*, i.e. we allow the decision making process to be more complex than the one captured by a standard Tobit model. The first part (being active) is portrayed with a Probit model, whereas the distribution of the positive counts is modeled with the help of a truncated Negative Binomial model since the observed density distribution of our dependent variable resembles the pattern of count data.

[33] The relevant peer group always consists of five age groups, namely the age group of the person to be evaluated and the four neighboring age groups.
[34] Our formula approximates our regression result which explains 93% of the variance of S.

The results of our regressions are shown in Table 2.1.1. Our hurdle model focuses on heterogeneity with respect to ability, i.e. we include dummy variables for each ability rank and also allow the life-cycle polynomials to differ across the ability ranks 5 (top researchers), 4 (accomplished researchers) and 1-3 (journeymen researchers).[35] Figures 2.1.4 and 2.1.5 visualize the fact that the time polynomials differ across ability ranks and that there are significant differences between the time polynomials of the *Probit* and *NegBin* part, thereby suggesting different forces governing the two respective processes. Our results indicate that the top-researchers manage to increase their publication incidence over time while their research productivity somewhat declines in the second half of their careers. It thus appears that the best researchers in the profession focus in the beginning of their careers on fewer research projects (articles) but execute them with more effort which gives rise to higher quality (better journals) and more extensive results (longer articles), and all this is achieved with fewer co-authors. Later on in their careers these researchers get involved in more projects that are, however, executed with less effort. The two processes (number of projects and research effort put into each project) neutralize each other and, in conjunction, give rise to the flat life cycles in overall research productivity already observed. Decomposing our measure of research productivity and regressing average quality, article length, and number of co-authors on our explaining variables indeed shows that older economists work together with more collaborators (co-authors), write shorter articles, and publish in lower quality journals. Interestingly, however, top researchers manage to maintain quality much more than their less gifted peers.[36]

As compared to the top-researchers, the "accomplished" researchers' publication incidence and research productivity declines more sharply over their life cycles. These life cycles are thus better in line with the predictions of the human capital approach to explaining labor productivity. The "journeymen" researchers, finally, have rather flat and nondescript life cycles.

The coefficients of the cohort dummies, not surprisingly, increase over time. This result is consistent with the joint hypothesis of more productive younger cohorts and a *constant* historical time effect. We admit, however, that it is not inconceivable that our regressions somewhat overestimate the identified vintage effects since the gradual substitution process towards publishing research results mainly in journals may still have

[35] It was necessary to bundle the first three ranks together because of the high degree of censoring within these ranks. Nevertheless, we still allow for different intercepts for each rank.
[36] See our companion paper: Rauber and Ursprung (2006).

been at work in the beginning of our period of observation. The estimated coefficients of the gender dummy variable indicate that female economists publish significantly less than their male peers. This negative effect, however, arises from the fact that female academic economists seem to be more likely not to engage in research at all. If female economists decide to be active researchers, then they are just as productive as their male peers. Our field dummies, finally, show that researchers specializing in macroeconomics are less likely to be active researchers, and active micro-economists publish more than their peers. Even though these effects appear to be relatively small and fragile, it might be worthwhile to bear these field effects in mind when evaluating individual economists.

In a second (standard Tobit) regression we focus on heterogeneity with respect to *cohort membership*. As in the hurdle model, we allowed the life cycle polynomials to differ, this time across our six cohorts. Figure 2.1.6 visualizes the cohort specific time polynomials. It can be seen with the naked eye that the shape of these life cycles differs across cohorts: younger cohorts have more hump-shaped life cycles than older cohorts. With respect to the other explaining variables nothing changes dramatically.

We thus arrive at the result that the life cycles of younger cohorts – as far as we can tell from the initial phases of these cycles – correspond more closely to the predictions of the standard human capital approach to explaining changes in labor productivity than the evidence we have for older economists. Various hypotheses lend themselves to explaining this result. The first and arguably most plausible one maintains that the academic environment has become increasingly more competitive over the last 35 years. In a more competitive work environment, employees who want to succeed are forced to optimize under the pertaining constraints. It is thus not surprising that their behavior more closely corresponds to the predictions of the human capital model that narrowly focuses on labor market incentives. An alternative hypothesis is that doctoral students of older cohorts have been exposed to different role models than the younger cohorts. This hypothesis relates to the preference formation process which works through sociological imprinting. The last hypothesis does not assume a change in preference formation but different preferences of the people who decide to pursue an academic career. Whether it is possible to empirically discriminate between the three hypotheses (that are, of course, not mutually exclusive), remains to be seen.[37]

[37] See Frank and Schulze (2000) for an experimental design to test a related set of hypotheses.

2.1.5 Persistence of research productivity

The *economics of science* literature has clearly demonstrated that an academic scientist's research productivity has a noticeable influence on his or her labor market success. First of all, research productivity varies positively with pay (cf. Kenny and Studley, 1996, and Moore et al., 2001, for empirical evidence relating to the economics profession). A strong research record has, moreover, also a positive influence on the obtainable job status in terms of the employing university's reputation (cf. Grimes and Register, 1997, and Coupé et al., 2003), and scientists with strong research records are more likely to be granted tenure and to be promoted to higher academic ranks (cf. Coupé et al., 2003). Tenure and promotion to the highest level of the academic hierarchy may, on the other hand, have detrimental effects on research productivity because these types of upgrading are irrevocable and thus reduce incentives to work hard. Backes-Gellner and Schlinghoff (2004), for example, have show that research productivity of German (business) economists increases before the only crucial career step (appointment to a professorship) and is reduced afterwards. An early study on the impact of tenure that arrived at similar results for the United States is Bell and Seater (1978).[38]

Precisely because irrevocable career steps are liable to have a certain influence on research productivity, it is important to know at what stage of the academic career the research potential of a scientist can be assessed with reasonable accuracy and to what extent this potential is liable to be used in the post-tenure period. In other words, it is (from a managerial point of view) important to possess firm information on the *persistence* of individual research productivity. Inspection of our aggregate and individual data has already revealed that research productivity in our sample of economists is characterized by a great deal of persistence. In this section, we focus on the question whether the traditional American policy to grant, postpone, or decline tenure after a review period of six years does make sense in the light of our empirical evidence. Many knowledgeable observers agree that young scientists have to wait too long to be promoted to a professorship in the German university system. On the average, the implicit probation period amounts to eight years (German economists obtain their doctoral degrees when they are about 30 years old and are, on the average, appointed to their first professorship at the age of 38). The objective of the investigation presented in this section is to inquire whether the review period could indeed be shortened without great loss in terms of evaluation accuracy.

[38] For a recent theoretical study of tenure and related incentive schemes in academia, see Dnes and Garoupa (2005).

As compared to *tenure-induced effects* on research productivity, the optimal *timing of the tenure decision* has not found a great deal of attention in the scientometric literature dealing with the economics profession. A notable exception is the study by Hutchinson and Zivney (1995). These authors regress the average annual post-tenure productivity (measured in numbers of journal articles) on the pre-tenure oeuvre of economists using two hypothetical review periods, namely the standard six years and four years. Their regression analysis leads them to concur with Bell and Seater's (1978) conclusion based on cross-sectional data "that granting of tenure seems to have negative effects on individual publishing performance" (p. 614). "Yet, because the negative effect is so small numerically, 0.01 articles per year, our results indicate that publishers maintain essentially constant pre- and post-sixth-year rates of publication over their post-doctorate years. Moreover, shortening the review period from six years after the doctorate to four, relying upon our 1969-1979 doctorates, only slightly reduces the ability to predict future journal publication rates based on existing journal publication information while also producing almost constant pre- and post-fourth-year rates of publication" (Hutchinson and Zivney, 1995, p. 74).

In order to check whether the German economists' academic standing reached by their sixth year after the doctorate is a good indicator for their mid-career reputation (at the approximate age of 42, i.e. in the twelfth year after the doctorate), we ranked all economists in our sample at career time t=6 according to the size of their oeuvres in relation to a special five year cohort for each class.[39] We then define quintile ranks and assigned each researcher the appropriate rank. Repeating this procedure for the career year t=12, we arrived at the mid-career ranking of the same economists and were then able to compute the probability of moving from one quintile rank to another within the observation period. These transition probabilities are shown separately in Table 2.1.2 for the older economists in our sample (classes of 1969 to 1980) and for the younger ones (classes of 1981 to 1992). Due to the inescapable problem of research-inactive scholars we had to group the first two quintiles together with the consequence that the probabilities in the columns do not add up to 100%.

The results summarized in Table 2.1.2 once more show that research production is indeed characterized by a great deal of persistence. The probabilities on the main diagonal are substantially larger than the off-diagonal probabilities, implying that marked changes in the academic standing are low probability events. Table 2.1.2, in particular, shows that appointing a young professor with a high reputation is a relatively safe bet these days. On

[39] Members of the class of 1981, for example, are ranked in the cohort comprising the classes of 1979 up to 1983.

the other hand, appointing a professor with a bad publication record and hoping (perhaps based on hearsay) for the best, is not much more than wishful thinking. The probability of a bottom group researcher making it in the first six years of his or her full professorship to the top 40% is nowadays not more than 4 out of 100.[40] Table 2.1.2 also documents that the research track record has become a better indicator of future research productivity over the years. The transition probabilities of the younger economists are more centered on the main diagonal than those of the older economists.

The evidence summarized in Table 2.1.2 documents that, currently, a six year review period provides ample evidence for an informed tenure decision. The question therefore arises as to whether the German method of appointing professors (i.e. after an average review period of eight years) is indeed significantly superior in terms of avoiding bad appointments to justify the cost (especially the attendant loss of appeal to pursue an academic career). To investigate this question, we have computed the transition probabilities of the *younger* German economists also for hypothetical review periods of eight and four years. The results are summarized in Table 2.1.3. Given that we work with stock variables, it is not surprising that the predictions become somewhat sharper when using an eight instead of a six year review period, and somewhat more diffuse when using a four year period. More interesting is the fact that reducing the review period from the German standard of eight years to the American standard of six years does not appear to come at an inordinate loss of information. Research excellence, in particular, can be detected after six years just as well as after eight years. In many cases of truly superior young scientists, a review period of four years may well be sufficiently long to make a reasonably safe appointment decision. Our conclusion is thus in line with the results derived for the United States by Hutchinson and Zivney.

2.1.6 Some new rankings for German economics departments

If one agrees that the evaluation of individual researchers should take career age and cohort affiliation into account, then these age dimensions should also be considered when ranking whole departments. After all, meaningful department rankings are supposed to reflect the research competence of its members and not the age structure of the departments' faculty. In this section we therefore present some rankings of German economics departments that

[40] Notice that the persistence documented in Table1 is, of course, to some extent predicated by the question we ask, i.e. by the fact that we use stock data that reflect reputation. Using flow data would certainly increase the inter-quintile transition probabilities.

reflect the life cycle dimension of the evaluated faculties. The objective is to demonstrate how, *in principle*, such rankings can be conceptualized and to show how rankings that incorporate life cycle information compare to traditional rankings that do not do so.

We decided to produce rankings that are comparable to the research rankings published by the *Centrum für Hochschulentwicklung* (CHE) because the CHE-rankings, even though criticized by an impressive number of knowledgeable observers of the German research landscape, nevertheless are quite influential. The reference groups of the CHE-rankings are the *tenured* professors of the respective departments. Whether this reference group constitutes a meaningful basis for an evaluation is questionable. Nevertheless we adopt here this approach in order to provide results that are easily comparable to an established German standard.

The rankings that are presented in Table 2.1.4 refer to 52 economics departments. All of these departments confer degrees in economics and belong to a German university; we thus do not consider economics departments of second-tier universities, the so-called *universities of applied sciences*. One of the main (but little appreciated) challenges of *current potential rankings* as compared to *work-done-at rankings* consists in the identification of the respective faculty members. Since some of the faculty lists used by the CHE are grossly at variance with a truthful representation, we decided to base our rankings on a revised set of faculty lists that is reproduced in the appendix of the working paper version of this article.

Our first ranking (see column A in Table 2.1.4) simply represents the mean of the individual research standings of the respective faculty members, where the individual research standing is defined via the percentile value of average life-time research productivity within a three years cohort comprising all economists who received their doctoral degrees in the same year as the evaluated individual or in a neighboring year. Since these overlapping three-year cohorts are rather small for some years, we also show a ranking using cohorts of five years (column B). The rankings appear to be quite insensitive to the chosen cohort size: only three out the 52 ranked departments move by three ranks and one (Lüneburg, one of the two smallest departments with three professors) by four ranks across the two rankings. The two first rankings are thus very similar which is confirmed by a rank-correlation coefficient amounting to 99.6%.

As far as the top-ranked departments are concerned, the results of the first two rankings confirm, in essence, the results of earlier studies and the assessment of informed

observers of the German economics profession.[41] Somewhat surprising is perhaps the fact that the LMU Munich is only placed 9th.[42]

The first two rankings do not take into account that the research standing of individual economists is sensitive to their respective field of specialization. As we have shown in section 4, the field of specialization has a statistically significant influence on our measure of research productivity. The ranking presented in column C of Table 2.1.4 therefore adjusts for these field-specific differences in publication behavior by aligning the field-specific means. This ranking is still closely correlated to the former ones: the rank-correlation coefficients amounting to 96.6% and 96.5%, respectively. Now we observe however quite a few larger deviations in individual rankings. Nevertheless, the *group* of leading departments does not change as compared to the baseline rankings.

Thus far our rankings were based on orderings of individual scientists within narrow peer groups. One could argue that relying exclusively on actual data of relatively small cohorts may, in some cases, bias the evaluation of individual scientists and thereby give rise to unfair rankings. If, for example, unusually many first-rate scientists happen to be of approximately the same age, scientists who have the "bad luck" to be their contemporaries appear to be mediocre even when their overall research record is quite good, simply because they are compared only to their immediate cohort peers who are, coincidentally, very good. This kind of bias can be avoided by using our formula presented in equation (3) - albeit at the cost of losing some information. The ranking presented in column D of Table 2.1.4 is based on the ranking of the respective faculty members according to our formula. Since the formula-based ranking in some instances does markedly differ from the baseline ranking that uses actual cohort data we conclude that the identified bias may have an undue effect even in the aggregate.

The last two rankings presented in Table 2.1.4 do not take the life cycle dimension of individual research productivity into account. They are based on a method that is similar to the method used by Combes and Linnemer (2003) in their "career" rankings, i.e. we compute the average research productivity of each department member and then either use the department-average of the respective percentile rankings (column E) or the average of the individual productivities (column F). Comparing these standard rankings with our baseline ranking demonstrates that life cycle effects are not only significant for the

[41] See, for example, Ursprung (2003).

[42] More important than the rank is of course the numerical value of the variable on which the ranking is based (these values are reported in the working paper version of this article, CESifo Working Paper No. 1673). In this respect *ratings* are more meaningful than *rankings*.

evaluation of individual scientists but also for the ranking of whole departments (the rank correlations between ranking E and F and the ranking A amount to 94% and 88%. Consider, for example, the department of the LMU. According to the standard ranking E, the LMU is ranked 6th while according to our life-cycle rankings A and B it is ranked only 9th. This drop is apparently due to the fact that the most productive members of the LMU department are relatively young; neglecting the fact that young economists are in general more productive than older ones thus gives rise to an overestimation of the department's research standing. The cases of Frankfurt a.M., the two small departments of the RWTH Aachen and Lüneburg, and Erfurt are similar. The departments of the FU and HU Berlin, Mannheim, Bielefeld, Frankfurt a.O. and Osnabrück represent the counterpart category. These departments do significantly better when life cycle effects are taken into account. In these departments it is thus the old guard that is more productive - at least in relative terms.

The last ranking (F) is more sensitive to outliers than ranking E because there is no upper bound for individual productivity. Extremely productive scientists thus give rise to a non-representative department average. Which of these two standard rankings is to be preferred depends of course on the context of the investigation. In any event, these two standard rankings clearly support our main argument: life cycle considerations also matter for research rankings of whole university departments.

2.1.7 Tables and Figures

Figure 2.1.1: Percentiles 1969-1998

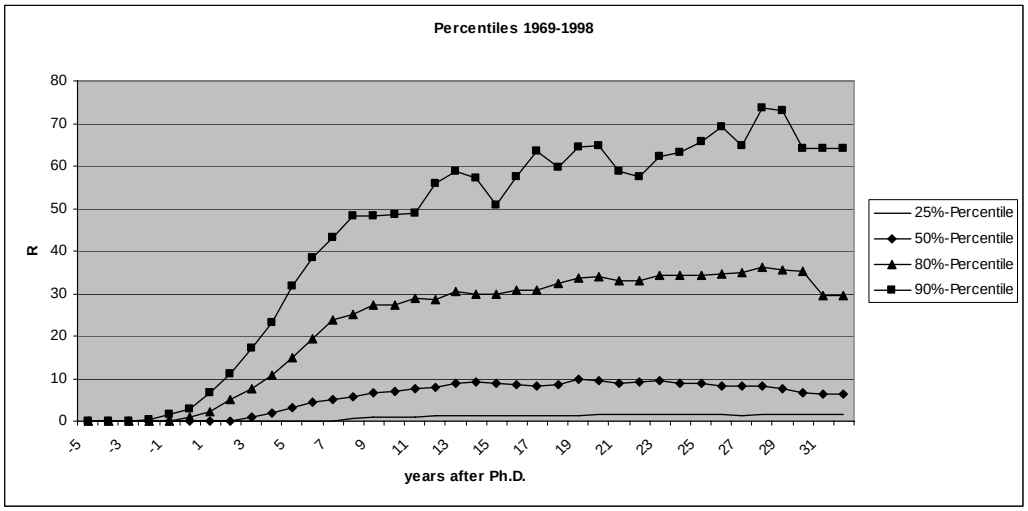


Figure 2.1.2: Percentiles for cohort 1969-1974

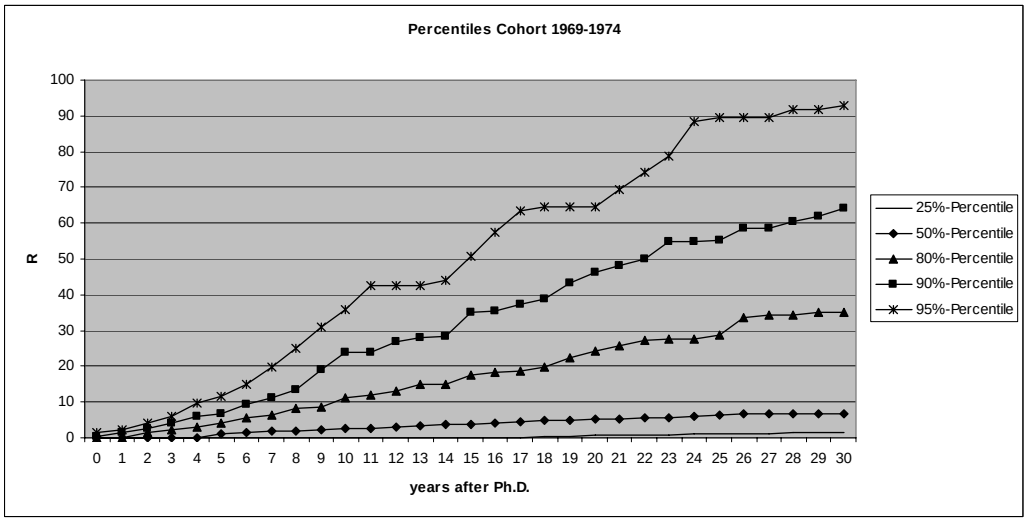


Figure 2.1.3: 80%- Percentile Lines for all cohorts

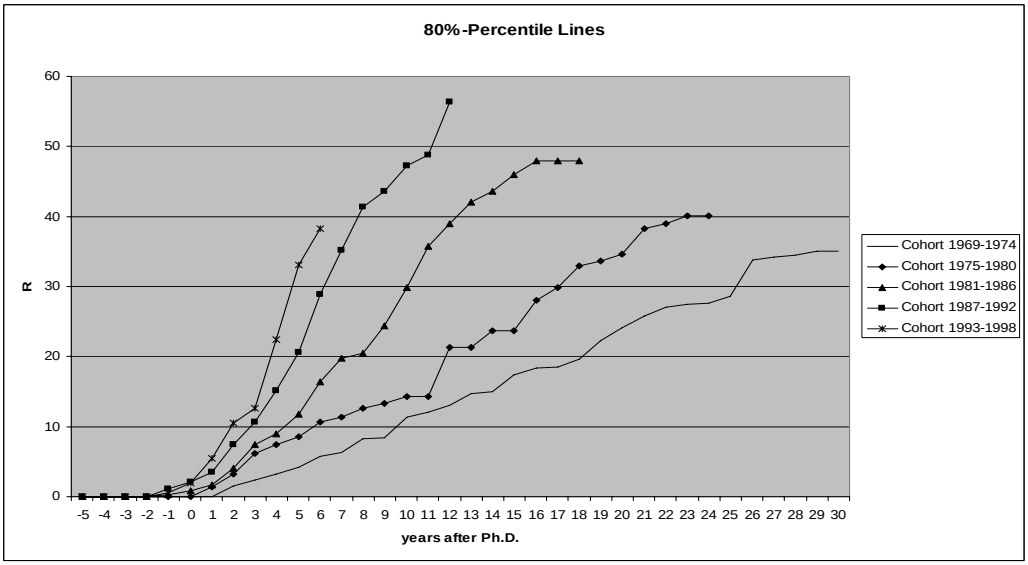


Figure 2.1.4: Probability of being active

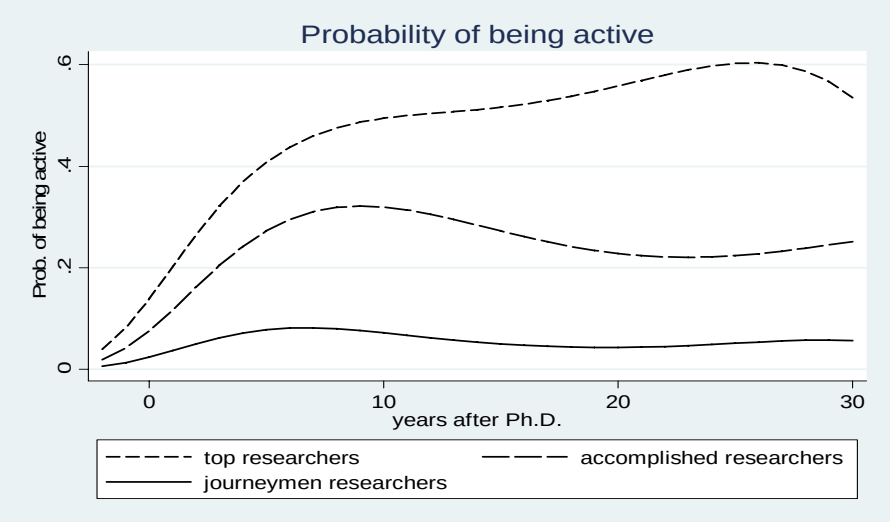


Figure 2.1.5: Conditional productivity by rank

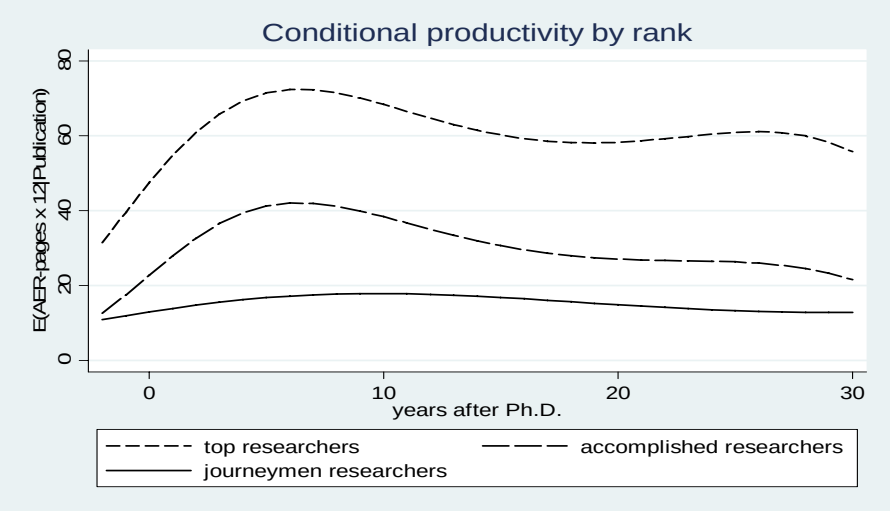


Figure 2.1.6: Tobit estimates by cohort (macroeconomist, rank 4)

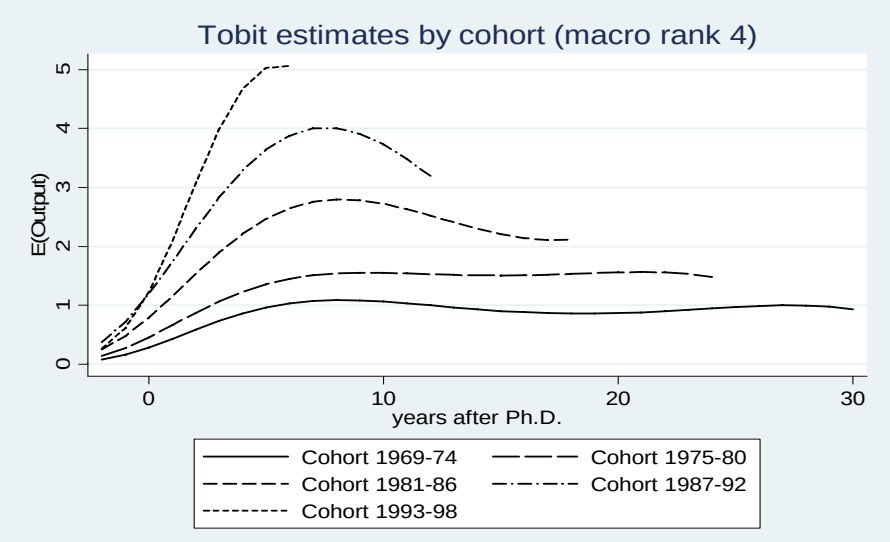


Table 2.1.1: Estimates

	Estimates by Ability				Estimates by Cohort	
	Probit		NegBin		Tobit	
	sideline researchers		*sideline researchers*		*cohort 69-74*	
T	0.2149***	(12.54)	0.0747***	(4.01)	0.2198***	(8.34)
T^2	-0.0254***	(9.39)	-0.0052***	(3.40)	-0.0233***	(6.71)
T^3	0.0011***	(7.08)	$-8.89E^{-5}$*	(2.66)	0.001***	(5.75)
T^4	$-1.44E^{-5}$***	(5.54)			-0.0001***	(5.16)
	accomplished researchers		*accomplished researchers*		*cohort 75-80*	
T	0.2698***	(14.12)	0.2328***	(6.32)	0.2148***	(8.61)
T^2	-0.0248***	(7.62)	-0.0284***	(5.46)	-0.0238***	(5.34)
T^3	0.0008***	(4.75)	0.0012***	(4.62)	0.0011***	(3.96)
T^4	$-9.70E^{-6}$***	(3.23)	$-1.73E^{-5}$***	(4.21)	$-1.83E^{-5}$***	(3.33)
	top researchers		*top researchers*		*cohort 81-86*	
T	0.2783***	(13.14)	0.1612***	(6.18)	0.2382***	(11.79)
T^2	-0.0268***	(8.38)	-0.02***	(5.84)	-0.0214***	(8.17)
T^3	0.0011***	(6.66)	0.0009***	(5.03)	0.0006***	(6.12)
T^4	$-1.71E^{-5}$***	(5.92)	$-1.29E^{-5}$***	(4.41)		
					cohort 87-92	
T					0.2622***	(15.31)
T^2					-0.0226***	(6.77)
T^3					0.0005***	(2.71)
					cohort 93-98	
T					0.3736***	(14.59)
T^2					-0.0194***	(4.08)
T^3					-0.0037***	(2.74)
T^4					0.0003***	(2.71)
C7580	0.2673***	(5.11)	0.1943**	(2.14)	0.2551**	(2.64)
C8186	0.6564***	(12.14)	0.263***	(3.26)	0.6917***	(7.64)
C8792	0.8478***	(15.84)	0.4847***	(6.42)	1.1374***	(13.43)
C9398	1.0608***	(19.83)	0.5546***	(7.84)	1.1818***	(13.43)
FEMALE	-0.202***	(2.96)	0.0329	(0.47)	-0.1457***	(2.67)
MICRO	0.1411*	(1.73)	0.0294	(0.35)	0.176***	(2.83)
MACRO	0.1041	(1.36)	-0.1325**	(1.97)	0.0672	(1.12)
PUBLIC	0.1674**	(2.29)	-0.1197*	(1.87)	0.1191**	(2.03)
ECONOMETRICS	0.1403*	(1.78)	-0.1748**	(2.34)	0.0651	(1.04)
RANK 2	0.9168***	(12.49)	0.5113***	(5.63)	1.3164***	(20.52)
RANK 3	1.3602***	(18.26)	0.8883***	(10.29)	2.3888***	(37.89)
RANK 4	1.4634***	(14.24)	1.0707***	(8.29)	3.7225***	(59.02)
RANK 5	1.8102***	(17.31)	1.8069***	(14.49)	6.4509***	(101.61)
CONST	-3.0002***	(28.60)	2.1868***	(17.26)	-3.1103***	(29.3)
Observations	15478		3834		15478	
Pseudo R^2	0.29				0.1469	
Log Likelihood	-6165		-18795		-16173	

Notes: Absolute t-values in parentheses. Tobit: marginal effects on unconditional expected value are reported.
* denotes significance at 10 percent, ** at 5 percent and *** at 1 percent level.

Table 2.1.2: Transition probabilities: year 6 –year 12

		1&2	3	4	5
1&2	Coh.1	**0.80**	**0.14**	**0.04**	**0.02**
	Coh.2	0.83	0.13	0.04	0.00
3	Coh.1	**0.30**	**0.41**	**0.24**	**0.05**
	Coh.2	0.28	0.44	0.23	0.05
4	Coh.1	**0.02**	**0.37**	**0.47**	**0.14**
	Coh.2	0.00	0.37	0.56	0.07
5	Coh.1	**0.00**	**0.00**	**0.19**	**0.81**
	Coh.2	0.00	0.00	0.14	0.86

Notes: Cohort A: 1969-1980, Cohort B: 1981-1992

Table 2.1.3: Transition probabilities of cohort B

		1&2	3	4	5
1&2	**4-12**	**0.80**	**0.13**	**0.05**	**0.02**
	6-12	0.83	0.12	0.05	0.00
	8-12	*0.88*	*0.11*	*0.01*	*0.00*
3	**4-12**	**0.30**	**0.39**	**0.26**	**0.05**
	6-12	0.28	0.44	0.23	0.05
	8-12	*0.24*	*0.62*	*0.14*	*0.00*
4	**4-12**	**0.07**	**0.35**	**0.45**	**0.13**
	6-12	0.00	0.37	0.56	0.07
	8-12	*0.00*	*0.19*	*0.70*	*0.11*
5	**4-12**	**0.00**	**0.02**	**0.19**	**0.79**
	6-12	0.00	0.00	0.14	0.86
	8-12	*0.00*	*0.00*	*0.12*	*0.88*

Notes: year 4 – year 12, year 6 – year 12, *year 8 – year 12*

Table 2.1.4: Department Rankings

	Life Cycle				Standard	
	A	B	C	D	E	F
FU Berlin	6	6	7	6	9	4
HU Berlin	5	5	1	5	8	5
HWP Hamburg	49	48	44	48	49	47
LMU München	9	9	6	7	6	3
RWTH Aachen	11	12	19	13	4	9
TU Berlin	25	24	27	27	23	29
TU Chemnitz	43	44	46	45	42	43
TU Dresden	18	16	16	14	14	16
Uni Augsburg	30	28	37	28	31	37
Uni Bamberg	37	37	42	31	35	38
Uni Bielefeld	7	7	10	9	11	13
Uni Bonn	1	1	2	1	1	1
Uni Bremen	46	47	40	50	51	50
Uni Dortmund	10	11	8	11	12	12
Uni Duisburg-Essen	44	43	45	44	44	42
Uni Erfurt	41	41	41	38	25	20
Uni Erlangen-Nürnberg	24	25	31	25	24	22
Uni Frankfurt / Main	13	13	11	12	10	10
Uni Frankfurt / Oder	8	10	12	10	13	18
Uni Freiburg	33	33	29	35	32	30
Uni Gießen	35	36	43	39	46	49
Uni Göttingen	34	34	34	29	26	28
Uni Halle-Wittenberg	40	39	33	41	40	41
Uni Hamburg	32	32	26	34	33	31
Uni Hannover	20	19	25	23	19	26
Uni Heidelberg	23	26	18	16	29	7

	Life Cycle				Standard	
	A	B	C	D	E	F
Uni Hohenheim	26	23	22	24	22	27
Uni Jena	52	51	47	49	47	48
Uni Karlsruhe	36	35	32	33	37	34
Uni Kiel	3	2	5	3	2	8
Uni Köln	27	27	36	26	36	33
Uni Konstanz	2	3	3	4	3	6
Uni Leipzig	48	49	48	47	48	46
Uni Lüneburg	12	8	9	8	5	11
Uni Magdeburg	19	20	20	21	15	25
Uni Mainz	16	18	15	19	21	17
Uni Mannheim	4	4	4	2	7	2
Uni Marburg	39	40	38	32	30	14
Uni Münster	42	42	35	42	38	36
Uni Oldenburg	21	21	21	20	20	15
Uni Osnabrück	14	14	14	18	28	35
Uni Paderborn	50	50	51	51	50	51
Uni Passau	31	30	30	36	34	32
Uni Potsdam	28	31	24	37	39	39
Uni Regensburg	17	17	13	17	18	24
Uni Rostock	45	46	49	46	45	40
Uni Siegen	29	29	28	30	27	23
Uni Stuttgart	47	45	52	43	43	45
Uni Trier	51	52	50	52	52	52
Uni Tübingen	15	15	17	15	16	19
Uni Würzburg	22	22	23	22	17	21
UniBW Hamburg	38	38	39	40	41	44

Notes:
A: Life Cycle 3 years
B: Life Cycle 5 years
C: Life Cycle 3 years with field correction (mean +3/-3 years)
D: Formula
E: Standard Approach: ranking within total Dataset
F: Standard Approach: simple average of productivity

References

Backes-Gellner, U. and A. Schlinghoff (2004): Careers, incentives and publication patterns of US and German (business) economists, *http://papers.ssrn.com/sol3/papers.cfm?abstract_id=616822*.

Baser, O. and E. Pema (2004): Publications over the academic life-cycle: Evidence for academic economists, *Economics Bulletin* 1, 1-8.

Bell, J. and J. Seater (1978): Publishing performance: Departmental and individual, *Economic Inquiry*, 16, 599-615.

Combes, P. and L. Linnemer (2003): Where are the economists who publish? Publication concentration and rankings in Europe based on cumulative publications, *Journal of the European Economic Association*, 1, 1250-1308.

Coupé, T., V. Smeets and F. Warzynski (2005): Incentives, sorting and productivity along the career: Evidence from a sample of top economists, *The Journal of Law, Economics and Organization*, 22, 137-167.

Dnes, A. and N. Garoupa (2005): Academic tenure, posttenure effort, and contractual damages, *Economic Inquiry*, 43, 832-839.

Goodwin, T.H. and R.D. Sauer (1995): Life cycle productivity in academic research: Evidence from cumulative publication histories of academic economists, *Southern Economic Journal*, 61, 728-743.

Frank, B. and G. Schulze (2000): Does economics make citizens corrupt?, *Journal of Economic Behavior and Organization*, 43, 101-113.

Grimes, P.W. and C.A. Register (1997): Career Publications and academic job rank: Evidence from the class of 1968, *Journal of Economic Education*, 28, 82-92.

Hall, B., J. Mairesse and L. Turner (2005): Identifying age, cohort and period effects in scientific research productivity: Discussion and illustration using simulated and actual data on French physicists, *NBER Working Paper No. 11739*.

Hartley, J.E., J.W. Monks and M.D. Robinson (2001): Economists' publication patterns, *The American Economist*, 45, 80-85.

Hutchinson, E.B. and T.L. Zivney (1995): The publication profile of economists, *Journal of Economic Education*, 26, 59-79.

Kalaitzidakis, P., T. Mamuneas and T. Stengos (2003): Ranking of academic Journals and Institutions in Economics, *Journal of the European Economic Association*, 1, 1346-1366.

Kenny, L. and R. Studley (1995): Economists' salaries and lifetime productivity, *Southern Economic Journal*, 65, 382-393.

Levin, S. and P. Stephan (1991): Research productivity over the life cycle: Evidence for academic scientists, *American Economic Review*, 81, 114-132.

Moore, W., R. Newman and G. Turnbull (2001): Reputational capital and academic pay, *Economic Inquiry,* 39, 663-671.

Oster, S.M. and D.S. Hamermesh (1998): Aging and productivity among economists, *Review of Economics and Statistics*, 80, 154-156.

Rauber, M. and H. Ursprung (2006): Life cycle and cohort productivity in academic economic research: Evidence for Germany, *mimeo, University of Konstanz*.

Stephan, P. (1996): The economics of science, *Journal of Economic Literature*, 34, 1199-1235.

Ursprung, H. (2003): Schneewittchen im Land der Klapperschlangen: Evaluation eines Evaluators, *Perspektiven der Wirtschaftspolitik*, 2, 177-190, 2003.

Life Cycle and Cohort Productivity in Economic Research: The Continental European Experience as exemplified by the case of Germany

2.2.1 Introduction

The science system in general and the economics profession in particular have in recent years become subjects of economic inquiry. Stephan (1996) surveys the *economics of science* literature of the first generation, and a survey summarizing the current literature dealing with the economics profession is to be found in Coupé (2004). Among the aspects that have received a great deal of attention are the methods of measuring research output, the identification of the determinants of research productivity, and the analysis of the job market consequences of research success.

As far as the job market consequences of research success are concerned, the available literature clearly indicates that research productivity as measured by publications and/or citations is a crucial determinant of salary (see, for example, Kenny and Studley, 1996, and Moore et al., 2001), tenure and academic rank (see Coupé et al., 2006), and the obtainable job status in terms of the employing university's reputation (see Grimes and Register, 1997, and Coupé et al., 2006). When it comes to identifying and explaining the pattern of research productivity over career time, the empirical evidence becomes less clear-cut. Human capital theory suggests a hump-shaped progression of individual research productivity since the stock of human capital (which drives high productivity) needs to be built up at the beginning of the career, and obsolescence of knowledge is likely to dominate the positive effect of increased experience towards the end of professional life.[43] A standard hump-shaped research productivity curve indeed emerges in some empirical studies investigating professional economists (see Kenny and Studley, 1996, Oster and Hamermesh, 1998, and Baser and Pema, 2004). It is, however, conceivable that the identified hump-shape represents an artifact of the *quadratic* specification of elapsed career time in the employed regressions of research productivity. Goodwin and Sauer (1995) identify a more complex career productivity profile that follows a fifth degree polynomial, whereas evidence uncovered by Hutchinson and Zivney (1995) and Hartley et al. (2001) do not indicate any significant decline in productivity as experience increases - a result that is compatible with the view that research behavior, rather than being determined by human capital considerations, can be explained by sociological factors related to social imprinting.

The social imprinting hypothesis suggests that significant variations in research behavior may be observed when comparing different cohorts of researchers. So far, however, the empirical studies have not uncovered strong cohort effects in the economics

[43] For a survey of the literature dealing with how life cycle productivity changes in response to changes in cognitive abilities, see Skirbekk (2003).

profession: Basar and Pema (2004) do not find any cohort effects at all, and Goodwin and Sauer (1995) report only marginally significant effects which, however, may well reflect the fact that the members of the analyzed cohorts differ in age, implying that the older cohorts are composed of academic survivors and thus liable to have been more productive on the average. Notice also that the hitherto available empirical evidence relates to the United States; studies relating to countries whose academic institutions were subject to major recent changes may exhibit substantially different cohort effects.

One of the objectives of our paper is to analyze how the traditional continental European science system has responded, on the one hand, to structural changes that provide more incentives for high research productivity, and, on the other hand, to the increased competition stemming from the fact that the European science system has become more and more exposed to the global standards set by the Anglo-Saxon system. We analyze this transformation process by scrutinizing the research behavior of different cohorts of scientists. For this purpose, the German economics profession appears to represent a very suitable example because in a large country such as Germany international competition was little noticed before the onset of European economic and political integration. Moreover, the German economic profession has for a long time been dominated by an idiosyncratic approach (the so-called historical school) that virtually decoupled the German profession from the emerging mainstream of economic research. The initial position of the German profession has thus been quite far removed from the mainstream. Moreover, the onset of the transition is reasonably close to the time period for which empirical evidence is available. Considering, finally, the success of the youngest batch of German economists in the global academic labor market, one can argue that the transition process is now coming to an end, implying that we can capture a substantial part of the whole adjustment process.

In this study we measure research output with the help of publications. There is general agreement that publications need to be adjusted for quality if they are used as indicators of research productivity. Two ways of controlling for publication quality have been employed in the literature: some scholars (for example Goodwin and Sauer, 1995) restrict themselves to articles published in a select list of highly reputable journals, whereas others (for example Kenny and Studley, 1996, and Coupé et al., 2006) base their measure of research productivity on a more encompassing list of journals and use explicit quality weights that are based on the respective journals' scientific impact. Hybrid approaches with two or more quality classes of journals are also quite common (see, for example, Grimes and Register, 1997, Oster and Hamermesh, 1998, and Moore et al., 2001).

Since research productivity consists of a quantity and a quality component, the identified career patterns can, in principle, be decomposed into a quantity and a quality cycle if the quality range of the considered journals is not overly restricted. Particularly interesting insights from quality-quantity decompositions refer to heterogeneity in research ability. It transpires that quality publishers are in general also quantity publishers (see Hutchinson and Zivney, 1995) and that the post-peak decline of the most prolific economists is much smaller than the decline of the less productive economists (see Grimes and Register, 1997). Oster and Hamermesh (1998) show that top producers keep on producing high-quality research, but at a slower rate, whereas the slowdown of second-rate economists leads them to publish in lower quality outlets. Truly creative economics at the highest level is, however, mainly undertaken by the young (see Oster and Hamermesh, 1998, and van Dalen, 1999).

A related strand of the literature investigates the impact of institutional features on the pattern of research productivity. Of special interest are the influence of entry barriers (such as the institution of the "habilitation" which is still exercised in some continental European countries), mid-career hurdles such as tenure and rank promotions, and also institutional provisions that affect the mobility of academic researchers between universities.[44] Entry and promotion barriers have typically been portrayed as contests designed to induce higher research effort via increased competition (see Backes-Gellner and Schlinghoff, 2004, Coupé et al., 2003, and Dnes and Garoupa, 2005). The empirical evidence indicates that these institutional provisions do indeed work as incentive schemes and thus influence the pattern of research productivity: those life-cycle studies that identify hump-shaped productivity patterns usually find that research productivity peaks about six years into the professional career, i.e. around the time when professors can apply for tenure. The post-tenure decline in productivity appears however to be rather small (see Bell and Seater, 1978, and Hutchinson and Zivney, 1995). Somewhat more informative results emerge from micro-econometric studies using information about when exactly the individual researchers were promoted: Backes-Gellner and Schlinghoff (2005) uncover strong evidence for the United States and Germany indicating that promotion tournaments give rise to an increase in research productivity before promotion and a lapse of productivity afterwards. Moreover, they show that the career profiles of German economists is

[44] Such provisions can either be designed to restrain mobility (examples are lock-ins via retirement benefits and German-type cartel agreements among university presidents or their superiors in the respective governments) or to increase mobility (international mobility of researchers is promoted, for example, with the help of the Marie Curie Actions organized and financed by the European Commission).

characterized by a more pronounced post-tenure decline than the profiles of their American colleagues, the reason being that the German university system lacks a second career step, namely promotion to full professor. Analyzing publication records of 650 economists who are members of the top-1000 group according to a world-wide ranking, Coupé et al. (2006) corroborate the result that promotions cause cyclical deflections in research productivity: pre-promoted economists are more productive than post-promoted ones, and tenure has an additional negative effect on research productivity.

The focus of our study is however not on the institutional features of the German academic labor market. We rather treat career steps as an endogenous to academic careers and relate individual research productivity to career age. The chapter unfolds as follows. In section 2.2.2 we describe our data set and in section 2.2.3 we present our base-line estimates of the life cycles in research productivity. We identify life cycles that are akin to, but rather flatter than the life cycles of American economists uncovered by Goodwin and Sauer (1995). Moreover, we arrive at the result that the German profession is characterized by significant cohort effects in research productivity. We also find that the shape of the life cycles depends on the individual researchers' ability. Studies focusing on aggregates thus miss an essential part of the story that relates to heterogeneity. In section 2.2.4 we then go on to investigate cycles in the constituent parts (quantity, quality, number of co-authors) of our measure of research productivity. Section 2.2.5 concludes.

2.2.2 The Data

2.2.2.1 The sample

Whereas many other bibliometric studies focus on researchers who publish frequently, our dataset compromises, in principle, all German academic economists. Our dataset encompasses 699 economists who received their doctoral degrees between 1963 and 1998 and who were employed by a German university in the year 2004 or had retired from a German university briefly before.[45] The youngest economists in our sample thus have a minimum of six years of post-Ph.D. experience.

Our study relies on the *EconLit* data base that contains journal publication records from 1969 onwards. In choosing the starting year of 1963 we thus lose only the first six years of the 1963-1968 cohort. We measure research output exclusively on the basis of the

[45] We gathered information on more than one thousand German economist. Our sample comprises however only those economists who obtained their doctoral degree after 1962 and for whom we could actually establish the exact year in which they obtained their doctoral degree. We thank Robert Hofmeister and Frieder Mokinski for valuable research assistance.

journal literature. This admittedly neglects other types of research outlets such as monographs and articles published in collected volumes and proceedings. We are, however, in accord with most scholars in the field who are confident that *EconLit* indexes the most important journals of the economics profession and that the articles published in these journals together constitute the lion's share of economic research (see, for example, Hartley et al., 2001, Combes and Linnemer, 2003, and Coupé, 2003).

We collected all *EconLit*-listed journal publications authored or co-authored by the economists included in our sample up to the year 2004 and linked the annual records to the year in which the author obtained his or her doctoral degree.[46] We were thus able to establish individual life cycles of research productivity for a large number of German economists. These life cycles represent the basic input for our empirical analysis.

Only 7 percent of the 699 economists in our sample are women. Fourteen percent specialize in microeconomics, 27 percent in macroeconomics and international economics, 35 percent in public economics and 16 percent in econometrics. Economists who could not be assigned to one of these fields were assigned to the field OTHER. Interestingly, 94 or about 13 percent of the economists in our sample have never published in an EconLit-listed journal.

2.2.2.2 The dependent variable: Individual annual research productivity

EconLit indexes these days over 800 journals. It is quite evident that the quality standards set by these journals are quite diverse. As a consequence, publication-based bibliometric measures need to control for journal quality. This can be done by restricting the set of journals. We do, however, not believe that this is a viable strategy of measuring research output because a robust research indicator needs to draw on all available information. Using, for example, only a relatively small number of top-journals would bias the indicator in favor of top-researchers specializing in hot topics. Moreover, life cycle patterns in research quality (as compared to cycles in overall output) can only be properly identified if the whole quality range of research products is taken into account.

To control for the quality of the journals indexed in *EconLit* we settled for a standard method proposed by Combes and Linnemer (2003).[47] Their "CLpn" scheme weighs quality according to the respective journal's reputation and impact, and converts research output in

[46] Whenever *EconLit* reported "et al." we identified the hidden co-authors by tracing the article.
[47] One disadvantage of using the CLpn scheme is that journal quality is kept constant over the period of investigation that covers, after all, a time-span of 36 years. Concerning this, the discrete nature of the scheme might be advantageous.

standardized units of AER-page equivalents by also taking into account the number (p) of pages and the number (n) of co-authors. The imputed quality weights lie between unity for top journals and one twelfth for journals with the lowest quality standards. The top-tiered journals are the *American Economic Review*, *Econometrica*, the *Journal of Political Economy*, the *Quarterly Journal of Economics* and the *Review of Economic Studies*. Sixteen journals receive a weight of two thirds. Weights then decline in discrete steps (one half, one third, one sixth) down to the minimum weight of one twelfth.

To construct our dependent variable, the number of pages of each article is multiplied by the respective CL journal weight and this product is then divided by the number of authors. Adding the scores calculated according to this rule over all articles published by researcher i in year t, we arrive at our basic research productivity measure. To check for the robustness of our results, we have, however, also used an alternative journal-quality weighting scheme. We will explicitly refer to this robustness check when we discuss the respective results.

2.2.2.3 The explanatory variables

To identify life cycle patterns in individual research productivity we regress our dependent variable, research productivity of researcher i at time t, on several independent variables, the most important one being experience or career-time.

Experience

In accordance with the literature we align all individual life cycles by using as the reference year the year in which the researchers obtained their doctoral degrees. In our regressions we do, however, also include the research output generated in the pre-Ph.D. years by letting the life cycles begins five years before the reference year zero. To estimate the shape of the lifecycles we include career-time polynomials of different orders in the regressions. Simple t-tests as well as likelihood ratio-tests were used to determine the optimal degree of the polynomial. In most cases a 5^{th} degree polynomial has proven to fit the data best.[48]

Individual heterogeneity

It cannot be ruled out that publication habits vary across different fields of research. A simple comparison of the average yearly per capita research productivity across different

[48] Goodwin and Sauer (1995) come to a similar conclusion using their data on US economists.

fields reveals that this conjecture cannot be easily dismissed: these productivities range between 2.04 AER-equivalent pages in microeconomics and 0.42 AER-equivalent pages in our remainder group OTHER. We decided therefore to include the *field of research* as a dummy variable to allow for different research cultures across fields. A second reason for including field dummies is that these variables would also capture any bias stemming from an uneven coverage of the research fields in the *EconLit* data base as a whole and/or within each quality-group of journals. The interpretation of field-specific effects on research productivity is therefore not straight-forward.

The gender issue has for a long time played a major role in labor economics and has, as a consequence, been taken up also in several studies of research productivity. We follow this tradition and include a gender variable that may capture gender specific differences in research productivity.

Cohorts and historical time

Research productivity may not only vary across different fields of economic research but also across historical time. To allow for vintage effects we include cohort dummy variables in our specification. They are constructed by using the reference year in which the researchers obtained their doctoral degrees, starting in 1963.

A second possibility is to include a time trend in the regressions. Just as cohort dummies, a time trend will capture changes of research behavior across historical time. Whereas cohort dummies portray changes in research behavior that are peer-group specific (they could, for example, portray different cultural imprinting patterns across time), a time trend indicates that individual research productivity does change over time for *all* researchers and this change is independent of experience. Such time trends might capture changes in publication customs, for example a substitution away from monographs and collected volumes towards journals. Unfortunately, a separate identification of linear cohort effects and a linear time-trend appears not to be possible since the difference between historical time and career age is used to assign the individual researcher to a cohort. Imposing specific functional forms to separate the two effects appears to be a rather dubious strategy because there are no obvious restrictions that could be imposed.[49] In the following

[49] Different restrictions and even small specification errors might have large effects on the estimates. For a discussion see Rodgers (1982). By including a time trend in addition to our cohort dummies we therefore would not gain much additional insights since the estimated effects would solely depend on the underlying functional forms and should therefore not be interpreted.

section we therefore present first our regression results that do not control for historical time effects, and then deal with the historical time problem explicitly in the subsection 2.2.3.2.

2.2.3. Results

2.2.3.1 Identifying life cycles in research productivity

The set of explaining variables of our base-line lifecycle regressions consists of a career time polynomial, the cohort dummies, the gender dummy, the field dummies and a constant. The dependent variable Y_{it} represents individual i's research productivity as measured by our productivity index at career-time t. Because of the high degree of censoring (about ¾ of our Y-observations are zeroes) we cannot apply OLS and have to rely on techniques which can properly accommodate heavily censored data sets.

The results are summarized in Table 2.2.1. In the first column we present Tobit estimates using a 5th degree polynomial for career time. Column two shows the results obtained from a model presupposing an exponential conditional mean function that is estimated via nonlinear least squares:

$$E(Y_{it} \mid x_{it}) = \exp(x_{it}'\beta).$$

In the following two columns we present estimates of a hurdle model. The hurdle model assumes that the decision to undertake research at all might be driven by other forces than the decision with respect to how much research effort is expended by an active researcher and is therefore parametrically richer than the Tobit model. We model the two stages of the decision making process as follows:

$$\Pr(Y_{it} = 0) = \exp(-\gamma_{it})$$

$$\Pr(Y_{it} = y \mid x_{it}, y > 0) = \frac{\lambda_{it}^{y}}{y!(1-\exp(-\lambda_{it}))}\exp(-\lambda_{it}),$$

$$\text{where } \lambda_{it} = \exp(x_{it}'\beta), \quad \gamma_{it} = \exp(x_{it}'v)$$

The occurrence of non zero counts is modeled via a Poisson probability specification and conditional output is described using a truncated Poisson density.[50] The Poisson model appears to be appropriate since the observed density distribution of our dependent variable

[50] We also estimated a Negative Binomial specification for the conditional output (see Pohlmeier and Ulrich, 1995, for an example of a complete NegBin hurdle specification). The resulting estimates are well in line with the estimates of the Poisson specification.

resembles the pattern of count data. This resemblance (spikes at steps of one twelfth) emerges because the CLpn-index is based on journal weights that are multiples of one twelfth. To arrive at proper count data we divided our dependent variable by one twelfth and rounded to the next integer. The transformed variable can then, of course, be analyzed by using a count data model in which one count can be interpreted as one twelfth of an AER-equivalent page or one page published in a journal of lowest quality.[51]

In the last column of Table 2.2.1 we present Tobit estimates if the weighting scheme of journal quality underlying the Combes and Linnemer (CL) measure of research output is replaced by the "KMS" weighting scheme proposed by Kalaitzidakis, Mamuneas and Stengos (2003) which covers only 159 Journals and gives top-journals a much larger weight than the CL scheme. The *Journal of International Economics* for example, the most highly regarded journal in its field, is allocated a relative weight of 2/3 (as compared to the *American Economic Review*) in the CL scheme, whereas it receives a relative weight vis-à-vis the AER of only about 8% in the KMS scheme.

The estimates presented in Table 2.2.1 are not obscenely at variance with the standard life cycle hypothesis. As can be seen from the panels in Figure 2.2.1, the estimated career-time polynomials imply in each case a hump-shaped curve of research productivity over career time. All models fit best with a life cycle polynomial of degree five which gives rise to a peak in research productivity which occurs around the eights career year, i.e. when German academic economists are usually promoted to full professor. Even though the standard life cycle hypothesis passes the test reasonably well, we do not find a marked and final decline in research productivity after the initial peak. Research productivity rather appears to remain quite constant over a substantial part of the lifecycle which implies that our estimates may just as well be construed to support the sociological hypothesis of imprinting. The increase in research productivity towards the end of the researchers' careers identified by both the exponential model and the hurdle model is in line with the results presented by Goodwin and Sauer (1995). Their estimates for American economists show however a more substantial decline in research productivity during the mid-career years. Interestingly, the hurdle specification indicates that the probability of undertaking research and conditional research output follows somewhat different time patterns.

[51] To check for the robustness of our results we additionally used ½ and ¼ of an AER-equivalent page as count units without obtaining significantly different results. However, since the underlying density has spikes at steps of one twelfth, the applied scheme appears to be more natural and precise. As an additional robustness check of our specification we estimated a hurdle model which assumes a lognormal distribution of the positive scores of Y (see Wooldridge, 2002). The results are similar to the ones obtained from the count data hurdle model presented above.

Our estimation results documented in the first four columns of Table 2.2.1 do not appear to depend on the employed CL method of measuring research output. When research output is measured with the help of the KMS method, the Tobit estimates survive the robustness check with flying colors (see column 5, Table 2.2.1, and Figure 2.2.1, panel 4). Also the hurdle model yields fairly similar results when the KMS measure is used (estimation results not shown); the career patterns documented in panel 5 of Figure 2.2.1 reinforce our conclusion that one can lose important information when estimating publication incidence and conditional output together. The more marked drop in the conditional output is a consequence of the more top-heavy KMS quality weighting scheme and represents a clear sign that research quality changes as the economists' careers progress. We will pursue this hypothesis further in section 4.

The coefficient of the gender dummy FEMALE indicates that female economists publish significantly less than their male peers. The hurdle model reveals however that this negative effect seems to be mainly due to the decision to engage in research activities at all rather than a consequence of a lower productivity of female economists who are active researchers. Tobit estimates of separate career time polynomials for male and female economists (estimation results not shown) show that female economists suffer a drop of research productivity beginning in their eight's career year when they are about 38 years old; they appear however to recuperate around the 18th career year when they are about 48 years old (see Figure 2.2.1, panel 6). This well squares with the interpretation of a "maternal leave from research", especially if one allows for a one or even two years publication lag.[52]

As expected, the coefficients of the cohort dummies increase over time.[53] We interpret this result to imply that members of younger cohorts are more productive researchers than their older peers. About the reasons for this phenomenon we can only speculate: the evidence certainly does not contradict the hypothesis that over the last thirty years the German economics profession has increasingly been exposed to the Anglo-Saxon research tradition that stresses the requirement to document one's research efforts on a continuous basis. Many economists who returned in the 1970s and 1980s from the UK and the US were instrumental in sharing their experience with their graduate students who

[52] Notice, that estimates of the career patterns of female economists are based on a rather small number of observations and need, therefore, to be interpreted with caution. Moreover, it is worth pointing out that even though the point estimates of the life cycle polynomials differ, this difference is statistically not significant. This might also be due to the small number of female economists.

[53] We estimated our baseline regression (column 1) also with ten year cohort dummies as well as a polynomial specification of the cohort effects. The results are in line with the results presented here.

internalized this research culture which nowadays characterizes the academic environment at German graduate schools and dominates the increasingly competitive hiring strategy employed by the leading departments. In order to scrutinize the determinants of the identified cohort effects, we analyze below cohort-specific lifecycles which will be more informative than the pooled life-time productivities presented so far.

Although the Tobit estimates seem to be well in line with the results of the other estimators, a test for heteroscedasticity and a Pagan and Vella (1989) conditional moment test on normality of the underlying disturbance reject the hypotheses, thereby casting doubt on the applicability of this estimator. This caveat probably does not come as a surprise, considering the count data character of the publication process. Since, however, the Tobit estimates are in accordance with to the other estimates this may be interpreted as a sign of the robustness of our results. Nevertheless, we now proceed to employ more robust econometric techniques.

2.2.3.2 Quantile regressions and cohort-specific life cycles

The semi-parametric censored quantile regression estimator for censored data developed by Powell (1984 and 1986) is more robust than the estimators used above because it allows the error terms to be heteroscedastic and non-normally distributed. Since we have to reject both of the hypotheses we now employ this estimator. We estimate the 75, 80, 85, 90 and 95 percentiles. The results are depicted in the first panel of Figure 2.2.2 and the estimates are presented in Table 2.2.2 for the 85 and 95 percentiles. The figure reveals that the most productive researchers are much more productive than the less productive ones; more precisely, the line-up of individual research productivities (from low to high) is heavily skewed to the right. This feature (which is reflected in the figure by the fact that the distance between the percentile lines becomes increasingly larger) is a stylized fact of all distributions of research productivities. More interesting is perhaps the fact that this skewness appears to be pretty stable over career time. Most important for our argument is however that our main results gleaned from the less robust estimation models presented in the previous subsection are confirmed.

Until now we used the whole sample of economists to estimate the shape of the productivity life cycles, allowing only for cohort-specific constant terms. It would, however, not be farfetched to assume that over the last thirty-five years the shape of the lifecycles may have been subject to significant changes. As we have argued above, increased competition between researchers or other institutional changes may have influenced

research behavior. To account for this possibility we allow for separate time polynomials for each cohort using censored quantile regressions. The results are documented in Table 2.2.2 columns 3-7, and in the second panel of Figure 2.2.2. This figure reveals that the productivity lifecycles of younger cohorts are – as far as one can tell from the initial phases of these cycles - more hump-shaped than those of earlier cohorts.[54] The research behavior of the younger German economists thus appears to be much more in line with the predictions of the standard human capital approach to explaining changes in labor productivity than the research track record their older peers. This evidence supports the hypothesis that the German academic environment has become increasingly competitive over the last 35 years, with the consequence that the academics who entered the market later were forced to adopt a more and more narrow optimizing behavior if they wanted to be successful. As a result, the younger economists' research behavior corresponds more closely to the predictions of the incentive-centered human capital approach to explaining labor market outcomes.

So far we have not exploited the panel structure of our dataset. To account for the multitude of fixed effects which are specific to the individual researchers, we use a quantile estimator due to Honoré (1992). It is semi-parametric and therefore robust with respect to distributional assumptions and generalizes our results in the sense that we now explicitly take individual specific fixed effects into account. The estimates are presented in Table 2.2.3.[55] As can be seen in Figure 2.2.3, the shape of the estimated career-time polynomial compares well with the pooled quantile estimates presented in Table 2.2.2.[56] Our previous results thus pass this robustness test with flying colors.

At this stage a caveat is called for. Over the last thirty-five years the publication habits in the German economics profession may have changed. If it is true that monographs and articles in collected volumes have become less important research outlets as compared to journal publications, this substitution process might be responsible for the large cohort effects we observe in our empirical analysis.

There is some evidence that the ratio of the number of active researchers and the number of journals slightly decreased over the last thirty years.[57] The crucial question thus is whether younger economists indeed publish significantly more than their older peers or

[54] In an earlier paper (see Rauber and Ursprung, 2006) we have identified the same pattern of cohort-specific lifecycles by including cohort-specific career-time polynomial in a standard Tobit regression.
[55] We used the PANTOB estimation program written by Bo Honoré and J. Campbell.
[56] Note, however, that the *levels* of the cycles cannot be interpreted because they are governed by the individual effects. The stacking of the curves in Figure 3 therefore serves only to illustrate the results.
[57] Goyal, van der Leu and Moraga-Gonalez (2004) count the number of authors in EconLit: 33770 in the 70's, 48608 in the 80's and 81217 in the 90's. The number of journals indexed in EconLit in 1975, 1985 and 1995 is 200, 311 and 535, respectively. Calculating author per journal ratios yields: 168.85, 156.3 and 151.

whether the cohort effects identified above are simply due to dramatic changes in the publication habits. To disentangle these effects we follow the general approach advocated by Rodgers (1982): Since identification by functional form is rather arbitrary and even small specification errors might lead to large differences in the estimates, we make use of a proxy variable for the prevailing publication habits.

To pin down the development of publication habits we roughly estimated the co-citation patterns of journal articles vis-à-vis other journal articles, articles in collected volumes, and monographs.[58] The percentage share of journal citations in journal articles indeed increases over time. The estimated time trend is depicted in Figure 2.2.4.[59] Since our measure of research output is truncated we cannot apply a trend correction before estimation. Such a procedure would not allow correcting for the marginal density of publication incidence. We therefore apply our correction after estimation, i.e. we rescale the estimated research life cycles of each cohort by multiplying each year's estimated output by the ratio of the journal citation share in 2004 (which amounts to 60%) and the journal citation share in the respective year. We thereby obtain a correction which represents an upper limit of the substitution of research output towards the learned journals.

The result of our correction exercise is depicted in the second panel of Figure 2.2.4. Even though we biased our test against the hypothesis of significant cohort effects in research productivity, this hypothesis survives the test easily. In other words, the younger German economists are so much more productive in producing journal articles than their older peers that the implied superiority in research productivity cannot be contested by any reasonable correction for the observed changes in publication habits.

2.2.3.3 Ability-specific life cycles

We now turn to analyzing to what extent the life cycles vary across groups of different academic achievements. This focus distinguishes our study from all those studies that investigate only a subset of highly productive individuals. Since we deal with many different types of researchers, we now relax the constraint of a uniform career-time polynomial for all individuals. We do so by applying a mixture model in the first stage of

[58] We based our investigation on a random sample of articles published in the *American Economic Review*, *The European Economic Review*, *Public Choice* and the *Jahrbücher für Nationalökonomie und Statistik* an calculated the citation-shares of journals, monographs, collected volumes, working papers, and statistical sources for the years 1969, 1978, 1987, 1996 and 2005. In total we classified 8824 citations.

[59] We approximated the original time series by fitting a quadratic polynomial.

the hurdle model.[60] The combined model performs two tasks: first, it endogenously allocates each researcher to an appropriate (ability-) group and, second, it estimates the life cycle parameters for each group. The likelihood function of the mixture model has the following appearance:

$$f(y_i \mid w_i) = \sum_{k=1}^{K} \Pr(C_i = k) \Pr(Y_i = y_i \mid C_i = k, W_i = w_i)$$

.

$W_i = (w_{i1} ..., w_{iT})$ represent all time-dependent covariates which are captured by a time polynomial up to order four. The probability of group affiliation is specified by a generalized logit function and the conditional output density follows a censored normal distribution. We estimated the mixture model separately for each cohort and assigned individuals to two different groups.[61] In each cohort, the model clearly identifies two different groups: *journeymen* researchers and *accomplished* researchers. About two thirds of all individuals are assigned into the journeymen group whereas one third is assigned into the accomplished group.

Since two different types of economists are identified, we re-estimated our hurdle model - this time allowing for separate lifecycles for the two groups. The results are depicted in Figure 2.2.5. We observe that the probability of publication as well as the conditional output clearly differs across the two groups. More productive researchers appear to have incentives to publish a steady stream of papers until the end of their career. Reputation or intrinsic motivation might be a reason. Moreover, the conditional output of the accomplished researchers stabilizes at a much higher level than the output of the journeymen researchers. Needless to say, that these differences cannot be identified by simply focusing on a subset of highly productive individuals.

2.2.4. An exercise in deconstruction: Quality, quantity and co-authorship

Up to now we have treated research productivity as measured by the CLpn index as a preordained unit of account. The shapes of the identified life cycles suggest however that the constituent parts of this productivity measure might follow quite different patterns that

[60] A different method which yields similar results is applied in Rauber and Ursprung (2006).

[61] Cohort specific estimation ensures that group assignment is not driven by the cohort effects which we identified above. We used the SAS estimation procedure TRAJ which maximizes the joint likelihood of the mixture model. For estimation details see Jones, Nagin and Roeder (2001). We settled here for two groups because more groups would give rise to an insufficient number of members in some of the groups and because such a division is also indicated by the Bayesian Information Criterion.

cannot be uncovered by an investigation at the aggregate level. In this section we therefore deconstruct the employed index and focus our investigation on the constituent parts thereof, namely on quality, quantity and the number of collaborators. In order to identify life-cycle patterns in these constituent parts of research productivity we "deconstruct" the density of our dependent variable in the following way:

$$f(Y) = f_E(E \mid \theta_E) \cdot f_N(N \mid E=1) \cdot f_C(C \mid N, E=1) \cdot f_Q(Q \mid C, N, E=1)$$

where $Y = E \cdot N \cdot C \cdot Q$.

The first factor on the RHS captures whether economist i has been involved in producing research output in year t or not. The second marginal density represents the *number of publications* given that at least one publication has been produced in t. The third factor denotes the average quantitative *contribution* per article (number of pages per coauthor) and the fourth factor the average *quality* of the articles authored or co-authored by economist i in year t. An exemplary deconstruction of the score can be found in the Appendix.

The first column in Table 2.2.5 presents the regression for the number of authored or co-authored journal articles. As can be seen from the first panel of Figure 2.2.6, this number reaches a first maximum approximately seven years after German economists are granted their doctoral degrees and remains thereafter more or less constant for about ten years. Around the middle of the career the number of publications begins to increase again and continues to do so until about five years before retirement. We will show below that this second increase is due to a higher co-authorship incidence of older economists. Whereas young economists appear to write most of their articles by themselves (and therefore publish only a few), older researchers tend to publish together with co-authors and therefore put their names on a larger number of papers. This increase in co-authorships might be either due to network effects, or to the fact that senior economists more often write joint papers with the doctoral students they supervise.

In the second column of Table 2.2.5 our dependent variable is the logarithm of average research quantity (number of pages per article divided by the number of authors). Explanatory variables are our usual independent variables and the number of articles authored or coauthored. An inspection of the second panel of Figure 2.2.6 reveals that the average contribution per paper declines after an early career peak. The increased incidence of co-authorships of course contributes to the decline after the first peak. It thus transpires

that at the beginning of their careers, economists, conceivably for reputation reasons, focus their research activity on relatively few projects that are pursued without collaborators, whereas at later stages they tend to spread themselves wider and prefer to engage more in collaborative research endeavors. The minor peak that can be found before retirement may be an artifact of the econometric specification, but it may also reflect a certain leaning of older economists to busy themselves with sweeping themes that require a lot of space to be developed.

The third and arguably most important constituent part of our measure of research productivity is (average) quality. Our regression results for the average research-quality variable are summarized in the third column of Table 2.2.5. We regress the logarithm of the average quality on our independent variables, the number of articles published and the logarithm of the average research quantity.[62]

As far as the "average economist" is concerned, it is fair to say that not only overall research *productivity* but also average research *quality* follows a hump-shaped lifecycle: the average quality sharply increases at the very beginning of the career as the budding economists become increasingly accomplished, but begins to decline already around the twelfth career year when the average German economist is about 42 years old. Since the average economist's lifecycle in research productivity is relatively flat as compared to the identified lifecycle in research quality, this indicates that quantity is substituted for quality as the economist's career progresses. In order to check whether this substitution process is ability specific, we estimated ability specific quality lifecycles by using a procedure advocated by Goodwin and Sauer (1995): We defined quintile ranks according to average lifetime productivity within each three years cohort of researchers. We then assigned each researcher the appropriate rank and included for the first, second, and the bottom three ranks separate career-time polynomials as well as researcher fixed effects in the regression.[63] In contrast to the endogenous grouping presented before, this procedure allows us to focus specifically on highly accomplished researchers. The results presented in panel 4 of Figure 2.2.6 indicates that top-performers are able to keep up research quality much more than their less gifted peers: the relative drop between the career years 10 and 30 amounting to about

[62] The density of the quality variable is centered around the discrete steps of the underlying weighting scheme. To check for the robustness of the results we also transformed our quality measure into a variable that can assume six different values that correspond to the original journal quality weights. We then applied an ordered probit model to estimate the underlying quality lifecycle. The results are in line with the linear regression results presented here.

[63] We bundled the bottom groups because of the high degree of censoring and because our main focus is on the high rate publishers. We excluded the oldest cohort because for these researchers we do not observe the first six post Ph.D. years and our ability indicator would therefore be biased.

13% for the top researcher and 30% for the accomplished and journeymen researchers.[64] These results lend strong support to our notion that when measuring research productivity over the lifecycle it is imperative to include all types of journals; employing bibliometric approaches that focus on a subset of prime-rate journals cannot detect patterns of research behavior that involve substitution of quantity for quality.

We now, finally, return to our hypothesis maintaining that co-authorship becomes more attractive as the average economist's career progresses.[65] To explore this hypothesis in more detail, we construct a co-author index measuring each economist's average number of collaborators (including him- or herself), by using the number of pages as the respective weight for each journal article published in the respective year. The regression explaining the number of co-authors is presented in the forth column of Table 2.2.5. The implied life-cycle is depicted in the fifth panel of Figure 2.2.6. This figure reveals that the number of co-authors is relatively high for graduate students and reaches a minimum about three years after economists are conferred their doctoral degrees. Afterwards the number of co-authors steadily increases over the whole life-cycle. This piece of evidence points towards network advantages of more mature economists and, as far as the odd early-career twist is concerned, to a high incidence of collaborative efforts between graduate students and supervisors.

The last regression presented in Table 2.2.5 re-estimates the impact of our explanatory variables on the average quality of research without conditioning on the length or number of articles. As compared to the former regression we also included here our index of the average number of co-authors. It transpires that quality indeed depends on the number of collaborators: working with other scholars appears to increase research quality.

2.2.5 Conclusions

In investigating the careers of German academic economists we have come across two characteristics that we regard to be essential for our understanding of the profession. First, we discovered that the pattern of research productivity over the life cycle is co-determined by economic incentives and by sociological factors. The influence of the economic incentives is reflected in the hump-shape of the identified life cycles, the sociological factors show up in the marked cohort effects. As compared to the lifecycles of their American peers, the life cycles of German economists turn out to be flatter and the level of

[64] Oster and Hamermesh (1998) arrive at a similar result.
[65] There is a small literature on the topic of co-authorship; see, for example, McDowell and Smith (1992), Hollis (2001) and Laband (2002).

research productivity appears to depend much more on cohort specific factors. We do, however, not interpret these finding as evidence supporting the hypothesis that the American profession is mainly driven by economic incentives and the German profession by sociological factors. Our results simply reflect the fact the academic environment in Germany has changed much more dramatically over the period of our investigation than the science system in the United States.

The second uncovered characteristic of the economics profession that deserves special attention is the fact that lifecycles in research productivity are ability and gender specific. Studies that attempt to identify the research behavior of the "representative" economist miss a large part of the story. The economics profession is very heterogeneous and neglecting this heterogeneity may give rise to severe misinterpretations. It is worth emphasizing that this heterogeneity in ability not only affects the variance of the *level* of individual research productivity (this we have known for a long time from various ranking exercises), heterogeneity also has distinct effects on the dynamic dimension of research productivity, i.e. on the *shape* of the individual life cycles. The ability-induced variation in life cycle patterns is especially striking when one compares life cycles in the *quality* of research.

As mentioned above, the fact that the life cycles in research productivity turn out to be rather flat in the German profession lends some support to the sociological imprinting hypothesis. This does, however, not imply that economic incentives are of second-order importance. Career hurdles, for example, may well provide incentives which have a great deal of influence: since we find early career peaks that appear to coincide with the timing of the only career hurdle in the traditional German science system, our results are certainly compatible with the existence of pre-tenure peaks and post-tenure kinks, and thus with the results derived by Backes-Gellner and Schlinghoff (2004). Moreover, we have found strong evidence suggesting that the marked increase in average research productivity across cohorts has been accompanied by a significant change in the career profiles: the research-productivity lifecycles of the youngest German economists closely resemble the lifecycles of their Anglo-Saxon peers. This implies that the observed process of catching-up with the most productive research systems is about to be accomplished not by changing the behavior of the profession at large but rather by letting the new generation of economists grow into an academic environment in which the behavior of the researchers is guided by economic incentives. Economic incentives thus do not appear to change accustomed behavioral patterns; incentives do however influence the behavior of the incoming generations of

scientists. We thus arrive at the conclusion that research behavior is co-determined by economic and sociological factors. If one attempts to make a national science system catch up with the frontier of research, a time span of two generations appears to be a minimum: one generation to transfer and implement the new spirit accompanied by the requisite institutional reforms, and one generation to overcome the acquired habits.

Finally, we would like to highlight a management consequence that arises from this study. Since life-cycle and cohort effects turn out to represent major determinants of research production in Germany, this information should be taken into account not only on the occasion of evaluating individual researchers, but also when one attempts to rank university departments, the reason being that the exogenous age and cohort structure of the departments significantly affects the observed research productivity. It therefore appears to be obvious that these effects should be deducted from the gross amount of research produced if one attempts to fairly represent a department's research standing. Even though adjustments for career-age have been made in the ranking literature (see, for example, Combes and Linnemer, 2003), these adjustments were up to now based on an ad hoc reckoning. Our empirical study provides the kind of information that would have to be used in more sophisticated rankings. Our companion paper (2006) presents a new ranking methodology which incorporates these lifecycle and cohort aspects.

2.2.6 Tables and Figures

Table 2.2.1: Pooled estimates

	Tobit	Hurdle Model			Tobit Kalaitzidakis weights
		Conditional Exponential Mean (NLS)	Probability	Cond. Mean	
T	1.9989***	0.6394***	0.3885***	0.1819***	93.952***
	(17.12)	(11.26)	(19.23)	(7.28)	(10.16)
$T^2/10$	-2.413***	-0.9498***	-0.5097***	-0.2608***	-124.101***
	(11.40)	(9.58)	(13.07)	(5.37)	(8.20)
$T^3/100$	1.2635***	0.5809***	0.2908***	0.1519***	70.560***
	(7.76)	(8.25)	(9.47)	(3.90)	(6.44)
$T^4/1000$	-0.2989***	-0.1550***	-0.0746***	-0.0386***	-18.142***
	(5.61)	(7.19)	(7.34)	(2.93)	(5.20)
$T^5/10000$	0.0260***	0.0149***	0.0070***	0.0035**	1.7090***
	(4.22)	(6.31)	(5.96)	(2.26)	(4.30)
C6974	0.6744	0.0874	0.1945	-0.1704	-7.5984
	(0.76)	(0.31)	(1.18)	(1.17)	(0.15)
C7580	3.1352***	0.4203	0.5938***	0.0202	106.020*
	(3.12)	(1.55)	(3.44)	(0.13)	(1.90)
C8186	5.3574***	0.9202***	1.0030***	0.1052	213.806***
	(5.23)	(3.48)	(5.68)	(0.72)	(3.62)
C8792	6.9861***	1.1961***	1.2231***	0.2166	276.024***
	(7.74)	(4.88)	(7.80)	(1.63)	(5.16)
C9398	8.5998***	1.3649***	1.4860***	0.2883**	325.609***
	(9.89)	(5.79)	(9.43)	(2.39)	(6.42)
MICRO	8.0980***	1.3866***	1.3753***	0.4562***	512.220***
	(6.52)	(5.05)	(6.65)	(2.78)	(6.20)
MACRO	5.9956***	1.0372***	1.1500***	0.1440	324.030***
	(5.45)	(3.89)	(5.78)	(0.93)	(5.15)
PUBLIC ECONOMICS	4.8836***	0.9315***	0.9688***	0.1329	273.39ß***
	(4.38)	(3.46)	(4.81)	(0.85)	(4.40)
ECONOMETRICS	3.6397***	0.6340**	0.7692***	-0.0122	264.470***
	(3.07)	(2.18)	(3.51)	(0.07)	(3.87)
FEMALE	-3.8694***	-0.5504**	-0.6591***	-0.1345	-182.122***
	(4.57)	(2.54)	(4.45)	(1.08)	(3.53)
CONSTANT	-18.911***	-2.4434***	-3.6847***	3.4968***	-1098.304***
	(13.57)	(6.98)	(15.64)	(18.85)	(10.21)
Observations	18610	18610	18610		18610
(Pseudo)-R^2	0.0597	0.1806			0.0338
Log Likelihood	-20229.9		-134262.8		-23384.8

Notes: Absolute t-values in parentheses, based upon a clustering robust covariance-matrix (on individual level). *** denotes significant at the 1 percent level, ** at the 5 percent level and * at the 10 percent level (significance levels apply to all tables).

Table 2.2.2: Quantile estimates

	Censored Quantile regressions		Censored Quantile Regressions: Estimates by Cohort (0.85- Percentile)				
	0.85	0.95	1969-74	1975-80	1981-86	1987-92	1993-98
T	2.1676***	2.8923***	0.853***	0.586***	1.519***	2.176***	2.707***
	(14.49)	(19.21)	(4.13)	(4.06)	(6.57)	(5.36)	(7.21)
$T^2/10$	-2.8526***	-3.7942***	-0.864***	-0.333**	-1.340***	-2.000***	-2.121***
	(10.04)	(10.03)	(3.51)	(2.39)	(4.83)	(3.03)	(4.88)
$T^3/100$	1.6422***	2.1279***	0.350***	0.054	0.353***	0.450	
	(7.61)	(6.51)	(3.18)	(1.53)	(3.85)	(1.60)	
$T^4/1000$	-0.4266***	-0.5312***	-0.050***				
	(6.16)	(4.84)	(3.04)				
$T^5/10000$	0.0404***	0.0482***					
	(5.15)	(3.83)					
C6974	0.0250	-0.8927**					
	(0.12)	(1.96)					
C7580	1.1243***	2.456***			0.887		
	(4.52)	(3.28)			(1.36)		
C8186	3..5250***	5.2910***			1.387		
	(11.10)	(7.22)			(1.87)*		
C8792	5.0167***	7.9526***			2.248		
	(10.96)	(9.31)			(2.87)		
C9398	6.8719***	9.4926***			2.559***		
	(12.80)	(13.17)			(3.25)		
MICRO	6.3848***	9.2334***			6.877***		
	(14.59)	(12.46)			(12.38)		
MACRO	4.4841***	4.8795***			4.804***		
	(12.41)	(8.48)			(10.44)		
PUBLIC ECONOMICS	3.8014***	4.2405***			4.610***		
	(10.72)	(8.35)			(10.40)		
ECONOMETRICS	3.1507***	2.7574***			3.882***		
	(8.32)	(5.60)			(8.34)		
FEMALE	-2.491***	-3.0367***			-2..032***		
	(6.26)	(3.89)			(5.27)		
CONSTANT	-7.870***	-5.6269***			-5.691***		
	(14.84)	(8.99)			(8.08)		
Observations	16003	16003			13555		

Notes: Absolute t-values in parentheses. We excluded researchers which never had a publication. For the cohort regressions we also excluded researchers with a Ph.D. before 1969. Estimates are based on the Fitzenberger (1997) algorithm. Standard Errors bootstrapped using Bilias et al. (2000) method (200 replications).

Table 2.2.3: Semiparametric Fixed Effects estimation (Honoré 1992)

	Cohort				
	1969-74	1975-80	1981-86	1987-92	1993-98
T	2.430***	1.696***	3.031***	3.342***	3.844***
	(4.18)	(5.09)	(3.99)	(7.37)	(8.90)
$T^2/10$	-2.472***	-1.764***	-3.915***	-3.587***	-4.775***
	(4.10)	(4.19)	(2.89)	(5.88)	(5.24)
$T^3/100$	0.989***	0.722***	1.945***	1.089***	1.715***
	(3.94)	(3.10)	(2.18)	(4.50)	(2.68)
$T^4/1000$	-0.136***	-0.102***	-0.331*		
	(3.82)	(2.31)	(1.69)		
Observations	13555				
Value of Loss Function	104162.908				

Notes: Quadratic Loss Function; Start values found through random search. Absolute t-values in parentheses. We excluded researchers who never published or received their Ph.D. before 1963.

Table 2.2.4: Ability groups

	Probability		Cond. Expected Output	
	Journeymen	Accomplished	Journeymen	Accomplished
T	0.370***	0.436***	0.115***	0.189***
	(11.58)	(14.44)	(5.60)	(7.30)
$T^2/10$	-0.479***	-0.585***	-0.083***	-0.245***
	(6.62)	(8.28)	(4.89)	(6.92)
$T^3/100$	0.259***	0.357***	0.015***	0.111***
	(3.76)	(5.58)	(3.99)	(6.24)
$T^4/1000$	-0.065**	-0.097***		-0.017***
	(2.40)	(4.07)		(5.69)
$T^5/10000$	0.006*	0.009***		
	(1.71)	(3.04)		
C7580	0.309***		0.167	
	(3.48)		(1.28)	
C8186	0.957***		0.316***	
	(10.62)		(2.73)	
C8792	0.883***		0.320***	
	(11.41)		(2.84)	
C9398	1.109***		0.330***	
	(13.97)		(3.28)	
MICRO	0.596***		0.180	
	(3.32)		(1.59)	
MACRO	0.593***		-0.031	
	(3.40)		(0.30)	
PUBLIC ECONOMICS	0.579***		-0.055	
	(3.32)		(0.52)	
ECONOMETRICS	0.518***		-0.165	
	(2.83)		(1.45)	
FEMALE	-0.310***		-0.028	
	(2.83)		(0.25)	
GROUP2	1.231***		0.714***	
	(11.62)		(6.83)	
CONSTANT	-3.543***		3.053***	
	(18.87)		(21.43)	
Number of Observations	15478		3834	
Log Pseudolikelihood	-6492.2		-102426.7	

Notes: Grouping was carried out using a mixture model as described in Section 3.3. Absolute t-values in parentheses, based upon a clustering robust covariance-matrix. We excluded researchers with Ph.D. before 1969 due to missing data at the onset of the career.

Table 2.2.5: Density Deconstruction

	Conditional number of articles	Conditional average quantity: Number of pages per article divided by number of authors	Conditional quality	Co-authorship	Quality and Co-authorship
T	0.2660***	0.0487***	0.0173***	-0.0496**	0.0180***
	(7.36)	(4.28)	(3.81)	(2.48)	(3.99)
$T^2/10$	-0.2933***	-0.066***	-0.0064***	0.1032***	-0.0064***
	(5.80)	(4.65)	(4.83)	(3.75)	(4.89)
$T^3/100$	0.1285***	0.0256***		-0.0389***	
	(5.18)	(3.92)		(2.90)	
$T^4/1000$	-0.0184***	-0.0031***		0.0044**	
	(4.90)	(3.26)		(2.14)	
C6974	0.121	0.0093	-0.2093**	0.1790	-0.2165**
	(0.58)	(0.15)	(2.44)	(1.41)	(2.46)
C7580	0.4895***	-0.0580	-0.1175	0.3210**	-0.1026
	(2.47)	(0.85)	(1.30)	(2.48)	(1.07)
C8186	0.8473***	-0.0868	-0.099	0.5360***	-0.0746
	(4.02)	(1.40)	(1.09)	(4.04)	(0.79)
C8792	0.8152***	-0.1379**	0.0256	0.8761***	0.0343
	(4.03)	(2.13)	(0.28)	(6.69)	(0.36)
C9398	1.0418***	-0.1118*	0.0191	1.0836***	0.0205
	(5.73)	(1.80)	(0.22)	(8.24)	(0.23)
MICRO	0.5096*	-0.2743***	0.3916***	0.2460	0.4612***
	(1.65)	(4.56)	(3.68)	(1.38)	(4.14)
MACRO	0.5086*	-0.1339**	0.0203	-0.0080	0.0747
	(1.70)	(2.32)	(0.22)	(0.05)	(0.74)
PUBLIC ECONOMICS	0.5586*	-0.1520***	0.0052	0.0489	0.0644
	(1.86)	(2.73)	(0.06)	(0.28)	(0.65)
ECONOMETRICS	0.4006	-0.2705***	0.0633	0.3408*	0.1175
	(1.30)	(3.87)	(0.63)	(1.85)	(1.08)
FEMALE	-0.4511**	0.0752	0.0259	-0.2811**	0.0017
	(2.16)	(1.15)	(0.30)	(2.49)	(0.02)
CONST	-1.8815***	2.6544***	-1.2858***		-1.9005***
	(5.35)	(34.99)	(9.54)		(15.33)
NUMBER OF ARTICLES		-0.0071	0.0882***		
		(0.74)	(7.20)		
Ln (CONTRIBUTION)			-0.2480***		
			(10.30)		
Ln (COAUTHORS)					0.2437***
					(6.05)
Observations	4295	4295	4295	4295	4295
(Pseudo)-R^2		0.03	0.12	0.05	0.08
Log Likelihood	-4486.2			-3758.3	

Notes: Absolute t-values in parentheses, based upon a clustering robust covariance-matrix. Estimation methods: Column (1): Zero-Truncated Poisson regression; (2): OLS on Logarithm of average contribution; (3), (5): OLS on Logarithm of average quality; (4) Ordered Probit

Figure 2.2.1: Pooled specification

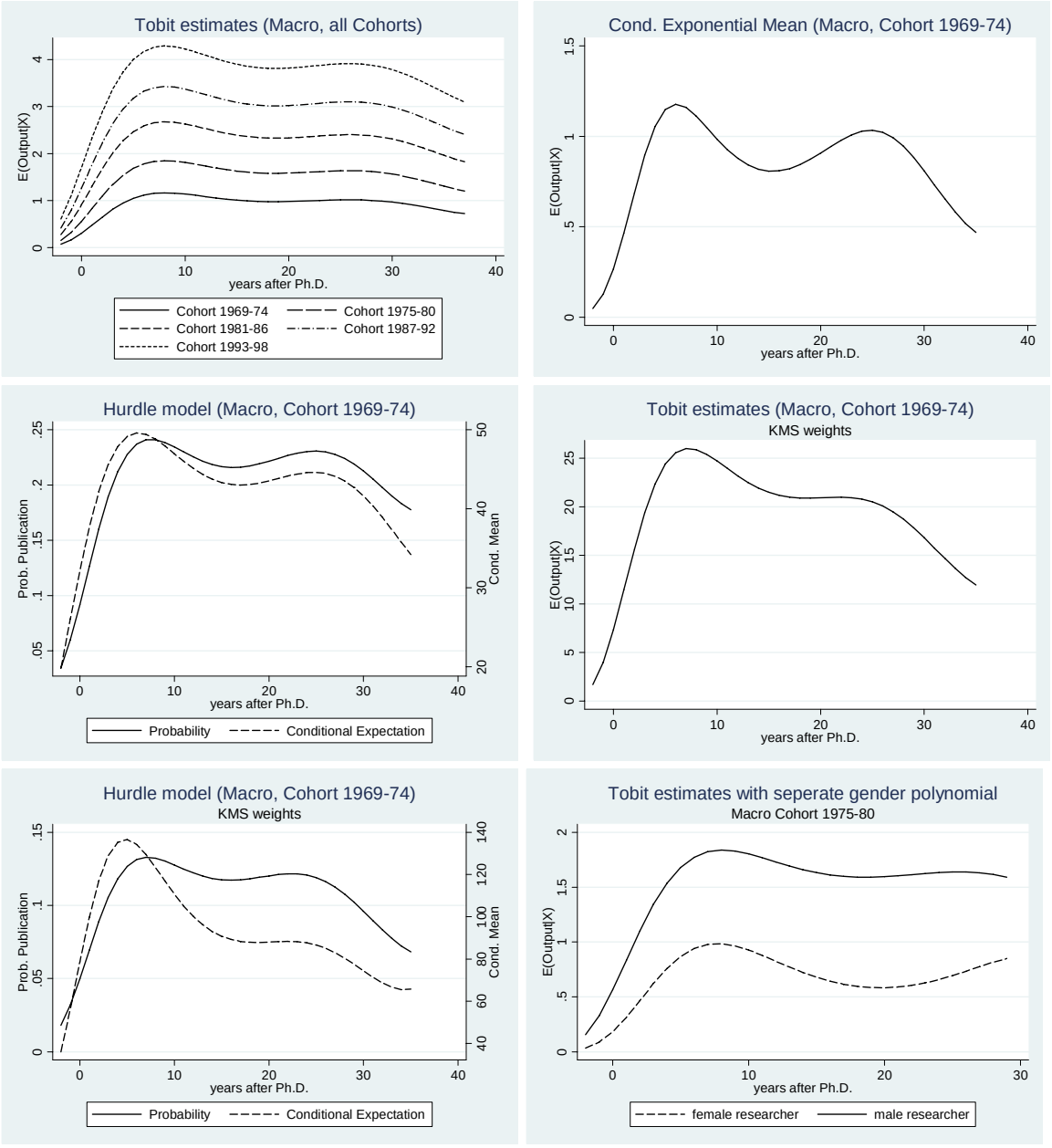


Figure 2.2.2: Censored Quantile regressions

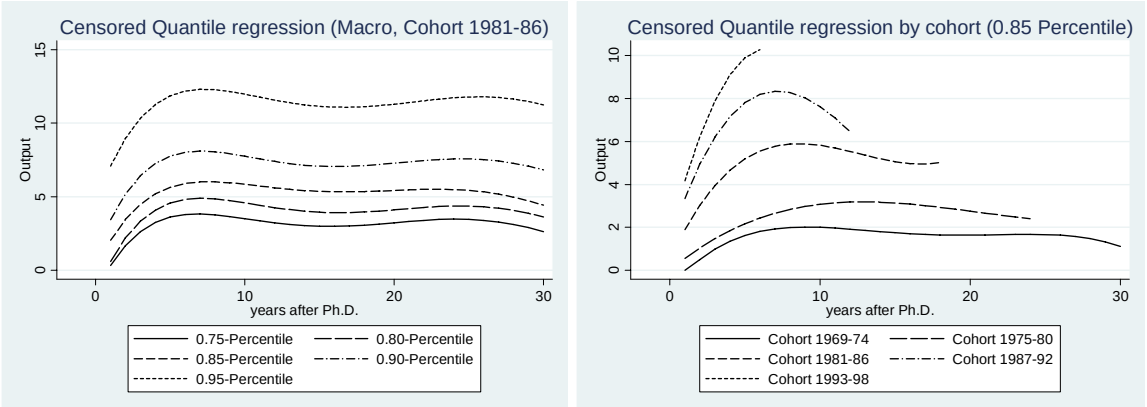


Figure 2.2.3: Semiparametric Fixed Effects estimation (Honoré, 1992)

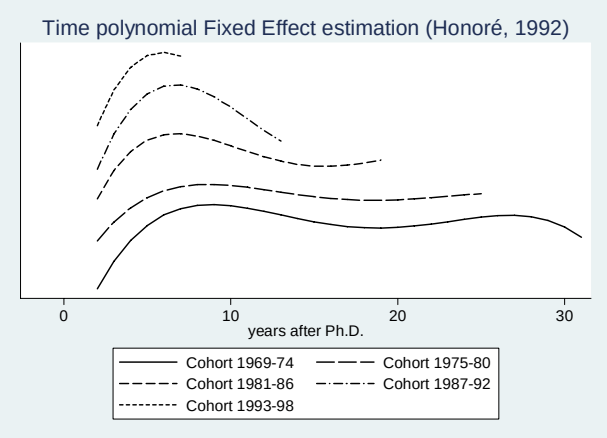


Figure 2.2.4: Historical time and cohort effects

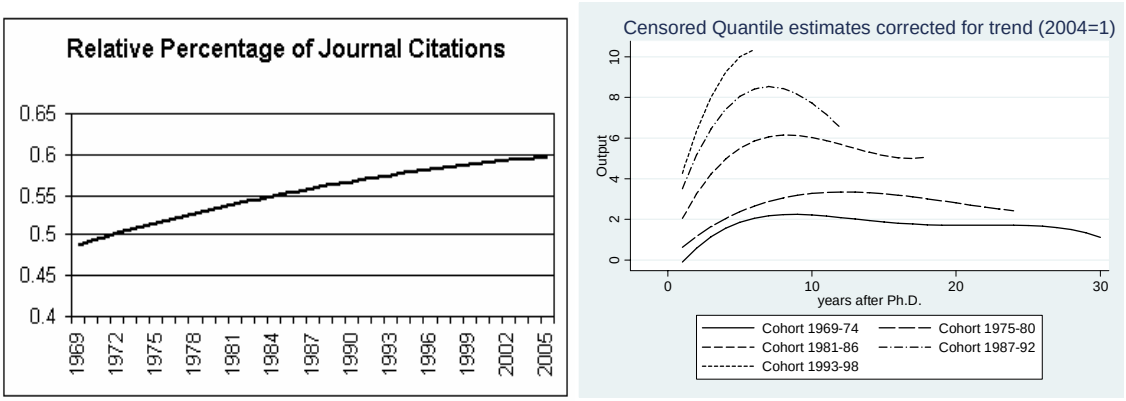


Figure 2.2.5: Ability specific regressions

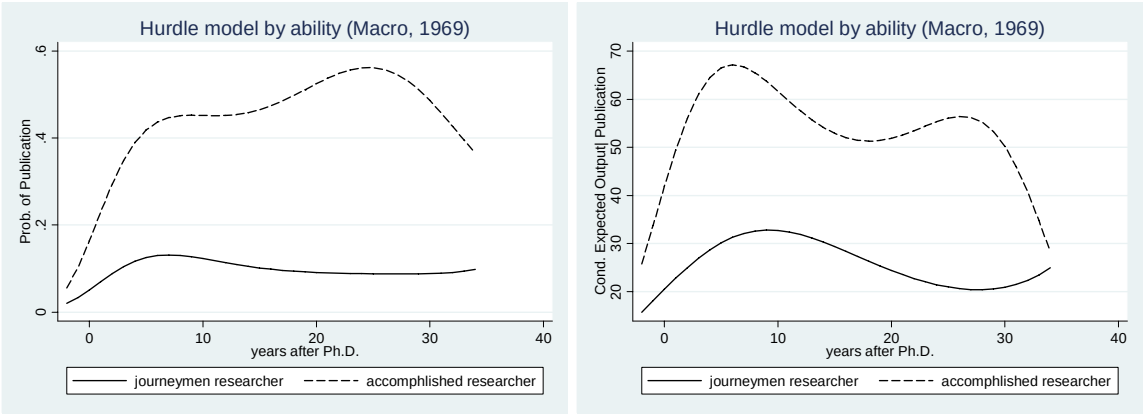
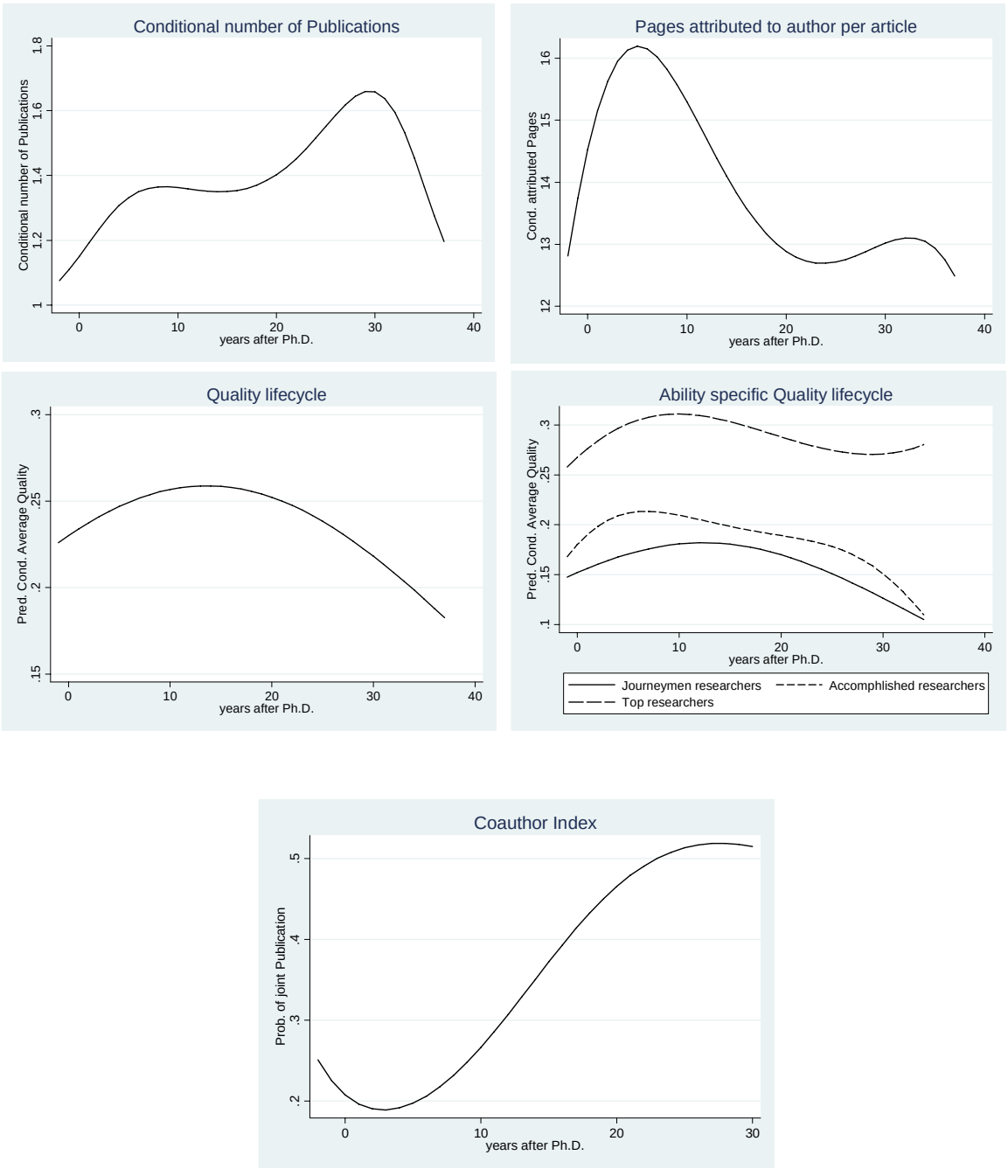


Figure 2.2.6: Density Deconstruction



2.2.7 Appendix

Table 2.2.6: Descriptive statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
FLOW_CL	18610	1.1367	3.4256	0	56.83333
FLOW_KA	18610	22.8178	140.5206	0	4181.14
PUBLD	18610	0.2307	0.42134	0	1
ARTICLES_ON	4295	1.5909	1.0327	1	14
LN_PAGES_PAPER	4295	2.4046	0.6216	-0.6931472	4.26268
LN_AVQUAL	4295	-1.6571	0.7464	-2.484907	0
COAUT	4295	1.5566	0.6452	1	4
T	18610	11.2948	10.7457	-5	41
C6974	18610	0.2882	0.45295	0	1
C7580	18610	0.1616	0.36817	0	1
C8186	18610	0.1229	0.32838	0	1
C8792	18610	0.1326	0.33916	0	1
C9398	18610	0.1262	0.33210	0	1
GENDER	18610	0.0591	0.23583	0	1
MICRO	18610	0.1349	0.34171	0	1
THEORY	18610	0.2670	0.44240	0	1
PUBLIC ECON.	18610	0.3538	0.47819	0	1
ECONOMETRICS	18610	0.1561	0.36300	0	1

Exemplary Density Deconstruction

Assume two papers are published in year t, one together with a coauthor (40 pages, quality $\frac{1}{12}$) and the other one without co-author (15 pages, quality $\frac{1}{2}$). The Output (*Y*) is then calculated as:

$$Y_{it} = \frac{1}{2} \cdot 40 \cdot 0.08 + 1 \cdot 15 \cdot 0.5 = 9.1$$

The number of pages attributed to the author is denoted by

$$P_{it} = \frac{1}{2} \cdot 40 + 1 \cdot 15 = 35$$

We then simply divide Y_{it} through P_{it} to arrive at the average quality for year t:

$$Q_{it} = \frac{9.1}{35} = 0.26 .$$

To compute the average (quantitative) contribution per paper we divide *Pages* through the *number* (here 2) of articles written. This is a measure of the *average contribution* (measured in pages; here 17.5) to each paper authored or co-authored.

References

Backes-Gellner, U. and A. Schlinghoff (2004): Careers, incentives and publication patterns of US and German (business) economists, *http://papers.ssrn.com/sol3/papers.cfm?abstract_id=616822*.

Baser, O. and E. Pema (2004): Publications over the academic life-cycle: Evidence for academic economists, *Economics Bulletin*, 1, 1-8.

Bell, J. and J. Seater (1978): Publishing performance: Departmental an individual, *Economic Inquiry*, 16, 599-615.

Bilias Y., S. Chen and Z.Ying (2000): Simple Resampling methods for censored Regression Quantiles, *Journal of Econometrics*, 99, 373-386.

Combes, P. and L. Linnemer (2003): Where are the economists who publish? Publication concentration and rankings in Europe based on cumulative publications, *Journal of the European Economic Association*, 1, 1250-1308.

Coupé, T. (2003): Revealed performances: Worldwide rankings of economists and economics departments, 1990-2000, *Journal of the European Economic Association*, 1, 1309-1345.

Coupé, T. (2004): What do we know about ourselves? On the economics of economics, K*yklos*, 57, 197-215.

Coupé, T., V. Smeets and F. Warzynski (2003): Incentives in economic departments: Testing Tournaments, *Aarhus School of Business Working Paper 03-25*.

Coupé, T., V. Smeets and F. Warzynski (2006): Incentives, sorting and productivity along the career: Evidence from a sample of top economists, *Journal of Law, Economics, and Organization*, 22, 137-167.

Dnes, A. and N. Garoupa (2005): Academic tenure, post-tenure effort, and contractual damages, *Economic Inquiry*, 43, 832-839.

Fitzenberger, B. (1997): A Guide to Censored Quantile Regressions, in G.S. Maddala and C.R. Rao (eds.), *Handbook of Statistics*, 15, 405-437.

Goodwin, T.H. and R.D. Sauer (1995): Life cycle productivity in academic research: Evidence from cumulative publication histories of academic economists, *Southern Economic Journal*, 61 728-743.

Goyal, S. and van der Leu, M. and J. L. Moraga-Gonzales (2004): Economics: An emerging small world?, *CESifo Working Paper No. 1287*.

Grimes, P.W. and C.A. Register (1997): Career Publications and academic job rank: Evidence from the class of 1968, *Journal of Economic Education*, 28, 82-92.

Hartley, J.E., J.W. Monks and M.D. Robinson (2001): Economists' publication patterns, *The American Economist*, 45, 80-85.

Hollis, A. (2001): Co-authorship and the Output of Academic Economists, *Labour Economics*, 8, 503-30.

Honoré, B.E.(1992): Trimmed Lad and Least Squares Estimation of Truncated and Censored Regression Models with Fixed Effects, *Econometrica*, 60, 533-565.

Hutchinson, E.B. and T.L. Zivney (1995): The publication profile of economists, *Journal of Economic Education*, 26, 59-79.

Jones, B.L. Nagin, D.S. and Roeder K. (2001): A SAS Procedure Based on Mixture Models for Estimating Development Trajectories, *Sociological Methods & Research*, 29, 374-393.

Kalaitzidakis, P., Mamuneas, T. and T. Stengos (2003): Rankings of academic Journals and Institutions in economics, *Journal of the European Economic Association*, 1, 1346-1366.

Kenny, L. and R. Studley (1995): Economists' salaries and lifetime productivity, *Southern Economic Journal*, 65, 382-393.

Laband, D. (2002): Contribution, Attribution and the Allocation of Intellectual Property Rights: Economics versus Agricultural Economics, *Labour Economics*, 9, 125-31.

McDowell, J. and J. Smith (1992): The Effect of Gender-Sorting on Propensity to Coauthor: Implications for Academic Promotion, *Economic Inquiry*, 30, 68-82.

Moore, W., R. Newman and G. Turnbull (2001): Reputational capital and academic pay, *Economic Inquiry*, 39, 663-671.

Oster, S.M. and D.S. Hamermesh (1998): Aging and productivity among economists, *Review of Economics and Statistics*, 80, 154-156.

Pagan, A., and F. Vella (1989): Diagnostic Tests for Models Based on Individual Data: A Survey, *Journal of Applied Econometrics*, 4, 29-59.

Pohlmeier W. and V. Ulrich (1995): An Econometric Model of the Two-Part Decisionmaking Process in the Demand for Health Care, *The Journal of Human Resources*, 30, 339-361.

Powell, J.L. (1984): Least Absolute Deviations Estimation for the Censored Regression Model, *Journal of Econometrics*, 25, 303-325.

Powell J. L. (1986): Censored Regression Quantiles, *Journal of Econometrics*, 32, 143-155

Rauber M. und H. Ursprung (2006): Evaluation of researchers: A life cycle analysis of German Academic Economists, *CESifo Working Paper No. 1673*.

Rodgers W. L. (1982): Estimable Functions of Age, Period, and Cohort Effects, *American Sociological Review*, 47, 774-787.

Skirbekk V. (2003): Age and Individual Productivity: A Literature Survey, *Max Planck Institute for Demographic Research, Working Paper WP 2003-028*.

Stephan, P. (1996): The economics of science, *Journal of Economic Literature*, 34, 1199-1235.

Wooldridge J. (2002): Econometric Analysis of Cross Section and Panel Data, *Cambridge, Massachusetts: The MIT Press*.

Van Dalen, H. (1999): The golden age of Nobel economists, *American Economist*, 19-35.

Chapter 3

Noncognitive Skills and Success in Life: The Importance of Motivation and Self-Regulation

3.1 Introduction

Recent econometric literature focuses on the importance of so called noncognitive skills on labour market outcomes and success (e.g. Heckman and Rubinstein, 2001; Heckman, Stixrud and Urzua, 2006; Flossmann, Piatek and Wichert, 2006).

Although no general definition of the term noncognitive skills exists in psychology, motivational psychology offers two important concepts that seem to build up a large part of what economists describe as noncognitive skills: motivation and self-regulation or volition (Achtziger and Gollwitzer, 2006). Motivation is responsible for goal setting whereas self-regulation is important for goal striving. Goal setting and goal striving are governed by different psychological processes and are both necessary for performance and achievement (Heckhausen and Gollwitzer, 1987).

Research is still preliminary but first results support the hypothesis that noncognitive skills play an important role for achievement and success in life. However, many studies dealing with this topic are hampered by identification problems. Most datasets available do not include measures of noncognitive skills; those that do usually miss information on the cognitive skills of individuals. For example, the 1999 wave of the German Social Economic Survey (GSOEP), a longitudinal household survey conducted since 1984, includes questions on Rotter's (1966) locus of control but no measure for cognitive skills. However, without simultaneous data on cognitive and noncognitive abilities identification has to rely on strong assumptions. Such assumptions are usually not testable and what is measured therefore remains somewhat unclear.

Furthermore, even if data are available causality may not be clear. If an individual is interviewed about noncognitive skills at age forty, it is obvious that the answer is partly determined by experience and success in life until then. To overcome this reverse causality one has to use instrumental variable estimation. However, even in longitudinal datasets there are often no or only weakly correlated instruments available.[66]

Finally, since psychologists have dealt with noncognitive skills for many years, it seems to be advisable to base the analysis of noncognitive skills on their well-developed theories and concepts. Only by doing this it is possible to give these skills a deeper meaning and interdisciplinary interpretation.

In this paper, the effects of motivation, self-regulation and social skills on various outcomes in life will be analysed with respect to the issues mentioned above. Performance

[66] Although the GSOEP includes a large number of questions on private and professional development it is quite difficult to find useful instruments for Rotter's (1966) locus of control (Göggel, 2007).

in school, achievement in professional life and happiness will be used as the three well defined indicators for success. Having data on both cognitive and noncognitive skills at the time of adolescence and early adulthood, the panel structure of the dataset will allow the explicit taking of timing and causality into account.

The results of the analysis indicate that self-regulation and attribution are important for success in all three areas. However, the effect differs between females and males and within different fields. Social skills do not prove to have a significant impact on success. Overall it seems fair to conclude that both cognitive and noncognitive skills, developed in early childhood, shape the individual path through life. A straightforward conclusion of the analysis is that parents' investment in the noncognitive skills of their children is worthwhile and the costs of doing so are rather modest.

The rest of this chapter unfolds as follows: Section 3.2 briefly describes the dataset used for the analysis. Section 3.3 introduces several important psychological concepts. Section 3.4 is devoted to a discussion of cognitive skills and the interaction of both types of skills. The results are presented in section 3.5. Section 3.6 focuses on sample section and section 3.7 finally concludes.

3.2 The Data

Data was used from a longitudinal panel study of 3240 10^{th} grade students attending 121 classes at 68 "Gymnasien" in North Rhine-Westphalia (Central Archive for Empirical Social Research, 2007; Meulemann, 2007). The testing and interviews were conducted at three different points in time: during the 10^{th} grade (1970), at age 30 (1984) and at age 43 (1997). The study was restricted to North Rhine-Westphalia and the high-school students were selected by a two-stage clustered sampling procedure.

The 10^{th} grade students were asked questions concerning their habits, success in school and also family related topics. They also participated in a psychometric test (Intelligence Structure Test, IST; Amthauer, 1953). Furthermore, parents and teachers were asked to answer questionnaires. For the present study only the parent and student questionnaires were evaluated, because the teacher questionnaire focused on the whole class and included no additional information on individual students. About ten years after the first interview the students' grades were collected from the schools and also recorded. Follow-up interviews were conducted 1984 and 1997. Questions concerning the professional and academic career as well as private development were asked. Overall, the sample size was reduced to about 1600 participating individuals at age 43, which is about 50 percent of the

initial sample size. Panel attrition is a common problem in longitudinal studies and I will deal explicitly with this issue in section 6.[67]

Finally, note that the sample consists only of students from grammar schools (the highest level of German school education), which helps to avoid confounding effects of education and social origin. In fact, by focusing on this group of students the initial conditions in the sample are much more homogenous and, therefore, it is less probable that the estimates suffer from any omitted variable bias.

3.3 Noncognitive skills

This section briefly reviews some important concepts from motivational psychology. Since motivation induces volition and self-regulative behavior, the initial discussion concerns the relationship between attribution and motivation and thereafter the concepts of self-regulation and volition (Heckhausen and Gollwitzer, 1987).

Motivation and Attribution: Weiner et al. (1971, 1972) developed a model of attribution and showed that success and failure are attributed by individuals to internal or external factors. Individuals can follow an internal attribution strategy by attributing success to effort and ability. External attribution factors are, for example, family and luck. The level of time for stability of attribution can also be considered (Weiner, 1971; Bierbrauer, 1996). Individuals can attribute success to unstable and task specific factors or to stable factors such as the family. A scheme of attribution can be found in the following Figure:

[67] Note that the degree of attrition seems not to be exceptionally high, given that over a time horizon of 20 years the SOEP lost about the same percentage of its initial participants. The SOEP-West started initially in the year 1984 with 12290 participants. In the year 2004 only 6811 individuals that participated in the first wave were still in the sample (SOEP 2007).

	INTERNAL	EXTERNAL
STABLE	**ABILITY**	**FAMILY**
UNSTABLE	**EFFORT**	**LUCK**

Figure 3.1: Attribution schemes , Source : Weiner et al. (1971). Note that Weiner et al. (1971) refer to "Task difficulty" in the upper right corner. Since family is time stable and external I include this factor.

The students in the sample were asked in the 10th grade how important they thought that ability, effort, luck and support by the family were for success in school. These questions refer to the attribution scheme introduced by Weiner et al. (1971) and therefore proxy the attribution of success. The answer possibilities were always on a scale from 0, implying not important, to 5 which denoted very important. Similar questions were also asked at both follow-up interviews. The questions at that time were slightly modified, dealing more generally with individual success in life. In the analysis the answers from the 10th grade and the first follow-up interview were focused upon to avoid reverse causality. An aggregation of the attribution factors with respect to both dimensions internal vs. external and stable vs. unstable was made since both attribution styles might have important motivational consequences. The resulting variables range between 0 and 10, a higher score indicates a stronger success attribution to the respective factor.

There are several psychological studies that document how the attribution style is linked to motivation. Weiner and Kukla (1970) formulate the hypothesis that attribution, locus of control (Rotter, 1966) and motivation are closely connected:

> *"In summary, an interaction is expected between locus of control, success and failure, and the resultant achievement motivation [need for achievement minus anxiety of failure]. The prototypic high-achievement-oriented individual is conceptualized as one who assumes responsibility for success, but relatively*

> *denies his liability for failure. On the other hand, the individual low in resultant achievement concerns is believed to assume the blame for failure, while denying himself the luxury of personal praise for success."* Weiner and Kukla (1970: p.9).

After a series of experiments they came to the conclusion that "…individuals classified as high in resultant achievement [need for achievement minus anxiety of failure] tend to attribute success to themselves more than individuals low in resultant achievement motivation." (Weiner and Kukla, 1970, p.12). Several studies from developmental psychology confirm this observation: The attribution style is formed during early childhood and a positive or depressing attribution style is fully developed at an age of 11 years (Heckhausen and Heckhausen, 2006, p.:431, Heckhausen, 1984). During adolescence both attribution styles lead to a different development of children (Heckhausen and Heckhausen, 2006, p.:431). Furthermore, there is evidence that self-esteem is also closely connected to attribution: Individuals with a high degree of self-esteem tend to frequently attribute success as being internal.[68]

However, the time stability of attribution can also have important motivational consequences as it influences success expectations: Meyer (1973) confronted individuals in an experiment with a series of failure experiences and divided the participants into two groups. One group was comprised of individuals that revealed an attribution pattern to time stable factors. Interestingly, this group significantly reduced their expectations of future success in each round of the experiment, whereas the other group had roughly constant success expectations over time. Abramson et al. (1989, p. 361) further confirm the importance of time stability of attribution with their research on hopelessness. Bad events can lead to a destruction of self-esteem if attribution is stable and global. Finally attribution to stable factors like family or ability might potentially lead to less task specific effort.

Self-regulation: The concept of self-regulation has a long history in psychological literature. Even William James (1890) named a chapter in Volume 2 of his "Principles of Psychology" *Will*. He therein discusses, among other things, why sometimes a certain action is not undertaken although there might be a motivation for it (obstructed will) and why at other

[68] See Stiensmeier-Pelster and Heckhausen (2006), p. 378 on this issue.

times actions not intended are undertaken (explosive will). A recent definition of self-regulation is given by Peterson and Seligman (2004):

> *"Self-regulation refers to how a person exerts control over his or her own responses so as to pursue goals and live up to standards. These responses include thoughts, emotions, impulses, performances, and other behaviors. The standards include ideals, moral injunctions, norms, performance targets, and the expectations of other people."* Peterson and Seligman (2004, p.: 500).

In motivational psychology the term volition describes such self-regulatory capacities in the context of goal striving. Since the introduced definition of noncognitive skills is derived from motivational psychology, it seems to be useful to explain the concept of volition in some detail. Volition is closely connected to the Rubicon-model of action phases (Heckhausen and Gollwitzer, 1987). The name Rubicon-model derives from a tiny river in Northern Italy: Crossing the Rubicon-River and entering Italy with a standing army was forbidden by an ancient Roman law and any general that did so was a traitor. This tiny stream therefore revealed Caesar's intentions and marked the point of no return (Achtziger and Gollwitzer, 2006). According to the Rubicon-model (Heckhausen and Gollwitzer, 1987) every action includes such a point of no return at which the individual moves from goal setting to goal striving:

> *"Once subjects move from planning and goal-setting to the implementation of plans, they cross a metaphorical Rubicon. That is, their goals are typically protected and fostered by self-regulatory activity rather than reconsidered or changed, often even when challenged."* Corno, (1993, p.:15).

Whereas motivational factors determine goal setting, the primary role of volition is the implementation of existing goals and therefore goal striving (Corno, 1993). Since self-regulation is the more general concept reference will be made to self-regulation in what follows but volition is, of course, a part it.

Self-regulation is scored by using parents' judgments on a scale of 1 (very gifted) to 3 (not gifted) of their child's ability to assert itself and to deal with difficult situations. Since this measure is from the parents' questionnaire at the time of adolescence, it is reasonably

exogenous to events later in life. Furthermore, parents are probably in the best position to judge on this ability of their offspring.[69]

Social skills: Social skills and especially friendships seem to be important for a child's developmental process. Observations in child care centers show that even among infants a considerable differentiation in social interaction exists (Hartup, 1992). Friendships support the child's developmental process within several dimensions:

> *"1.These relationships are contexts in which basic social skills are acquired or elaborated (e.g. social communication, cooperation, and group entry skills);*
> *2. They are information sources for acquiring self-knowledge, knowledge about others, and knowledge about the world;*
> *3. They are emotional and cognitive resources (both for "having fun" and adapting to stress); and*
> *4. They are forerunners of subsequent relationships (modeling the mutual regulation and intimacy that most close relationships require)."*
> Hartup, (1992), p. 184.

However, there is also evidence for negative peer group effects in the literature: Urberg, Degirmencioglu and Pilgrim (1997) found that close friends and friendship groups had a significant influence on the use of cigarettes and alcohol among adolescents.[70]

Social skills were evaluated here by using parents' answers on a scale from one (very gifted) to three (not gifted) on how talented their child was in getting new friends. Because of the possibly negative influences during adolescence the direction of the effect seems to be ex ante unclear.

3.4 Cognitive skills

Cognitive skills are measured by well developed standardized tests of mathematical and language capabilities. The results of such a standard psychometric test (Intelligence Structure Test, IST; Amthauer, 1953) are available for the high-school students in the sample because the students participated in class room tests during the 10th grade. The IST

[69] Note that parents' judgement might not always be comparable. Therefore this variable is almost certainly measured with some noise.
[70] Close friends and friendship groups e.g. both independently contributed to the prediction of drinking to intoxication (Urberg, Degirmencioglu and Pilgrim, 1997).

is multidimensional and consists of nine separate tests. However, only a short version of the IST with four subtests was carried out.[71] The distribution of the number of correctly solved questions is depicted in the Appendix Figure 3.3. The average number of correctly solved questions is 40.5. Because all students attended the same school track, differences in the test performance should reflect underlying differences in intelligence and not different educational attainment.[72]

There is some evidence in the psychological literature which shows that cognitive skills and motivation are roughly uncorrelated (Gagné and St. Père, 2002).[73] Furthermore, Corno (1993, p.15) reports that "Conative (motivational and volitional) aptitudes are considered by most theorists to be conceptually and empirically distinct from general or specialized intellectual abilities". Nevertheless it seems important to include a measure of pure cognitive skills into the analysis to disentangle the effect of both types of skills.

To test for the correlation between the measures of cognitive and noncognitive skills, simple bivariate correlation coefficients between the variables using the answers from the 10th grade questionnaire were calculated. The results are reported in Table 3.2. Interestingly, both types of skills are roughly uncorrelated which confirms the already mentioned psychological evidence. A similar result was also reported by Heckman, Stixrud and Urzua (2006, p.9) for their sample of NLSY pupils.

3.5 Results

In this section an analysis of whether motivation, self-regulation and social skills formed during childhood have an impact on outcomes later in life is made. Three different measures of success will be used: Performance in school, professional achievement and happiness.

3.5.1 School performance

The first question addressed is how noncognitive skills affect academic performance. In Germany, the school leaving grades directly influence later academic career opportunities because universities do not select their students on the basis of admission tests. The

[71] The tests carried out were Wortauswahltest, Analogietest, Zahlenreihentest and Würfeltest. I decided to include the number of correctly solved questions into the regression because all pupils attend the same track and class and are therefore comparable. Furthermore, mean age in the 10th grade is 15.4 years with a standard deviation of 0.88, so two thirds of all pupils are between 14.5 and 16.2 years old which is also reasonably close. Finally, I was advised to use the raw scores by HOGREFE (publisher of the IST).

[72] Note that there might be differences with respect to schools and teachers. However, every federal state has a standardized curriculum and therefore these differences should be of minor importance.

[73] Also creativity seems to be uncorrelated with cognitive skills, see Sen and Hagtvet (1993).

permission to study several popular subjects like medicine or psychology depends solely upon the grades.

Grades for the 13th class are available for most pupils in the sample that stayed in school until then.[74] These grades were collected directly from the schools for the majority of students so there was essentially no possibility of misreporting. A natural starting point for the analysis is therefore to disentangle the effect of cognitive and noncognitive skills on success in school. The regression results are presented in Table 3.3: In columns 1, 2, 5 and 6 the dependent variable is the average grade and in column 3 and 4 the two subjects mathematics and German are analyzed separately.

There seems to be no significant difference between the school performance of boys and girls in the sample (FEMALE). The psychometric test score (IST) is highly significant with a negative sign in all regressions, implying that higher cognitive ability improves (lowers) the grade. Note, that the t-values are the highest among all included variables. The variable IST is bounded between 12 and 70 correctly solved questions; this translates into a difference of the average grade of roughly -0.72.

Self-regulation shows up significantly in the regression for the average grade (columns 1 and 2) and increases achievement in school: Students with a high self-regulatory capacity (SELFREG_H) are about one-eighth grade better than the students in the lowest category. The students in the middle category (SELFREG_M) still have a one-twelfth better average grade. There is no significant difference between a stable and unstable attribution style (column 1). However, individuals who rank high in internal attribution (ATTR.INTERNAL) and low in external attribution (ATTR.EXTERNAL) have better grades as can be seen by inspection of column 2. This result is in line with Heckhausen and Heckhausen (2006) and confirms that attribution styles formed during childhood have a significant impact on performance. Finally, the variable measuring social skills (FRIEND) is insignificant in all regressions.

Generally, the results show that noncognitive skills are nearly as important as cognitive skills for performance in school. Although the effect does not seem to outperform the effect of cognitive skills, as in Duckworth and Seligman (2005), it is nevertheless quite substantial.[75]

[74] Note that some students left school before the Abitur.

[75] Ranking highest in internal attribution and having a high self-regulatory capacity improves the grade by -0.37 as compared to -0.72 for cognitive skills.

Of course one has to be cautious because of potential reverse causality. However, the explanatory variables seem to be reasonably exogenous because the grades were collected in the 13th class which was 3 years after the initial interview. Furthermore, the information on social skills and self-regulation was from the parents' questionnaire and therefore probably relates to the whole development of the child.

The variable REPEAT is a dummy variable which is coded one for students that repeated an academic year at least once. It is highly significant and has a positive sign, which indicates that students that repeated performed worse even though they had more time to deal with the same curriculum. The educational attainment of the mother (MOTHER EDU) is significant at the 1 percent level and shows the expected sign.[76] As can be seen in column 1, the mother's educational attainment is also by far more important than that of the father (FATHER EDU).[77] Interestingly, the educational achievement of the mother is highly significant for the average and the German grade but not for the mathematical performance.

Mathematics and German (columns 3 and 4 of Table 3.3) both arguably differ a lot in their requirement profile. Cognitive skills are probably more important for mathematics whereas e.g. debating skills should contribute more to the German grade. This hypothesis is easily confirmed by comparing the coefficient of the psychometric test (IST) in columns 3 and 4: The point estimate and t-statistic is larger for mathematics than for German. The noncognitive skills variables also reflect these differences impressively: Self-regulation (SELFREG_H) is highly significant in the regression for the German grade but insignificant for the mathematics grade. This confirms the hypothesis outlined above because German requires class room discussions and therefore verbal performance is an integral part of the teacher's evaluation. Obviously, a high degree of willpower is helpful for that kind of task.

For mathematics (column 3) things are quite different: an internal attribution style improves grades whereas self-regulation has no significant impact on achievement. Therefore, motivation seems to be much more important for the mathematics performance. Because mathematics requires sedulous learning and practice, it seems to be plausible that individuals who attribute success to effort and ability are more motivated than those that attribute success to luck or family. Some anecdotal evidence for this hypothesis is given by Stiensmayer-Pelster and Heckhausen (2006), who refer in an (hypothetical) example exactly

[76] Whether this effect is indeed due to a higher educational attainment of the mother or due to genetic reasons is not important for the present study. See e.g. Plug (2004) for a discussion. Note that the educational attainment is coded from 1 (elementary school without further training) to 13 (university degree) in the dataset. I also used several different scales and a full set of dummy variables to check for the robustness of the results.
[77] The correlation coefficient between the father's and the mother's educational attainment is 0.62.

to this situation. All this supports the insight that different noncognitive skills beneficially unfold in different subjects.

As is well known, boys and girls differ in their development during adolescence. Therefore the effect of self-regulation was estimated separately for boys and girls: With respect to the German grade, self-regulation significantly improves the performance of boys and girls but the effect is larger for girls. For the average grade a positive effect of self-regulation can only be identified for girls.[78] This result might be due to the development edge of girls over boys during adolescence.

Finally to check for the robustness of the results several robustness tests were performed. To test for omitted variable bias the number of evenings spent at home was additionally included and also a residential school dummy as explanatory variable (column 5). As can be seen, EVENINGS has a negative coefficient and is significant at the 10 percent level, indicating that the more evenings spent at home during the 10th grade improved the average grade three years later significantly. With respect to the measures of noncognitive skills nothing changed very much. Furthermore, the baseline regression with classroom fixed effects was estimated because there might have been differences between teachers and schools in grading. This even increased the t-values of self-regulation slightly as can be seen in column 6.

3.5.2 Professional career

3.5.2.1 Wage and Income

To investigate the impact of noncognitive skills on the professional career several wage and earnings equations were estimated and included in the measures for cognitive and noncognitive skills. The logarithm of net hourly wage at the time of the second follow-up interview was used as a dependent variable in the baseline regressions. It was computed by dividing net monthly income by the average hours worked per month.[79] The distribution of the wage variable is depicted in Appendix, Figure 3.2. and the results are shown in Table 3.4. Dummy variables for the highest school degree, successful completed studies and occupation as a civil servant are included in the baseline specification in column 1.

[78] German grade: SELFREG_H: girls: -0.33 (t-value: 2.57); boys: -0.22 (1.83); Average grade: SELFREG_H: girls: -0.17 (1.98); boys: -0.08 (0.98).

[79] Gross monthly income is not available in the dataset. Note that average hours worked is measured by actual and not by contractual specified hours. Average hours worked per month is calculated as weekly hours times four.

Work experience (EXPERIENCE), measured in years and calculated as the sum of all actual periods of employment, is highly significant and has a positive coefficient as can be seen by inspection of column 1. The female dummy is negative and insignificant. Given that we have a sample of high-school graduates and control for a large number of cognitive and noncognitive skills, this is not surprising.

The psychometric test score (IST) has the expected positive coefficient and is significant at the 10 percent level, indicating that a higher level of cognitive skills increases the wage, even when controlled for the highest school degree obtained.

The measures for noncognitive skills show up remarkably in the baseline regression: Self-regulation is highly significant for males and has a positive coefficient (SELFREG MALE). Interestingly, for women the effect of self-regulation is not statistically different from zero. A possible explanation for this will be discussed later. Finally, the coefficients of stable and unstable attribution styles taken from the teenager's questionnaire have the right sign but are insignificant in the baseline regression. The same is true for internal and external attribution (not shown).[80]

The second column presents the estimates of a reduced form where most control variables are not included. Since many career choices during life are probably directly and indirectly influenced by cognitive and noncognitive skills, this regression better identifies the total effect on the wage. As expected, the psychometric test score (IST) now becomes highly significant. Also, the importance of self-regulation increases and the variables are now significant at the 1 percent level for males.

In column 3 instrumental variable estimation is used for the attribution style (ATTR. STABLE and ATTR.UNSTABLE). The reason is that in the 10th grade interview the students were only asked about attribution of success at school. These questions might therefore not be appropriate when dealing with overall personal development. More precise questions concerning personal success in life were asked at both follow-up interviews and therefore instrumental variable regression seems to be suitable for the analysis.[81]

A two-stage least squares regression was used for estimation of the wage equation. The instruments were the attribution scores from the 10th grade and first follow-up interview. A twelve year time lag between the first and the second follow-up ensures

[80] Note that 4 individuals reported a very high hourly wage above 600 DM. If these observations are excluded, stable and unstable attribution becomes significant at the 5 percent level.

[81] Furthermore, the attribution variable from the first interview was taken from the teenager's questionnaire. I therefore suspect that this variable is measured less precisely than the other variables on noncognitive skills which were collected from the parents' questionnaire.

validity of the instruments (Hansen J-statistic p-value 0.46). The first stage partial R^2 of 0.09 and 0.08 for both endogenous variables were remarkably high and allow precise identification.[82] The results are presented in column 3. Most strikingly, unstable attribution has a much larger impact on wages; the coefficient is positive and significant at the 1 percent level.[83] Stable attribution has a negative but insignificant coefficient. Individuals who attribute success task specific are probably more successful because unstable attribution increases task specific effort. However, as Meyer (1973) pointed out, a reduction of future success expectations due to stable attribution after a series of failures could also lead to this result. Finally, stable attribution can lead to hopelessness and resignation (Abramson et al., 1989).

In column 4 the same exercise is undertaken with respect to internal and external attribution. Internal attribution has a positive coefficient and is borderline significant at the 10 percent level whereas external attribution is insignificant. This verifies the result from our previous analysis of school performance, but some caution seems to be appropriate because of the low significance level. However, combining this result with that from column 3, internal and unstable attribution seems to be a preferable strategy for success in the job market.

Column 5 presents an instrumental variable regression as previously shown, however the independent variable here is the natural logarithm of income. As can be seen, the female dummy variable is now highly significant and has a negative coefficient because the hours worked are not controlled. Female volition (SELFREG_H FEMALE) now has a positive and significant effect on income. Combining this result with the one from column 1 leads to the conclusion that the effect of self-regulation seems to work for women mainly due to the labor supply decision. Women on average work less hours than males but those with a high degree of goal striving decide to participate more actively in the labor market and therefore earn a higher income. Probably because of their high degree of self-regulation they can deal with the burden of housework, child care and employment more effectively. However, conditional on hours worked there is no additional effect of self-regulation.

The final hypothesis to be addressed is whether the significant influence of noncognitive skills is entirely due to the already identified effect of the high-school leaving

[82] Using only the teenager's attribution style as instrument leads to a weak instrument problem (e.g. first stage F-value for unstable attribution: 1.01)

[83] I also estimated the reduced form equation presented in column 2 using two-stages least squares as a robustness check. This increases, as we should expect, the coefficients and t-values of the cognitive and noncognitive skills variables somewhat.

grades. As already mentioned, the grades have a strong influence on later career opportunities in Germany because universities select students on the basis of grades. Furthermore, grades are also used by employers to select the best candidates for vocational training. I therefore condition on the average grade in the regression presented in column 6 of Table 3.4. As can be seen, this reduces the sample size to nearly one half of the initial observations. Most variables including the average grade become insignificant. However, the measure for self-regulation (SELFREG_H MALE) remains as significant at the 5 percent level as unstable attribution does. Motivation and self-regulation developed in early childhood and adolescence therefore seem to influence wages not only through their initial effect on grades.

3.5.2.2 Occupational Prestige

The wage is only one possible measure for professional success. Another factor, more often used in sociology, is occupational prestige. Treiman (1977) developed an occupational prestige scale using data from about 60 countries. The scale ranges between 18 for unskilled workers and 78.9 for physicians and professors. The Treiman (1977) prestige score was used as a dependent variable to analyze the effect of motivation and self-regulation on occupational prestige. The distribution of the Treiman (1977) score in the present sample of individuals is depicted in the Appendix. Note that the relatively high mean of about 55 is due to the fact that the sample consists of former high-school students.

The results of the analysis are shown in Table 3.5. Strikingly, self-regulation has a highly significant impact on occupation prestige for men *and* women. The effect is about 3.6 additional prestige points for males with a high level of goal striving and therefore nearly as large as the influence of the cognitive skills (IST), which lead to a maximum of 3.9 additional scale points.[84]

Stable and unstable (column 3) as well as internal vs. external attribution (column 4) also show up significantly in the instrumental variable regressions with the same pattern as in the prior wage regressions. Self-regulation and motivation seem to lead to the selection of professions with relatively high occupational prestige for both men and women.

Finally, it remains to be tested whether the wage differences identified previously are only due to this selection effect of more prestigious jobs. Therefore the Treiman (1977) score is included in the baseline OLS wage regression. The results are depicted in column 5.

[84] 0.0666*[70-12] since 70 was the maximum and 12 the minimum number of correctly solved questions.

As can be seen, Treiman's (1977) prestige score is highly significant in the wage regression. However, the self-regulation variables still remain significant for males, indicating that the higher wage is not only due to the selection of more prestigious jobs. Note, however, as should be expected, that the coefficients of the noncognitive skill variables are reduced and therefore indicate that indeed part of the effect on wages is due to selection.[85]

3.5.3 Happiness

Professional success is only one area where one would expect noncognitive skills to be of importance. This section undertakes an analysis of happiness. Using the same sample of individuals, Meulemann (2001) finds that success evaluation at age 43 is more important than objective success for happiness and that more recent successes have larger impacts on satisfaction than past ones.

However, whereas Meulemann (2001) looks backward from age 43, this analysis includes the measures of noncognitive skills from the students' questionnaire and analyzes happiness later in life. Happiness in private and professional life was ranked by the individuals answer on a scale from zero (not satisfied) to ten (maximum satisfaction) of how happy they were with their private and professional development at the time of the second follow-up interview. Since private and professional happiness were separately coded both answers were analysed separately. The results are shown in Table 3.5. Because of obvious causality problems, instrumental variable estimation was used for the attribution style as previously explained.

Women seem to evaluate their private life more positively than men, as can be seen in column 2. The psychometric test score (IST) only influences happiness in private life significantly. Furthermore, a high degree of self-regulation (SELFREG MALE) increases happiness with private development significantly for males.[86]

Internal attribution increases happiness at work significantly (column 1). This is consistent with the previous results, where internal attribution also increased the wage and occupational status. Interestingly, for private happiness external attribution seems to be a

[85] A further estimation was made of the instrumental variable wage regression (Table 3.4, column 3) including Treiman's (1977) prestige score. The results of this exercise show that unstable attribution remains highly significant and therefore confirm the conclusion drawn above.

[86] To test whether this effect is due to stable personal characteristics I additionally controlled for the answers in the first follow-up interview. The coefficients remained significant.

better strategy of coping. However, the variable is only at the 10 percent level significant and therefore this result should be interpreted with some caution.[87]

The level of the educational attainment of the father decreases the happiness with one's own professional life. This effect could be due to a status effect. Having a father with a high educational status probably devaluates the personal achievement and increases the aspiration level.

Finally, completed studies also increases satisfaction in private as well as professional life. Whereas there is a direct link between income and completed studies, the private satisfaction could increase e.g. because of better mating possibilities.

3.6 Selection

There was a substantial decline in the number of participants especially during the first follow-up interview at the age of thirty. This decline in sample size is a common methodological issue in studies with a longitudinal design.[88] Meulemann, Wieken-Mayser and Wiese (1987) analysed the sample composition of the dataset and compared the 10th grade sample with the participants of the first follow-up interview. They found no differences in composition with respect to gender, age, religion and only small differences for socioeconomic background and residence. However, individuals without "Abitur" are significantly underrepresented in the sample. The reason for this might be the early experience of failure and therefore the reluctance to admit this by participating in the study again (Meulemann, Wieken-Mayser, Wiese, 1987). Since the selection mechanism is observable and a large number of background characteristics can be controlled, including cognitive skills, the use of inverse probability weighting is appropriate (Wooldridge 2002). Further details on inverse probability weighting and possible improvements can be found in Flossmann (2006).[89]

A further estimation was made of the reduced form wage and prestige regressions as seen earlier (Table 3.4 and 3.5, columns 2). A dummy variable on whether "Abitur" grades are available in the dataset is included in the selection equation. The reason is the fore-

[87] Meulemann (2001) includes contemporaneous attribution into the happiness regressions and finds also significant effects.
[88] Even the LAU, a longitudinal study of high school students in Hamburg officially supported by the ministry lost a large number of its initial participants after two years. LAU 5 started 1996 with more than 12000 students but due to movements, class repetitions etc. only about 9800 could be interviewed again in 1998 (LAU, 2007).
[89] I profited strongly from a discussion with Anton Flossmann on this issue.

mentioned selection mechanism.[90] The results are depicted in Table 3.7. As can be seen, nothing much changes. Especially the effects of the variables measuring noncognitive skills remain almost unchanged and the t-statistics even increase.

3.7 Conclusion

The analysis confirms recent insights that noncognitive skills formed during childhood have a significant and lasting impact on success in life and influence performance in school, professional career and happiness.

Several basic concepts from motivational psychology were focused upon. Attribution, self-regulation and social skills were the three dimensions of noncognitive skills built into the analysis. Motivation and self-regulation are important for performance in school. The identified effect of self-regulation is larger for girls than for boys, which is probably due to their early development during adolescence.

The analysis of the labour market shows that for males, wages, income and occupational prestige increase with the level of self-regulation. The effect on wages is partly due to different initial conditions and selection into more prestigious jobs. For women it can be observed that a high degree of self-regulation leads to more active participation in the labour market (more hours worked) and therefore a higher income. Probably coping with family stress and employment pressure is easier with sound self-regulatory capacities. Women with a high degree of volition also work in more prestigious jobs, but an additional effect on wages cannot be identified. Finally, individuals that attribute success to internal and unstable factors seem to be more successful than others. As discussed, these results are in accordance with several other motivational theories.

Since self-regulation and motivation are important for success, a crucial question is how the development of these skills can be supported and fostered. Several psychological studies stress the importance of the relationship between the parents and child within the first years of life: Branson (2000) discusses beneficial parental interventions that support the development of self-regulation within several development phases (e.g. 0-3 months, 4-12 months, 12-36 months): Young infants need predictable interactions and rhythms for development and the presence of trusted persons give them some control over new situations. Preschool and Kindergarten children should be given reasons for existing rules

[90] Furthermore, in some rare cases, the schools were not cooperative and did not pass the exam results and student files to the researchers, which also lower the probability of location and therefore participation in the follow-up interview.

and the parental focus should not be on punishment but rather on problem solving and the explanation of appropriate behaviour (Branson, 2000). Furthermore, children confronted with excessive demands by their mothers seem to develop often a pronounced fear of failure later on (Heckhausen and Heckhausen, 2006, p.: 417). To summarize the evidence so far, there is the possibility for parents to facilitate the development of noncognitive skills especially in early childhood, thereby increasing the success of their children later in life.

3.8 Tables and Figures

Table 3.1: Descriptive statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
FEMALE	3385	0.4715	0.4993	0	1
IST	3191	40.479	8.9614	12	70
SELFREG_H MALE	2608	.10429	.30570	0	1
SELFREG_M MALE	2608	.35621	.47897	0	1
SELFREG_L MALE	2608	.07131	.25740	0	1
SELFREG_H FEM	2608	.13573	.34257	0	1
SELFREG_M FEM	2608	.26955	.44381	0	1
SELFREG_L FEM	2608	.06288	.24279	0	1
UNSTABLE_S	3222	6.3386	1.9332	0	10
STABLE_S	3222	5.6245	2.0122	0	10
UNSTABLE_FI	1969	6.6166	1.8653	0	10
STABLE_FI	1962	6.3028	1.9541	0	10
UNSTABLE_SI	1591	6.8228	1.5654	0	10
STABLE_SI	1590	6.3572	1.6178	1	10
INTERNAL_S	3224	7.6421	1.5601	0	10
EXTERNAL_S	3222	4.3246	2.4258	0	10
INTERNAL_FI	1964	6.9287	1.672	0	10
EXTERNAL_FI	1966	5.998	2.2112	0	10
INTERNAL_SI	1589	7.0755	1.3171	1	10
EXTERNAL_SI	1591	6.0987	1.8709	0	10
FRIEND	2616	1.8341	0.6264	1	3
AV GRADE	1783	2.9682	0.5396	1.0833	4.1
AGE	3286	15.418	0.8879	13	19
EDUC FATHER	3223	5.8787	4.2744	1	13
EDUC MOTHER	3245	4.2148	3.537	1	13
LN WAGE	1748	3.1083	0.6072	0.0202	6.6924
EXPERIENCE	1861	13.002	5.9471	0	26.417
EXPERIENCE2	1861	204.4	148.68	0	697.84
STUDY	1801	0.6824	0.4657	0	1
CIV_SERVANT	1853	0.4614	0.4986	0	1
MITTLERE REIFE	1987	0.1741	0.3793	0	1
FACHOBERSCHULE	1987	0.1007	0.3009	0	1
ABITUR	1987	0.7121	0.4529	0	1
TREIMAN SCORE	1841	55.1189	11.6635	18	78.9
REPEAT	3182	0.3654	0.4816	0	1

Notes: The descriptive statistics are calculated for the total sample. Of the 3385 students 145 were not present at the first interview and therefore the net sample size is 3240.

Table 3.2: Correlations between variables

	IST	M1	M2	M3	F1	F2	F3	IA	EA	SA	UA	FR
IST	1.00											
M1	0.12	1.00										
M2	0.13	-0.26	1.00									
M3	0.04	-0.09	-0.21	1.00								
F1	-0.08	-0.14	-0.30	-0.11	1.00							
F2	-0.14	-0.21	-0.45	-0.16	-0.24	1.00						
F3	-0.10	-0.09	-0.19	-0.07	-0.10	-0.16	1.00					
IA	-0.04	0.01	-0.09	-0.03	0.06	0.06	0.01	1.00				
EA	-0.10	-0.02	-0.07	-0.05	0.05	0.08	0.00	0.20	1.00			
SA	-0.07	-0.02	-0.08	-0.05	0.05	0.08	0.01	0.57	0.67	1.00		
UA	-0.08	0.00	-0.08	-0.04	0.06	0.07	-0.01	0.46	0.72	0.26	1.00	
FR	0.06	-0.15	0.06	0.15	-0.18	0.02	0.13	-0.02	-0.06	-0.02	-0.06	1.00

IST: psychometric test score; M1-M3: Self-regulation male (1-high); F1-F3: Self-regulation female (1-high); IA/EA: internal/ external attribution; SA/UA: stable/ unstable attribution; FR: Social Skills (Friendship)

Table 3.3: School performance

	1 Average Grade	2 Grade	3 Math	4 German	5 Av. Grade	6 Av. Grade
FEMALE	0.0107	0.0127	-0.0415	-0.0800	0.0149	-0.0464
	(0.27)	(0.33)	(0.62)	(1.17)	(0.38)	(0.85)
IST	-0.0124***	-0.0121***	-0.0362***	-0.0179***	-0.0122***	-0.0118***
	(7.08)	(6.90)	(11.87)	(6.11)	(6.95)	(6.97)
SELFREG_H	-0.1260**	-0.1245**	0.0842	-0.2696***	-0.1257**	-0.1371**
	(2.11)	(2.09)	(0.80)	(2.95)	(2.14)	(2.28)
SELFREG_M	-0.0841	-0.0869*	-0.0288	-0.0654	-0.0880*	-0.0945*
	(1.61)	(1.67)	(0.33)	(0.81)	(1.70)	(1.82)
ATTR. INTERNAL		-0.0249***	-0.0372**	-0.0049	-0.0238**	-0.0203**
		(2.71)	(1.99)	(0.36)	(2.57)	(2.22)
ATTR. EXTERNAL		0.0070	0.0291**	-0.0050	0.0071	0.0104*
		(1.14)	(2.52)	(0.56)	(1.15)	(1.66)
FRIEND	0.0099	0.0116	-0.0763	-0.0108	0.0171	0.0155
	(0.36)	(0.42)	(1.43)	(0.22)	(0.60)	(0.51)
REPEAT	0.2243***	0.2176***	0.4975***	0.2828***	0.2050***	0.1679***
	(6.27)	(6.02)	(7.20)	(4.98)	(5.46)	(4.88)
FATHER EDU	-0.0057	-0.0057	-0.0065	0.0072	-0.0053	-0.0027
	(1.54)	(1.53)	(0.87)	(0.97)	(1.42)	(0.66)
MOTHER EDU	-0.0175***	-0.0190***	-0.0075	-0.0345***	-0.0186***	-0.0181***
	(3.29)	(3.57)	(0.76)	(3.89)	(3.45)	(3.34)
CONSTANT	3.6431***	3.7518***	5.2438***	4.3376***	3.845***	3.7047***
	(30.55)	(28.12)	(20.75)	(21.66)	(27.30)	(27.29)
ATTR_UNSTABLE	-0.0037					
	(0.46)					
ATTR_STABLE	-0.0033					
	(0.40)					
EVENINGS					-0.0208*	
					(1.83)	
RES_SCHOOL					-0.1063	
					(1.06)	
Method	OLS	OLS	OLS	OLS	OLS	FE
Observations	1385	1385	1383	1384	1384	1385
R^2	0.10	0.11	0.14	0.09	0.11	0.25

Notes: * denotes significance at the 10 percent, ** at the 5 percent and *** at the 1 percent level, respectively. The t-values are based upon a clustering robust covariance-matrix on class level. Dependent variables are untransformed grades. The average grade was calculated as arithmetic mean of all available grades.

Table 3.4: Professional career: Wage

	1 ln(wage)	2 ln(wage)	3 ln(wage)	4 ln(wage)	5 ln(income)	6 ln(wage)
EXPERIENCE	0.0492*** (3.32)		0.0537*** (3.33)	0.0507*** (3.39)	0.0623** (3.05)	0.0595*** (2.96)
EXPERIENCE_2	-0.0007 (1.29)		-0.0010* (1.76)	-0.0008 (1.52)	-0.0008 (1.19)	-0.0013* (1.67)
FEMALE	-0.0348 (0.49)	-0.0730 (0.94)	-0.0613 (0.72)	-0.0673 (0.88)	-0.4994*** (6.65)	-0.1015 (1.10)
IST	0.0027* (1.82)	0.0041*** (2.72)	0.0040*** (2.58)	0.0034** (2.30)	0.0022 (1.33)	0.0039* (1.91)
SELFREG_H MALE	0.2015** (3.06)	0.2261*** (3.21)	0.2013*** (2.69)	0.1891*** (2.71)	0.2507*** (3.40)	0.2031** (1.97)
SELFREG_M MALE	0.1369*** (3.25)	0.1628*** (3.45)	0.0945* (1.93)	0.1156** (2.53)	0.1647*** (3.35)	0.0132 (0.23)
SELFREG_H FEMALE	-0.0174 (0.21)	-0.0045 (0.05)	-0.0446 (0.51)	-0.0504 (0.62)	0.1859** (2.44)	-0.0241 (0.24)
SELFREG_M FEMALE	-0.0109 (0.15)	0.0249 (0.33)	0.0086 (0.11)	0.0149 (0.21)	0.0591 (0.84)	0.0170 (0.19)
ATTR. STABLE	-0.0109 (1.50)	-0.0090 (1.16)	-0.0262 (0.78)		-0.0204 (0.57)	-0.0413 (0.87)
ATTR. UNSTABLE	0.0133 (1.59)	0.0127 (1.45)	0.1132*** (3.13)		0.0959** (2.37)	0.1032** (2.34)
FRIEND	-0.0319 (1.19)	-0.0210 (0.73)	-0.0216 (0.78)	-0.0313 (1.14)	-0.0199 (0.70)	0.0151 (0.45)
CONSTANT	2.1928*** (11.47)	3.0644*** (26.89)	1.6219*** (3.92)	1.6312*** (4.13)	6.463*** (13.90)	1.9922*** (3.35)
ATTR. INTERNAL				0.0598* (1.72)		
ATTR. EXTERNAL				0.0270 (0.95)		
ABITUR GRADE						-0.0179 (0.43)
Control Dummies						
School Degree	✓	-	✓	✓	✓	✓
CIVIL SERVANT	✓	-	✓	✓	✓	✓
STUDY	✓	-	✓	✓	✓	✓
Method	OLS	OLS	TSLS	TSLS	TSLS	TSLS
Observations	1048	1051	1033	1033	1036	640
R^2	0.19	0.06	0.10	0.17	0.39	0.04
Hansen Stat. / p-value			1.54 (0.46)	0.44 (0.80)	0.93 (0.63)	2.95 (0.23)

Notes:* denotes significance at the 10 percent, ** at the 5 percent and *** at the 1 percent level respectively. The t-values are based on a heteroscedasticity robust covariance-matrix.

Table 3.5: Professional career: Treiman (1977) Occupational Prestige

	1 Treiman Score	2 Treiman Score	3 Treiman Score	4 Treiman Score	5 ln (wage)
EXPERIENCE	0.3337		0.3700	0.4214	0.0466***
	(1.03)		(1.07)	(1.22)	(3.23)
EXPERIENCE_2	-0.0108		-0.0139	-0.0149	-0.0006
	(0.95)		(1.15)	(1.23)	(1.16)
FEMALE	-1.0755	-0.8100	-1.5175	-1.3224	-0.0287
	(0.72)	(0.44)	(0.97)	(0.83)	(0.41)
IST	0.0666**	0.1227***	0.0788**	0.0726**	0.0017
	(1.98)	(3.09)	(2.23)	(2.06)	(1.15)
SELFREG_H MALE	3.579**	5.2710***	3.5738**	3.3580**	0.1629**
	(2.33)	(2.92)	(2.31)	(2.12)	(2.48)
SELFREG_M MALE	2.470**	3.9583***	1.8827	1.8723	0.1219***
	(2.03)	(2.78)	(1.51)	(1.48)	(2.93)
SELFREG_H FEMALE	2.9008**	3.7700**	2.7577**	2.6318*	-0.0345
	(2.26)	(2.33)	(2.05)	(1.90)	(0.42)
SELFREG_M FEMALE	2.254**	3.2503**	2.193*	2.1402*	-0.0087
	(1.98)	(2.20)	(1.87)	(1.78)	(0.12)
ATTR. STABLE	0.0030	0.0136	-0.1978		-0.0106
	(0.02)	(0.08)	(0.35)		(1.45)
ATTR. UNSTABLE	0.0256	-0.0123	1.2834*		0.0131
	(0.17)	(0.07)	(1.83)		(1.56)
FRIEND	-0.1519	0.5805	0.0650	-0.1006	-0.0309
	(0.29)	(0.96)	(0.12)	(0.18)	(1.17)
CONSTANT	40.1937	47.4366	32.993***	29.54***	1.9134
	(10.33)	(16.79)	(4.52)	(3.81)	(9.09)
ATTR. INTERNAL				1.5302**	
				(2.17)	
ATTR. EXTERNAL				-0.1346	
				(0.25)	
TREIMAN SCORE					0.0075***
					(4.80)
Control Dummies					
School Degree	✓	-	✓	✓	✓
CIVIL SERVANT	✓	-	✓	✓	✓
STUDY	✓	-	✓	✓	✓
Method	OLS	OLS	TSLS	TSLS	OLS
Observations	1106	1108	1091	1091	1038
R^2	0.30	0.03	0.29	0.27	0.21
Hansen Stat. / p-value			0.25/ 0.88	2.09/ 0.35	

Notes: * denotes significance at the 10 percent, ** at the 5 percent and *** at the 1 percent level, respectively. The t-values based on a heteroscedasticity robust covariance-matrix. The dependent variables (column 1-4) are the untransformed Treiman (1977) scores.

Table 3.6: Happiness

	1 Happiness work	2 Happiness private
FEMALE	-0.1970	0.7318**
	(0.69)	(2.48)
IST	0.0076	0.0132**
	(1.32)	(2.44)
SELFREG_H MALE	-0.0439	0.5542*
	(0.18)	(1.93)
SELFREG_M MALE	-0.0733	0.4961**
	(0.36)	(2.17)
SELFREG_H FEMALE	0.0562	0.1007
	(0.20)	(0.41)
SELFREG_M FEMALE	-0.1378	-0.0488
	(0.56)	(0.22)
ATTR. INTERNAL	0.2705**	0.0871
	(2.33)	(0.80)
ATTR. EXTERNAL	0.1471	0.1781*
	(1.41)	(1.82)
FRIEND	-0.2332**	0.0132
	(2.42)	(0.14)
FATHER EDU	-0.0381**	-0.0249
	(2.24)	(1.50)
MOTHER EDU	0.0100	-0.0096
	(0.46)	(0.46)
MITREIFE	0.3624	1.2162
	(0.85)	(2.39)**
FACHOB	0.6121	1.3272
	(1.38)	(2.53)**
ABITUR	0.3272	1.0775
	(0.77)	(2.13)**
STUDY	0.3801***	0.2502*
	(2.61)	(1.74)
CONSTANT	4.0426***	3.8039***
	(3.16)	(3.14)
Method	TSLS	TSLS
Observations	1238	1242
R2	0.07	0.05

Notes: * denotes significance at the 10 percent, ** at the 5 percent and *** at the 1 percent level, respectively. The t-values are based on a heteroscedasticity robust covariance-matrix.

Table 3.7: Selection effects

	1 ln(wage)	2 Treiman Score
FEMALE	-0.0674	-1.0998
	(0.80)	(0.62)
IST	0.0050***	0.1230**
	(3.28)	(2.73)
SELFREG HIGH MALE	0.2301***	5.6619***
	(3.30)	(3.11)
SELFREG MED MALE	0.1730***	4.2444***
	(3.54)	(3.11)
SELFREG HIGH FEM.	-0.0146	4.0898**
	(0.16)	(2.87)
SELFREG MED FEM.	0.0189	3.6897**
	(0.22)	(2.44)
ATTR. STABLE	-0.0089	0.0813
	(1.14)	(0.50)
ATTR. UNSTABLE	0.0118	-0.0479
	(1.43)	(0.27)
FRIEND	-0.0201	0.7968
	(0.05)	(1.17)
CONSTANT	3.0232***	46.4672***
	(25.07)	(16.29)
Method	IPW	IPW
Observations	2468	2468

Notes: * denotes significance at the 10 percent, ** at the 5 percent and *** at the 1 percent level, respectively. Mother's education level, father's education level, a dummy variable whether the Abitur grade is available, self-regulation, friendship, female and IST are included in the selection equation. All standard errors are bootstrapped (100 replications).

Appendix

Definitions

FEMALE: Dummy variable which is one for women and zero otherwise.

IST: Number of correctly solved questions in the Intelligence Structure Test (IST; Amthauer, 1953). The tests carried out were Wortauswahltest, Analogietest, Zahlenreihentest and Würfeltest.

SELFREG: Self-regulation is taken from the parents' questionnaire and measures on a scale from 1 (very gifted) to 3 (not gifted) the child's ability to assert themselves and to deal with difficult situations. Self-regulation enters the regressions as a set of dummy variables.

UNSTABLE ATTRIBUTION: Unstable attribution is constructed as sum of attribution to EFFORT and LUCK. It ranges between zero and ten and higher scores indicate stronger attribution to unstable factors.

STABLE ATTRIBUTION: Stable attribution is constructed as sum of attribution to ABILITY and FAMILY. It ranges between zero and ten and higher scores indicate stronger attribution to stable factors.

INTERNAL ATTRIBUTION: Internal attribution is constructed as sum of attribution to ABILITY and EFFORT. It ranges between zero and ten and higher scores indicate stronger attribution to internal factors.

EXTERNAL ATTRIBUTION: External attribution is constructed as sum of attribution to FAMILY and LUCK. It ranges between zero and ten and higher scores indicate stronger attribution to external factors.

FRIEND: Friend is taken from the parents' questionnaire and measures on a scale from 1 (very gifted) to 3 (not gifted) the child's ability to get new friends.

AVERAGE GRADE: Average over all available grades (13^{th} class).

EDUC FATHER / EDUC MOTHER: The educational attainment is coded from 1 (elementary school without further training) to 13 (university degree) in the dataset (Central Archive for Empirical Social Research, 2007). I use the original variable for the analysis.

LN (WAGE): Is the natural logarithm of the wage. Wage is computed as net monthly income divided by average hours worked per month for all individuals that worked at the time of the second follow-up. Note that average hours worked is measured by actual and not by contractual specified hours. Average hours worked per month is calculated as weekly hours times four.

EXPERIENCE: Sum of all actual periods of employment (measured in years).

STUDY: Dummy variable for successful completed studies.

CIVIL SERVANT: Dummy variable for occupation as civil servant.

REPEAT: Dummy variable for students that at least once repeated a class.

TREIMAN SCORE: Treiman's (1977) occupational prestige score.

Figure 3.2: Distribution of net hourly wage

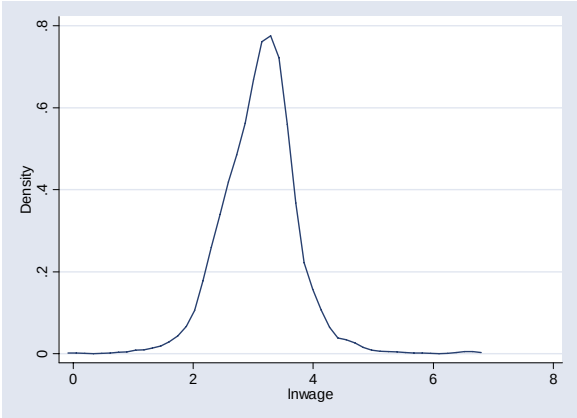


Figure 3.3: Distribution of IST

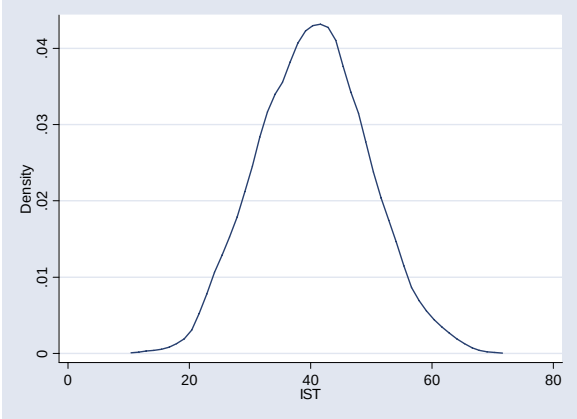
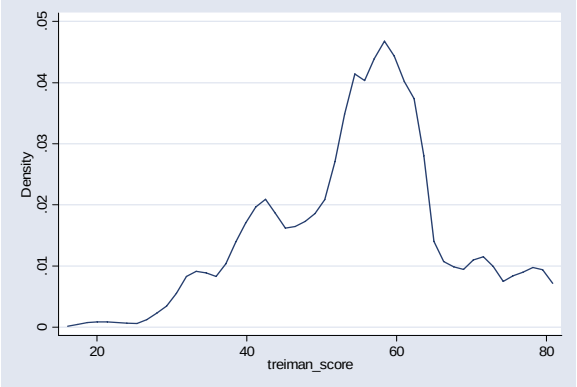


Figure 3.4: Distribution of Treiman's (1977) Occupational Prestige



References

Abramson, L.Y., Seligman, M.E.P. and J.D. Tesdale (1989): Hopelessness depression: A theory-based subtype of depression, *Psychological Review,* 96, 358-372.

Achtziger A. and P.M. Gollwitzer (2006): Motivation und Volition im Handlungsverlauf*, in Motivation und Handeln, Eds. J. Heckhausen and H. Heckhausen, Heidelberg: Springer, 277-302*.

Amthauer, R. (1953): Intelligenz-Struktur-Test. 2, erweiterte Auflage, *Göttingen: Verlag für Psychologie, Dr. C.J. Hogrefe*.

Bierbrauer, G. (1996): Sozialpsychologie*, in Grundriß der Psychologie: eine Reihe in 22 Bänden, hrsg. von Herbert Selg und Dieter Ulich; Stuttgart: Kohlhammer*.

Birkelbach, K., Hellwig, J.O., Hemsing, W. and H. Meulemann (2000): Success and Satisfaction in the Occupational Career from Youth to Middle Age in a Cohort of German Gymnasium Pupils, 1969, 1984, 1997, *in Scheuch, E.K. and Sciulli, D. (eds.) The annals of the IIS. Nouvelle Serie Volume VII*.

Branson, M. B. (2000): Self-Regulation in early childhood, Nature and Nurture*, New York: The Guilford Press*.

Central Archive for Empirical Social Research (2007): Dataset Gymnasiastenstudie, Köln: Zentralarchiv für Empirische Sozialforschung.

Corno, L (1993): The Best-Laid Plans. Modern Conceptions of Volition and Educational Research, *Educational Researcher*, 22 (2), 14-22.

Duckworth A.L. and M.E.P. Seligman (2005): Self-Discipline outdoes IQ in Predicting Academic Performance of Adolescents, *Psychological Science*, 16, 939-944.

Flossmann, A. (2006): Accounting for Missing Data in M-Estimation- A General Matching Approach, University of Konstanz: Working Paper.

Flossmann, A., Piatek, R. and L. Wichert (2006): Going beyond Returns to Education: The Role of Noncognitive Skills on Wages in Germany, University Konstanz: Working Paper.

Gagné F. and F. St. Père (2002) : When IQ is controlled, does motivation still predict achievement?, *Intelligence,* 30, 71-100.

Göggel K. (2007): Economic analysis of self-regulation among the German youth, *Presentation at the University Konstanz, Seminar Skill Formation*.

Hartup, W. W. (1992): Friendships and their Developmental Significance. *in Childhood Social Development: Contemporary Perspectives. Edited by H. McGurk; Hove: Lawrence Erlbaum Associates Ltd. Publishers*.

Heckhausen H. (1984): Attributionsmuster für Leistungsergebnisse: Individuelle Unterschiede, mögliche Arten und deren Genese, *in: E.F. Weinert & R.H. Kluwe (eds.) Metakognition, Motivation und Lernen (133-164). Stuttgart: Kohlhammer*.

Heckhausen H. and P.M. Gollwitzer (1987): Thought contents and cognitive functioning in motivational versus volitional states of mind, *Motivation and Emotion*, 11, 101-120.

Heckhausen J. and H. Heckhausen (2006): Motivation und Entwicklung, *in Motivation und Handeln, Eds. J. Heckhausen and H. Heckhausen, Heidelberg: Springer*.

Heckman J. and Y. Rubinstein (2001): The importance of noncognitive skills: Lessons from the GED testing program, *American Economic Review***,** 91,145-149.

Heckman, J., Stixrud J. and S. Urzua (2006*): The Effects of Noncognitive abilities on Labor Market outcomes and social behaviour*, *Journal of Labor Economics, 24(3) 411-82*.

James, William (1890): The principles of psychology (Vol. 2), *New York: Dover Publications*.

LAU (2007): Aspekte der Lernausgangslage und der Lernentwicklung. Lehmann R.H., HamburgerBildungsserver, *http://lbs.hh.schule.de/welcome.phtml?unten=/schulentwicklung/lau/lau5/*.

Meulemann, H. (2007): Projektbericht zur Vorlage bei der DFG*, Az.Me 577/7-1; -2*.

Meulemann, H. (2001): Life satisfaction from late adolescence to mid-life, *Journal of Happiness Studies*, 2, 445-465.

Meulemann, H. Wieken-Mayser M. and W. Wiese (1987*):* Lebensplanung und Lebenserfolg in privilegierten Lebensläufen. Erster Teil: Arbeitsbericht, *Köln: Zentralarchiv für Empirische Sozialforschung*.

Meyer, W.U. (1973): Leistungsmotiv und Ursachenerklärung von Erfolg und Misserfolg, *Stuttgart: Klett*.

Peterson C. and M.E.P. Seligman (2004): Character strengths and virtues: a handbook and classification. *American Psychological Association, Oxford University Press*.

Plug, E. (2004): Estimating the Effect of Mother's Schooling on Children's Schooling using a sample of Adoptees, *American Economic Review*, 94, 358-368.

Rotter J. (1966): Generalized Expectancies for Internal versus External Control of Reinforcement, *Psychological Monographs*, 80, 1-28.

Sen, A. K. and K.A. Hagtvet (1993): Correlations among creativity, intelligence, personality, and academic achievement, *Perceptual and Motor Skills*, 77, 497-498.

Stiensmeyer-Pelster J. and H. Heckhausen: Kausalattribution von Verhalten und Leistung, *in Motivation und Handeln, Eds. J. Heckhausen and H. Heckhausen, Heidelberg: Springer*.

SOEP (2007): Das sozio-ökonomische Panel, Stichprobenentwicklung, *http://www.diw.de/deutsch/sop/uebersicht/index.html*.

Treiman D.J. (1977): Occupational Prestige in Comparative Perspective, *New York: Academic Press*.

Urberg, K.A., Degirmencioglu, S.M. and C. Pilgrim (1997): Close Friend and Group Influence on Adolsecent Cigarette Smoking and Alcohol Use, *Developmental Psychology*, 33, 834-844.

Weiner B., and A. Kukla (1970): An Attributional Analysis of Achievement Motivation, *Journal of Personality and Social Psychology*, 15, 1-20.

Weiner, B., Frieze, I., Kukla, A., Reed, L., Rest, S. and R. Rosenbaum (1971): Perceiving the causes of success and failure, *Morristown, N. J.: General Learning Press*.

Weiner, B. (1972): Theories of motivation. *Chicago: Markham Publishing Company*.

Wooldridge J. (2002): Econometric Analysis of Cross Section and Panel Data, *Cambridge, Massachusetts: The MIT Press*.

References

Abramson, L.Y., Seligman, M.E.P. and J.D. Tesdale (1989): Hopelessness depression: A theory-based subtype of depression, *Psychological Review,* 96, 358-372.

Achtziger A. and P.M. Gollwitzer (2006): Motivation und Volition im Handlungsverlauf*, in Motivation und Handeln, Eds. J. Heckhausen and H. Heckhausen, Heidelberg: Springer, 277-302*.

Aizenman J. (1989a): Country Risk and Contingencies, *International Economic Journal,* 3, 81-102.

Aizenman J. (1989b): Country Risk, Incomplete Information and Taxes on International Borrowing, *Economic Journal,* 99, 147-161.

Amthauer, R. (1953): Intelligenz-Struktur-Test. 2, erweiterte Auflage, *Göttingen: Verlag für Psychologie, Dr. C.J. Hogrefe*.

Arellano M. and O. Bover (1995): Another Look at the Instrumental-Variable Estimation of Error-Components Models, *Journal of Econometrics*, 68, 29-52.

Arellano M. and S. Bond (1991): Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations, *Review of Economic Studies,* 58, 277-297.

Asiedu E. and A.P. Villamil (2002): Imperfect enforcement, foreign investment, and foreign aid, *Macroeconomic Dynamics*, 6, 476-495.

Backes-Gellner, U. and A. Schlinghoff (2004): Careers, incentives and publication patterns of US and German (business) economists, *http://papers.ssrn.com/sol3/papers.cfm?abstract_id=616822*.

Barro, R.J. and J-W. Lee (2004): IMF programs: Who is chosen and what are the effects?, *mimeo, Harvard University*.

Baser, O. and E. Pema (2004): Publications over the academic life-cycle: Evidence for academic economists, *Economics Bulletin*, 1, 1-8.

Bell, J. and J. Seater (1978): Publishing performance: Departmental and individual, *Economic Inquiry*, 16, 599-615.

Bierbrauer, G. (1996): Sozialpsychologie*, in Grundriß der Psychologie: eine Reihe in 22 Bänden, hrsg. von Herbert Selg und Dieter Ulich; Stuttgart: Kohlhammer*.

Bilias Y., S. Chen and Z.Ying (2000): Simple Resampling methods for censored Regression Quantiles, *Journal of Econometrics*, 99, 373-386.

Bird, G. and D. Rowlands (2002): Do IMF Programmes Have a Catalytic Effect on Other International Capital Flows?, *Oxford Development Studies,* 30, 229-249.

Birkelbach, K., Hellwig, J.O., Hemsing, W. and H. Meulemann (2000): Success and Satisfaction in the Occupational Career from Youth to Middle Age in a Cohort of German Gymnasium Pupils, 1969, 1984, 1997, *in Scheuch, E.K. and Sciulli, D. (eds.) The annals of the IIS. Nouvelle Serie Volume VII*.

Blundell R. and S.R. Bond (1998): Initial conditions and moment restrictions in dynamic panel data models, *Journal of Econometrics,* 87, 115-143.

Bond S.R. (2002): Dynamic Panel Data Models: A Guide to Micro Data Methods and Practice, *CEMMAP Working Paper,* No. CWP09/02.

Branson, M. B. (2000): Self-Regulation in early childhood, Nature and Nurture, *New York: The Guilford Press*.

Bruno, G.S.F. (2005): Approximating the bias of the LSDV estimator for dynamic unbalanced panel data models, *Economics Letters,* 87, 361-66.

Cantor R. and F. Packer (1996): Determinants and Impact of Sovereign Credit Ratings, *FRBNY Economic Policy Review*, October 1996.

Central Archive for Empirical Social Research (2007): Dataset Gymnasiastenstudie, Köln: Zentralarchiv für Empirische Sozialforschung.

Chang C., E. Fernandez-Arias, and L. Serven (1998): Measuring Aid Flows: A New Approach, *World Bank, Development Research Group, Washington D.C.*

Ciocchini F. and E. Durbin and D. Ng (2003): Does Corruption Increase Emerging Market Bond Spreads?, *Journal of Economics and Business,* 55, 503-28.

Combes, P. and L. Linnemer (2003): Where are the economists who publish? Publication concentration and rankings in Europe based on cumulative publications, *Journal of the European Economic Association*, 1, 1250-1308.

Corno, L (1993): The Best-Laid Plans. Modern Conceptions of Volition and Educational Research, *Educational Researcher*, 22 (2), 14-22.

Coupé, T. (2003): Revealed performances: Worldwide rankings of economists and economics departments, 1990-2000, *Journal of the European Economic Association*, 1, 1309-1345.

Coupé, T. (2004): What do we know about ourselves? On the economics of economics, K*yklos*, 57, 197-215.

Coupé, T., V. Smeets and F. Warzynski (2005): Incentives, sorting and productivity along the career: Evidence from a sample of top economists, *The Journal of Law, Economics and Organization*, 22, 137-167.

Coupé, T., V. Smeets and F. Warzynski (2003): Incentives in economic departments: Testing Tournaments, *Aarhus School of Business Working Paper 03-25*.

Cunningham, A., L. Dixon, and S. Hayes (2001): Analysing yield spreads on emerging market sovereign bonds, *Bank of England Financial Stability Review,* December 2001, 175-186.

Dnes, A. and N. Garoupa (2005): Academic tenure, post-tenure effort, and contractual damages, *Economic Inquiry*, 43, 832-839.

Duckworth A.L. and M.E.P. Seligman (2005): Self-Discipline outdoes IQ in Predicting Academic Performance of Adolescents, *Psychological Science*, 16, 939-944.

Easterly W. (2003): Can Foreign Aid Buy Growth?, *Journal of Economic Perspectives,* 17, 23-48.

Easterly W., R. Levine, and D. Roodman (2004): Comment: Aid, Policies, and Growth, *American Economic Review*, 94, 774-780.

Eaton J., M. Gersovitz, and J. Stiglitz (1986): The Pure Theory of Country Risk, *European Economic Review,* 30, 481-513.

Eichengreen B. and A. Mody (1998): What explains changing Spreads on Emerging Market Debt: Fundamentals or Market Sentiment?, *NBER Working Paper No. 6408*.

Fitzenberger, B. (1997): A Guide to Censored Quantile Regressions, in G.S. Maddala and C.R. Rao (eds.), *Handbook of Statistics*, 15, 405-437.

Flossmann, A. (2006): Accounting for Missing Data in M-Estimation- A General Matching Approach, University of Konstanz: Working Paper.

Flossmann, A., Piatek, R. and L. Wichert (2006): Going beyond Returns to Education: The Role of Noncognitive Skills on Wages in Germany, University Konstanz: Working Paper.

Frank, B. and G. Schulze (2000): Does economics make citizens corrupt?, *Journal of Economic Behavior and Organization*, 43, 101-113.

Gagné F. and F. St. Père (2002) : When IQ is controlled, does motivation still predict achievement?, *Intelligence*, 30, 71-100.

Gelos, R.G., R. Sahay, and G. Sandleris (2004): Sovereign Borrowing by Developing Countries: What Determines Market Access?, *IMF Working Paper No. 04/221*.

Göggel K. (2007): Economic analysis of self-regulation among the German youth, *Presentation at the University Konstanz, Seminar Skill Formation*.

Goodwin, T.H. and R.D. Sauer (1995): Life cycle productivity in academic research: Evidence from cumulative publication histories of academic economists, *Southern Economic Journal*, 61 728-743.

Goyal, S. and van der Leu, M. and J. L. Moraga-Gonzales (2004): Economics: An emerging small world?, *CESifo Working Paper No. 1287*.

Grimes, P.W. and C.A. Register (1997): Career Publications and academic job rank: Evidence from the class of 1968, *Journal of Economic Education*, 28, 82-92.

Hadi A.S.(1994): A Modification of a Method for the Detection of Outliers in Multivariate Samples, *Journal of the Royal Statistical Society, Series B*, 56, 393-396 .

Hall, B., J. Mairesse and L. Turner (2005): Identifying age, cohort and period effects in scientific research productivity: Discussion and illustration using simulated and actual data on French physicists, *NBER Working Paper No. 11739*.

Hansen H. and F. Tarp (2000): Aid effectiveness disputed, *Journal of International Development*, 12, 375-398.

Hansen H. and F. Tarp (2001): Aid and growth regressions, *Journal of Development Economics,* 64, 547-570.

Hansen L. (1982): Large Sample Properties of Generalized Method of Moments Estimators, *Econometrica,* 50, 1029-1054.

Haque U. N., M. Kumar, N. Mark, and D. Mathieson (1996): The Economic Content of Indicators of Developing Country Creditworthiness, *IMF-Staff Papers,* 43, 688-724.

Haque U. N., N. Mark, and D. Mathieson (1998): The Relative Importance of Political and Economic Variables in Creditworthiness Ratings, *IMF Working Paper WP/98/46*.

Harms P. and M. Lutz (2004): The Macroeconomic Effects of Aid, *University of St. Gallen Discussion Paper 2004-11*.

Harms P. and M. Lutz (2006): Aid, Governance, and Private Foreign Investment: Some Puzzling Findings for the 1990s, *The Economic Journal*, 116, 773-790.

Hartley, J.E., J.W. Monks and M.D. Robinson (2001): Economists' publication patterns, *The American Economist*, 45, 80-85.

Hartup, W. W. (1992): Friendships and their Developmental Significance. *in Childhood Social Development: Contemporary Perspectives. Edited by H. McGurk; Hove: Lawrence Erlbaum Associates Ltd. Publishers*.

Heckhausen H. (1984): Attributionsmuster für Leistungsergebnisse: Individuelle Unterschiede, mögliche Arten und deren Genese, *in: E.F. Weinert & R.H. Kluwe (eds.) Metakognition, Motivation und Lernen (133-164). Stuttgart: Kohlhammer*.

Heckhausen H. and P.M. Gollwitzer (1987): Thought contents and cognitive functioning in motivational versus volitional states of mind, *Motivation and Emotion*, 11, 101-120.

Heckhausen J. and H. Heckhausen (2006): Motivation und Entwicklung, *in Motivation und Handeln, Eds. J. Heckhausen and H. Heckhausen, Heidelberg: Springer*.

Heckman J. and Y. Rubinstein (2001): The importance of noncognitive skills: Lessons from the GED testing program, *American Economic Review***,** 91,145-149.

Heckman, J., Stixrud J. and S. Urzua (2006*): The Effects of Noncognitive abilities on Labor Market outcomes and social behaviour*, *Journal of Labor Economics, 24(3) 411-82*.

Hollis, A. (2001): Co-authorship and the Output of Academic Economists, *Labour Economics*, 8, 503-30.

Honoré, B.E.(1992): Trimmed Lad and Least Squares Estimation of Truncated and Censored Regression Models with Fixed Effects, *Econometrica*, 60, 533-565.

Hutchinson, E.B. and T.L. Zivney (1995): The publication profile of economists, *Journal of Economic Education*, 26, 59-79.

James, William (1890): The principles of psychology (Vol. 2), *New York: Dover Publications*.

Jones, B.L. Nagin, D.S. and Roeder K. (2001): A SAS Procedure Based on Mixture Models for Estimating Development Trajectories, *Sociological Methods & Research*, 29, 374-393.

Judson R. and A. Owen (1999): Estimating Dynamic Panel Data Models: A Guide for Macroeconomists, *Economics Letters*, 65, 9-15.

Kalaitzidakis, P., Mamuneas, T. and T. Stengos (2003): Rankings of academic Journals and Institutions in economics, *Journal of the European Economic Association*, 1, 1346-1366.

Kaminsky G. and C. Reinhart (1999): The Twin Crises: The Causes of Banking and Balance-of-Payments Problems, *American Economic Review,* 89, 473-500.

Kenny, L. and R. Studley (1995): Economists' salaries and lifetime productivity, *Southern Economic Journal*, 65, 382-393.

Kiviet, J.F. (1995): On bias, inconsistency, and efficiency of various estimators in dynamic panel data models, *Journal of Econometrics,* 68, 53-78.

Laband, D. (2002): Contribution, Attribution and the Allocation of Intellectual Property Rights: Economics versus Agricultural Economics, *Labour Economics*, 9, 125-31.

Larrain F., H. Reisen, and J. von Maltzan (1997): Emerging Market Risk and Sovereign Credit Ratings?, *OECD Development Centre, Technical Paper No. 124*.

LAU (2007): Aspekte der Lernausgangslage und der Lernentwicklung. Lehmann R.H., HamburgerBildungsserver, *http://lbs.hh.schule.de/welcome.phtml?unten=/schulentwicklung/lau/lau5/*.

Lee S. (1993): Are the credit ratings assigned by bankers based on the willingness of LDC borrowers to repay?, *Journal of Development Economics*, 40, 349-359.

Levin, S. and P. Stephan (1991): Research productivity over the life cycle: Evidence for academic scientists, *American Economic Review*, 81, 114-132.

McDowell, J. and J. Smith (1992): The Effect of Gender-Sorting on Propensity to Coauthor: Implications for Academic Promotion, *Economic Inquiry*, 30, 68-82.

Meulemann, H. (2001): Life satisfaction from late adolescence to mid-life, *Journal of Happiness Studies*, 2, 445-465.

Meulemann, H. (2007): Projektbericht zur Vorlage bei der DFG, *Az.Me 577/7-1; -2*.

Meulemann, H. Wieken-Mayser M. and W. Wiese (1987*):* Lebensplanung und Lebenserfolg in privilegierten Lebensläufen. Erster Teil: Arbeitsbericht, *Köln: Zentralarchiv für Empirische Sozialforschung*.

Meyer, W.U. (1973): Leistungsmotiv und Ursachenerklärung von Erfolg und Misserfolg, *Stuttgart: Klett*.

Mody, A. and D. Saravia (2003): Catalyzing capital flows: Do IMF-supported programs work as commitment devices?, *IMF Working Paper 03/100*.

Moore, W., R. Newman and G. Turnbull (2001): Reputational capital and academic pay, *Economic Inquiry*, 39, 663-671.

Nickell S. (1981): Biases in Dynamic Models with Fixed Effects, *Econometrica*, 49, 1417-26.

OECD (2005): CRI online database, *Development Assistance Committee*.

Oster, S.M. and D.S. Hamermesh (1998): Aging and productivity among economists, *Review of Economics and Statistics*, 80, 154-156.

Pagan, A., and F. Vella (1989): Diagnostic Tests for Models Based on Individual Data: A Survey, *Journal of Applied Econometrics*, 4, 29-59.

Peterson C. and M.E.P. Seligman (2004): Character strengths and virtues: a handbook and classification. *American Psychological Association, Oxford University Press*.

Plug, E. (2004): Estimating the Effect of Mother's Schooling on Children's Schooling using a sample of Adoptees, *American Economic Review*, 94, 358-368.

Pohlmeier W. and V. Ulrich (1995): An Econometric Model of the Two-Part Decisionmaking Process in the Demand for Health Care, *The Journal of Human Resources*, 30, 339-361.

Political Risk Services: *International Country Risk Guide*, various issues.

Powell J. L. (1986): Censored Regression Quantiles, *Journal of Econometrics*, 32, 143-155

Powell, J.L. (1984): Least Absolute Deviations Estimation for the Censored Regression Model, *Journal of Econometrics*, 25, 303-325.

Przeworski, A. and J.R. Vreeland (2000): The effect of IMF programs on economic growth, *Journal of Development Economics*, 62, 385-421.

Ramcharan, R. (2003): Assessing IMF program effectiveness, *IMF Research Bulletin,* 4(2).

Rauber M. und H. Ursprung (2006): Evaluation of researchers: A life cycle analysis of German Academic Economists, *CESifo Working Paper No. 1673*.

Rauber, M. and H. Ursprung (2006): Life cycle and cohort productivity in academic economic research: Evidence for Germany, *mimeo, University of Konstanz*.

Reinhart, C. (2002): Default, Currency Crises and Sovereign Credit Ratings, *NBER Working Paper No. 8738*.

Reinhart, C.M., K.S. Rogoff, and M.S. Savastano (2003): Debt Intolerance, *Brookings Papers on Economic Activity*, 2003(1), 1-74.

Rodgers W. L. (1982): Estimable Functions of Age, Period, and Cohort Effects, *American Sociological Review*, 47, 774-787.

Roodman, D. (2004): The Anarchy of Numbers: Aid, Development, and Cross-country Empirics, *Center for Global Development Working Paper 32*.

Rose, A. (2005): One reason countries pay their debts: renegotiation and international trade, *Journal of Development Economics,* 77, 189-206.

Rotter J. (1966): Generalized Expectancies for Internal versus External Control of Reinforcement, *Psychological Monographs*, 80, 1-28.

Sen, A. K. and K.A. Hagtvet (1993): Correlations among creativity, intelligence, personality, and academic achievement, *Perceptual and Motor Skills*, 77, 497-498.

Skirbekk V. (2003): Age and Individual Productivity: A Literature Survey, *Max Planck Institute for Demographic Research, Working Paper WP 2003-028*.

SOEP (2007): Das sozio-ökonomische Panel, Stichprobenentwicklung, *http://www.diw.de/deutsch/sop/uebersicht/index.html*.

Stephan, P. (1996): The economics of science, *Journal of Economic Literature*, 34, 1199-1235.

Stiensmeyer-Pelster J. and H. Heckhausen: Kausalattribution von Verhalten und Leistung, *in Motivation und Handeln, Eds. J. Heckhausen and H. Heckhausen, Heidelberg: Springer*.

Treiman D.J. (1977): Occupational Prestige in Comparative Perspective, *New York: Academic Press*.

Urberg, K.A., Degirmencioglu, S.M. and C. Pilgrim (1997): Close Friend and Group Influence on Adolsecent Cigarette Smoking and Alcohol Use, *Developmental Psychology*, 33, 834-844.

Ursprung, H. (2003): Schneewittchen im Land der Klapperschlangen: Evaluation eines Evaluators, *Perspektiven der Wirtschaftspolitik*, 2, 177-190, 2003.

Van Dalen, H. (1999): The golden age of Nobel economists, *American Economist*, 19-35.

Weiner B., and A. Kukla (1970): An Attributional Analysis of Achievement Motivation, *Journal of Personality and Social Psychology*, 15, 1-20.

Weiner, B. (1972): Theories of motivation. *Chicago: Markham Publishing Company*.

Weiner, B., Frieze, I., Kukla, A., Reed, L., Rest, S. and R. Rosenbaum (1971): Perceiving the causes of success and failure, *Morristown, N. J.: General Learning Press*.

Windmeijer, F. (2005): A finite-sample correction for the variance of linear efficient two-step GMM estimators, *Journal of Econometrics,* 126, 25-51.

Wooldridge J. (2002): Econometric Analysis of Cross Section and Panel Data, *Cambridge, Massachusetts: The MIT Press*.

World Bank (1998): Assessing aid: what works, what doesn't, and why, *A World Bank policy research report, Oxford University Press*.

World Bank (2005): World Development Indicators on CD-ROM, *Washington DC, The World Bank*.

Zusammenfassung

Die vorliegende Dissertation besteht aus vier eigenständigen Forschungsarbeiten, die ich in der Zeit von April 2004 bis Juli 2007 an der Universität Konstanz verfasst habe. Gemeinsam ist allen Arbeiten eine empirische Ausrichtung und die Erkenntnis, dass die empirische Analyse für das Verständnis ökonomischer Zusammenhänge unverzichtbar ist. Mittels ökonometrischer Methoden werden in den einzelnen Kapiteln ökonomische Fragestellungen untersucht, die einen empirischen Zugang notwendig machen. Diese Zusammenfassung gibt einen kurzen Überblick über die einzelnen Kapitel und stellt die zentralen Ergebnisse der Arbeiten vor.

Kapitel 1 basiert auf dem Arbeitspapier *Foreign Aid and Developing Countries' Creditworthiness*. In dieser Arbeit, welche aus meiner Diplomarbeit hervorgegangen ist, untersuchen Prof. Philipp Harms (RWTH Aachen) und ich die These, dass sich Entwicklungshilfezahlungen positiv auf die Kreditwürdigkeit von Schwellen- und Entwicklungsländern auswirken. Wir formulieren ein einfaches theoretisches Modell und prüfen dessen Vorhersagen anschließend empirisch. Im Modell senken Transferzahlungen den Nutzen, der bei Nichtrückzahlung eines Darlehens für den Schuldner entsteht. In Folge dessen sinkt das Ausfallrisiko für den Gläubiger.

Als Maß für die Kreditwürdigkeit eines Landes ziehen wir in der empirischen Analyse die Beurteilung des Institutional Investor Magazine heran. Der verwendete Längsschnittdatensatz beinhaltet Informationen für eine große Zahl Entwicklungsländer von Anfang der 80er bis Ende der 90er Jahre. Unsere Ergebnisse belegen, dass sich Entwicklungshilfezahlungen positiv auf die Kapitalmarktbeurteilung auswirken. Die Stärke des Effekts variiert allerdings für unterschiedliche Formen der Entwicklungshilfe und Ländergruppen. Während wir einen positiven Effekt für Maßnahmen der technischen Unterstützung und Hilfszahlungen messen, erhöhen Darlehen die Kreditwürdigkeit nicht.

Das zweite Kapitel besteht aus zwei Arbeiten, welche sich inhaltlich mit der Forschungsproduktivität in Deutschland tätiger Ökonomen beschäftigen. Die Arbeit

Evaluation of Researchers: A Life Cycle Analysis of German Academic Economists, gemeinsam verfasst mit Prof. Heinrich Ursprung (Universität Konstanz), ist Inhalt des ersten Abschnitts. Ausgangspunkt dieser Arbeit ist die Feststellung, dass eine Evaluation von Leistung nur dann sinnvoll ist, wenn Individuen verglichen werden, die bestimmte Charakteristika teilen. Für Forscher stellt das (Karriere-) Alter das wichtigste Charakteristikum dar, da die Forschungsproduktivität einerseits von Fähigkeiten und Humankapital, und andererseits von Erfahrung und Aktualität des angeeigneten Wissens abhängt. Das Alter eines Forschers spiegelt diese Dimensionen wider und sollte daher in einer Evaluation die Bezugsgruppe definieren.

Anhand eines eigens von uns erstellten Datensatzes, der fast alle in Deutschland tätigen Ökonomen enthält, zeigen wir, dass sich Publikationsverhalten und Forschungsproduktivität in den letzten Jahrzehnten deutlich verändert haben. Kohorteneffekte und ökonomische Anreize waren wesentlich für diese Umgestaltung.

Diese Beobachtung führt zu dem notwendigen Schluss, dass eine Berücksichtigung des Merkmals Alter bei Evaluationen unabdingbar ist. Wir schlagen das Kohortenranking als eine praktikable Methode vor, die verzerrungsfreie Beurteilungen ermöglicht, und erstellen exemplarisch ein entsprechendes Ranking. Ferner entwickeln wir auf der Basis von Promotionsjahr und Forschungsleistung eine einfache Formel, die die relative Leistung eines Forschers innerhalb seiner Alterskohorte approximiert. Mittels dieser Formel können unsere Überlegungen in der Praxis auch ohne Informationen über die jeweilige Alterskohorte angewandt werden. Ein weiteres Ergebnis der Arbeit ist schließlich die Erkenntnis, dass sich die zukünftige Forschungsproduktivität bereits wenige Jahre nach der Promotion relativ gut abschätzen lässt und von hoher Persistenz gekennzeichnet ist.

Der zweite Abschnitt des Kapitels geht inhaltlich auf die ebenfalls mit Prof. Heinrich Ursprung gemeinsam verfasste Arbeit *Life Cycle and Cohort Productivity in Economic Research: The Continental European Experience as exemplified by the case of Germany* zurück. In diesem Abschnitt stehen, anders als im vorangegangenen, ökonometrische Produktivitätsanalysen im Mittelpunkt der Betrachtung. Wir untersuchen unter Zuhilfenahme verschiedener ökonometrischer und statistischer Methoden das Publikations- und Forschungsverhalten von in Deutschland tätigen Ökonomen. Dabei stellen wir deutliche Veränderungen während der letzten Jahrzehnte fest: Die Publikationstätigkeit junger Kohorten ähnelt im Vergleich zu Älteren sehr viel stärker dem Publikationsmuster von US-amerikanischen Forschern und folgt den Vorhersagen der Humankapitaltheorie. Neben

diesen ökonomischen Einflussfaktoren sind aber auch soziologische Faktoren, insbesondere Kohorteneffekte wirksam und erhöhen die Produktivität. Ein weiteres Ergebnis unserer Analyse ist, dass Publikationsqualität, Quantität und Kooperation über den Lebenszyklus hinweg variieren und unterschiedliche Entwicklungen beschreiben. Schließlich geht aus unseren Auswertungen hervor, dass Publikationsverhalten und Produktivität in erster Linie durch Heterogenität gekennzeichnet sind und daher das Bild eines „repräsentativen" Ökonomen überdacht werden sollte.

Kapitels 3 dieser Dissertation besteht aus der Forschungsarbeit *Noncognitive Skills and Success in Life: The Importance of Motivation and Self-Regulation*. Diese Arbeit untersucht den Einfluss verschiedener aus der Motivationspsychologie bekannter nichtkognitiver Fähigkeiten auf den Lebenserfolg. Insbesondere die Bedeutung von Attributionsverhalten und Willensstärke bzw. Selbstkontrolle stehen im Mittelpunkt der Untersuchung.

Unter Attribution versteht man das von Individuen offenbarte Zurechnungsverhalten beim Erfahren von Erfolg und Misserfolg: So kann Erfolg einerseits sich selbst oder exogenen Faktoren, andererseits zeitkonstanten oder variablen Faktoren zugeschrieben werden. Zahlreiche psychologische Studien belegen, dass ein Zusammenhang zwischen dem beobachtbaren Attributionsverhalten und zugrunde liegenden Motivationstyp - Fehlervermeidung oder Erfolgssuche- besteht. Selbstkontrolle und Willensstärke schließlich sind wichtig um bereits motivierte Ziele auch beim Auftreten von Widerständen verwirklichen zu können.

Als Datenbasis für die Untersuchung verwende ich eine Längsschnittstudie, die Informationen über 3240 Gymnasiasten aus dem Bundesland Nordrhein-Westfalen enthält. Die Schüler wurden in der 10. Klasse im Jahr 1970 das erste Mal befragt. Wiederbefragungen wurden 1984 und 1998 durchgeführt. Da die nichtkognitiven Fähigkeiten schon im Jugendlichenalter erhoben wurden, können Wirkungszusammenhänge und Kausalität berücksichtigt werden. Die Schüler haben in der 10. Klasse an einem Intelligenztest teilgenommen, so dass der Einfluss von kognitiven und nichtkognitiven Fähigkeiten separat betrachtet werden kann.

Lebenserfolg wird in der Untersuchung mehrdimensional durch schulische Leistung, berufliche Entwicklung und Zufriedenheit im Alter von etwa 43 Jahren definiert.

Wesentliches Ergebnis der Arbeit ist, dass nichtkognitive Fähigkeiten alle Lebensbereiche stark beeinflussen. Für den schulischen Erfolg sind das Attributionsmuster und die Fähigkeit zur Selbstkontrolle wichtig, allerdings gibt es Unterschiede zwischen

einzelnen Schulfächern. Für Lohn, Einkommen und den beruflichen Status spielen Selbstkontrolle und Attributionsverhalten ebenfalls eine wichtige Rolle: Ein hohes Maß an Selbstkontrolle und ein internales, nicht stabiles Attributionsmuster wirken sich positiv auf die berufliche Entwicklung aus. Während für Männer ein direkter Effekt auf den Lohnsatz existiert, scheint für Frauen insbesondere die Arbeitsmarktpartizipation durch nichtkognitive Fähigkeiten beeinflusst zu werden.

Aufgrund der gefundenen signifikanten Einflüsse wird deutlich, dass eine frühzeitige Förderung von kognitiven *und* nichtkognitiven Fähigkeiten im Kindesalter eine wichtige Investition der Eltern in die Zukunft ihrer Kinder darstellt.

Erklärung

Ich erkläre hiermit, dass ich die vorliegende Arbeit mit dem Titel

Four Essays on Empirical Economics

ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus anderen Quellen direkt oder indirekt übernommenen Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet. Weitere Personen, insbesondere Promotionsberater, waren an der inhaltlich materiellen Erstellung dieser Arbeit nicht beteiligt.[91] Die Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

Konstanz, den 5. August 2007 _____

(Michael Rauber)

[91] Siehe hierzu die Erklärung zur Abgrenzung auf der folgenden Seite.

Abgrenzung

Das erste Kapitel dieser Dissertation ist einer gemeinsamen Arbeit mit Herrn Prof. Philipp Harms (RWTH Aachen) entnommen. Der Aufsatz ist aus meiner Diplomarbeit an der Universität Konstanz entstanden. Die einzelnen Leistungen sind wie folgt abzugrenzen:
Das theoretische Modell ist hälftig Herrn Prof. Harms und mir zuzurechnen. Die Aufbereitung der Daten und Auswahl der geeigneten ökonometrischen Schätzmethoden geht auf mich zurück. Die Spezifikation des empirischen Models und Einbettung in die Entwicklungshilfeliteratur sind Herrn Prof. Harms zuzuschreiben. Die Schätzungen wurden von uns beiden durchgeführt. Die Verfassung von Einleitung und Schluss sind zu zwei Dritteln Herrn Prof. Harms und zu einem Drittel mir zuzurechnen.

Das zweite Kapitel entstammt zwei gemeinsamen Arbeiten mit Herrn Prof. Heinrich Ursprung (Universität Konstanz). Die individuellen Leistungen sind für beide Arbeiten wie folgt zuzuordnen: Die Idee für beide Aufsätze stammt von Herrn Prof. Ursprung. Die Durchführung und insbesondere die empirischen Analyse sind mir anzurechnen. Herr Prof. Ursprung hat im Laufe der Arbeiten an beiden Papieren meine Ergebnisse kommentiert und Anregungen für weitere Untersuchungen gegeben. Bei der Datenerfassung wurde ich von wissenschaftlichen Hilfskräften unterstützt.

Das dritte Kapitel der vorliegenden Dissertation habe ich ohne die Hilfe Dritter angefertigt.

Konstanz, den 5. August 2007 _____

(Michael Rauber)