THE BONN JOURNAL OF ECONOMICS

Volume III(1) July 2014

Department of Economics

University of Bonn

Legal Information

Publisher	The Bonn Journal of Economics Adenauerallee 24-42 53113 Bonn http://bje.uni-bonn.de bje@uni-bonn.de
V. i. S. d. P.	Prof. Dr. K. Sandmann Head of Department of Economics University of Bonn
Place of Publication	Bonn, Germany
ISSN	2195-3449

The Bonn Journal of Economics

Founded in 2012

VOLUME 3

ISSUE 1

Founded by

Johannes Hermle Justus Inhoffen Tobias Ruof

Department of Economics
University of Bonn

The Bonn Journal of Economics

Legal Reform of Corporate Governance and Shareholder Orientation in Germany

Mikko Bayer *

Introduction

Corporate governance (CG) has been a much debated topic in German business, finance and academia since at least the 1990s. Designed to manage and control the firm in the presence of separation between ownership and control, a distinct national CG system had emerged over time. This framework featured many institutional peculiarities which were largely stipulated by national commercial law. While the system was firmly in place until very recently, external developments such as financial globalization and European integration called for extensive reforms which were gradually enacted by the German legislator. Arguably, these reforms have shifted the system towards greater consideration for the interests of investors in the capital markets, putting shareholder value creation on the agenda of German businesses.

Employing a German-specific CG Indicator, a comprehensive panel fixed effects econometric model and a large panel database on over 120 exchange-listed German firms, this article tries to investigate whether we can find evidence for a

*Mikko Bayer received his degree in Economics (M.Sc.) from the University of Bonn in February 2013. The present article refers to his master thesis under the supervision of Prof. Dr. Rainer Haselmann, which was submitted in February 2013.

statistically significant and positive relationship between shareholder-friendly reforms of CG laws and firms' shareholder value generation, as well as associated market expectations. The empirical results show some evidence in support of the research hypothesis. However, this is only true if market measures are employed, reflecting market expectations rather than actual shareholder value generation.

Functions and Systems of Corporate Governance

Straddling the topics of legal science, corporate finance and organizational theory, CG is broadly defined as the system organizing the management and control of corporations on behalf of various stakeholders' interests. Any CG system is a unique combination of mutually reinforcing governance instruments in a specific regulatory framework, with the common purpose of minimizing agency costs arising from asymmetric information and associated issues of adverse selection and moral hazard due to hidden action (Witt, 2006; Berndt, 2002; Heinrich, 2002). Generally, so-called insider and outsider systems of CG are distinguished.

Outsider systems are often associated with Anglo-Saxon countries with capital market-oriented financial systems where equity is a preferred source of corporate financing, and shares are spread among dispersed shareholders. Stock corporations are led by a single board of directors, which is appointed and controlled by shareholders who can exercise voting rights with relative ease and power. Corporate takeovers occur frequently, providing an important tool for disciplining managers' behavior and protecting shareholders' interests. Accounting and disclosure standards are relatively transparent, and executives are commonly incentivized to act in line with shareholders' interests, e.g. through stock option compensation schemes. (Matthes, 2000; Heinrich, 2002; Witt, 2006). By contrast, insider systems serve the interests of various stakeholders in addition to equity shareholders, such as creditors, employees, the state or even society as a whole. Hence, shareholder value is merely one among many firm objectives. Insider systems are characterized by bank-based financial systems, and bank loans are the mainstay of corporate financing. What is more, financial intermediaries are also active players in CG. In Germany, this was evidenced by the omnipresence of bank representatives on stock corporations' supervisory boards, the body in charge of controlling the management board on behalf of a firm's owners. Share ownership was highly concentrated in a complex network of capital cross-holdings colloquially known as "Deutschland AG". Individual shareholders had little say on corporate affairs, and their powers were heavily curtailed by legal provisions such as voting rights limitations. Hostile takeovers of German firms were next to impossible due to a lack of adequate legislation. Reporting standards were strongly geared towards creditors' interests, reflected by conservative accounting methods and low managerial transparency. Furthermore, there were no legal provisions on management incentives worth mentioning. (Boehmer, 2003; Hackethal, Schmidt, and Tyrell, 2005; Heinrich, 2002; Matthes, 2000; Witt, 2006).

As legislative interventions at the national level, such as provisions on property rights, contract and civil law, enforcement devices and financial market regulation may influence CG at the firm level Berndt (2002), it is worth investigating whether the recent reforms in Germany are associated with improved shareholder value creation. However, there is no straightforward approach as empirical research typically focuses on the relationship between either law and finance, or CG and economic performance. Law and finance is typically concerned with cross-country studies and employs standardized measures of investor protection and accounting standards while research on CG and economic performance mostly looks at firm-specific CG measures and relates them to firm-specific profitability and associated market valuation.

In a seminal contribution on law and finance, López de Silanes, La Porta, Shleifer, and Vishny (1998) find that countries with common-law systems exhibit high investor protection relative to civil-law systems (typically associated with Anglo-Saxon and Continental European countries, respectively), and a negative relationship between shareholder concentration and investor protection. Relevant in terms of CG, they include an investor protection index, and an accounting standards score. In a related study, La Porta, Lopez-de Silanes, Shleifer, and Vishny (1997) suggest a positive relationship between countries' investor protection standards and the relative size of stock markets, the number of listed firms and initial public offerings. While these results are enlightening, the authors strongly focus on voting rights while there may be other, country-specific aspects relevant for investor protection. Similarly, measuring the mere quantity of included accounting items may not accurately reflect the quality of the system. Furthermore, countries' legal investor protection and accounting standards are likely to be endogenous, raising concerns about causality. It may thus be more revealing to investigate the effects of legal change, rather than legal origin.

Addressing this issue, Haselmann, Pistor, and Vig (2010) investigate the effects of legal changes of creditor rights on loan supply in twelve Eastern European transition economies during the 1990s. They employ fixed effects panel regression methodology with bank level data as explained variables, while their creditor rights index serves as the explanatory variable of interest, reflecting legal standards at the country level. The idea is to tackle the endogenous nature of legal institutions which has hampered previous research on law and finance and also captures the effects of unobserved heterogeneity. Generally, it has become customary to measure countries' financial development in terms of indices. A well-known example is the World Bank's annual Doing Business report, which measures countries' investor protection standards in a score of select provisions on managerial transparency liability, and shareholders' ease of filing lawsuits World Bank (2013). In a similar vein, the Heritage Foundation's Index of Economic Freedom measures the "rule of law" in countries in terms of property rights protection, provisions against corruption, and "regulatory efficiency" (Miller, Holmes, Feulner, Kim, and Riley, 2013).

There is also a growing literature on the relationship between CG and economic performance, but little work exists on the effects of relevant legal standards on firm performance. Employing a panel fixed effects model to control for country and industry-specific, time-invariant heterogeneities, Rajan and Zingales (1998) find that industries with high external financing needs grow relatively faster in countries with higher financial development and related CG standards, notably accounting standards. Using panel data on firms from 38 countries, Himmelberg, Hubbard, and Love (2002) find that firms' cost of capital increases with share ownership concentration, which in turn decreases in countries' investor protection standards, and further suggest that a strengthening of investor protection laws and their enforcement improves capital allocation and associated growth. While reasonably addressing legal change, they isolate ownership concentration as one particular agency conflict affecting firms, thus neglecting interdependencies with other relevant aspects of CG. High share ownership concentration is another common criticism about the country's traditional system, along with excessive bank control and associated transparency issues. Their effects on firm performance are analyzed by various authors who come to contradictory conclusions, possibly due to a lack of industry-specific variables and a lacking focus on the change of control over time (see Boehmer (2003)).

Using panel data on 100 exchange listed German firms covering the years 1996-99, Tuschke, Sanders, et al. (2003) find that corporate divestitures (i.e. the reduction of ownership concentration) are positively associated with subsequent firm performance, as are the adoption of stock incentive plans and international accounting standards. However, this holds only for market measures of firm performance like market capitalization, but not for book measures such as return on sales. While reasonably addressing specific German CG issues and managerial incentives, the focus on voluntary firm-level CG reforms raises new concerns of endogeneity as it is not a priori clear whether such reforms actually cause higher shareholder returns and expectations, or whether increased market orientation leads to the gradual adoption of adequate CG standards. The focus of this article, however, is on the effects of changes in CG legislation at the national level on shareholder value at the firm level. These reforms and how they are used to construct a legal indicator will be discussed in the following section.

Legal Reforms and The Corporate Governance Indicator

The traditional German CG system was challenged by several economic and political developments which surfaced simultaneously during the 1990s. Notably, European political and economic integration gained momentum during that time and numerous European Union legal initiatives called for greater capital market integration, with particular focus on CG policies starting from the 2000s (Beckmann, 2007; Heinrich, 2002). These initiatives resulted in a series of policy guidelines and directives which had to be translated into law by the national governments. In this course, many provisions of the German commercial laws were found at odds with these guidelines and were consequently altered or abolished. At the same time, and more importantly perhaps, remarkable changes took place both in the German firms and financial institutions which were so closely linked in the old insider system. Those banks and insurances (notably Deutsche Bank and Allianz) which had been at the center of the old "Deutschland AG" shifted the focus of their business away from industrial holdings and loan supply towards more profitable activities such as investment banking, thereby retreating from their central role in Germany's CG system (Beyer, 2003; Hackethal, Schmidt, and Tyrell, 2005). Consequently, large German corporations increasingly turned towards (international) equity investors in order to meet their financing needs. Not only did a growing number of German firms seek listing in overseas stock exchanges, but international investors also showed increased interest in the German equity market (O'Sullivan, 2003). Among these were so-called institutional investors such as large pension funds, insurances, and other investment firms, often from Anglo-Saxon countries. These investors allocate substantial funds to equity titles, creating pressure on firms to pursue shareholder value generation and to adopt capital market friendly measures of CG (Beckmann, 2007). This was also thought to improve the market valuations of German corporations' stocks which had been low compared to their Anglo-Saxon counterparts, in part due to their relatively lower profitability which was in turn attributed to Germany's seemingly outdated insider system of CG (Erlen, 2002; Streeck and Höpner, 2003).

The viability of the system was further challenged by a severe economic downturn in the early 1990s (O'Sullivan, 2003) and a series of prominent corporate crises and mismanagement scandals throughout the decade. The demise of large corporations such as the steel producer Metallgesellschaft, construction firm Philipp Holzmann, or electrical equipment producer AEG sparked a public debate about the control efficiency of German supervisory boards, the role of banks, deficits of internal risk management and conflicts of interests due to extensive personal interrelationships within the insider system (Albers, 2002; Beckmann, 2007; Ipsen and Pfitzinger, 2003). In sum, the above developments called for extensive reforms of the CG system, aiming at greater capital market orientation and shareholder value generation, which was subsequently addressed both at the firm and at the legislator levels.

To reflect the quality of the most relevant legal standards and their change over time, I propose a corporate governance indicator (CGI). Following mainstream approaches, the CGI is also structured along the dimensions of investor protection and transparency. Investor protection reflects legal provisions protecting shareholders' rights (e.g. in the annual meeting), their claims (e.g. of small against large blockholders), and the enforcement of standards (e.g. through supervisory authorities). Transparency, on the other hand, tries to capture the most relevant disclosure requirements on managerial activities and financial reporting. In addition, the CGI also reflects legal incentives supposed to align managerial actions with investors' best interests. German examples include the introduction of a law regulating executive compensation with stock options and a tax reform exempting firms from capital gains taxes in the case of divestitures of corporate holdings (which were often unprofitable but maintained for tax reasons). The CGI focuses on twelve legal reform acts passed between 1994 and 2009, encompassing 19 legal reform measures of interest which explicitly address those issues of investor protection, transparency, and incentives that were so frequently criticized about the old system.

Landmark legal reform acts include the Second Financial Market Development Act of 1994, the 1998 Corporate Transparency and Control Act, the Transparency and Disclosure Act of 2002, the Accounting Law Reform Act of 2004, and the 2005 Corporate Integrity and the Modernization of Shareholder Action Act, to name just a few. These reforms altered many existing laws, particularly the Stock Corporations Law and the Commercial Code, but also introduced new laws such as the Securities Trading Law. All legal reform measures and how they are included in the CGI (either as investor protection (IP), transparency (T), or incentive (INC) items), are summarized in table 1, along with the most important laws they altered, the year they came into effect, a description of the factual change they brought about and how this addresses specific German CG issues. A dummy representation of the CGI can be found in table 2.

Panel Database and Empirical Strategy

I construct an panel database on over 120 German corporations listed in the various segments of the Frankfurt stock exchange, covering a timespan of twenty years (1992-2011, where available). Any financial intermediaries such as banks and insurances are excluded as they are subject to a much more particular regulatory framework that goes beyond general CG issues. Firms headquartered overseas and incorporated in foreign legal forms are also excluded.

To measure shareholder orientation and profitability, both market and accounting variables were obtained. The market variable employed is average historical market capitalization, calculated as the product of the share price and the number of shares outstanding, on an annualized basis. This reflects a firm's valuation by investors towards whose interests the legal reforms were geared in the first place. As an accounting variable I use basic earnings per share (EPS), computed as net income available to common shareholders divided by the weighted average shares outstanding and including the effects of all one-time, non-recurring gains and losses (e.g. resulting from corporate divestitures). It should be noted that market capitalization is a forward-looking measure reflecting investors' expectations on firm profitability, while EPS reflects actual earnings from the past period, as reported in annual statements. I further include several macro and firm level control variables to account for alternative explanations. Macro control variables include the annual averages of the IFO institute's monthly business situation index for models of EPS, while the IFO business climate index is used for models of market capitalization. These indices serve as macro analogues for GDP growth and growth expectations, respectively. To reflect inflation, I include the consumer price index on Germany from the World Bank's World Development Indicators database. Firm-specific control variables include firms' effective interest and tax rates, as well as firm size in terms of the (logged) number of employees. All firm specific variables were downloaded from the Bloomberg database. Only fourteen firm-year observations were excluded as outliers. In total, there are 1,606 firm-year observations of market capitalization, and 1,989 observations of EPS, covering 129 firms.

To investigate the research question, a panel data regression model with fixed effects is applied. Using firm-level data, the following specification is tested:

$$y_{it} = a + \beta X_{it} + \gamma CGI_{t-1} + \delta_i + \epsilon_{it} \tag{1}$$

where i denotes firms and t denotes years. The dependent variable, y, alternatively stands for EPS, or the log of market capitalization. Macro and firm-specific control variables are included in the vector X. CGI is the legal indicator , and γ is the associated coefficient of interest, capturing the response of the dependent variables to the legal change. δ_i denotes firm fixed effects, that is, unobserved and time-invariant firm-specific characteristics affecting profitability. Finally, ϵ_{it} denotes the error term.

One major reason for using panel data is to resolve the issue of omitted variables (Wooldridge, 2002). In my model, this unobserved variable is represented by the term δ_i, and is assumed to differ between firms, but to be constant over time. In terms of my study, one could imagine that some firms are systematically more profitable than others due to the superior productivity of their workforce, or due to firm-specific voluntary CG measures that go beyond legal requirements. However, as the parameter of interest is γ, the undesired residual needs to be taken care of. This is done by the so-called fixed effects or "within" transformation (Wooldridge, 2002). Assuming that the explanatory variables are strictly exogenous conditional on the unobserved effect, i.e. $(u_it \mid X_i, CGI, \delta_i) = 0, t = 1, 2, ..., T$, equation (1) is first averaged to obtain the cross section equation

$$\overline{y}_i = a + \beta\overline{X}_i + \gamma\overline{CGI} + \delta_i + \overline{\epsilon}_i \tag{2}$$

where $\overline{y}_i = T^{-1}\sum_{t=1}^{T} y_{it}$, $\overline{X}_i = T^{-1}\sum_{t=1}^{T} X_{it}$, $\overline{CGI} = T^{-1}\sum_{t=1}^{T} CGI$ and $\overline{\epsilon}_i = T^{-1}\sum_{t=1}^{T} \epsilon_{it}$.

Subtracting equations (1) and (2) yields the fixed effects transformed equation

$$(y_{it} - \overline{y}_i) = \beta(X_{it} - \overline{X}_i) + \gamma(CGI_{t-1} - \overline{CGI}) + (\epsilon_{it} - \overline{\epsilon}_i) \tag{3}$$

where the fixed effect term has been wiped out, and the coefficients of interest can be consistently estimated. Due to the large size of the sample, I use heteroskedasticity-robust standard errors in all specifications to account for possibly different variance structures across firms (Wooldridge, 2002) and serial correlation in the idiosyncratic error terms. Specifying robust standard errors is equivalent to clustering standard errors on the firm level and produces consistent estimators even when disturbances are not identically distributed over firms and in the presence of autocorrelation, as long as there are many firm observations, and disturbances are uncorrelated across firms (StataCorp, 2009).

Alternatively, I test the hypothesis by running the dynamic panel regression specification

$$y_{it} = a + \alpha y_{i,t-1} + \beta X_{it} + \gamma CGI_{t-1} + \delta_i + \epsilon_{it} \tag{4}$$

where one lag of the dependent variable is included as an explanatory variable. The idea is that previous period EPS or market capitalization may be highly explanatory for the observed current period value. Model coefficients are estimated from the first differenced equation

$$\Delta y_{it} = \alpha \Delta y_{i,t-1} + \beta \Delta X_{it} + \gamma \Delta CGI_{t-1} + \epsilon_{it} \tag{5}$$

where the undesired panel effect has again been eliminated. More obviously than in specification (3), the coefficient of interest, γ, captures the effects of legal change. However, this specification is prone to dynamic panel bias arising from the systematic correlation between the lagged dependent variable and the idiosyncratic error term. Therefore, an estimator is employed that constructs GMM type instrumental variables from the lagged levels of the dependent variable to produce consistent estimators, and works in the presence of low-order serial correlation in the error terms (StataCorp, 2009). Post-estimation tests for serial correlation were conducted to ensure the validity of the used moment conditions and to determine the appropriate lags for constructing the GMM-type instruments.

Regardless of the specification, endogeneity is an important concern, and is thus addressed in various ways. First and foremost, the dependent variables are at the individual firm level, while most explanatory variables are at the macro level. Notably, the variable of interest (CGI) reflects changes in codified national legal standards. Firms are involuntarily subject to these standards unless they relocate to other jurisdictions, an eventuality I have controlled for by excluding such firms from the sample. Furthermore, the legal reforms were largely driven by external trends such as financial globalization and European integration. The specification further exhibits a lagged design on the variable of interest, to alleviate concerns of potential reverse causality. Finally, in specifications featuring the lagged dependent variable as an explanatory factor, I employ an instrumental variable technique to address the dynamic panel bias.

Results

The results for the models of market capitalization and EPS are presented in tables 3 and 4, respectively. In both tables, column 1 shows that the coefficient on the CGI is positive and highly significant for both models. Next, I include macro and firm control variables as described in the previous section. Column 2 shows that the legal variable is still positive and highly significant for market capitalization, while the coefficient loses its significance for the model of EPS. In both models, the coefficient on the business climate/situation is positive and significant as expected, while the coefficients on the other control variables are mostly insignificant or very small. Column 3 shows the results from the dynamic specification, where one lag of the dependent variable is included, and the coefficient estimates are obtained from equation (5). The coefficient on the legal variable is positive and significant at the 95 percent confidence level for EPS, and at the 99 percent level for market capitalization. What is more, the signs of the coefficients on virtually all control variables are highly consistent with expectations. The overall model fit is significantly better in models of market capitalization.

Next, the regressions from columns 1 to 3 are repeated, but the legal variable is now disaggregated into components of investor protection, transparency, and incentives. The results of the respective horserace regressions are found in columns 4 to 6 of tables 3 and 4. In models of market capitalization, the coefficient on investor protection is positive and highly significant in specifications with and without control variables (columns 4 and 5), and also in the dynamic specification (column 6). However, the signs of the transparency and incentive coefficients are negative unless the dynamic specification is employed (although only the coefficient on incentives in the specification without control variables is significant). Furthermore, previous period market capitalization appears to be highly explanatory for the current period's value. Overall, the dynamic specification in column 6 exhibits high significance for nearly all coefficients of interest, and the signs of all parameter estimates are highly consistent with expectations.

Table 4 shows that the regression results of the disaggregated CGI on EPS are less consistent with expectations. Only the specification without control variables (column 4) yields statistically significant parameter estimates, however the sign for incentive is again negative and thus inconsistent with expectations. This is not the case in the specifications with control variables and in the dynamic model; however, all parameter estimates of interest are statistically insignificant. Overall, it seems difficult to find evidence for a positive effect of CG reforms on firms' accounting profitability. This also resembles the findings of Tuschke, Sanders, et al. (2003).

Conclusions and Outlook

Generally, the empirical results are most consistent with expectations when market capitalization is employed. This suggests that investors in the capital market reward shareholder-oriented reforms in CG legislation by allocating funds to German equity titles, thus pushing up market prices. This is true for all three specified dimensions of shareholder orientation in CG, namely investor protection, transparency, and incentives. However, when EPS are employed, the results also suggest a positive relationship between legal standards and firm profitability, but significance of the results is highly subject to the model specification. A possible explanation for these results may be the forward and backward looking natures of the respective market and accounting measures of shareholder value generation. This may suggest that, on average, investors are overly optimistic about the economic effects of governance reform.

In any case, it is worth some thought whether shareholder orientation should actually be the central paradigm of "good" CG that it seems to be today or whether the focus should be on other, more long-term objectives. While the old German insider CG system is unlikely to be revived, its comprehensive stakeholder approach certainly has its merits which may not be adequately reflected by mainstream empirical research. It thus deserves thorough review when designing the corporate governance systems of the future.

References

ALBERS, M. (2002): "Corporate governance in Aktiengesellschaften: Entscheidungsprozess und Wirkungsanalyse zum Gesetz zur Kontrolle und Transparenz im Unternehmensbereich (KonTraG)," Europäische Hochschulschriften / 5: Volks- und Betriebswirtschaft, pp. 20–21, 77–79. Lang.

BECKMANN, M. (2007): "Das Finanzkapital in der Transformation der europäischen Ökonomie," pp. 100–114, 134–142. Westfälisches Dampfboot.

BERNDT, M. (2002): "Global Differences in Corporate Governance Systems: Theory and Implications for Reforms," Ökonomische Analyse des Rechts Series. Deutscher Universität.

BEYER, J. (2003): "Alle Macht dem Markt?: Fallstudien zur Abwicklung der Deutschland AG," Schriften des Max-Planck-Instituts für Gesellschaftsforschung, Köln, pp. 118–146. Campus.

BOEHMER, E. (2003): "Germany," *Corporate Governance and Economic Performance*, pp. 118–146.

ERLEN, H. (2002): "Corporate governance: deutsche Unternehmenskultur im Wandel ; Gespräche mit führenden Köpfen aus Wirtschaft und Wissenschaft," pp. 45–49. Ernst & Young.

HACKETHAL, A., R. H. SCHMIDT, AND M. TYRELL (2005): “Banks and German Corporate Governance: On the Way to a Capital Market-Based System?,” *Working Paper Series Finance & Accounting*, No. 146.

HASELMANN, R., K. PISTOR, AND V. VIG (2010): “How Law Affects Lending,” *The Review of Financial Studies*, vol. 23, no. 2, 549–580.

HEINRICH, R. (2002): “Complementarities in Corporate Governance,” Kieler Studien - Kiel Studies, pp. 9–12, 184–193. Springer.

HIMMELBERG, C., R. HUBBARD, AND I. LOVE (2002): “Investor protection, ownership and the cost of capital,” Policy research working papers. World Bank, Development Research Group, Public ServiceFinance.

IPSEN, D., AND J. PFITZINGER (2003): “Alle Macht dem Markt?: Fallstudien zur Abwicklung der Deutschland AG,” Schriften des Max-Planck-Instituts für Gesellschaftsforschung, Köln, pp. 60–92. Campus.

LA PORTA, R., F. LOPEZ-DE SILANES, A. SHLEIFER, AND R. W. VISHNY (1997): “Legal determinants of external finance,” *The Journal of Finance*, vol. 52, no. 3, 1131–1150.

LÓPEZ DE SILANES, F., R. LA PORTA, A. SHLEIFER, AND R. VISHNY (1998): “Law and finance,” *Journal of political economy*, 106, 1113–1155.

MATTHES, J. (2000): “Das deutsche Corporate-governance-System: Wandel von der Stakeholder-Orientierung zum Shareholder-Value-Denken,” Beiträge zur Wirtschafts- und Sozialpolitik. Division, Dt. Inst.-Verlag.

MILLER, T., K. HOLMES, E. FEULNER, A. KIM, AND B. RILEY (2013): “Index of Economic Freedom 2013,” INDEX OF ECONOMIC FREEDOM. Heritage Foundation.

O’SULLIVAN, M. (2003): “The political economy of comparative corporate governance,” *Review of International Political Economy*, 10(1), 23–72.

RAJAN, R. G., AND L. ZINGALES (1998): “Financial Dependence and Growth,” *Review of International Political Economy*, 88(3), 559–586.

StataCorp (2009): “Stata Longitudinal-Data/Panel-Data Reference Manual. Release 11,” .

STREECK, W., AND M. HÖPNER (2003): “Alle Macht dem Markt?: Fallstudien zur Abwicklung der Deutschland AG,” Schriften des Max-Planck-Instituts für Gesellschaftsforschung, Köln, pp. 25–28. Campus.

TUSCHKE, A., G. SANDERS, ET AL. (2003): “Antecedents and consequences of corporate governance reform: The case of Germany,” *Strategic Management Journal*, 24(7), 631–649.

Witt, P. (2006): "Corporate Governance-Systeme im Wettbewerb," pp. 1, 62–72, 78–90. Dt. Univ.-Verlag.

Wooldridge, J. (2002): "Econometric Analysis of Cross Section and Panel Data," MIT Press.

World Bank (2013): "Doing Business 2013: Smarter Regulations for Small and Medium-Size Enterprises," Doing Business. World Bank Publications.

Appendix

2. FFG: Zweites Finanzmarktförderungsgesetz (Second Financial Market Development Act)

4. FFG: Viertes Finanzmarktförderungsgesetz (Fourth Financial Market Development Act)

AktG: Aktiengesetz

ARUG: Gesetz zur Umsetzung der Aktionärsrichtlinie

BilReG: Bilanzrechtsreformgesetz

HGB: Handelsgesetzbuch

KapAEG: Kapitalaufnahmeerleichterungsgesetz

KonTraG: Gesetz zur Kontrolle und Transparenz im Unternehmensbereich

KStG: Körperschaftssteuergesetz

StSenkG: Steuersenkungsgesetz

UMAG: Gesetz zur Unternehmensintegrität und Modernisierung des Anfechtungsrechts

VorstAG: Gesetz zur Angemessenheit der Vorstandsvergütung

VorstOG: Vorstandsvergütungs-Offenlegungsgesetz

WpHG: Wertpapierhandelsgesetz

WpÜG: Wertpapererwerbs- und Übernahmegesetz

	Reform Act	Laws Changed/Introduced	Effective	Description of Measure and Relevance for Corporate Governance
Anti-Insider Legislation (IP)	2.FFG	§§ 12-20, esp. § 14 WpHG	1994 (mostly)	Laws are in place prohibiting the use of insider information. This addresses the potential abuse of informational asymmetries against outsiders.
Enforcement/ Supervision Authority (IP)	2.FFG*	§§ 3-11, 16, 21, 29 (et al.) WpHG	1995	A federal authority is in place supervising and enforcing the compliance with anti-insider and other corporate governance laws (e.g. the ad-hoc disclosure of facts affecting the stock price).
Voting Rights Disclosure (T)	2.FFG*	§§ 21-30 WpHG	1995	Public disclosure is required if a block shareholder attains voting rights above specified threshold percentages. Addresses the lack of transparency of capital holdings.
International Reporting Standards (INC)	KapAEG	§ 292a HGB	1998-2004	Firms may file consolidated annual statements according to accepted international instead of German accounting standards, as an incentive to shift to capital market-friendly reporting.
"Anti-Bank" Voting Rights (IP)	KonTraG	§§ 12, 128, 134, 135 AktG	1998	Bank' proxy voting power is curtailed and multiple/maximum voting rights are banned. Addresses the protection of small shareholders' interests in the general meeting.
Supervisory Board (IP)	KonTraG	§§ 90, 110, 111, 147, 171, 328 AktG	1998	The duties of the supervisory board are clarified as an institution of delegated shareholder control. Addresses the control deficits and conflicts of interest in the old insider system.
Risk/Audit Standards (IP)	KonTraG	§ 91 AktG, §§ 189, 317, 321, 322, 323 HGB	1998 (mostly)	Comprehensive audit requirements and provisions on risk management measures are in place, aiming at shareholder protection.
Stock Option Compensation (INC)	KonTraG	§§ 192, 193 AktG	1998	The supervisory board may decide on stock option compensation schemes as an incentive to align managerial and market interests.
Tax Exemption on Divestitures	StSenkG	§ 8b KStG	2001	Profits associated with firms' divestitures of corporate holdings are exempt from capital gains tax as an incentive to reduce ownership concentration.
Market for Corporate Control (IP)	WpÜG	WpÜG	2002	A legal framework is in place regulating the takeover of capital shares and control over corporations via the capital market, including the potential replacement of management.
Directors' Dealings Disclosure (T)	4.FFG*	§ 15a AktG	2002	Management/supervisory board members' sales and purchases of the own company's stock must be disclosed. Increases managerial transparency.
CG Code/ Best Practice (INC)	TransPuG	§ 161 AktG	2002	A widely accepted code of corporate governance best practices is in place. Compliance is voluntary bust must be disclosed, creating an incentive to act in line with market interests.
International Reporting Standards (T)	BilReG	§ 315a HGB	2005	International reporting standards are compulsory for all capital-market oriented stock corporations. Introduces a market-friendly disclosure standard.
Shareholder Lawsuits (IP)	UMAG	§§ 93, 117, 147-149, 243, 246a (et al.) AktG	2005	Shareholders may easily sue for damages against management and supervisory boards and appeal against general meeting decisions. Improves shareholder protection.
No Share Deposit (IP)	UMAG	§ 123 AktG	2005	Shareholders are not required to deposit their shares prior to and during the general meeting. Aims at increasing coring rights exercise.
Executive Compensation Disclosure (T)	VorstOG	§§ 285, 286, 289, 314, 315 HGB	2006	The disclosure of individual board members' compensation packages is compulsory, increasing managerial transparency.
Executive Cooling-off Period (IP)	VorstAG	§ 100 AktG	2009	Management board members may not move directly to the supervisory board after the expiry of their terms. Reduces insider bias.
Executive Compensation Adequacy (INC)	VorstAG	§§ 87, 107, 116, 193 AktG	2009	Supervisory board has to set management compensation in line with performance and sustainable growth. Serves to align managerial incentives with shareholders' (long-term) interests.
Proxy Voting by Mail/ Electronic (IP)	ARUG	§§ 118, 121 AktG	2009	Shareholders may personally exercise voting rights without physical presence in the annual meeting, improving conditions for voting rights exercise.

Table 1: Timeline of Legal Reform Measures of Corporate Governance

	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Anti-insider Legislation	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Enforcement Supervision Authority	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Anti-Bank Voting Rights	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Supervisory Board	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Risk Audit Standards	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Market for Corporate Control	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1
Shareholder Lawsuits	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1
No Share Deposit	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1
Proxy Voting by Mail Electronic	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1
Executive Cooling-off Period	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1
INVESTOR PROTECTION	0	0	1	2	2	2	5	5	5	5	6	6	6	8	8	8	8	10	10	10
Voting Rights Disclosure	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Directors' Dealings Disclosure	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1
International Reporting Standards	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1
Executive Compensation Disclosure	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1
TRANSPERENCY	0	0	0	1	1	1	1	1	1	1	2	2	2	3	4	4	4	4	4	4
International Reporting Standards	0	0	0	0	0	0	1	1	1	1	1	1	1	0	0	0	0	0	0	0
Stock Option Compensation	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Tax Exemption on Divestitures	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1
CG Code Best Practice	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1
Executive Compensation Adequacy	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1
INCENTIVE	0	0	0	0	0	0	2	2	2	3	4	4	4	3	3	3	3	4	4	4
CORP. GOVERNANCE INDICATOR (TOTAL)	0	0	1	3	3	3	8	8	8	9	12	12	12	14	15	15	15	18	18	18

Table 2: Corporate Governance Indicator (Dummy Representation)

	(1)	(2)	(3)	(4)	(5)	(6)
Corp. Governance Indicator (t-1)	**0.098**	**0.060**	**0.062**			
	(0.008)*	**(0.015)***	**(0.008)***			
IFO Business Climate		**0.026**	**0.022**		**0.020**	**0.023**
		(0.002)*	**(0.001)***		**(0.002)***	**(0.002)***
Inflation		0.001	**−0.029**		−0.007	**−0.028**
		(0.009)	**(0.006)***		(0.006)	**(0.006)***
Effective Interest Rate		0.001	−0.000		0.001	−0.000
		(0.001)	(0.000)		(0.001)	(0.000)
Effective Tax Rate		**−0.001**	**−0.001**		**−0.001**	**−0.001**
		(0.000)*	**(0.000)***		**(0.000)***	**(0.000)***
Employees (log)		**0.450**	0.084		**0.443**	0.085
		(0.086)*	(0.081)		**(0.085)***	(0.082)
Market Cap. (log, t-1)			**0.592**			**0.579**
			(0.040)*			**(0.039)***
Investor Protection (t-1)				**0.259**	**0.163**	**0.054**
				(0.024)*	**(0.031)***	**(0.016)***
Transparency (t-1)				−0.056	−0.011	**0.069**
				(0.047)	(0.044)	**(0.024)***
Incentive (t-1)				**−0.100**	−0.028	**0.071**
				(0.027)*	(0.026)	**(0.016)***
Constant	**5.714**	−0.541	**2.094**	**5.615**	0.663	**2.034**
	(0.092)*	(1.171)	**(0.733)***	**(0.087)***	(0.863)	**(0.775)***
Adjusted R^2	0.3732	0.4725	-	0.4091	0.4807	-
Observations	1,606	1,530	1,415	1,606	1,530	1,415

Robust standard errors in parentheses. Significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$
Columns 1 and 2 report the regression results from specification (1), with and without firm and macro-level control variables. Column 3 reports the regression results from specification (4). Columns 4 and 5 report the regression results on the disaggregated legal variable from specification (1), with and without control variables. Column 6 reports the regression results on the disaggregated legal variable from specification (4).

Table 3: Firm-level regressions (Market Capitalization (log))

	(1)	(2)	(3)	(4)	(5)	(6)
Corp. Governance	**0.140**	0.043	**0.106**			
Indicator (t-1)	**(0.019)*****	(0.028)	**(0.046)****			
IFO Business Situation		**0.042**	**0.039**		**0.043**	**0.042**
		(0.007)***	**(0.008)*****		**(0.009)*****	**(0.009)*****
Inflation		0.039	-0.015		0.033	-0.014
		(0.025)	(0.031)		(0.039)	(0.038)
Effective Interest Rate		0.001	**0.002**		0.001	**0.002**
		(0.001)	**(0.001)***		(0.001)	**(0.001)***
Effective Tax Rate		**-0.007**	**-0.005**		**-0.007**	**-0.005**
		(0.001)***	**(0.001)*****		**(0.001)*****	**(0.001)*****
Employees (log)		0.174	-0.244		0.175	-0.192
		(0.324)	(0.261)		(0.326)	(0.283)
Earnings Per Share (t-1)			**0.305**			**0.302**
			(0.077)***			**(0.080)*****
Investor Protection (t-1)				**0.237**	0.014	0.045
				(0.073)***	(0.068)	(0.66)
Investor Protection (t-1)				**0.210**	0.110	0.167
				(0.101)**	(0.176)	(0.183)
Investor Protection (t-1)				**-0.136**	0.072	0.157
				(0.078)*	(0.092)	(0.112)
Constant	-0.136	**-8.431**	-0.210	-0.170	**-7.963**	-0.933
	(0.193)	**(2.975)*****	(3.134)	(0.199)	**(4.096)***	(3.051)
Adjusted R^2	0.1229	0.1620		0.1330	0.1612	
Observations	1989	1827	1762	1989	1827	1762

Robust standard errors in parantheses. Significance levels: * $p<0.1$; ** $p < 0.05$; *** $p < 0.01$
Columns 1 and 2 report the regression results from specification (1), with and without firm and macro-level control variables. Column 3 reports the regression results from specification (4). Columns 4 and 5 report the regression results on the disaggregated legal variable from specification (1), with and without control variables. Column 6 reports the regression results on the disaggregated value from specification (4).

Table 4: Firm-level regressions (Earnings per Share)

On the nature of private information in corporate leniency

Franziska Heinicke *

Introduction

Since the US Department of Justice introduced new leniency policies in 1993 (the Corporate Leniency Policies) to improve the effectiveness of successfully detecting and prosecuting cartels, leniency programs have become a crucial part of cartel prosecution worldwide. While differing within certain parameters and guidelines, leniency programs all offer a reduced fine to one or more firms that are willing to provide evidence on an illegal collusion they are involved in. This reduced fine serves as an incentive for colluding firms to report to the competition agency (CA).

The implementation of leniency programs was accompanied by a broad range of literature on the behavior of firms under the new conditions. The first paper directly focusing on corporate leniency and proposing a first model to identify incentives and motives of involved parties was Motta and Polo (2003). This first model only allowed for leniency application as an collusive decisions of cartels. The following literature then introduced models where spontaneous reporting can occur. These results are collected in Spagnolo (2008). In most of these models,

*Franziska Heinicke received her degree in Economics (B.Sc.) from the University of Bonn in July 2013. The present article refers to her bachelor thesis under the supervision of Prof. Dr. Dennis Gärtner, which was submitted in June 2013.

firms have complete information on the leniency program, their own profits and all related probabilities. As a result, they will either not form a cartel at all because cartel profits are not high enough anymore, or if they form will never use leniency programs because they have no incentive to risk the high cartel profit.
Harrington (2008) aims to solve this contradiction by allowing the probability of getting caught and convicted to change over time. Cartels might then establish at one point but apply for leniency later. This however results in the so called 'Rush to the Courthouse-Effect' where either no firm applies or all firms try to outpace each other in applying.
The theoretical work on corporate leniency is accompanied by experiments that aim at getting an undisturbed view on firm behavior. One experiment conducted by Bigoni et al. (2010) suggests that players who are presented with a constant product of a probability of conviction and a fine, tend to apply more often when facing a high fine. This means that, independent from the likelihood of a successful prosecution, players fear the high fine. In Harrington (2012) this behavior is interpreted as firms being afraid of a possible leniency application of other cartel members and modeled by allowing firms to have private information on the likelihood of getting caught and convicted by the CA. Already in Pinna (2010), Sauvagnat (2010) and Silbye (2010) private information is included in a model for either the CA or the colluding firms. Harrington's approach differs in that it concludes on a Bayes-Nash equilibrium where one firm might turn its partner in without this partner 'rushing' behind to get a chance to apply as well.
While this thesis aims at a similar equilibrium that allows for asymmetric behavior of firms, it is based on a different interpretation of private information in cases of collusion. Firms will have private information on the amount and quality of evidence, which they collected during the period of collusion. The model of Harrington (2012) will then be adjusted to the new assumptions.

Private information in corporate leniency

Harrington (2012) supposes that firms receive a private signal that gives them information on the likelihood of the CA convicting the cartel without any firm applying for leniency. Given their own signal, they then form expectations on the signal of the other firm. This definition implies that firms' signals are based on information on CA behavior. Only having different information on the CA will give firms different signals on the likelihood of conviction.

The CA generates information on all its activities and their probable impact on the success of prosecution. In many countries the CA not only waits for cartel members to come forward, but also actively investigates certain sectors to detect possible cartels and, if there is sufficient evidence, also starts an investigation. [1] Therefore the CA does indeed create information that might affect firms' leniency decision.

However CAs are public organizations and, following the nature of public organizations, CAs underlie certain guidelines of transparency and publicity. [2] Every individual person can inform himself/herself about CA activities and firms that are involved in cartels naturally have an even higher interest in getting to know about every detail of the CA's work. Therefore the following part of this thesis will argue that there is no private or secret information generated by the CA but all its activities are openly known by all cartel members.

In addition to the information on the CA, firms themselves will also create information in the form of evidence over the period of collusion. This evidence includes all the material that in case of court proceedings could be used against them. All colluding firms know how much evidence could have been collected at

[1] Germany as an example conducts sector inquiries and collects information by other market participants that can lead to further investigation, in Bundeskartellamt (2010). *The Bundeskartellamt in Bonn – Organisation, Tasks and Activities.* p. 22 and 42

[2] Again referring to the example of Germany, the CA is only able to search the premises of the involved firms after instituting fine proceedings which is a public procedure, in Bundeskartellamt (2012). *Effective cartel prosecution - Benefits for the economy and consumers.* p. 20

a maximum but it is privately known by each firm how much evidence it actually created and kept. The idea of giving firms private information on their own evidence was already introduced by Silbye (2010) who proposed a model where the CA offers a leniency rate based on this evidence to encourage the high-evidence firm to apply first. While the leniency rate itself is kept fixed in the following model, certain other features of Silbye's model will be included in the model set up by Harrington (2012).

Modeling private evidence

General set-up

The following model describes a situation where a cartel between two firms has already ended and firms have to decide if they apply for a leniency program or not. If a firm applies it will have to pay the leniency share θ of the total fine F. If no firm applies for leniency the CA might still be able to uncover the cartel and make both firms pay a fine F. At this point firms learn about their collected evidence λ_i, $\lambda_i \in [\underline{\lambda}, \bar{\lambda}] \subseteq [0, 1]$.

In line with the work of Silbye (2010) this model will assume that the evidence one firm provides when applying will determine the likelihood of the other firm getting convicted, meaning that if firm i decides to make use of the leniency program firm j will be convicted with probability λ_i.

Having stated this, the range of $[\underline{\lambda}, \bar{\lambda}]$ can be described more specifically. The upper end of this range, namely $\bar{\lambda}$, will be very close or equal to 1 because the highest amount of evidence possibly collected will lead to a very high probability of conviction if revealed to the CA. The lower end $\underline{\lambda}$ on the other hand is very unlikely to be 0 as for colluding firms it is not possible to communicate without leaving any kind of evidence.

The evidence described in the variable λ will be positively correlated among

firms. The more evidence they find for themselves, the more evidence the other firm could have collected. Taking the position of firm i firm j's evidence will follow a cumulative distribution function of $G(\lambda_j|\lambda_i)$ for which the following holds:
AI *$G(\lambda_j|\lambda_i)$ $(j \neq i)$ is continuously differentiable in λ_i and λ_j. If $\lambda'' > \lambda'$ then $G(\bullet|\lambda_i = \lambda'')$ strictly first-order stochastically dominates (FOSD) $G(\bullet|\lambda_i = \lambda')$.*
The probability of conviction, λ_0, derives from a degree of evidence that is available to the CA without any firm applying. Colluding firms will not be able to control and hide all evidence. Obvious changes in prices or the arrangement of meetings among partners in neutral places are examples for this openly available evidence. In addition, the CA can also gain information from other market participants who might be able to observe collusive patterns earlier than the CA. Firms refer to this risk of open evidence as $E[\lambda_0|\lambda_1, ..., \lambda_n]$, or $E[\lambda_0|\lambda_1, \lambda_2]$ for the case of $n = 2$ that is analyzed in the model. For this expectation on openly available evidence, assume that firms expect a high λ_0 for a high level of own evidence and by applying for leniency firms will contribute an amount of evidence greater than λ_0:
AII *$E[\lambda_0|\lambda_1, ..., \lambda_n]$ is continuously differentiable for every $i \in [1, ..., n]$, $\frac{\partial E[\lambda_0|\lambda_1,...,\lambda_n]}{\partial \lambda_i} \geq 0 \ \forall \ i$, and $E[\lambda_0|\lambda_1, ..., \lambda_n] \leq \lambda_i \ \forall \ i$.*

Public evidence

When it is public knowledge to both firms how much evidence the other firm holds, then $E[\lambda_0|\lambda_1, \lambda_2]$ is obviously the same for both firms. The one stage decision problem of firms can be illustrated in a static game structure, as in Table 1 where firm 1 plays rows and firm 2 plays columns. In a first step, which is not displayed here, the actual values of λ_1 and λ_2 are realized. It is a dominant strategy for both firms to apply for leniency whenever $E[\lambda_0|\lambda_1, \lambda_2] > \theta$, which means that for a sufficiently high risk of getting convicted firms always strictly prefer to apply[3]. It is also a dominant strategy for firm i to remain silent whenever $E[\lambda_0|\lambda_1, \lambda_2] \leq \theta$ and $\lambda_j < \theta$.

For $E[\lambda_0|\lambda_1, \lambda_2] \leq \theta$ and $\lambda_1, \lambda_2 > \theta$ there are two equilibria and firms are essentially interested in choosing the same action as the other firm. If at least one's firm evidence is below θ, the other firm will strictly prefer to remain silent which accordingly leads to both firms rejecting leniency as they want to copy each other's actions.

So only for the case limited to $E[\lambda_0|\lambda_1, \lambda_2] \leq \theta$ and $\lambda_1, \lambda_2 > \theta$ is the behavior of firms not determined by dominant strategies. The tie between the two equlibria will be broken by selecting the Pareto-efficient equilibrium, which minimizes the expected fines for firms. Note that in the considered case

$$E[\lambda_0|\lambda_1, \lambda_2]F \leq \theta F < \frac{\lambda_2 + \theta}{2} F$$

for firm 1, which means that the expected fine from the equilibrium of both applying is higher than the expected fine from the equilibrium of none applying. The same is true for firm 2. So the Pareto-efficient equilibrium is given by the set of all pairs of (λ_1, λ_2) for which $E[\lambda_0|\lambda_1, \lambda_2] \leq \theta$, in other words it is Pareto-efficient for both to remain silent whenever this symmetric behavior is an equilibrium.

Summarizing these considerations for public evidence, the symmetric optimal strategy profile for firms is

$$\phi(\lambda_1, \lambda_2) = \begin{cases} \text{Apply} & \text{if } E[\lambda_0|\lambda_1, \lambda_2] > \theta \\ \text{Not apply} & \text{if } E[\lambda_0|\lambda_1, \lambda_2] \leq \theta. \end{cases} \tag{1}$$

[3]The second condition for this being a dominant strategy is actually $\lambda_2 > \theta$ for firm 1 (and $\lambda_1 > \theta$ for firm 2), but due to AII this condition is automatically fulfilled.

Private evidence

Under evidence being private each firm faces a higher degree in uncertainty when evaluating the risk of being caught. Following Harrington (2012) this model will assume them to solve this uncertainty by acting according to a cut-off strategy. They do not apply for low values of evidence but prefer to apply when they observe high values of evidence because they deem it too risky to stay silent. The symmetric strategy profile can be formulated as

$$\phi(\lambda_i) = \begin{cases} \text{Apply} & \text{if } \lambda_i \in (x, \bar{\lambda}] \\ \text{Not Apply} & \text{if } x \in [\underline{\lambda}, x]. \end{cases} \tag{2}$$

Given this strategy the set of symmetric cut-off Bayes-Nash equilibria will be defined by certain values of x for which it is optimal for firms to follow the strategy ϕ whenever the other firm does the same.

If firm 2 commits to the cut-off strategy, applying for leniency will imply an expected fine for firm 1 as follows:

$$G(x|\lambda_1)\theta F + (1 - G(x|\lambda_1)) \left(\frac{E[\lambda_2|\lambda_1, \lambda_2 > x] + \theta}{2} \right) F. \tag{3}$$

The first part of this expected fine is the probability of firm 2 not applying, which is the case whenever $\lambda_2 \leq x$, multiplied by the leniency reduced fine that firm 1 will receive for sure in this case. In the second part firm 2 is applying as well and both firms have an equal chance of being the first and of receiving leniency.

On the other hand, the expected fine from not applying can be calculated in the same way and is

$$G(x|\lambda_1)E[\lambda_0|\lambda_1, \lambda_2 \leq x]F + (1 - G(x|\lambda_1))E[\lambda_2|\lambda_1, \lambda_2 > x]F. \tag{4}$$

Again the first part displays the expected fine if firm 2 does not apply for leniency, in which case both firms only face the risk of getting convicted due to openly available evidence, and the second part is the expected fine if firm 2 applies for leniency, which means that firm 1 will be convicted with probability $E[\lambda_2|\lambda_1, \lambda_2 > x]$.

Summing up the previous equations, firm 1 strictly prefers to apply whenever the expected fine from applying is lower than the expected fine from not applying, which is true whenever expression (4) is greater than expression (3) or

$$E[\lambda_0|\lambda_1, \lambda_2 \leq x] - \theta > -\left(\frac{E[\lambda_2|\lambda_1, \lambda_2 > x] - \theta}{2}\right)\left[\frac{1 - G(x|\lambda_1)}{G(x|\lambda_1)}\right]. \tag{5}$$

The difference between the risk due to openly available evidence and the leniency rate must be sufficiently large to make applying an optimal choice. According to this condition, applying is the optimal choice if the left hand side of the equation is positive (then the right side is negative due to AII) or if both sides are negative but the difference between the risk due to available evidence and the leniency rate is sufficiently small. For all other cases firm 1 will prefer to remain silent, especially for the case $E[\lambda_2|\lambda_1, \lambda_2 > x] < \theta$. So the optimal strategy profile can be rewritten as

$$\phi(\lambda_i) = \begin{cases} \text{Apply} & \text{if } \lambda_i \in (x, \bar{\lambda}] \text{ and } E[\lambda_2|\lambda_1, \lambda_2 > x] \geq \theta \\ \text{Not Apply} & \text{if } x \in [\underline{\lambda}, x] \text{or } E[\lambda_2|\lambda_1, \lambda_2 > x] < \theta. \end{cases} \tag{6}$$

For further analysis on the relevant cut-off values define

$$\Delta(\lambda_1, x) \equiv E[\lambda_0|\lambda_1, \lambda_2 \leq x] - \theta + \left(\frac{E[\lambda_2|\lambda_1, \lambda_2 > x] - \theta}{2}\right)\left[\frac{1 - G(x|\lambda_1)}{G(x|\lambda_1)}\right]. \tag{7}$$

Now it will be optimal to for firms apply for $\Delta(\lambda_1, x) > 0$. This function is increasing in λ_1[4] and therefore it is possible to determine x through the function $\Delta(x, x)$. If $\Delta(x, x) \geq 0$ then firm 1 will apply for every $\lambda_1 > x$ because then $\Delta(\lambda, x)$ will be positive for all values of $\lambda_1 > x$. With the same argumentation it holds that the firm will not apply for every value of $\lambda_1 < x$ iff $\Delta(x, x) \leq 0$. Therefore, x will be an equilibrium cut-off iff $\Delta(x, x) = 0$. For a better interpretation of the relevant values of x consider

$$\begin{aligned} \Phi(x) &= G(x|x)\Delta(x,x) \qquad (8) \\ &= G(x|x)(E[\lambda_0|\lambda_1, \lambda_2 \leq x] - \theta) + \frac{E[\lambda_2|\lambda_1, \lambda_2 > x] - \theta}{2}(1 - G(x|x)). \end{aligned}$$

Again, applying will be preferred for $\Phi(x) > 0$, but this expression has the advantage that it gives a good illustration for when x takes the extreme values of $\underline{\lambda}$ or $\bar{\lambda}$. For the lower end $G(x|x)$ will be 0 which means that the difference between $E[\lambda_2|\lambda_1, \lambda_2 > x]$ and θ is essential when determining whether $\underline{\lambda}$ is an equilibrium cut-off. When the highest possible amount of evidence is considered as a possible threshold, then the opposite is the case and the difference between $E[\lambda_0|\lambda_1, \lambda_2 \leq x]$ and θ becomes essential for determining an equilibrium cut-off. In conclusion, the behavior of firms can be summarized as follows:

- If $\Phi(x) = 0$, x is an optimal symmetric cut-off value for firms when they follow a symmetric cut-off strategy. They will apply when they observe evidence with a value above this threshold and they will refrain from applying when their evidence is below this threshold.

- $\underline{\lambda}$ will be an equilibrium cut-off iff $E[\lambda_2|\lambda_1, \lambda_2 > \underline{\lambda}] \geq \theta$. Under public evidence firms apply whenever $E[\lambda_0|\lambda_1, \lambda_2] > \theta$. Due to AII the condition

[4]The sign of the derivative actually depends on the term $E[\lambda_2|\lambda_1, \lambda_2 > x] - \theta$, but after adjusting the strategy profile only the case of a positive derivative remains relevant for the further analysis of when firms are applying for leniency.

stated here under private signals is a weaker one. Note that if $\underline{\lambda}$ is not an equilibrium cut-off the only feasible equilibrium threshold will be $x = \bar{\lambda}$.

- $\bar{\lambda}$ will be an equilibrium cut-off iff $E[\lambda_0|\lambda_1, \lambda_2 \leq \bar{\lambda}] \leq \theta$, which is similiar to the condition for remaining silent under public evidence.

At this point it is not possible to strictly determine whether public or private evidence lead to more leniency applications. This would depend on how firms form their expectations on λ_0 and on the evidence of the other firm as can be seen by comparing $E[\lambda_0|\lambda_1, \lambda_2] - \theta > 0$ to condition (5). These inequalities capture the condition for applying and public and private evidence respectively. If firms strictly underestimate the evidence of the other firm ($E[\lambda_2|\lambda_1, \lambda_2 > x]$) the case of private evidence might even lead to less leniency applications. Nonetheless, the model under private evidence captures firms' fear of the other firm applying first and delivering its evidence to the CA. It also holds the possibility for an asymmetric behavior of firms, where one firm wishes to apply and the other to stay silent, if they receive different signals.

Conclusion and discussion

The main purpose of this thesis was to get a better understanding on the kind of information that is private to each firm and to shed some light on what impact this might have on firms' behavior. Similar to the results of Hamilton (2012) this model does not make quantitative predictions on the effect of private evidence but incorporates the possibility of an asymmetric behavior of firms given their evidence. Furthermore the introduction of private private evidence holds some potential for future research.

Private evidence provides a good starting point when thinking about transferring the model into a dynamic setting with several periods. The amount of evidence firms possess will naturally build up during periods of collusion, which means that actually an increase in λ_i can be expected until it reaches a critical point. Such a setting has the potential to explain firms' behavior as observed in the real world. Besides this, the model can also capture actions of the CA to influence firms' behavior as have been suggested in the literature.

In Silbye (2010) the CA makes use of private evidence by conditioning the leniency rate on the amount of evidence. In the present model it would also be possible to introduce a leniency scheme based on firms' evidence to reward high evidence firms. This would not only encourage firms to keep evidence but would also encourage high evidence firms and discourage others. That would enable the CA to reach the cartel member with the highest evidence and hence to increase its chances on convicting all cartel members.

Sauvagnat (2010) suggests the CA to have some private information on the probability of a successful conviction. He allows the CA to 'bluff' in some cases where conviction without leniency is highly unlikely. This would affect firms' expectations on λ_0 because the start of an investigation then means that with some probability the CA already has enough evidence for a conviction.

Finally Harrington (2012) proposes a more active behavior of the CA, in which it purposely leaks information to only one firm to actively create private information among firms. Here this would result in different values of $E[\lambda_0|\lambda_1, \lambda_2]$ among firms.

These examples show that the model developed in this thesis might serve as a framework for future leniency analysis. It enables the incorporation of several real-world features, that can potentially lead to well-founded policy implications.

References

BIGONI, M., C. FRIDOLFSSON, L. COQ, AND G. SPAGNOLO (2010): "'Trust, salience, and deterrence: evidence from an antitrust experiment," SSE/EFI Working Paper Series in Economics and Finance No. 696.

BUNDESKARTELLAMT (2010): "The Bundeskartellamt in Bonn Organisation, Tasks and Activities," `www.bundeskartellamt.de/wEnglisch/download/pdf/1011Infobroschuereenglisch.pdf`.

——— (2012): "Effective cartel prosecution - Benefits for the economy and consumers," `www.bundeskartellamt.de/wEnglisch/download/pdf/BKartAKartellverfolgungeng2020webbf.pdf`.

HARRINGTON, J. J. (2008): "Optimal corporate leniency programs," *Journal of Industrial Economics*, 56, 215–46.

——— (2012): *Corporate leniency with private information: an exploratory example*chap. Recent advances in the analysis of competition policy and regulation, pp. 28–48. Cheltenham: Edward Elgar Publishing Ltd.

MOTTA, M., AND M. POLO (2003): "Leniency programs and cartel prosecution," *International Journal of Industrial Organization*, 21, 347–79.

PINNA, A. (2010): "Optimal leniency programs in antitrust," Universita di Cagliari and Universita di Sassari, CRENoS Working Paper 2010/18.

SAUVAGNAT, J. (2010): "Prosecution and leniency programs: a fool's game," Toulouse School of Economics Working Paper Series 10-188.

SILBYE, F. (2010): "Asymmetric evidence and optimal leniency programs," Ph.D. thesis, University of Copenhagen.

SPAGNOLO, G. (2008): *Handbook of Antitrust Economics*chap. Leniency and whistleblowers in antitrust, pp. 259 – 303. Camebridge, MA: MIT Press.

Appendix

Tables

	apply	not apply
apply	$\frac{\lambda_2+\theta}{2}F,\ \frac{\lambda_1+\theta}{2}F$	$\theta F,\ \lambda_1 F$
not apply	$\lambda_2 F,\ \theta F$	$E[\lambda_0\vert\lambda_1,\lambda_2]F,\ E[\lambda_0\vert\lambda_1,\lambda_2]F$

Table 1: Game illustration for public signals in the modified model

Two-Sided Platforms with Negative Externalities and Quality Investments - An Application to Media Markets

Nima Jouchaghani *

Introduction

Two sided platforms have a wide range of interesting application possibilities, where the results of classic economy theory models like Cournot, Hotelling (1929) or Salop (1979) fail to explain some certain characteristics, which are important for the understanding of the choices made by the participating agents. The most important characteristics which are considered in the case of two sided platforms, are network effects and the existence of a third party in addition to consumers and producers, the platform owner. Network effects can be observed on many platforms, where consumers and/or producers interact with each other. For example there are network effects in social network platforms like Facebook, in specific, the utility of a Facebook member is growing with the number of his friends and other people who are also using Facebook. So there exists a positive externality, which members exert on each other.

One of the first applications of two sided markets with positive externalities can

*Nima Jouchaghani received his degree in Economics (B.Sc.) from the University of Bonn in March 2013. The present article refers to his bachelor thesis under supervision of JProf. Dr. Eugen Kovac, which was submitted in March 2013.

be found in Rochet and Tirole (2003). They find out that the structure of credit card and video game markets fits into the theory of two sided platforms with positive externalities. In the case of credit cards, consumers using a credit card have a higher utility, if there are many stores, which accept the card. On the other side the stores benefit from a growing number of consumers who are using the credit card which is accepted by them. So the agents in this model exert a positive externality on each other.

Up to this point we have only mentioned positive externalities, but there are also platforms in which we have the case, that one side benefits from a growing number of agents on the other side, while the other side exerts negative externalities. Consider the Facebook example again, companies are able to put advertisement banners on the website. Since members use social media networks mostly to interact with friends and other people, they dislike a growing number of advertisement. Therefore the companies exert a negative externality on the users, while they have benefits from a growing number of users at the same time.

In this work we deal with an application of two sided platforms with negative externalities, in specific we analyze the structure of competition in television stations. An early contribution to this topic can be found in Anderson and Coate (2005). They analyze broadcasting markets in order to find out if there is too much advertisement in matters of welfare, such that regulation is needed. Another paper to this topic is given by Reisinger (2004), whos work has a new feature given by viewers being able to choose the time they want to spend on a platform, so the utility they gain by using the platform is not given by a constant as in Anderson and Coate (2005).

This work is based on the model which is used by Reisinger (2004), but there are two modifications in case of competition structure. While Reisinger (2004) models two competing platforms, which can set a price for advertisement slots in order to maximize profits, we consider one monopoly platform owner. The reason for the modification is that the goal of this work is to develop a model, which will be compared with aggregated stylized facts of the television market in Germany.

The next central feature of this model refers to the decision variables of the platform owner. While in recent models the owners always choose a price for advertisement slots, we consider a new variable, the quality investment. The program of television stations mostly contains movies, series, talk shows, sports, documentaries and news. Since television stations benefit from a rising number of viewers[1], they try to offer a broad range of high quality content. This means in the case of movies for example, they try to buy licenses for the newest or most popular movies in order to attract more viewers.

The structure of this work is organized as follows: First the reader will be shown data for the German television market where we point out on some developments. Afterwards the model which is used to describe the market will be presented and solved for subgameperfect Nash equilibria. Finally we try to find a connection between the empirical facts and the predictions of the model.

The German Television Market

The German television market can be divided into the private and public television stations. Figure 1 in Appendix B shows that in the case of private television stations there are two mentionable companies which represent almost the whole private sector, the RTL Group and ProSiebenSat1 Media. In contrast to the public sector, the only way for private television stations to finance their programs,

[1] The more viewers they attract to the television station, the more valuable it is for advertisers to put advertisement on it. Since in Germany there is a limit of 12 minutes of advertisement per hour (see § 45 RStV,NI (1) for private television stations, § 16 RStV,NI (3) and § 16 RStV,NI (1) for public television stations),the only way to make more profit is to rise the value of advertisement time.

is by selling advertisement slots. The public broadcasting sector in Germany is represented by ARD and ZDF, who are mostly financed by fees.

In this work, we focus on the previously listed television stations due to the fact, that they represent over 87% of the market share of the German television market, which is sufficient for the comparison to a model, in which a monopoly platform owner sets the amount of advertisement and the level of quality investment.

The problem occurring at this point is, what kind of data to consider for the level of quality investment. Since there is a limit of advertisement per hour, both have an incentive to increase the value of advertisement time. This can be achieved by purchasing licenses or with good "in-house productions", in order to attract more viewers. So the factors we consider for the level of quality investment are: purchase/leasing of licenses, personal and material costs.

Figure 6 in Appendix B pictures the total amount of quality investments for the years 2001-2011. One can see that there is an obvious trend of increasing quality investments over the years, with a structural discontinuity at 2009, which might be a delayed consequence of the financial crisis. So we observe an almost constant growth of quality investments.

The next development refers to the total amount of advertisement in German television. Figure 2 shows that the amount of advertisement started climbing steadily, with a structural discontinuity at 2009, which might also be a consequence of the financial crisis[2]. Nevertheless there is a clearly defined pattern to this figure, showing that the amount of advertisement has been growing steadily, in specific it almost doubled from 2001-2011.

Figures 3-5 refer to developments on the viewers side. While figure 3 shows, that the average viewing time per day is increasing almost continuously, figure

[2]Consider that usually advertisement slots are bought by companies from the industry. Since the financial crisis brought many companies into difficulties, it might be that there has been a scarceness of demand for slots due to liquidity shortfall

4 points out that the percentage of television spectators started slightly growing at 1994, peaking at 2004 and then began dropping again.
Another development, which is quite important for the decision making of television stations and the viewers, is illustrated in figure 5. The basic assumption of our work is given by viewers disliking advertisement, since they watch television mainly to enjoy the offered content. But there might be changes in matters of how negative the effect of advertisement is received from the viewers perspective. Figure 5 pictures a questionnaire with a sample of 31447 persons, representing the German population. The main result is given by both general and television advertisement beeing received less negative over time. An explanation might be that over time, there could be a "learning effect" on the advertiser's side. In order to maximize the sales, advertisers need information about the viewers, in specific they need to know about the preferences and the behavior of viewers. So by advertising and afterwards observing the sales, advertisers can operate more efficiently in matters of on which television stations to advertise on, at which point in time and with what type of advertisement. Therefore, our assumption at this point is given by viewers aversion for advertisement decreasing over the time.

Model

The model used in this work, is based on the model given by Reisinger (2004). Besides the changes, including only one platform owner and the extension with quality investment, we also consider a concrete function for the utility of participating on the platform, given by $v(t)$.

Platform

There is a monopoly platform owner, which on the one side can set a price p for advertising on the platform, but on the other side can't exclude viewers from using it. Before setting the price, the platform makes a quality investment I, in order to attract more viewers. The profit function is given by

$$\Pi = np - I \tag{1}$$

where n is the amount of advertisement. The platform owners' object is to maximize profits.

Viewers

There is a mass M of viewers, who are utility maximizing individuals. They are uniformly distributed on the interval [0,1], where the platform is located at point 1. Viewers can choose the time t they want to spend on the platform. The more time they spend on the platform, the higher is the utility they gain. We consider $v(t)$ as an increasing, continuous and concave function, given by $\sqrt{t}$. Viewers also gain utility from the investment I of the platform, given by I^{α}. For now we only assume that I^{α} is also an increasing, continuous and concave function. The utility of the time, a viewer spends on the platform is decreasing by the amount n of advertisement on the platform.

Viewers have total time T to spend on the platform, or on other activities. Given t, the time a viewer is spending for other activities is given by $T - t$. The utility a viewer gains for spending time on other activities is normalized on 1 per time unit.

The maximization problem of a viewer who is located at x is given by:

$$\max_{t} \quad U(t,x) = (T - t) + \sqrt{t} I^{\alpha} - \gamma t n - \mathbb{1}_t \tau (1 - x)$$

where τ represents the transportation costs as given in a classic Hotelling model and γ is a parameter which measures the negative externality which advertisement exerts on the viewer. The indicator function $\mathbb{1}_t$ has value 1 for $t > 0$ and 0 for $t = 0$.

Consider that viewer gain no utility from the platform, if either t, I or both are zero.

Solving for t yields:

$$t(n, I) = \frac{I^{2\alpha}}{4(1+\gamma n)^2} \tag{2}$$

The next step is to find the marginal viewer, who is indifferent between participating on the platform and not. Viewers, who do not spend any time on the platform gain utility T. So we search for $\hat{x}$ such that $\mathrm{U}(t, \hat{x}) \overset{!}{=} T$, in specific :

$$(T - t) + \sqrt{t}I^{\alpha} - \gamma tn - \tau(1 - \hat{x}) \overset{!}{=} T \Longleftrightarrow \hat{x} = 1 - \frac{\sqrt{t}I^{\alpha} - \gamma tn - t}{\tau}$$

Since Viewers are uniformly distributed on [0,1], the demand of the platform is given by

$$(1 - \hat{x}) = \frac{\sqrt{t}I^{\alpha} - t(1+\gamma n)}{\tau} = \frac{I^{2\alpha}}{4(1+\gamma n)\tau} \tag{3}$$

The aggregated demand is given by $M(1 - \hat{x})$.

Advertisers

Advertisers are monopoly producers of differentiated products, for which a fraction β of users has a reservation value of K, while a fraction of $(1 - \beta)$ has a reservation value of 0.

Since advertisers are monopolists, they set the price for their product on K in order to gain the whole surplus from viewers[3].

With buying an advertisement slot, they can inform the viewers about their product. The profit of an advertiser, who decides to buy a slot on the platform is given by

$$\pi = MK\beta(1 - \hat{x})t - p \tag{4}$$

[3]We assume that viewers always have an income $\geq$ K

where $MK\beta(1-\hat{x})t$ is the gross value of advertisement on the platform and p is the price an advertiser has to pay for a slot. For simplicity, we stick to Reisinger (2004) with the assumption, that advertisers do not have any production costs for both the product and advertisement, considering that this does not change the qualitative results.

Timing and structure of the game

In the first stage, the platform owner decides about how much he wants to invest into the platform quality and the price for advertising slots. Since we have a monopoly platform owner, he can choose the optimal amount of advertisement instead of the price, without changing the results in equilibrium. In the next stage, the price for advertising is determined and viewers decide about how much time they want to spend on the platform.

Solving for Equilibrium

In this section we solve the described game for subgame perfect Nash equilibria. We assume that advertisers who do not put any advertisement on the platform, have a profit of 0. Hence the monopoly platform owner sets a price in equilibrium, such that advertisers are indifferent between participating on the platform and not. Therefore in equilibrium we have

$$p \overset{!}{=} MK\beta(1-\hat{x})t = \frac{MK\beta I^{4\alpha}}{16\tau(1+\gamma n)^3} \Longleftrightarrow np - I = \frac{MK\beta I^{4\alpha} n}{16\tau(1+\gamma n)^3} - I \qquad (5)$$

where the right term is equal to (1), which represents the profit function of the platform.

In order to derive the optimal amounts of advertisement and quality investment, we solve (5) for the maximization problem of the platform given by :

$$\max_{n,I} \quad \Pi = \frac{MK\beta I^{4\alpha} n}{16\tau(1+\gamma n)^3} - I \tag{6}$$

Solution

In equilibrium,we assume $\alpha=\frac{1}{8}$ such that the optimal amounts of n and I are given by

$$n^* = \frac{1}{2\gamma} \ \ and \ \ I^* = \frac{(MK\beta)^2}{\tau^2\gamma^2 46656} \tag{7}$$

which are sufficient for maximum[4].

Next, with this expressions, we can derive the optimal amount of time, spent by the viewers and the demand function of the platform in equilibrium. They are given by:

$$\begin{aligned} t^* &= \frac{(I^*)^{\frac{1}{4}}}{4(1+\gamma n^*)^2} = \frac{\sqrt{MK\beta}}{\sqrt{\tau}\sqrt{\gamma}\sqrt{17496}} \\ and \ (1-\hat{x})^* &= \frac{(I^*)^{\frac{1}{4}}}{4(1+\gamma n^*)\tau} = \frac{\sqrt{MK\beta}}{\tau^{\frac{3}{2}}\sqrt{\gamma}\sqrt{7776}} \end{aligned} \tag{8}$$

For our discussion, it is important to observe that increasing transportation cost, has a stronger impact on the demand function, than on the viewing time[5].

Discussion

In this section we are going to compare the results, given by our model and the empirical data which was presented before. We try to find explanations for the developments which took place on the German television market in connection with the prognostication of our model. Furthermore, we make assumptions about the long time behavior of the exogenous parameters in our model and explain why

[4]See appendix A, we have to assume $\alpha \leq \frac{1}{4}$ in order for (n^*, I^*) to be sufficient for maximum

[5]We observe that $(1-\hat{x})^*$ is decreasing faster in τ than t^*

there should have been a change. Consider that we restrict our explanations on the time interval of 2001-2011 due to limited data.

First recall the developments of the German television market. We observed an increasing amount of quality investment, viewing time and amount of advertisement, while the viewing rate remained constant. Therefore we are going to discuss, what possible explanations there could be for the viewing time increasing, while the percentage of television spectators stays almost constant, as shown in figures 3 and 4.

First reconsider the argument, that the nuisance parameter γ is decreasing over the time. The results in equilibrium of our model forecast, that investments I, viewing time t , percentage of television spectators $(1-\hat{x})$ and amount of advertisement n should increase. So in order to find an explanation for $(1-\hat{x})$ staying constant we need to make assumptions about the development of the other parameters. The percentage of television spectators, or in our model the demand of the platform in equilibrium given by (3) is decreasing in τ, considered as the transportation cost. In ordinary Hotelling models, the transportation cost represents the level of product differentiation, meaning that the higher τ, the more difficult it is for firms to attract new consumers. Since viewers can only choose between joining the platform and spending time on other activities, we have to explain why τ should increase in our model. For an explanation of this development we need to consider that "other activities" also includes the use of internet.

Since the beginnings of the internet, there have been lots of technical improvements, which offer new possibilities for internet users. In the present, users can watch almost all the television contents online at no cost[6] , so the internet has become a kind of substitute for television. Reisinger (2004) brings up the argu-

[6]Even if we do not consider the illegal possibilities, all television channels in Germany do have websites where the program can be watched at any time for a specific time period

ment, that over time, "people form habits", such that they will not switch easily from one platform to another. This could be a possible explanation for τ rising in our model, since a viewer who watches television gets used to this process more and more over time, which implies that he will not switch that easily to the internet as a platform.

Consider that we could also argue that τ should be decreasing over the time, since the contents of the internet and television became similar, which should be an indicator for a decreasing level of product differentiation. But we have to take into account that over time, the internet even over exceeded the possibilities of television. In specific there is more content available on the internet, than on television. Online platforms like Youtube even allow users to create own channels such that there is much more variety for the viewers, than given by television. Regarding figure 4, we can see that percentage of television spectators reaches the maximum at 2004, and then begins to decrease slowly until 2011. There might be a relation between the developments in the case of Youtube[7], which was founded in 2005 and grew constantly since then, and the decreasing percentage of television spectators. If we also consider that over the time the download speed has been increasing, the possibilities of enjoying media on the internet have also grown.

So assuming that τ begins increasing at 2005, this might be an explanation for the stagnation of the percentage of spectators. If we take a look at $(1 - \hat{x})$ in equilibrium, we see that it is the only variable, where the effect of τ rising is stronger than the effect of the decreasing nuisance parameter γ. So if we consider γ decreasing stronger then τ increasing, there is still a possibility for the

[7]We could assume that Youtube represents the internet as a media platform. Consider that there are several other ways to enjoy media in internet, like video on demand websites or even the websites of the television stations, but they all require almost the same technology, so we could take the development of Youtube as an approximation for the development of the whole media sector in internet

negative effect of the transportation cost to dominate over the decreasing nuisance parameter. This means in specific that in the case of our model, there is the possibility for I and t to increase, while $(1 - \hat{x})$ is almost constant or even decreasing.

But what about the remaining parameters? Consider the mass of viewers M. The match for M in reality can be assumed as the total population. Since there has not been any big change in matters of the total population in Germany, considering our time interval, we can assume that M is constant.

Next we take a look at the reservation value K. Since most advertisement in television is made for consumption goods, the inflation rate could be a good approximation for the changing of K. For Germany the average inflation rate each year[8] is approximately given by 1,6%, so for the time interval of 2001-2011, K has been increasing about 19%. So besides the decreasing of γ, we have a second parameter which has an increasing effect on the variables.

The remaining parameter β was defined as the fraction of users, who have a reservation value of K for the advertised products. If we take a look at Figure 5 in Appendix B, we see that the rate of viewers suggesting "advertisement helpful for the consumers" is increasing, the same goes for the rate of viewers claiming that "Television advertisement is quite informative". A growing affirmation for this two statements could be an evidence for advertisement to have an increasing influence on the viewers. Therefore there is evidence for β increasing over time.

Next we are going to discuss, if our model is able to explain the developments on the German television market, under the assumptions which we made considering the parameters. Reconsider that the optimal amount of advertisement in equilibrium was given by $n^* = \frac{1}{2\gamma}$. Since we argued that γ is decreasing, our model predicts that n^* should be increasing. If we take a look at figure 2, we see that

[8] For the years 2001-2011

this is the case until 2008, where we argued that the reason for this structural discontinuity could be a delayed consequence of the financial crisis.

The amount of investments in equilibrium was given by $I^* = \frac{(MK\beta)^2}{\tau^2\gamma^2 46656}$. The changes in γ, β and K increase investments in equilibrium, while τ decreases them. So the positive effect could overweight the negative effect of τ, such that our model forecasts that I^* should be increasing. Regarding figure 4, we see that on average the quality investments are increasing until 2008. The reason why there is a structural discontinuity, beginning in 2009, could also be an effect of the financial crisis.

Viewing time and demand of the platform in equilibrium were given by $t^* = \frac{\sqrt{MK\beta}}{\sqrt{\tau}\sqrt{\gamma}\sqrt{17496}}$ and $(1-\hat{x})^* = \frac{\sqrt{MK\beta}}{\tau^{\frac{3}{2}}\sqrt{\gamma}\sqrt{7776}}$. As mentioned before, we observe that for $(1-\hat{x})^*$ the impact of τ is stronger than any other parameter, while for t^* we have the same properties as for I^*. So in matters of our model, there exists a situation, in which an increase of β,K and a decrease of γ dominate over the effect of growing τ for I^* and t^*, while $(1-\hat{x})^*$ is decreasing. By looking at figures 3, 4 and 5 we observe, that this is exactly the case for the years 2005 - 2011, if we exclude the effects of the financial crisis on the investments.

Conclusion

The goal of this work has been to develop a model in order to describe the developments of the German television market. We argued that the theory of two sided platforms with negative externalities provides a good approximation of the structure of television markets. It has been shown that our model can explain the developments, if we make certain assumptions about the parameters. Therefore the application possibilities are restricted. Extensions of the model could be given by relaxing the fact that there is only one monopoly platform. Considering a model with two or more competing platforms would be a more realistic assumption concerning the german television market and could lead to much more accurate statements.

References

AGF (2012a): "Marktanteile 2012," http://www.agf.de/daten/zuschauermarkt/marktanteile, Accessed: March 26, 2013.

——— (2012b): "Seher 1988-2012," http://www.agf.de/daten/zuschauermarkt/seher, Accessed: March 26, 2013.

——— (2012c): "Verweildauer 1988-2012," http://www.agf.de/daten/zuschauermarkt/verweildauer, Accessed: March 26, 2013.

Anderson, S., and S. Coate (2005): "Market Provision of Broadcasting: A Welfare Analysis," *The Review of Economic Studies*, 72(2), 947–972.

Hotelling, H. (1929): "Stability in Competition," *Economic Journal*, 39(153), 41–57.

Reisinger, M. (2004): "Two-Sided Markets with Negative Externalities," http://epub.ub.uni-muenchen.de/478/.

Rochet, J., and J. Tirole (2003): "Platform Competition in Two-Sided Markets," *Journal of the European Economic Association*, 1(4), 990–1029.

Salop, S. (1979): "Monopolistic Competition with Outside Goods," *Bell Journal of Economics*, 10(1), 141–156.

STA (2011): "TV-Werbeminuten in Deutschland 2002-2011," http://de.statista.com/statistik/daten/studie/4769/umfrage/tv-werbesekundenin-deutschland-seit-2000/, Accessed: March 26, 2013.

ZdW (2010): "Werbeklima 2007-2010," http://www.zaw.de/index.php?menuid=66, Accessed: March 26, 2013.

Appendix

Proof of the Solution

The problem of the platform was given by

$$\max_{n,I} \quad \Pi = \frac{MK\beta I^{4\alpha} n}{16\tau(1+\gamma n)^3} - I$$

By deriving the first order conditions we get

$$\frac{\partial \Pi}{\partial n} = \frac{\frac{MK\beta}{16\tau} I^{4\alpha}[(1+\gamma n)^3 - 3\gamma n(1+\gamma n)^2]}{(1+\gamma n)^6} \stackrel{!}{=} 0 \tag{9}$$

$$\Leftrightarrow (1+\gamma n) - 3\gamma n \stackrel{!}{=} 0 \Longleftrightarrow n^* = \frac{1}{2\gamma}$$

and

$$\frac{\partial \Pi}{\partial I} = \frac{\frac{MK\beta}{16\tau} 4\alpha I^{4\alpha-1} n}{(1+\gamma n)^3} - 1 \stackrel{!}{=} 0 \tag{10}$$

$$\Leftrightarrow I^* = \left(\frac{(1+\gamma n)^3}{n \frac{MK\beta}{16\tau} 4\alpha} \right)^{\frac{1}{4\alpha-1}}$$

using $n^* = \frac{1}{2\gamma}$ we have

$$I^* = \left(\frac{27}{\frac{MK\beta\alpha}{\gamma\tau}} \right)^{\frac{1}{4\alpha-1}}$$

The next step is to proof that (n^*, I^*) is sufficient for maximum. Therefore we need to show, that the Hessian matrix of the the second-order partial derivations of Π is negative definite in (n^*, I^*). By simplifying (15) we have

$$\frac{\partial \Pi}{\partial n} = \frac{MK\beta}{16\tau} I^{4\alpha} \left(\frac{1-2\gamma n}{(1+\gamma n)^4} \right)$$

Deriving the second-order partial derivation yields

$$\frac{\partial^2 \Pi}{\partial n^2} = \frac{MK\beta}{16\tau} I^{4\alpha} \left(\frac{-2\gamma(1+\gamma n)^4 - (1-2\gamma n)4\gamma(1+\gamma n)^3}{(1+\gamma n)^8} \right)$$

$$\Leftrightarrow \frac{\partial^2 \Pi}{\partial n^2} = \frac{MK\beta}{16\tau} I^{4\alpha} \left(\frac{-2\gamma(1+\gamma n) - (1-2\gamma n)4\gamma}{(1+\gamma n)^5} \right)$$

using $n^* = \frac{1}{2\gamma}$ yields

$$\left.\frac{\partial^2 \Pi}{\partial n^2}\right|_{n=\frac{1}{2\gamma}} = \frac{MK\beta}{16\tau} I^{4\alpha} \left(\frac{-2\gamma - \gamma - 4\gamma + 4\gamma}{(\frac{3}{2})^5} \right) = -\frac{2MK\beta I^{4\alpha}\gamma}{81\tau} \tag{11}$$

In the next step we need to derivate (16) with respect to I

$$\frac{\partial^2 \Pi}{\partial I^2} = \frac{MK\beta 4\alpha(4\alpha-1)I^{4\alpha-2}n}{16\tau(1+\gamma n)^3}$$

using $n^* = \frac{1}{2\gamma}$ yields

$$\left.\frac{\partial^2 \Pi}{\partial I^2}\right|_{n=\frac{1}{2\gamma}} = \frac{MK\beta 4\alpha(4\alpha-1)I^{4\alpha-2}}{32\tau\gamma(\frac{3}{2})^3} = \frac{MK\beta\alpha(4\alpha-1)I^{4\alpha-2}}{27\tau\gamma} \tag{12}$$

At this point the reader might see that we need to assume $\alpha \leq \frac{1}{4}$ in order for the Hessian matrix to be negative definite[9]. For simplicity we assume $\alpha = \frac{1}{8}$ such that (17) and (18) are given by

$$\left.\frac{\partial^2 \Pi}{\partial n^2}\right|_{n=\frac{1}{2\gamma}} = -\frac{2MK\beta\sqrt{I}\gamma}{81\tau} \quad and \quad \left.\frac{\partial^2 \Pi}{\partial I^2}\right|_{n=\frac{1}{2\gamma}} = -\frac{MK\beta I^{-\frac{3}{2}}}{432\gamma}$$

[9] Furthermore if we take a look at (16), we see that this inequation has to be strict in order for I^* to be an inner solution

The last step is to derive the cross derivatives of Π. Using the symmetry of second derivatives and $\alpha = \frac{1}{8}$ we have

$$\frac{\partial^2 \Pi}{\partial I \partial n} = \frac{\partial^2 \Pi}{\partial n \partial I} = \frac{\frac{MK\beta}{16\tau}(1-2\gamma n)}{2(1+\gamma n)^4 \sqrt{I}} \overset{n=\frac{1}{2\gamma}}{=} 0$$

Using the second order derivatives for the Hessian matrix we have

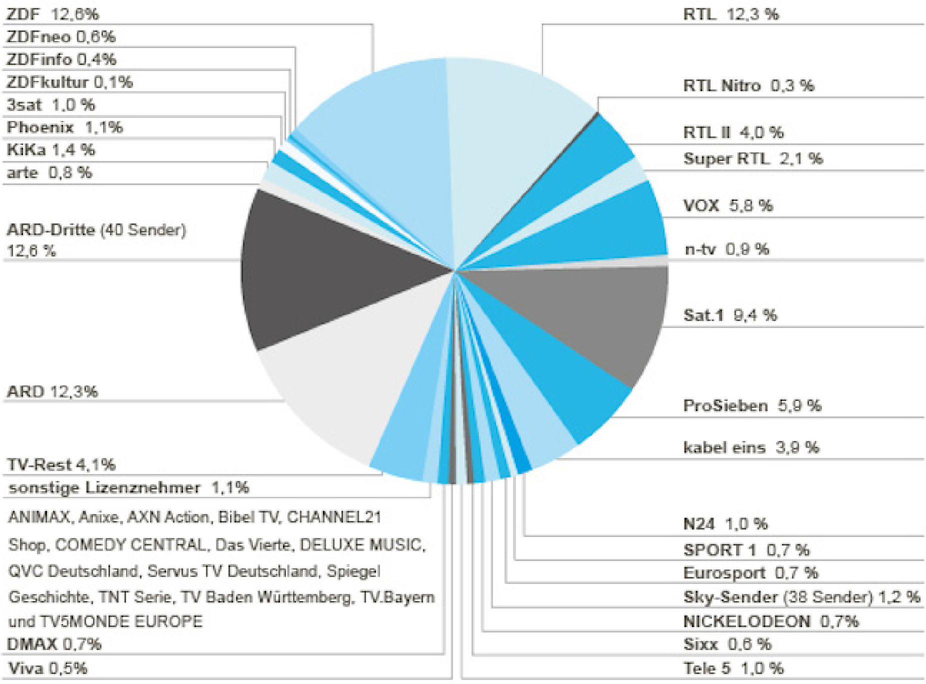
$$\mathcal{H} = \begin{pmatrix} -\frac{2MK\beta\sqrt{I}\gamma}{81\tau} & 0 \\ 0 & -\frac{MK\beta I^{-\frac{3}{2}}}{432\gamma} \end{pmatrix}$$

In order for (n^*, I^*) to be a maximum, it has to be shown that for all real valued column vectors $\mathbf{x} \in \mathbb{R}^2$ we have $\mathbf{x}^T \mathcal{H} \mathbf{x} < 0$, therefore $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ yields

$$\mathbf{x}^T \mathcal{H} \mathbf{x} = -x_1^2 \frac{2MK\beta\sqrt{I}\gamma}{81\tau} - x_2^2 \frac{MK\beta I^{-\frac{3}{2}}}{432\gamma}$$

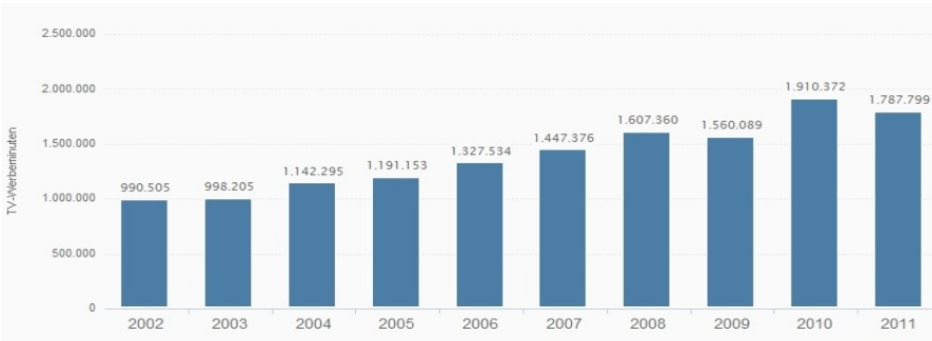
which is < 0 since $I^* > 0$ for the given parameters. $\square$

Figures



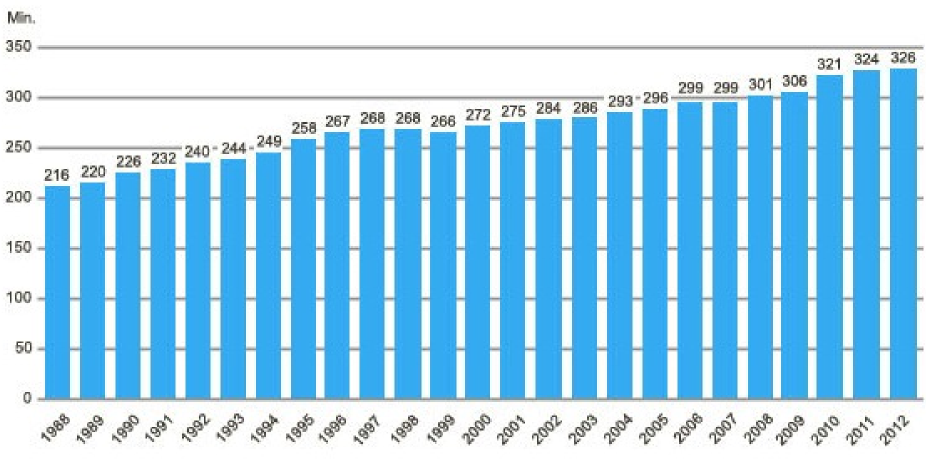
Source: Arbeitsgemeinschaft Fernsehforschung, AGF (2012a)

Figure 1: Market share of German television stations , based on daily average for 2012



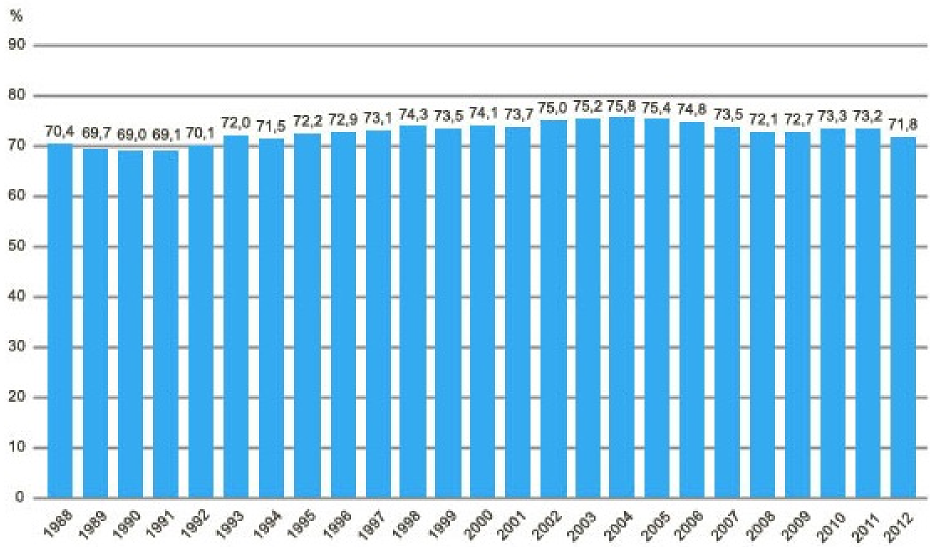
Source: Statista

Figure 2: Total amount of advertisement minutes in German television for 2002-2011



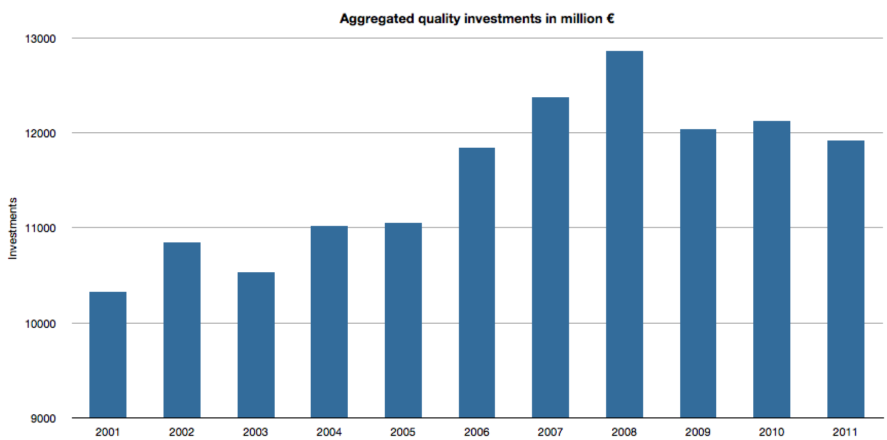
Source: Arbeitsgemeinschaft Fernsehforschung, AGF (2012c)

Figure 3: Average viewing time of a viewer per day in minutes for 1988-2012



Source: Arbeitsgemeinschaft Fernsehforschung, AGF (2012b)

Figure 4: Percentage of television spectators on an average day for 1988-2012



Source: Own illustration and calculation

Figure 5: Aggregated quality investments of German television stations in million €

Tables

Statement	2007	2008	2009	2010
Advertisement gives useful hints for new products	53,2%	57,0%	60,8%	63,2%
Sometimes advertisement is helpful for consumers	45,3%	51,9%	58,6%	61,2%
Mostly advertisement is amusing	35,9%	41,1%	43,6%	45,5%
I like watching television advertisement	33,3%	35,6%	37,2%	40,6%
Television advertisement is quite informative	36,6%	40,6%	43,2%	46,2%

Source: Zentralverband der deutschen Werbewirtschaft, ZdW (2010); Own illustriation

Table 1: Attitude towards advertisement, the percentages represent the relative amount of interviewed persons saying: "I totally agree" or "I mostly agree"

Sources and Calculation of Quality Investments

Company	2001	2002	2003	2004	2005	2006
ProSiebenSat1.Media	1908,7	1902,3	1695,5	1576,5	1620,2	1672,4
RTL Group	3458	3716	3702	4092	4306	4732
ZDF	1251,5	1359,9	1273,3	1364,3	1326,7	1443
ARD	3705,4	3870,3	3860,2	3981,9	3796,6	3994
Σ	10323,6	10848,5	10531,0	11014,7	11049,5	11841,4

Company	2007	2008	2009	2010	2011
ProSiebenSat1.Media	2335	2850,9	2310,8	2341,7	2159,2
RTL Group	4737	4738	4535	4382	4537
ZDF	1386,4	1490,1	1431,3	1508	1419,8
ARD	3920,7	3785,6	3759,5	3892,7	3806,5
Σ	12379,1	12864,6	12036,6	12124,4	11922,5

Table 2: An overview of the individual quality investments for each company for the years 2001 − 2011

Since the accounting method differs from company to company, we give a separate overview about the sources and the calculation methods. In general we used the annual financial statements of the companies for collecting the data we needed to calculate the quality investments. Values were rounded up to one decimal place.

ProSiebenSat1.Media

The annual financial statement for each year is available online at

`http://www.prosiebensat1.com/de/medialounge/downloads/publikationen/2012` [accessed 26.03.2013]

The underlying variables for calculating the quality investments were:

Programm- und Materialaufwand + Personalaufwand + Abschreibungen[10]+ Sonstige betriebliche Aufwendungen, for 2001 (p.71)[11], 2002 (p.46), 2003 (p.42)

Herstellungskosten + Vertriebskosten + Verwaltungskosten, for 2004 (p.62), 2005 (p.68), 2006 (p.172), 2007 (p.68), 2008 (p.90)

Umsatzkosten + Vertriebskosten + Verwaltungskosten, for 2009 (p.114), 2010 (p.115), 2011 (p.130)

RTL Group

[12] The annual financial statement for each year is available online at

`http://www.rtlgroup.com/www/htm/annualreport.aspx` [accessed 26.03.2013]

The underlying variables for calculating the quality investments were:

Consumption of current program rights + other operating expense, for 2001 (p.72), 2002 (p.72), 2003 (p.72), 2004 (p.80), 2005 (p.84), 2006 (p.110), 2007 (p.110), 2008 (p.112), 2009 (p.105), 2010 (p.139), 2011 (p.164)

---

[10]The depreciation was only considered due to the fact, that for the years 2004 - 2011 it was integrated in "Herstellungskosten" and "Umsatzkosten". So in order to make the different underlying variables comparable, we needed to include the depreciation for 2001-2003

[11]Example: Can be found in the annual financial statement of 2001 on page 71

[12]Consider that the RTL Group also owns television and radio stations in other countries than Germany, but since we are more interested in changes of investments than in absolute values, it should still be a good approximation for the german television channels

ARD

The annual financial statement for each year is available online at

`http://www.kef-online.de/inhalte/berichte.html` [accessed 26.03.2013]

The underlying variables for calculating the quality investments were:

Programmaufwendungen + Personalaufwendungen

for 2001 – 2003 see Report 15 vol.1, p.26 and p.35[13]

for 2004 – 2007 see Report 17, p.64 and p.77

for 2008 – 2011 see Report 18, p.50 and p.77

ZDF

The annual financial statement for each year is available online at

`http://www.kef-online.de/inhalte/berichte.html` [accessed 26.03.2013]

The underlying variables for calculating the quality investments were:

Programmaufwendungen + Personalaufwendungen

for 2001 – 2003 see Report 15 vol.1, p.29 and p.35

for 2004 – 2007 see Report 17, p.67 and p.77

for 2008 – 2011 see Report 18, p.52 and p.77

[13]Example: Can be found in the 15. Report, volume 1, page 26 for "Programmaufwendungen", page 35 for "Personalaufwendungen"

Variance bounds tests for the hypothesis of efficient stock market

Marco Maisenbacher*

Introduction

The theory of efficient financial markets was regarded as inviolable in academic literature for a long time. An efficient financial market is characterized by the complete reflection of all relevant information in the market prices, which means that permanent deviations from the fundamental justified valuation are impossible (Fama, 1970). Miller and Modigliani (1961) set up a first model to capture the idea of efficient stock markets. In this model the value of a stock should equal the rational expected, discounted value of all future dividends of the stock.

In the most recent financial crisis the validity of efficient financial markets was brought into question. Long time before the financial crisis Shiller (1981) recognized the phenomenon that especially stock price indices like the Standard and Poor's 500 Composite Stock Price Index are much too volatile to be explained by the traditional fundamental value model of Miller and Modigliani. Based on this observation Shiller develops a variance bound for efficient stock markets. He concludes the violation of the fundamental value model since the variance of

*Marco Maisenbacher received his degree in Economics (B.Sc.) from the University of Bonn in March 2013. The present article refers to his bachelor thesis under the supervision of Prof. Dr. Jörg Breitung, which was submitted in January 2013.

the S&P 500 stock price index is five times higher as suggested by his variance bound.

Shiller (1981) initiates the discussion about excess volatility on financial markets. In the following years especially Flavin (1983), Marsh and Merton (1986) and Kleidon (1986) criticize Shiller's conclusion from his variance bound test due to undesired small finite sample properties in his test.

In response to Shiller (1981) and the rising critic against his variance bound test, Mankiw, Romer and Shapiro (1985) develop a modified variance bound relation which holds in finite samples. After some further refinements of their test, Mankiw et al. (1991) obtained more differentiated results with respect to the validity of the efficient market hypothesis. According to their results, there is a overall tendency to reject the hypothesis of efficient stock markets but this finding is less pronounced compared to the earlier study of Shiller (1991).

In this paper the framework of Mankiw et al. (1991) is applied to an updated data set in order to test the hypothesis of efficient stock markets.

The original variance bound for efficient stock markets

The idea of a variance bound in efficient financial stock markets bases upon the present value model by Miller and Modigliani (1961). This model can be characterized by the following two equations:

$$P_t^* = \sum_{i=0}^{\infty} \left(\frac{1}{1+r} \right)^{i+1} D_{t+i} \tag{1}$$

$$P_t = E_t P_t^*, \tag{2}$$

where D_t denotes the dividend of a stock for period t and r is the expected return assumed to be constant. $E_t(\cdot)$ denotes the expectation conditional on information available at time t, especially the present and all past prices of the stock.

Therefore P_t^* is the unobservable ex post rational price of the stock and P_t is the rational forecast of P_t^* at time t. Equations (1) and (2) describe the fundamental value model, where the price of a stock equals the expected, discounted value of all future dividends. According to this model fluctuations of the market price of a stock should be fully explained by the emergence of new information about future dividends.

Shiller (1981) challenged the validity of the fundamental value model. The (ex post observable) series P_t^* appeared to be much smoother than the series of the actual market prices P_t. Shiller (1991) questioned that the disparity of the volatilities can be adequately explained by the mere emergence of new information. In order to analyze this issue empirically, Shiller derives the first variance bound for efficient stock markets. Equation (1) can be reformulated as

$$P_t^* \quad = \quad P_t + \epsilon_t \ , \tag{3}$$

where ϵ_t denotes the rational forecast error at time t with zero mean. Using the fact that under rational expectations ϵ_t has to be uncorrelated with all known information at time t , the variance of P_t^* simplifies to

$$Var(P_t^*) \quad = \quad Var(P_t) + Var(\epsilon_t) \tag{4}$$

and since $Var(\epsilon_t) \geq 0$, it follows that

$$Var(P_t^*) \geq Var(P_t). \tag{5}$$

Equation (5) states the simplest form of a variance bound in efficient stock markets and indicates that the variance of the ex post rational prices has to be at least as large as the variance of the market prices. Shiller (1991) compares the sample variance of P_t^* and P_t computed from the Standard and Poor's 500 Composite Stock Price Index for 1871–1979 and finds the variance of the market prices P_t to be five times higher than the one obtained from their rational ex post counterpart P_t^*. To ensure the variances of both series to be finite, Shiller assumes both series to be trend stationary. Obviously, this result questions the validity of the fundamental value model for stock markets.

In response to Shiller's work many authors criticize the assumptions and interpretation of his test. Flavin (1983) notes that both sample variances of P_t and P_t^* underestimate the true population variances. However the negative bias for the variance of P_t^* is larger, which means that the understimation for $Var(P_t^*)$ is stronger than for $Var(P_t)$. This negative bias results from replacing the unknown expectation by sample means.

Marsh and Merton (1986) challenge the entire interpretation of Shiller's results. According to them the violation of Shiller's variance bound in equation (5) does not necessarily imply a violation of the present value model. In contrast Shiller's method is a test of the joint hypothesis of efficient financial stock markets with constant return and a trend stationary dividend process. Following this argument a violation of the variance bound in (5) might occur due to a violation of the assumed constant return or the trend stationary dividend process even though the assumption of an efficient stock market is fulfilled.

Kleidon (1986) reveals a second weakness in Shiller's interpretation. He shows that from a single time series of the realized observations of P_t^* and P_t nothing can be concluded in terms of the validity of Shiller's variance bound. Kleidon explains this seemingly counterintuitive argument as follows. The variance bound (5) is based on repeated samples of the process $\{P_1^*, \ldots, P_T^*\}$ because different realizations of future dividends result in different sequences of $\{P_1^*, \ldots, P_T^*\}$. Shiller's variance bound implies that among all possible realizations the variance of P_t^* is expected to be larger than the variance of P_t. Accordingly the variance bound represents a cross-sectional relation between different states of an economy at time t and no relation between the time series variances.

The modified variance bound test

In response to Shiller's first trial, Mankiw et al. (1991) develop a modified test that is more accurate in finite samples. In order to derive their test statistic, some new definitions need to be introduced. Mankiw et al. (1991) define the ex post present value P_t^{*h} for the strategy of buying a stock at time t and holding it for h periods as

$$P_t^{*h} = \sum_{j=0}^{h-1} \left(\frac{1}{1+r}\right)^{j+1} D_{t+j} + \left(\frac{1}{1+r}\right)^{h} P_{t+h} \tag{6}$$

where D_{t+j} denotes the dividend in period $t+j$ and P_{t+h} is the market price in period $t+h$. Under the assumptions of the present value model it holds that

$$P_t = E_t P_t^{*h}. \tag{7}$$

The investment horizon h can be chosen as variable or constant. In the variable case, the investment horizon coincides with the end of the sample such that $h_t = T - t$. Alternatively h can be chosen as constant for every observation such that P_t^{*h} displays the ex post present value for the strategy of holding the stock until period $t+h$ and selling it for the market price. This new definition of P_t^{*h} is a basic component for the modified test. The derivation of this test is based on the ideas of Mankiw et al. (1985). Let P_t^0 be an arbitrary ("naive") forecast of the fundamental value of the stock:

$$P_t^0 = \sum_{k=0}^{\infty} \rho^{k+1} \widehat{D}_{t+k} \tag{8}$$

where $\widehat{D}_{t+k}$ denotes naive forecast for the dividend D_{t+k} at period t and ρ is the known (constant) discount factor. The naive forecast does not have to be rational (i.e. the forecast may neglect available information). It is important however that the market participants may have access to the naive forecast at period t, that is, the naive forecast is entailed in the investor's information set. In order to derive Mankiw et al.'s (1985) modified test the following identity serves as starting point:

$$P_t^{*h} - P_t^0 \quad = \quad (P_t^{*h} - P_t) + (P_t - P_t^0). \tag{9}$$

From equation (7) it follows that $P_t^{*h} - P_t$ displays the rational forecast error ϵ_t which is independent of any available information at period t. Therefore it holds that:

$$E_t[(P_t^{*h} - P_t)(P_t - P_t^0)] \quad = \quad 0. \tag{10}$$

Squaring equation (9), using expectations and substituting equation (10) yields:

$$E_t(P_t^{*h} - P_t^0)^2 = E_t(P_t^{*h} - P_t)^2 + E_t(P_t - P_t^0)^2. \tag{11}$$

Equation (11) will remain valid if the conditional expectations are normalized with any scaling variable W_t known at t. Equation (11) can be reformulated as

$$E_t\left(\frac{P_t^{*h} - P_t^0}{W_t}\right)^2 \quad = \quad E_t\left(\frac{P_t^{*h} - P_t}{W_t}\right)^2 + E_t\left(\frac{P_t - P_t^0}{W_t}\right)^2. \tag{12}$$

In a last step define

$$q_t = \left(\frac{P_t^{*h} - P_t^0}{W_t}\right)^2 - \left[\left(\frac{P_t^{*h} - P_t}{W_t}\right)^2 + \left(\frac{P_t - P_t^0}{W_t}\right)^2\right]. \tag{13}$$

Equation (12) implies that $E_t(q_t) = 0$ holds. However, by the law of iterative expectation, this implies $E_{t-s}(q_t) = 0$ for all $s \geq 0$. Therefore a test of the hypothesis of efficient stock markets implies that

$$H_0: \quad \alpha = 0, \tag{14}$$

in the regression $q_t = \alpha + \varepsilon_t$, where ε_t is an error term with expectation zero. Since the expectation of q_t is zero for all t, the expectation of the mean $\overline{q}$ is zero as well. In order to construct a test statistic for the null hypothesis (14), asymptotically valid standard errors have to be constructed. Mankiw et al. (1991) stress the issue of autocorrelation in the errors. For constant holding periods h and under the assumptions of efficient markets q_t and q_{t-j} are correlated for $j < h$ but uncorrelated for $j \geq h$. In the case of variable holding periods h_t the correlation does not vanish after a fixed lag. To account for the autocorrelation in the error terms Mankiw et al. (1991) use Newey-West standard errors which are asymptotically valid. In the case of a constant holding periods h, the truncation lag is set to $h - 1$ since the autocorrelation vanishes after this lag. In the case of variable holding periods the rule of thumb of Newey-West is chosen for the truncation lag.[1] The final test statistic is the square of the t-statistic $\widehat{\alpha}^2/\widehat{Var}(\widehat{\alpha})$, which is a two-sided Wald-statistic with a χ^2 distribution with one degree of freedom. Note that the estimated variance of $\widehat{\alpha}$ is calculated with the Newey-West standard errors.

[1]Newey-West's rule of thumb for the truncation lag for unknown autocorrelation is $P = int[4(T/100)^{\frac{2}{9}}]$.

Finally there are two more useful implications of Mankiw et al.'s (1985) test which can be derived from equation (12):

$$E\left(\frac{P_t^{*h}-P_t^0}{W_t}\right)^2 \quad \geq \quad E\left(\frac{P_t^{*h}-P_t}{W_t}\right)^2 \tag{15}$$

and

$$E\left(\frac{P_t^{*h}-P_t^0}{W_t}\right)^2 \geq E\left(\frac{P_t-P_t^0}{W_t}\right)^2. \tag{16}$$

The first upper bound in (15) claims that the expected squared error with a naive forecast (P_t^0) should be at least as large as the expected squared error with the optimal forecast (P_t). The upper bound in (16) states that the volatility of P_t^* around P_t^0 should be at least as large as the volatility of P_t around P_t^0. If the null hypothesis is rejected both upper bound relations can be helpful to detect the source of rejection.

In contrast to Shiller's test, Mankiw et al. (1985) show that their upper bound relations in (15) and (16) are unbiased regardless of the sample size and the underlying dividend process. This is achieved by centering the variances around a naive forecast and not around the sample mean.

Empirical Analysis

In this section the modified test of Mankiw et al. (1991) is applied to real data in order to test the hypothesis of efficient stock markets. All time series are annual data from 1871 to 2011. The stock price series consists of data of the Standard and Poor's 500 Composite Price Index, where the price of a year is represented by the average of the daily closing prices for January. The dividend series consists of dividends per stock, added over 12 months and adjusted to the index for the fourth quarter of each year. Both series are converted to real units with the

Consumer Price Index. In order to conduct the test of efficient stock markets the naive forecast P_t^0 has to be specified. In a first version of the test Mankiw et al. (1991) specify P_t^0 derived from the no-change forecast of future dividends so that the expected dividends are identical to the last observed value (D_{t-1}). Therefore P_t^0 can be expressed as:

$$P_t^0 = \frac{1}{r} D_{t-1}. \tag{17}$$

Table 1 presents the results of the test using the naive forecast in (17) and constant returns of 5%, 6% and 7% for different holding periods. The Columns (ii), (iii) and (iv) display the sample mean of $[(P_t^{*h} - P_t^0)/P_t]$, $[(P_t^{*h} - P_t)/P_t]^2$ and $[P_t - P_t^0)/P_t]^2$. The scaling variable W_t is P_t in order to diminish the issue of heteroskedasticity since the variables are growing over time. Column (v) shows the result of the test statistic, which is the squared difference of column (ii) with the sum of column (iii) and (iv), divided by the variance of this difference. The variance is calculated with the Newey-West standard errors. Under the null hypothesis (14) the test statistic is asymptotically χ^2 distributed with one degree of freedom, implying a critical value of 3.84 for a significance level of 0.05. Column (vi) shows the respective p-values. The theory of efficient stock markets predicts that the entries in column (ii) should equal the sum of the entries in column (iii) and (iv) which is tested by the statistic in column (v). Furthermore the entries in column (ii) should be larger than the entries in column (iii) and (iv), as presented in the upper bound relations in (15) and (16).

In terms of the validity of inequality (15) Table 1 shows that the relation holds except for the case of $r = 5\%$ with variable holding periods $h = T - t$. Only in this case the naive forecast in (17) is a better forecast than the market price in terms of the forecast error variance. The inequality (16) is stable as well since column

(ii) almost always exceeds column (iv). The only exception is the case of $r = 7\%$ and variable holding periods. Accordingly the volatility of P_t^{*h} around the naive forecast in (17) is almost always larger than the volatility of P_t around the naive forecast. Therefore, the market prices are not excessively volatile around the naive forecast. The p-values of the test statistic for the hypothesis that column (ii) equals the sum of column (iii) and (iv) show a tendency for accepting the null hypothesis of efficient stock markets. Only the p-values of 5 and 10 year holding periods with an expected return of 5% yields a rejection of the null hypothesis at the 10% significance level. However the picture for the variable holding periods is different. The p-values imply a significant rejection of the null hypothesis for every expected return. The present value model with constant expected return is not supported for variable holding periods and the naive forecast defined in (17).

Table 2 shows the results of a similar test as before but with a different naive forecast. The alternative naive forecast consists of a thirty year moving average of the dividends and can be written as:

$$P_t^0 = \frac{1}{r}\left[\frac{1}{30}\sum_{i=1}^{30} D_{t-i}\right]. \tag{18}$$

Mankiw et al. (1991) choose this particular forecast to smooth the series P_t^0. The smoothed naive forecast should help to detect the excess volatility of the market prices. Furthermore, the scaling variable W_t is set to P_t^0, which is supposed to avoid a bias in the test results due to possible excess volatility in the series P_t used above. The test results in Table 2 indeed display excess volatility of the market prices around the naive forecast defined in (18). The values in column (iv) exceed the values in column (ii) for every holding period and every expected return. However the second upper bound relation always holds since column (ii) always exceeds column (iii). This implies that the market price P_t is a better forecast for the ex post rational price P_t^{*h} than the naive forecast in (18). The p-values imply acceptance of the null hypothesis for constant holding periods below 10 years for every expected return. However for the 10 year holding periods the null hypothesis has to be rejected. In the case of variable holding periods the rejections are much weaker compared to the tests presented in Table 1. In the case with $r = 5\%$ the null can be accepted at the 5% level. In general the test with the modified naive forecast displays a stronger tendency to accept the hypothesis of efficient stock markets except for the 10 year holding period.

Next we relax the assumption of constant expected returns. We follow Mankiw et al. (1991) and construct time varying expected return as the sum of a variable, riskless interest rate (r_t^*) and a constant risk premium (ϕ). Therefore the one period nominal discount factor is given by $\rho_t = 1/(1 + r_t^* + \phi)$. Under the hypothesis of efficient stock markets it holds that:

$$\begin{aligned} P_t &= E_t \left[\sum_{j=0}^{h-1} \rho_t^{j+1} D_{t+j} + \rho_t^h P_{t+h} \right] \\ &\equiv E_t P_t^{*h} \ . \end{aligned} \tag{19}$$

For the naive forecast Mankiw et al. (1991) assume that the dividends grow with the riskless interest rate. Therefore the naive forecast can be expressed as $P_t^0 = \frac{1}{\phi} D_{t-1}$. According to Mankiw et al. (1991) a test with variable discount factors is especially suitable if changes in the interest rates are considered as important drivers for market price volatility. In this case one would expect a rejection of the null hypothesis of efficient stock markets in tests with constant expected return.

Table 3 on page 11 shows the test results for different risk premiums $(4\%, 5\%, 6\%)$ and the same holding periods as in the tests before. The data for the riskless interest rate (r_t) are annual commercial paper rates of "triple A" ranked companies which are taken from the Federal Reserve Bank of St.Louis. Furthermore all data are in nominal terms now since the dynamic of the inflation is captured by the riskless interest rate. The results of the new tests in Table 3 do not support the validity of the present value model with variable discount rates. Even though the null hypothesis cannot be rejected for low holding periods of one and two years, there is a strong tendency towards rejection for longer holding periods. Especially the variable holding periods lead to strong rejections. The upper bound for the volatility of the market prices is, like in the case of constant returns, almost always met. The upper bound of the forecast error variance holds for constant holding periods only. In general the test with variable discount rates exhibits a stronger tendency for rejecting the null hypothesis compared to the test with constant returns.

Table 4 shows the test results for variable discount rates based on the smoothed naive forecast $P_t^0 = \frac{1}{\phi}\left[\frac{1}{30}\sum_{i=1}^{30} D_{t-i}\right]$. The scaling factor W_t is again P_t^0. The test results tend to be similar to those of the analog test with constant returns. The upper bound of the volatility is almost always violated, whereas the upper bound for the forecast error variance is always met. The null hypothesis is accepted more often with the smoother naive forecast compared to the test before. The results of the tests with variable discount rate show that variation in the interest rates does not seem to play an important role for the volatility of the stock market prices. The tests do not yield more evidence in favor of the null hypothesis but tend to stronger rejections.

Conclusion

In this paper two different variance bounds tests were considered. Shiller's (1981) approach leads to a strong rejection of the hypothesis of efficient stock markets. However, due to the undesired small sample properties and the sensitivity in terms of the underlying dividend process, his results are not reliable. The results of the modified variance bound test of Mankiw et al. (1991) yields a more differentiated picture. Using this test, the hypothesis of efficient stock markets cannot be generally accepted, however it cannot be definitely rejected neither. For low holding periods there is a tendency for accepting the null hypothesis. In contrast, especially variable holding periods lead to a clear rejection. The degree of significance of the rejections depends on the choice of the naive forecast. The smoothed naive forecast yields weaker rejections of the null hypothesis. Furthermore it was shown that variation in the interest rates is not a significant driver of price volatility on stock markets since the test with variable discount rates leads even to stronger rejections.

To sum up, the modified test of Mankiw et al (1991) with an updated sample neither delivers clear evidence against the fundamental value model nor does it generally support the hypothesis .

References

(2012): "6-Month Commercial Paper Rate," `http://research.stlouisfed.org/fred2/series/WCP6M?cid=120`, Federal Reserve Bank of St. Louis, Economic Research.

Fama, E. F. (1970): "Efficient Capital Markets: A Review of Theory and Empirical Work," *The Journal of Finance*, 25(2), 383–417.

Flavin, M. A. (1983): "Excess Volatility in the Financial Markets: A Reassessment of the Empirical Evidence," *Journal of Political Economy*, 91(6), 929–956.

Friedman, M., and A. J. Schwartz (1982): "Monetary Trends in the United States and United Kingdom: The Relation to Income, Prices, and Interest Rates, 1867- 1971," *University of Chicago Press*, p. Chapter 4: The Basic Data.

Kleidon, A. W. (1986): "Variance Bounds Tests and Stock Price Valuation Models," *Journal of Political Economy*, 94(5), 953–1001.

Mankiw, G. N., D. Romer, and M. D. Shapiro (1985): "An Unbiased Reexamination of Stock Market Volatility," *The Journal of Finance*, 40(3), 677–687.

——— (1991): "Stock Market Forecastability and Volatility: A Statistical Appraisal," *Review of Economic Studies*, 58(3), 455–477.

Marsh, T. A., and R. C. Merton (1986): "Dividend Variability and Variance Bounds Tests for the Rationality of Stock Market Prices," *The American Economic Review*, 76(3), 483–498.

Miller, M. H., and F. Modigliani (1961): "Dividend Policy, Growth, and the Valua- tion of Shares," *The Journal of Business*, 34(4), 411–433.

Shiller, R. J. (1981): "Do Stock Prices Move Too Much to Be Justified by Subsequent Changes in Dividends?," *The American Economic Review*, 71(3), 421–436.

——— (2012): "ONLINE DATA ROBERT SHILLER," `www.econ.yale.edu/~shiller/data.htm`.

Appendix

Tables

(i)	(ii)		(iii)		iv)	(v)	(vi)
h	$E\left[\frac{P_t^{*h}-P_t^0}{P_t}\right]^2$	$=$	$E\left[\frac{P_t^{*h}-P_t}{P_t}\right]^2$	$+$	$E\left[\frac{P_t-P_t^0}{P_t}\right]^2$	χ^2	p-value
					$r = 5\%$		
1	0.1401		0.0295		0.1203	1.1560	0.2823
2	0.1542		0.0628		0.1186	1.8879	0.1694
5	0.2038		0.1583		0.1146	3.9917	0.0457
10	0.2993		0.2933		0.1015	3.2229	0.0726
$T-t$	0.3119		0.3450		0.1224	10.7373	0.0010
					$r = 6\%$		
1	0.1670		0.0284		0.1420	0.1213	0.7276
2	0.1854		0.0588		0.1398	0.3596	0.5487
5	0.2345		0.1358		0.1344	0.6263	0.4287
10	0.3005		0.2192		0.1191	0.7466	0.3876
$T-t$	0.1886		0.1364		0.1445	7.9729	0.0047
					$r = 7\%$		
1	0.2113		0.0276		0.1880	0.1463	0.7021
2	0.2274		0.0556		0.1857	0.3009	0.5833
5	0.2614		0.1193		0.1798	0.5654	0.4521
10	0.2893		0.1707		0.1640	1.3156	0.2514
$T-t$	0.1235		0.0921		0.1906	15.8151	0.0001

Note: column (i): holding periods; column (ii)-(iv): sample estimator (sample means) of the expectation from (12, weighted with the market price; column (v): $\chi^2(1)$ test statistic for the hypothesis, that column (ii) equals the sum of (iii) and (iv) ; column (vi): p-values of the test statistic. The naive forecast is defined in (17).

Table 1: Test with a naive forecast, constant expected return (r)

(i)	(ii)		(iii)		(iv)	(v)	(vi)
h	$E\left[\frac{P_t^{*h}-P_t^0}{P_t^0}\right]^2$	$=$	$E\left[\frac{P_t^{*h}-P_t}{P_t^0}\right]^2$	$+$	$E\left[\frac{P_t-P_t^0}{P_t^0}\right]^2$	χ^2	p-value
					$r = 5\%$		
1	1.3686		0.1058		1.3911	0.8206	0.3650
2	1.3167		0.2347		1.3740	0.6623	0.4158
5	1.1597		0.4299		1.2506	0.8374	0.3601
10	0.9174		0.6797		0.9729	11.7948	0.0006
$T-t$	0.9960		0.4948		1.4181	2.8938	0.0889
					$r = 6\%$		
1	2.2756		0.1514		2.3654	1.1434	0.2849
2	2.1420		0.3340		2.3358	0.8392	0.3596
5	1.7711		0.6000		2.1363	0.9293	0.3350
10	1.2679		0.9023		1.6894	10.2238	0.0014
$T-t$	1.1505		0.6311		2.4104	4.1781	0.0409
					$r = 7\%$		
1	3.4241		0.2060		3.6384	1.5866	0.2078
2	3.1554		0.4541		3.5933	1.0804	0.2986
5	2.4531		0.8164		3.3005	1.0698	0.3010
10	1.5932		1.2031		2.6466	9.9416	0.0016
$T-t$	1.2360		1.0670		3.7056	5.9329	0.0149

Note: See Table 1. The naive forecast is defined in (18).

Table 2: Test with naive forecasts based on smoothed dividends, constant expected return

(i)	(ii)		(iii)		(iv)	(v)	(vi)
h	$E\left[\frac{P_t^{*h}-P_t^0}{P_t}\right]^2$	$=$	$E\left[\frac{P_t^{*h}-P_t}{P_t}\right]^2$	$+$	$E\left[\frac{P_t-P_t^0}{P_t}\right]^2$	χ^2	p-value
					$\phi = 4\%$		
1	0.1944		0.0293		0.1820	2.0742	0.1498
2	0.1962		0.0616		0.1815	4.2748	0.0387
5	0.1993		0.1353		0.1811	5.9018	0.0151
10	0.2653		0.2508		0.1743	1.6808	0.1948
$T-t$	0.3150		0.4310		0.1829	8.4567	0.0036
					$\phi = 5\%$		
1	0.1401		0.0286		0.1206	1.0703	0.3009
2	0.1494		0.0585		0.1189	2.5016	0.1137
5	0.1633		0.1200		0.1149	4.4654	0.0346
10	0.2178		0.1995		0.1017	3.9402	0.0471
$T-t$	0.2023		0.2125		0.1227	12.7049	0.0004
					$\phi = 6\%$		
1	0.1602		0.0280		0.1422	1.1447	0.2847
2	0.1688		0.0562		0.1400	1.8632	0.1723
5	0.1743		0.1096		0.1346	3.3706	0.0664
10	0.2018		0.1679		0.1193	4.9080	0.0267
$T-t$	0.1410		0.1549		0.1448	20.8802	0.0000

Note: See Table 1. The naive forecast is $P_t^0 = \frac{1}{\phi} D_{t-1}$. The data are in nominal terms.

Table 3: Test using naive forecast with recent dividends and variable interest rates

(i)	(ii)		(iii)		(iv)	(v)	(vi)
h	$E\left[\frac{P_t^{*h}-P_t^0}{P_t^0}\right]^2$	$=$	$E\left[\frac{P_t^{*h}-P_t}{P_t^0}\right]^2$	$+$	$E\left[\frac{P_t-P_t^0}{P_t^0}\right]^2$	χ^2	p-value
					$\phi = 4\%$		
1	2.1453		0.1325		2.1594	0.5018	0.4787
2	2.1070		0.2954		2.1419	0.3577	0.5498
5	1.9917		0.5315		1.9918	0.5275	0.4676
10	1.7463		0.8426		1.6128	11.2890	0.0008
$T-t$	1.7607		0.6465		2.1874	2.1543	0.1422
					$\phi = 5\%$		
1	3.8768		0.2056		3.9888	0.8075	0.3689
2	3.7248		0.4556		3.9551	0.5225	0.4698
5	3.3027		0.8007		3.6903	0.7522	0.3858
10	2.6312		1.2010		3.0285	13.0424	0.0003
$T-t$	2.3323		0.8698		4.0405	3.5394	0.0599
					$\phi = 6\%$		
1	6.1238		0.2959		6.4341	1.2280	0.2678
2	5.7631		0.6545		6.3794	0.7409	0.3894
5	4.8122		1.1483		5.9690	1.0024	0.3167
10	3.5014		1.6839		4.9491	13.7785	0.0002
$T-t$	2.8715		1.4906		6.5161	5.1177	0.0237

Note: See Table 1. The naive forecast is $P_t^0 = \frac{1}{\phi}\left[\frac{1}{30}\sum_i^{30} D_{t-i}\right]$. The data are in nominal terms.

Table 4: Test with smoothed naive forecast and variable interest rates

Numerical Approximation of an Optimum Growth Program

Simon Hoof*

Introduction

Economics of growth deals with optimal intertemporal allocation of scarce resources. To consider it a problem of dynamic optimization one has to realize the dual character of consumption in an intertemporal framework. If present output can either be consumed or saved, which implies a transfer to future resources, a more of consumption today will lessen future output and hence future consumption possibilities. Less consumption today instead will imply forgone instantaneous utility for the sake of investment. Hence solving for the optimum amount of present consumption is a key issue of economic modeling.

The problem was first treated by Ramsey (1928). Koopmans (1965) and especially Cass (1965) explored Ramsey's classic approach by applying the *maximum principle*, which is a powerful dynamic optimization tool for economic problems (Shell, 1969). This work lead to the now well-known textbook model of optimal growth (Acemoglu, 2009; Chiang, 1992).

Since the existence of an optimum consumption trajectory is shown, we actually want to approximate a numerical solution. For this purpose we use a

* Simon Hoof received his degree in socioeconomics (B. A.) from the University of Hamburg in August 2013. The present article refers to his bachelor thesis under the supervision of Dr. Thorsten Pampel, which was submitted in October 2013.

technique named *eigendecomposition of a matrix* (Novales, Fernández, and Ruiz, 2009). The solution algorithm enables us to simulate optimum growth trajectories for different initial conditions or exogenous shocks.

Model

Consider the dynamical system described by (1)→(4)

$$\max_{c} \; \mathcal{W}(k(0)) = \int_0^\infty u(c)e^{-rt}\mathrm{d}t \tag{1}$$

$$\text{s.t.} \quad \frac{\mathrm{d}k(t)}{\mathrm{d}t} =: \dot{k} = f(k) - c - (n+\delta)k \tag{2}$$

$$k_0 = k(0) \tag{3}$$

$$c \in [0, f(k)] \tag{4}$$

where t is the time index. The objective functional (welfare integral) is denoted by $\mathcal{W}(\cdot)$. It is maximized over the control variable c, which represents per worker consumption. The society's utility function $u(\cdot)$ is twice continuous differentiable with respect to c and strictly concave, i.e. $u'(c) > 0$, $u''(c) < 0$. Future consumption utilization is discounted by the time preference rate $r > 0$. The law of motion for capital accumulation is given by (2), where k is the capital labor ratio. The linear homogeneous production function $f(\cdot)$, which represents the output per worker $f(k) = y$, is twice continuous differentiable with respect to k and strictly concave, i.e. $f'(k) > 0$, $f''(k) < 0$. The growth rate of the labor force[1] is denoted by n and $\delta \in [0, 1]$ represents the depreciation rate of capital. The initial condition is given by $k(0) = k_0$. The control variable is restricted by (4). The problem of the social planer is to maximize society's utility from present and future consumption flows by controlling the state k over c.

[1] For simplicity we assume that labor force equals population.

Dynamic Optimization The dynamic optimization problem given in (1) can be solved by the maximum principle of Pontrjagin, Boltjanski, Gamqrelidse, and Mishchenko (1964). This method refers to the Hamiltonian approach to dynamic economics (Cass and Shell, 1976).

Definition 1 (Hamiltonian). $\mathcal{H}(t,k,c,\mu) := e^{-rt}\{u(c) + \mu[f(k) - c - (n+\delta)k]\}$

where μ is the current value costate variable, which expresses the shadow price for an extra unit of capital at time t.[2] The shadow price therefore "translates" a change of the capital stock in a change of utility. Thus the Hamilton function measures total utility flows, which are derived directly by instantaneous utility $u(\cdot)$ and indirectly by the change of capital $\mu\dot{k}$. The maximum principle provides the following first order conditions[3]

$$\mathcal{H}_c = u'(c) - \mu = 0, \tag{5}$$

$$\dot{\mu} = -\mathcal{H}_k e^{rt} + r\mu = -\mu[f'(k) - (n+\delta+r)]. \tag{6}$$

where $\mathcal{H}_c$ is the partial derivative of the Hamiltonian with respect to c and $\mathcal{H}_k$ with respect to k respectively.[4] For an economic reasoning of these conditions see Dorfman (1969). Differentiating (5) with respect to t yields

$$\dot{\mu} = u''(c)\dot{c}. \tag{7}$$

[2]While the present value multiplier $\lambda e^{rt} := \mu$ measures the shadow price of an extra unit of capital at time $t = 0$.

[3]The conditions of the maximum principle are necessary and sufficient for a maximization, if the maximized Hamiltonian, defined as $\mathcal{H}^0(t,k,\mu) := \max_c \mathcal{H}(t,k,c,\mu)$, is concave in k as stated by Arrow (1968) and formally proved by Seierstad and Sydsæter (1977).

[4]For completion we have to impose two transversality conditions $\lim_{t\to\infty} \mu(t)e^{-rt} = \lim_{t\to\infty} \mathcal{H}(t) = 0$.

Combining and rearranging (5), (6) and (7) finally gives the *Euler equation*

$$\dot{c} = \frac{c}{\sigma}[f'(k) - (n + \delta + r)] \tag{8}$$

where $\sigma := -cu''(c)/u'(c) > 0$[5] is the measure of intertemporal risk aversion.

System Dynamics and Phase Diagram The differential equations (2) and (8) are known to be a two dimensional system of differential equations which describe the dynamics of the state and control variable in the (k, c)-plane. The possible phase transitions are described by the following cases

$$\dot{k} \left\{ \begin{matrix} > \\ = \\ < \end{matrix} \right\} 0 \quad \text{if} \quad c \left\{ \begin{matrix} < \\ = \\ > \end{matrix} \right\} f(k) - (n + \delta)k \tag{9}$$

$$\dot{c} \left\{ \begin{matrix} > \\ = \\ < \end{matrix} \right\} 0 \quad \text{if} \quad f'(k) \left\{ \begin{matrix} < \\ = \\ > \end{matrix} \right\} n + \delta + r \tag{10}$$

Figure 1(a) refers to the general system dynamics. We call the two loci $\dot{k} = 0$ and $\dot{c} = 0$ phase boundaries, which seperate the $\mathbb{R}^2_+ := \{(k, c)|k, c \geq 0\}$ in single phase regions where phase transitions occur. The arrows indicate the transition direction of capital and consumption over time. A dynamic equilibrium is located at the intersection of $\dot{k} = \dot{c} = 0$, which is formally defined as a fixed point.

Definition 2. *A fixed point $E(\tilde{k}, \tilde{c})$ is an equilibrium point in the (k, c)-plane, such that $\lim_{t\to\infty} \dot{k}(t) = \dot{c}(t) = 0$ holds.*

Figure 1(b) shows phase lines for different initial choices of $c(0)$, where $k(0)$ is given exogenously. There exists one pair of phase lines which flow towards the

[5] The strict concavity of $u(\cdot)$ implies $\lim_{c\downarrow 0} u'(c) \to \infty$, i.e. $c > 0$.

fixed point and is indicated by the broken lines $(- \cdot -)$. The "stable branch" is known to be the *balanced growth path*. With a given $k(0)$ there exists only one optimal choice of $c^*(0)$ which leads on the optimal growth path. A different choice of $c(0)$ will lead to an explosive growth of either c or k while the other variable shrinks to zero.[6]

Computation of an Optimum Growth Path

The fixed point property of $\dot{c} = 0$ yields $f'(\tilde{k}) = n + \delta + r$. Consider a Cobb-Douglas production function $f(k) = k^\alpha$ with $\alpha \in (0, 1)$ representing the prodcution elasticity of capital. The steady state value for capital is then given by

$$\begin{aligned} \alpha\tilde{k}^{\alpha-1} &= n + \delta + r \\ \Rightarrow \tilde{k} &= \left(\frac{\alpha}{n + \delta + r}\right)^{\frac{1}{1-\alpha}}. \end{aligned} \tag{11}$$

In addition a fixed point implies $\dot{k} = 0$, that is

$$\tilde{c} = \tilde{k}^\alpha - (\delta + n)\tilde{k}. \tag{12}$$

The $\ln(\cdot)$ of (2) and (8) are

$$\frac{\mathrm{d}\ln k}{\mathrm{d}t} = e^{(\alpha-1)\ln k} - e^{\ln c - \ln k} - (n + \delta) =: \Pi(k, c) \tag{13}$$

$$\frac{\mathrm{d}\ln c}{\mathrm{d}t} = \frac{1}{\sigma}\left[\alpha e^{(\alpha-1)\ln k} - (n + \delta + r)\right] =: \Psi(k, c) \tag{14}$$

[6]Formally one cannot neglect the complementary slackness condition eq. (4), i.e. consumption is only feasible if there is a positive capital stock. Hence the critical value $c = f(k)$ leads the growth path instantly to the repellor point $(0, 0)$.

A functional relation between consumption and capital is given by (see technical appendix p. 95)

$$\ln c(t) = \frac{\xi}{\omega_2}(\ln k(t) - \ln \tilde{k}) + \ln \tilde{c}, \quad \forall t \in [0, \infty). \tag{15}$$

where $\xi := \Psi(\tilde{k}, \tilde{c})_{\ln k} = \dfrac{(\alpha - 1)(n + \delta + r)}{\sigma}$ is an auxiliary variable and $\omega_2 = 0.5 \cdot \left(r - \sqrt{r^2 - 4\psi\xi}\right)$ is the second characteristic root with the second auxiliary variable $\psi := -\Pi(\tilde{k}, \tilde{c})_{\ln c} = \dfrac{(1-\alpha)(n+\delta) + r}{\alpha}$. Notice that consumption is now determinated by capital and a given $k(0) = k^*(0)$ thus yields $c^*(0)$. By the law of motion for capital (2) one might iterate the optimum trajectories for the model variables.

Simulation

In the spirit of Kendrick and Taylor (1971) and Islam (2001) we now compute optimal trajectories for model economies. The structural parameters of the benchmark economy are calibrated as follows $\alpha = 0.3$, $\delta = 0.2$, $\sigma = 2.0$, $r = 0.04$ and $n = 0.0$.[7]

Different Initial Conditions Figure 2(a) shows that two different initial capital stocks $k_1(0) > k_2(0)$ are given. First we analyse the general pattern, which both trajectories refer to. Assume the given capital stock $k(0)$ is lower than the long term optimal stock $\lim_{t\to\infty} k^*(t) = \tilde{k}$, which is derivable by the production function and the given structural parameters (see eq. (11)). The corresponding per capita consumption $\tilde{c}$ is obtained by (12). To force the economy on the stable growth path, which tends towards the dynamic equilibrium $(\tilde{k}, \tilde{c})$, we derive optimal initial consumption $c^*(0)$ by (15). Figure 2 shows the transition dynamics for

[7] The matlab source code for the considered simulations is available at `//www.bje.uni-bonn.de/.../volume-iii.../code-for-hoof-2014/`

capital, consumption, production, investments, instantaneous utility and current value shadow price.Initial consumption is lower than production $c(0) < y(0)$, i.e. gross investment is positive $i(0) > 0$. If gross investment outweighs depreciated capital, we have positive net investment, i.e. $i(0) - \tilde{i} > 0$ and the economy is accumulating capital. Hence the production in the next period will increase. The forgoing consumption today will increase future consumption potential and utility as well. Since the utility function is concave, the marginal utility of an additional consumption unit decreases as well as the shadow price of capital. That is the social planer is willing to give up consumption for an extra unit of capital if the consumption level is relatively high. The process of capital accumulation continues until the economy reaches the fixed point, where the social planer has no incentive to increase capital, because the future utility gain does not compensate present consumption abstinence. To keep production constant the planer uses a fraction of output to cover just the depriciated capital, i.e. the economy is in steady state. In addition it can be shown that the two transversality conditions $\lim_{t\to\infty} \mu(t)e^{-rt} = \mathcal{H}(t) = 0$ hold, as stated by the maximum principle.

Figure 2 shows a counterintuitively result. One may wonder that an economy with a relatively large initial capital stock will converge to steady state more rapidly than an economy with a relatively lower one. However figures 2(a)→2(f) show that those trajectories adjust to each other over time and reach steady state at the same time. This phenomenon is describable as follows: a given capital stock instantly determines initial consumption by (15). Production is given by $f(k(0)) = y(0)$ and dued to $i(0) = y(0) - c(0)$ initial investment is given as well. Figure 2(d) shows that investment of economy *2* excess economy *1*, i.e. $i_1(0) < i_2(0)$. Furthermore the extra amount of capital will be used more efficiently, since marginal productivity is relatively higher on a lower capital stock $f'(k_1(1)) < f'(k_2(1))$. As shown in figure 2(c) we might figure that by the steeper slopes of $\dot{y}_1 < \dot{y}_2$. Consequently consumption and corresponding utility increase faster (cf. fig. 2(b) and 2(e)). At the end the investment dynamic results in a successive adjustment over time until the fixed point is reached.

Impulse Response Consider a steady state economy in which all macroeconomic variables are quasi-constant. We now simulate a shock by fixing all structural parameters but increase the time preference ($r \uparrow$) at $t = 10$. The planer now values present consumption more than future consumption. Figure 3(b) shows that short-run consumption increases as well as instantaneous utility (cf. 3(e)). The more of consumption is possible at the expense of investment, which explains the slump in figure 3(d). The heavy decline of investment in the short-run is known as overshooting, since the reaction is stronger than long term development. It follows that future capital stock and simultaneously production shrink (cf. fig. 3(a) and 3(c)). On the other hand a higher consumption level leads to a lower shadow price for capital, since it is measured in marginal utility (cf. fig. 3(f)). This effect balances the shrinking process such that investment recovers and production is stabilized. In the long run the steady state values for capital, consumption, output, investment and utility are lower than before the shock. That is just the result of valuing present consumption at the expense of capital accumulation and future consumption potential.

Conclusion

The present paper shows an easy-to-use solving technique for dynamic optimization problems. The solution algorithm enables the user to approximate optimal trajectories for the state, costate and control variable. One might adjust and/or add structural parameters to solve various problems.

References

ACEMOGLU, D. (2009): *Introduction to Modern Economic Growth.* Princeton University Press.

ARROW, K. J. (1968): "Applications of Control Theory to Economic Growth," in *Mathematics of the Decision Sciences*, ed. by G. B. Dantzig, and A. F. Veinott, pp. 85–119. American Mathematical Society.

CASS, D. (1965): "Optimum Growth in an Aggregative Model of Capital Accumulation," *The Review of Economic Studies*, 32(3), 233–240.

CASS, D., AND K. SHELL (eds.) (1976): *The Hamiltonian Approach to Dynamic Economics.* Academic Press.

CHIANG, A. C. (1992): *Elements of Dynamic Optimization.* McGraw-Hill.

DORFMAN, R. (1969): "An Economic Interpretation of Optimal Control Theory," *The American Economic Review*, 59(5), 817–831.

ISLAM, S. M. N. (2001): "Computation of non-linear continuous optimal growth models: experiments with optimal control algorithms and computer programs," *Economic Modelling*, 18(4), 551–568.

KENDRICK, D. A., AND L. TAYLOR (1971): "Numerical Methods and Nonlinear Optimizing Models for Economic Planning," in *Studies in Development Planning*, ed. by H. B. Chenery, chap. 1, pp. 11–28. Harvard University Press.

KOOPMANS, T. C. (1965): "On the Concept of Optimal Economic Growth," in *Study week on the econometric approach to development planning*, Pontificiae Academiae Scientiarum scripta varia 28, pp. 225–300. North Holland.

NOVALES, A., E. FERNÁNDEZ, AND J. RUIZ (2009): *Economic Growth. Theory and Numerical Solution Methods.* Springer.

PONTRJAGIN, L. S., W. BOLTJANSKI, R. GAMQRELIDSE, AND E. F. MISHCHENKO (1964): *Mathematische Theorie optimaler Prozesse.* Oldenbourg.

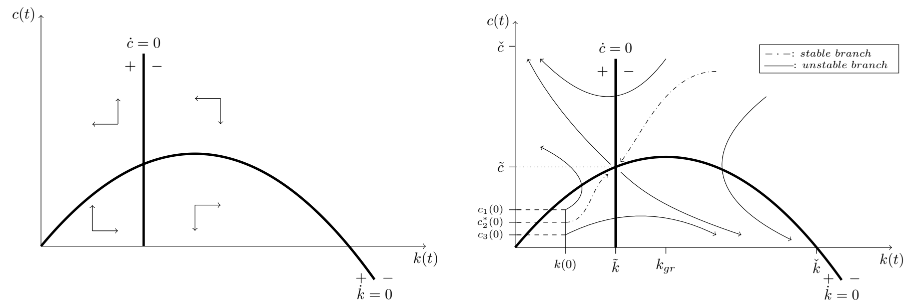
RAMSEY, F. P. (1928): "A Mathematical Theory of Saving," *The Economic Journal*, 38(152), 543–559.

SEIERSTAD, A., AND K. SYDSÆTER (1977): "Sufficient Conditions in Optimal Control Theory," *International Economic Review*, 18(2), 367–391.

SHELL, K. (1969): "Applications of Pontryagin's Maximum Principle to Economics," in *Mathematical Systems Theory and Economics I*, ed. by H. Kuhn, and G. Szegö, pp. 241–292. Springer.

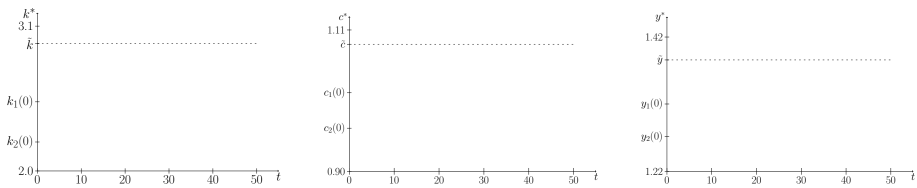
Appendix

Figures

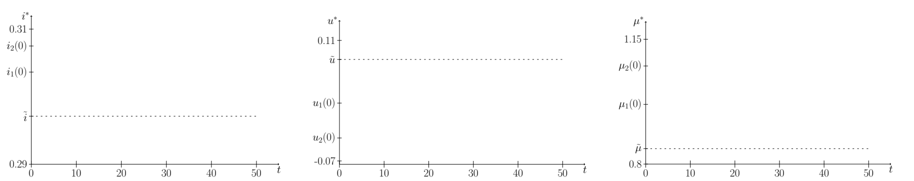


(a) (b)

Figure 1: Phase Diagram for Consumption and Capital Dynamics



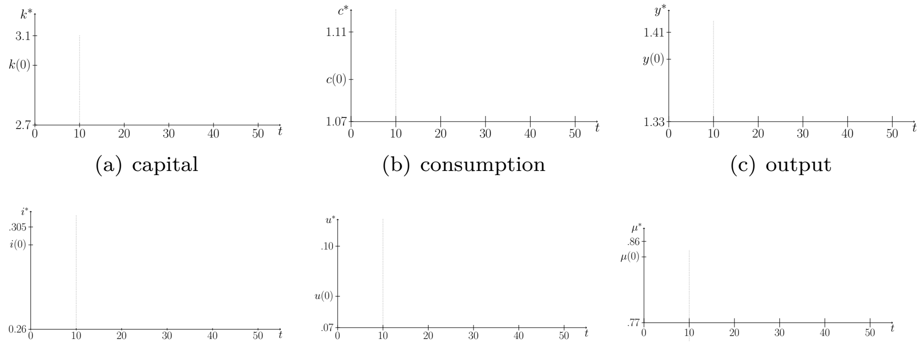
(a) capital (b) consumption (c) output



(d) investment (e) utility (f) shadow price $\mu(t)$

Note: These figures show the optimal trajectories of the considered macroeconomic variables for $t \in [0, 50]$ and for two different initial conditions $k_i(0) < \tilde{k}$, $i = 1, 2$. From $t > 50$ both model economies have adjusted to the fixed point, i.e. all variables are constant, unless one or more structural parameters are varied.

Figure 2: Transition to Fixed Point for different $k(0) < \tilde{k}$



(a) capital (b) consumption (c) output

(d) investment (e) utility (f) shadow price $\mu(t)$

Note: These figures show the optimal impulse response of the considered macroeconomic variables due to an permanent increase of the time preference rate. The economy is adjusting to a new dynamic equilibrium when the shock is imposed at $t = 10$.

Figure 3: Optimal Impulse Response on Permanent Change in r

Eigendecomposition of a Matrix

We can establish following fixed point properties by setting (14) and (13) equal to zero

$$
\begin{aligned}
e^{(\alpha-1)\ln\tilde{k}} &= \frac{n+\delta+r}{\alpha} \\
e^{\ln\tilde{c}-\ln\tilde{k}} &= e^{(\alpha-1)\ln\tilde{k}} - (n+\delta) = \frac{n+\delta+r}{\alpha} - (n+\delta) \\
&= \frac{(1-\alpha)(n+\delta)+r}{\alpha} := \psi > 0
\end{aligned}
$$

(13) and (14) can be considered as a two dimensional differential equation system, where its linearization around the fixed point is provided by a first order Taylor series, i.e.

$$
\begin{bmatrix} \Pi \\ \Psi \end{bmatrix}_E \cong \begin{bmatrix} \Pi_{\ln k} & \Pi_{\ln c} \\ \Psi_{\ln k} & \Psi_{\ln c} \end{bmatrix}_{(\tilde{k},\tilde{c})} \begin{bmatrix} \ln k(t) - \ln\tilde{k} \\ \ln c(t) - \ln\tilde{c} \end{bmatrix}
$$

The partial derivatives at the fixed point are

$$
\begin{aligned}
\frac{\partial \Pi(\tilde{k},\tilde{c})}{\partial \ln k} &= (\alpha-1)e^{(\alpha-1)\ln \tilde{k}} + e^{\ln \tilde{c}-\ln \tilde{k}} \\
&= \frac{(\alpha-1)(n+\delta+r)+(1-\alpha)(n+\delta)+r}{\alpha} = r \\
\frac{\partial \Pi(\tilde{k},\tilde{c})}{\partial \ln c} &= -e^{\ln \tilde{c}-\ln \tilde{k}} = -\psi < 0 \\
\frac{\partial \Psi(\tilde{k},\tilde{c})}{\partial \ln k} &= \frac{\alpha(\alpha-1)}{\sigma}e^{(\alpha-1)\ln \tilde{k}} = \frac{(\alpha-1)(n+\delta+r)}{\sigma} =: \xi < 0 \\
\frac{\partial \Psi(\tilde{k},\tilde{c})}{\partial \ln c} &= 0
\end{aligned}
$$

Such that finally results

$$
\underbrace{\begin{bmatrix} \Pi \\ \Psi \end{bmatrix}}_{\dot{h(t)}} \cong \underbrace{\begin{bmatrix} r & -\psi \\ \xi & 0 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} \ln k(t) - \ln \tilde{k} \\ \ln c(t) - \ln \tilde{c} \end{bmatrix}}_{h(t)} \tag{16}
$$

With eigenvalues ω_1 und ω_2

$$
\omega_1,\ \omega_2 = \frac{r \pm \sqrt{r^2 - 4\psi\xi}}{2}.
$$

Since $\omega_1\omega_2 = \det(A) = \psi\xi < 0$ the equilibrium point is a saddlepoint. For later reasoning we note $\omega_1 > r > 0$ and $\omega_2 < 0$. Let $\mathbf{j}' = (j_1, j_2)'$ the eigenvector of ω_1 and $\mathbf{m}' = (m_1, m_2)'$ of ω_2 respectively, i.e.

$$
A\begin{bmatrix} j_1 \\ j_2 \end{bmatrix} = \omega_1 \begin{bmatrix} j_1 \\ j_2 \end{bmatrix} \qquad \text{and} \qquad A\begin{bmatrix} m_1 \\ m_2 \end{bmatrix} = \omega_2 \begin{bmatrix} m_1 \\ m_2 \end{bmatrix}.
$$

Normalizing $j_1 = m_1 = 1$ yields for j_2 and m_2

$$j_2 = \frac{\omega_1 - r}{-\psi} = \frac{\xi}{\omega_1} \qquad \text{and} \qquad m_2 = \frac{\omega_2 - r}{-\psi} = \frac{\xi}{\omega_2}.$$

Let G be a matrix which contains as its columns the eigenvecotrs of A, G^{-1} its inverse and D a diagonal matrix, which elements are the eigenvalues of A, i.e.

$$G = \begin{bmatrix} j_1 & m_1 \\ j_2 & m_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ \frac{\xi}{\omega_1} & \frac{\xi}{\omega_2} \end{bmatrix}$$

$$G^{-1} = \begin{bmatrix} j_1 & m_1 \\ j_2 & m_2 \end{bmatrix}^{-1} = \frac{1}{j_1 m_2 - j_2 m_1} \begin{bmatrix} m_2 & -m_1 \\ -j_2 & j_1 \end{bmatrix} = \frac{\omega_1 \omega_2}{\xi(\omega_1 - \omega_2)} \begin{bmatrix} \frac{\xi}{\omega_2} & -1 \\ -\frac{\xi}{\omega_1} & 1 \end{bmatrix}$$

$$D = \begin{bmatrix} \omega_1 & 0 \\ 0 & \omega_2 \end{bmatrix}$$

We now factorize A to $A = GDG^{-1}$

$$A = \frac{\omega_1 \omega_2}{\xi(\omega_1 - \omega_2)} \begin{bmatrix} 1 & 1 \\ \frac{\xi}{\omega_1} & \frac{\xi}{\omega_2} \end{bmatrix} \begin{bmatrix} \omega_1 & 0 \\ 0 & \omega_2 \end{bmatrix} \begin{bmatrix} \frac{\xi}{\omega_2} & -1 \\ -\frac{\xi}{\omega_1} & 1 \end{bmatrix}$$

The compact notation of (16) is

$$\dot{h(t)} \cong Ah(t). \tag{17}$$

Since D is the Jordan canonical form of A, (17) can be solved with regard to the matrix exponential and a given initial condition $h(0)$

$$h(t) \cong e^{At} h(0) = e^{GDG^{-1}t} h(0) = G e^{Dt} G^{-1} h(0)$$

that is

$$\begin{bmatrix} \ln k(t) - \ln \tilde{k} \\ \ln c(t) - \ln \tilde{c} \end{bmatrix} \cong \frac{\omega_1 \omega_2}{\xi(\omega_1 - \omega_2)} \begin{bmatrix} 1 & 1 \\ \frac{\xi}{\omega_1} & \frac{\xi}{\omega_2} \end{bmatrix} \begin{bmatrix} e^{\omega_1 t} & 0 \\ 0 & e^{\omega_2 t} \end{bmatrix} \begin{bmatrix} \frac{\xi}{\omega_2} & -1 \\ -\frac{\xi}{\omega_1} & 1 \end{bmatrix} \begin{bmatrix} \ln k(0) - \ln \tilde{k} \\ \ln c(0) - \ln \tilde{c} \end{bmatrix}$$

alternatively written as a system of equations

$$\begin{aligned} \ln k(t) - \ln \tilde{k} &= e^{\omega_1 t} b_{11} + e^{\omega_2 t} b_{12} \\ \ln c(t) - \ln \tilde{c} &= e^{\omega_1 t} b_{21} + e^{\omega_2 t} b_{22} \end{aligned} \tag{18}$$

with

$$\begin{aligned} b_{11} &= \frac{\omega_1}{\xi(\omega_1 - \omega_2)}[\xi(\ln k(0) - \ln \tilde{k}) - \omega_2(\ln c(0) - \ln \tilde{c})], \\ b_{12} &= -\frac{\omega_2}{\xi(\omega_1 - \omega_2)}[\xi(\ln k(0) - \ln \tilde{k}) - \omega_1(\ln c(0) - \ln \tilde{c})], \\ b_{21} &= \frac{1}{\omega_1 - \omega_2}[\xi(\ln k(0) - \ln \tilde{k}) - \omega_2(\ln c(0) - \ln \tilde{c})], \\ b_{22} &= -\frac{1}{\omega_1 - \omega_2}[\xi(\ln k(0) - \ln \tilde{k}) - \omega_1(\ln c(0) - \ln \tilde{c})]. \end{aligned}$$

The left hand side of (18) provides the deviation of the current value from the steady state. For an optimal choice of $c(0)$ the systems tends for $t \to \infty$ on the stable growth path towards its fixed point. By definition the deviation equals zero at the fixed point and hence stability requires

$$\lim_{t \to \infty} = \ln k(t) - \ln \tilde{k} = \ln c(t) - \ln \tilde{c} = 0$$

With $\omega_2 < 0$ we get

$$\lim_{t \to \infty} = e^{\omega_2 t} b_{12} = e^{\omega_2 t} b_{22} = 0.$$

And since $\omega_1 > 0$, we set $b_{11} = b_{21} = 0$

$$
\begin{aligned}
\xi(\ln k(0) - \ln \tilde{k}) &= \omega_2(\ln c(0) - \ln \tilde{c}) \\
\Rightarrow \ln c(0) &= \frac{\xi}{\omega_2}(\ln k(0) - \ln \tilde{k}) + \ln \tilde{c}.
\end{aligned} \tag{19}
$$

This equation determines the optimum initial consumption $c^*(0)$ for a given stock of capital. Substitute (19) in b_{12} and b_{22} yields

$$
\begin{aligned}
b_{12} &= -\frac{\omega_2}{\xi(\omega_1 - \omega_2)} \left\{ \xi(\ln k(0) - \ln \tilde{k}) - \omega_1 \left[\frac{\xi(\ln k(0) - \ln \tilde{k})}{\omega_2} \right] \right\} \\
&= \ln k(0) - \ln \tilde{k}, \\
b_{22} &= -\frac{1}{\omega_1 - \omega_2} \left\{ \xi(\ln k(0) - \ln \tilde{k}) - \omega_1 \left[\frac{\xi(\ln k(0) - \ln \tilde{k})}{\omega_2} \right] \right\} \\
&= \frac{\xi}{\omega_2}(\ln k(0) - \ln \tilde{k}).
\end{aligned}
$$

The solution of (18) for a determined choice of $c(0)$ by $k(0)$ is given by

$$
\begin{aligned}
\ln k(t) - \ln \tilde{k} &= e^{\omega_2 t}(\ln k(0) - \ln \tilde{k}) \\
\ln c(t) - \ln \tilde{c} &= e^{\omega_2 t}\frac{\xi}{\omega_2}(\ln k(0) - \ln \tilde{k})
\end{aligned}
$$

The stability condition for the initial consumption (19) is presumed to be binding, then for every instant of time the relation between consumption and capital equals

$$
\ln c(t) = \frac{\xi}{\omega_2}(\ln k(t) - \ln \tilde{k}) + \ln \tilde{c}, \quad \forall t \in [0, \infty).
$$

The Allocation of and the Returns to Talent: An Empirical Model

JProf. Dr. Michael Boehm *

Introduction

Inequality has returned with a vengeance to center-stage of the economic debate.[1] Not least since the great success of Thomas Piketty's book "Capital in the Twenty-First Century" (Piketty, 2014) do we know that income and wealth inequality in developed countries have increased massively since the 1970s. To be more precise, research conducted over the last decade shows that wages in the upper third of the earnings distribution have risen handsomely while top shares of income have sky-rocketed (Machin and Van Reenen, 2008; Alvaredo, Atkinson, Piketty, and Saez, 2013). At the same time, there was hardly any real wage growth at the middle of the income distribution and a even a decrease of wages at the bottom of the income distribution in some countries and time periods, for example in the US during the 1980s (Acemoglu and Autor, 2011).

There exist several views why these developments occurred as they did, ranging from increased rent-extraction by the privileged (Bivens and Mishel, 2013)

*Department of Economics, Institute for Applied Microeconomics, University of Bonn, Adenauerallee 24-42, D-53113 Bonn, Germany, e-mail: mboehm1@uni-bonn.de, phone: +49 228 73-9208 or +49 228 73-9238 (secretariat)

[1] This was already noted by Nico Pestel in the previous issue of this journal (Pestel, 2013). While Pestel focused on the role of household composition, marital sorting, and female labor supply for inequality in Germany, I will concentrate my article on a method for better understanding the changing returns' to skills role in increasing inequality.

to the relatively high returns on capital (Piketty, 2014) to "efficient" market forces governed by supply and demand. In fact, the prevailing view among many economists is still that rising inequality in the overall population as well as at the top reflect increasing returns to skill or talent (Kaplan and Rauh, 2013). That is, if new technologies and increasing market sizes due to globalization shift out the demand curve for skill, the rewards for the talented will rise compared to the rest of the population.

But what do we mean by talent? Just calling the high earners the most talented would not be helpful since changes in demand, supply and many other factors would be able to explain the rising inequality in this case. Therefore, researchers have long been using individuals' education as a proxy for skill or even talent.

However, there are some problems with equating education and talent. First, formal educational attainment has increased substantially over the decades. Since it is rather unlikely that "talent" in society has improved accordingly, educational attainment does not enable us to compare like with like over time. Second, educational attainment is a one-dimensional measure of skill while there is now ample evidence that individuals' productive attributes are multidimensional and that they include cognitive as well as non-cognitive dimensions (Heckman, Stixrud, and Urzua, 2006, for example)

Fortunately, much better data has recently become available that get us closer to observing multidimensional talents of workers. In particular, survey datasets such as the National Longitudinal Survey of Youth in the United States and the Socio-Economic Panel in Germany measure respondents in terms of several cognitive and non-cognitive attributes. Moreover, detailed social security data in the Nordic countries provide cognitive and non-cognitive scores from military enlistment tests together with workers' detailed employment histories in firms that they have worked for. Thus, we are now in a substantially better position than in the past to analyze the sorting of talents into occupations, industries, and firms as well as the returns that go in hand with it.

In this article, I will outline a theoretical framework that I developed in my Ph.D. dissertation and that can be used to examine such questions (Boehm, 2013). The framework is based on the Roy model (Roy, 1951) of occupational choice and it can readily be brought to the data, either via simple comparison of its comparative statics with the actual evidence or via direct estimation of key parameters.

Skill Demand

Suppose overall output in the economy is produced using inputs from different industries, firms, or occupations k.[2] For example, production could be determined by a CES function of the form

$$Y_t = A_t \left[\sum_{k=1}^{K} \alpha_{kt} \left(S_{kt} \right)^{\rho-1/\rho} \right]^{\rho/\rho-1}, \tag{1}$$

where S_{kit} is the amount of k-*specific* skill employed in job k, α_{kt} is the productivity of job k in contributing to final output, and ρ the elasticity of substitution between jobs.

Differentiating (1) gives the return to S_{kt} in a competitive economy

$$R_{kt} = MPS_{kt} = \alpha_{kt} A^{\rho/\rho-1} Y_t^{1/\rho} S_{kt}^{-1/\rho}. \tag{2}$$

The relative return

$$\frac{R_{kt}}{R_{\tilde{k}t}} = \frac{\alpha_{kt}}{\alpha_{\tilde{k}t}} \left(\frac{S_{kt}}{S_{\tilde{k}t}} \right)^{-1/\rho} \tag{3}$$

[2] For brevity, I refer to all of these as "jobs" from now on.

to two skills k and $\tilde{k}$ thus depends positively on the relative productivity $\frac{\alpha_{kt}}{\alpha_{\tilde{k}t}}$ of the jobs that they are employed in and negatively on the extent that they are employed in these jobs $\frac{S_{kt}}{S_{\tilde{k}t}}$.

It is widely believed that changes that the economy has experienced in recent decades constituted shifts in the productivity of occupations, industries, firms, or tasks. For example, skill-biased technical change (or globalization) in the production framework (1) would constitute a rise of α_{kt} in the more highly skilled occupations or industries such as professional services, finance, or IT. Alternatively, routine-biased technical change (or offshoring of production) would be the provision of routine jobs k, e.g. in manufacturing or clerical work, at a lower price than R_{kt} by computers or foreign workers.

The next section shows what will happen to employment in different jobs and the wages of workers with different skills and talents in such cases.

Skill Supply

Suppose each worker i takes the job $k \epsilon \{1, ...K\}$ that offers him the highest potential wage:

$$W_{it} = max\{W_{1it}, W_{2it}, ..., W_{Kit}\}. \tag{4}$$

These potential wages are composed of the product of i's skill to carry out work in job S_{Kit} and the wage rate that prevails for that work in point in time t (R_{Kt}).

The crucial point in this analysis is that workers are heterogeneous in terms of their skills. Some workers have more skill than the average person in almost all jobs (absolute advantage) while others may be particularly productive in some dimensions (relative advantage).

Where may such differences in workers' skills stem from? We could think of them as arising from differences in endowments of talents.[3] For example,

[3]There could also be investments that may depend on the endowments of talents and on the

workers differ in cognitive mathematical and verbal ability, in physical strength and agility, but also in non-cognitive traits such as persistence and motivation. All of these talents combine to make individuals more and less skilled in different jobs, that is $S_{kit} = f_k(math_{it}, strength_{it}, persistence_{it}, ...)$.

Thus, potential log wages (logged variables are denoted in lower-case letters) in job k may become:

$$w_{kit} = r_{kt} + s_{kit} = r_{kt} + \beta_k x_{it} + u_{kit}, \tag{5}$$

where the vector $x_{it} = [x_{1it}, ..., x_{jit}, ..., x_{Jit}]'$ contains the observed components of talents such as the cognitive, physical and non-cognitive abilities mentioned above and the β_{Kj}s are the corresponding linear projection coefficients. The regression error u_{kit} is the unobserved component of worker i's skill in firm k, i.e. talents that are unobserved by the econometrician and their corresponding β_ks.[4]

Supply's reaction to changing demand

What can we learn from this model about the allocation of and the returns to talents? Suppose for example that for technological reasons there is an increase in the productivity of high-skill sectors such as IT or professional services.[5] To keep it simple, consider only two sectors in the following with sector 2 being the high-skilled sector for which demand rises. The increase in the productivity of this sector implies that α_{2t} increases compared to α_{1t} in production function (1). From equation (3) this also implies that for given quantities of skill hired, the relative wage rate offered for working in the high-skill sector rises

prevailing or expected returns to skills. I abstract from these here.

[4]I am writing this in logs because then it becomes additive and it corresponds to log wage regressions that are usually run in empirical work.

[5]As an alternative, we could a analyze a drop in the demand for routine occupations.

$$\triangle(r_{2t} - r_{1t}) > 0. \tag{6}$$

From a labor supply perspective, this triggers two interrelated changes. First, consider figure 1 which plots the indifference line between working in sector 1 and sector 2 into the skill space (s_{2it}, s_{1it}). With $\triangle(r_{2t} - r_{1t}) > 0$ this line shifts to the bottom right so that more workers (the area **C**) now prefer working in sector 2. Employment in the high-skill sector will thus rise.[6]

The second effect that (3) has is that it will raise the wages of those workers who have *relatively* high skills in the sector 2. By extension, this means that the returns to talents that are important for producing skill s_{2it} will also increase compared to talents that are important for producing s_{1it}.

We can see this effect more directly by exploiting the information about the allocation of talents. Consider a worker i's change in wage under a marginal shift in wage rates

$$dw_{it} = dr_{1t} + H_{it}\, d(r_{2t} - r_{1t}), \tag{7}$$

where $H_{it} = \mathbf{1}\left(s_{2it} - s_{1it} > -(r_{2t} - r_{1t})\right)$ is an indicator for i prefering to work in the high-skill sector.[7] Thus, workers who have high relative skill in sector 2 $(s_{2it} - s_{1it})$ benefit from a marginal increase of the relative wage rate $(r_{2t} - r_{1t})$ in that sector.

This result also persists for discrete changes in the wage rates. Integrating (7) gives

[6]While they have *relatively* lower skills in the high-skill sector, it is not obvious that the new entrants **C** into sector 2 are on average less skilled for that sector. It depends on the exact distribution of skills for *both* sectors in the population and can thus not generally be answered. Therefore, it can also not unambiguously be answered whether relative average wages in sector 2 will rise. In fact, the conditions under which the distribution of skill improves or deteriorates across sectors receives a lot of attention in textbook treatments of the Roy (1951) model.

[7]Note that by the optimality of the initial sector choices and the envelope theorem, H_{it} doesn't change for marginal shifts in wage rates.

$$\triangle w_{it} = \triangle r_{1t} + \int_{\widetilde{r}_0}^{\widetilde{r}_1} H_{it} d\widetilde{r}_t, \tag{8}$$

where $\widetilde{r}_t \equiv r_{2t} - r_{1t}$. We see in equation (8) that workers who start out in sector 2 or, to a lesser extent, who switch from sector 1 to sector 2 early have higher wage increases than workers who stay in sector 1. This is because the former benefit (more) from the positive additional wage rate increase in sector 2 via the integral in equation (8).

The workers in sector 2 or who switch into sector 2 early are also the ones who possess talents to which the returns rise. For example, math ability is probably important for producing in the high-skill sector, that is β_{2math} is high. Thus, workers who are more talented in math will have a higher relative skill in sector 2 ($s_{2it} - s_{1it}$) and their relative wages will rise. Moreover, the returns to math talent will rise when productivity in the high-skill sector increases.[8]

Estimation

Having shown theoretically how wages, the allocation of talents, and the returns to talents are related, we can also estimate key parameters of this model. First, given the fact that we often have data in the form of repeated cross-sections and thus don't observe the exact same worker in different points in time, let us consider the same "type" of workers according to their observable talents. Conditioning on observable talents x_{it} and taking expectations on both sides of equation (8) we get

$$E(\triangle w_{it}|x_i) = \triangle r_{1t} + \int_{\widetilde{r}_0}^{\widetilde{r}_1} p_H(x_{it}, \widetilde{r}_t) d\widetilde{r}_t, \tag{9}$$

where $p_H(x_{it}, \widetilde{r}_t) = Pr\left(s_2(x_{it}, u_{it}) - s_1(x_{it}, u_{it}) > -(r_{2t} - r_{1t})\right)$ is the prob-

[8]My CEP discussion paper (Boehm, 2013) shows more formally that the returns to talents in a linear wage regression change according to the sector they are productive in.

ability to enter the high-skill sector as a function of the observable talents x_{it}, unobservables u_{it}, and relative wage rates $\widetilde{r}_t$. Linearly approximating the integral in equation (9) between points in time $t = 0$ and $t = 1$ by

$$p_H(x_{it}, \widetilde{r}_t) \approx p_H(x_{it}, \widetilde{r}_t) + \frac{p_H(x_{it}, \widetilde{r}_1) - p_H(x_{it}, \widetilde{r}_0)}{\widetilde{r}_1 - \widetilde{r}_0}(\widetilde{r}_t - \widetilde{r}_0), \tag{10}$$

directly yields an estimable equation:

$$E(\triangle w_{it}|x_i) = \triangle r_{1t} + \frac{p_H(x_{it}, \widetilde{r}_1) + p_H(x_{it}, \widetilde{r}_0)}{2}\triangle(r_{2t} - r_{1t}). \tag{11}$$

Now run first-stage regressions that measure the allocation of talents in each period, i.e. that estimate $p_H(x_{it}, \widetilde{r}_t)$ as a function of x_{it} in $t = 0$ and $t = 1$. For example, this could be done using probit or logit type sorting regressions. From these regressions we learn which talents are associated with relatively more skill in sector 2 compared to sector 1 and vice versa. In the second stage we can then estimate the shifting wage rates $(\triangle r_{1t}, \triangle r_{2t})$ across sectors by regressing the changes in wages on the predicted probabilities $\hat{p}_H(x_{it}, \widetilde{r}_t)$ from the first stage as in equation (11).

Thus, the procedure outlined in this section demonstrates that the allocation of talents $(p_H(x_{it}, \widetilde{r}_t))$ is inherently linked to changing wages $(E(\triangle w_{it}|x_i))$ when demand in the economy evolves. At the same time this changes the returns to talents themselves. Moreover, the quantities discussed here are empirically measurable. Hence, we can estimate the shifts in wage rates that are offered across jobs, the sorting of talents, and the returns to talents that are due to these shifts.[9]

[9] We can also test restrictions of the model such as the hypothesis that all changes in returns to talents are due to shifts in wage rates across jobs (Boehm, 2013).

Conclusion

In my Ph.D. dissertation I have applied a framework similar to the one outlined above in order to estimate the wage effects of job polarization and routine-biased technical change. I found that job polarization plays an important role in holding down middle-skill workers' wages since the 1980s and that it may have generated a substantial part of the change in the overall wage distribution that we observe during this period.

As already mentioned, the framework can also be used to examine other hypothesized changes of the demand for skill in the economy, such as a rise in IT- or finance-related services. Of course there are several alternative frameworks with which one can approach the important questions about talent, skill, productivity, and inequality facing economics today. What seems certain, however, is that the improved data from surveys and administrative sources that have recently become available offer very promising new avenues to answering such questions.

References

ACEMOGLU, D., AND D. AUTOR (2011): "Chapter 12 - Skills, Tasks and Technologies: Implications for Employment and Earnings," vol. 4, Part B of *Handbook of Labor Economics*, pp. 1043 – 1171. Elsevier.

ALVAREDO, F., A. B. ATKINSON, T. PIKETTY, AND E. SAEZ (2013): "The Top 1 Percent in International and Historical Perspective," *The Journal of Economic Perspectives*, pp. 3–20.

BIVENS, J., AND L. MISHEL (2013): "The pay of corporate executives and financial professionals as evidence of rents in top 1 percent incomes," *The Journal of Economic Perspectives*, 27(3), 57–77.

BOEHM, M. J. (2013): "Has Job Polarization Squeezed the Middle Class? Evidence from the Allocation of Talents," *Centre for Economic Performance Discussion Paper*, 1215.

Heckman, J. J., J. Stixrud, and S. Urzua (2006): "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior," *Journal of Labor Economics*, 24(3), 411–482.

Kaplan, S. N., and J. Rauh (2013): "It's the Market: The Broad-Based Rise in the Return to Top Talent," *The Journal of Economic Perspectives*, pp. 35–55.

Machin, S., and J. Van Reenen (2008): "wage inequality, changes in," in *The New Palgrave Dictionary of Economics*, ed. by S. N. Durlauf, and L. E. Blume. Palgrave Macmillan, Basingstoke.

Pestel, N. (2013): "Economic Inequality in Germany and the Role of Household Context," *The Bonn Journal of Economics*, II(2), 91.

Piketty, T. (2014): *Capital in the Twenty-First Century*. Harvard University Press.

Roy, A. D. (1951): "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, 3(2), 135–146.

Appendix

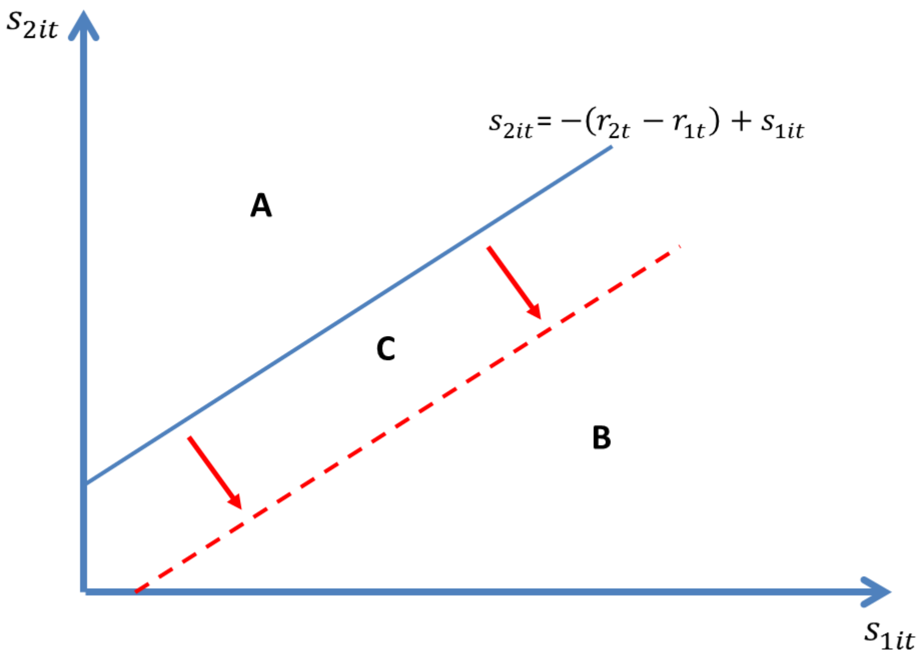


Figure 1: Skill selection into the two sectors.

Econometric tests for speculative bubbles

Prof. Dr. Jörg Breitung *

Introduction

Speculative bubbles have a long history in stock, commodity or real estate markets. Famous historical examples include the Dutch Tulipmania (1634-1637) and the Mississippi Bubble (1719–1720). Such bubbles may lead to severe economic crises such as the speculative excess on stock prices prior to the great depression from 1930–1933 or the recent financial crisis of 2007–2009 that was preceded by the US housing bubble.There is a broad consensus that bubbles are characterized by an explosive path of the underlying market prices, whereas in "normal times", speculative prices are well approximated by a random walk process. A standard economic model to motivate the occurrence and persistence of speculative bubbles is the framework of rational bubbles (see, e.g., Blanchard and Watson (1982) and Flood and Hodrick (1990)). In such models it is economically rational to invest in an obviously overpriced asset as long as the investor asserts that the price continues to rise exponentially.

The theory of rational bubbles commonly starts from a simple present value model for asset prices (see, e.g. Campbell, Lo, and MacKinlay (1997)). If in-

*Faculty of Management, Economics and Social Sciences , Institute for Econometrics and Statistics, University of Cologne, Meister-Ekkehart-Str. 9 Building: 113, D-50923 Cologne, Germany, e-mail: breitung@statistik.uni-koeln.de, phone: +49-221-470-4266

vestors are not averse to risk, all assets would have the same constant expected real return R in equilibrium that is

$$1 + R = E_t \left(\frac{P_{t+1} + D_{t+1}}{P_t} \right), \tag{1}$$

where $E_t(\cdot)$ denotes the expected value given the information in period t. This difference equation can be solved by forward iteration yielding

$$P_t = P_t^f + B_t, \tag{2}$$

$$\text{with } P_t^f = E_t(P_{t+1}^*) \equiv \sum_{i=1}^{\infty} \left(\frac{1}{1+R} \right)^i E_t\left(D_{t+i}\right). \tag{3}$$

$$\text{and } E_t\left(B_{t+1}\right) = (1+R)B_t. \tag{4}$$

The bubble component B_t arises from the fact that the solution of a difference equation is not unique. The component P_t^f is often referred to as the fundamental (stock) price which equals the expected present value of the future dividend payments, whereas equation (4) is a no arbitrage condition for the so-called bubble component of the stock price. If a bubble is present in the stock price, (4) requires that a rational investor, who is willing to buy that stock, must expect the bubble to grow at a rate equal to R. Excluding negative stock prices one can infer that $B_t \geq 0$ for all t. Whenever $B_t > 0$, a rational investor is willing to buy an "overpriced" stock, since (s)he believes that through price increases (s)he will be sufficiently compensated for the extra payment due to the bubble component (B_t). If the bubble component is a large part of the price, then the expectation that it will increase at rate R means that investors expect price increases that have nothing to do with changes in fundamentals. If enough investors have this expectation and buy shares, the stock price will indeed go up and complete a loop of a self-fulfilling prophecy. Obviously, it does not make sense to assume a bubble to continue infinitely long, since in this case the price of the asset will tend to infinity at an exponential rate. Brock (1982) analyzed the maximization problem of a competitive, representative, infinitely-lived investor and obtained a terminal condition (known as a transversality condition) that allows to exclude rational bubbles. Another theoretical challenge against rational bubbles is provided by Diba and Grossman (1988) who note that bubbles in real stock prices can never be negative. Since the fundamental value with nonzero dividend payment must grow with a lower rate, a negative bubble would imply that stock prices eventually become negative within a finite time horizon. Hence, negative bubbles are inconsistent with rational expectations and an infinite investment horizon but we cannot rule out (rational) bubbles when speculative investors focus on a limited time horizon. The simplest example of a process that satisfies (4) is the deterministic bubble, given by $B_t = (1+R)^t B_0$, where B_0 is an initial value. A somewhat more realistic example, in which the bubble does not necessarily grow forever was suggested by Blanchard and Watson (1982). Their bubble process is given by

$$B_{t+1} = \begin{cases} \pi^{-1}(1+R)B_t + \mu_{t+1}, & \text{with probability } \pi \\ \mu_{t+1}, & \text{with probability } 1-\pi \end{cases} \tag{5}$$

where $\{\mu_t\}_{t=1}^{\infty}$ is an independent and identically distributed (iid) sequence of random variables with zero mean. In each period, the bubble generated as in equation (5) will continue with probability π, or collapse with probability $1-\pi$. As long as the bubble does not collapse, the realized bubble return exceeds the interest rate R as a compensation for the risk of financial losses during a collapse of the bubble.

Time series properties of stock prices and dividends

Starting with Bachelier (1900) it is typically assumed that stock prices can be well approximated by a martingale process of the form

$$E_t(P_{t+1}^f) \;=\; P_t^f .$$

There are two possible rationales to motivate this assumption. First, assuming that D_t is a random walk with drift

$$D_{t+1} = \mu + D_t + \varepsilon_{t+1}$$

it follows that

$$P_t^f \;=\; \frac{(1+R)\mu}{R^2} + \frac{1}{R} D_t \ .$$

Thus, the fundamental value of the stock is also characterized by a random walk with drift. If $\mu \approx 0$, then $E_t(P_{t+1}^f - P_t^f) \approx E(\epsilon_{t+1})/R = 0$. Although empirical studies typically find that dividends are well approximated by a random walk, there is no convincing reason why this should always be the case.

In many applications the sampling frequency of stock prices is higher than the dividend period. Assume for example that dividends are payed out annually, whereas stock prices form a monthly series. Accordingly, the dividend series D_t is zero in eleven out of 12 months. During these 11 months the "perfect forecast present value" of future dividends $P_{t+1}^* = \sum_{i=1}^{\infty} \left(\frac{1}{1+R}\right)^i D_{t+i}$ does not change during periods without dividend payment (i.e. $P_{t+1}^* = P_t^*$). Therefore, the change in stock price during periods without dividend payment is solely due to the update of expectations about the future stream of dividends, that is,

$$P_t^f - P_{t-1}^f = E_t(P_{t+1}^*) - E_{t-1}(P_{t+1}^*).$$

If it is assumed that investors' expectations are rational, it follows that updates of expectations from period t to $t+1$ form a martingale difference sequence with

$$E_t(P_{t+1}^f - P_t^f) = 0$$

and, thus, P_t^f can be represented by a random walk. In time periods with dividend payment the change of the fundamental value results as

$$P_t^f - P_{t-1}^f = E_t(P_{t+1}^*) - E_{t-1}(P_{t+1}^*) + D_t \ .$$

Accordingly, the dividend adjusted price series $P_t - D_t$ (where D_t is zero for periods without dividend payment) can be represented by a random walk.[1]

Diba and Grossman (1988) and Campbell and Shiller (1987) argue that in the absence of bubbles, prices and dividends are cointegrated, that is, prices and dividends are driven by a common stochastic trend. This can be seen by noting that

$$\begin{aligned} S_t \equiv P_t^f - \frac{1}{R} D_t &= E_t(S_{t+1}^*) \\ \text{with} \quad S_{t+1}^* &= \sum_{i=1}^{\infty} \left(\frac{1}{1+R} \right)^i \Delta D_{t+i}. \end{aligned}$$

If ΔD_t is stationary and $R > 0$, then S_t^* is stationary as well. Since the (rational expectation) forecast of a stationary variable must also be stationary, it follows that under the assumptions of the present value model for stock prices $S_t = P_t^f -$

[1] Note that on exchange markets the stock price immediately adjusts for dividends so that the $P_t - D_t$ represents the *ex dividend* price of the share.

$R^{-1}D_t$ is a stationary linear combination of stock prices and dividends. As shown by Campbell and Shiller (1987) the model implies that also ΔP_t^f is a stationary process and, thus, the vector (P_t^f, D_t) is cointegrated in the terminology of Engle and Granger (1987).

Diba and Grossman (1988) apply unit root tests to prices and dividends from the S&P Composite Stock Price Index from 1871 to 1986 and find that prices and dividends are both $I(1)$. Furthermore, cointegration tests suggest that stock prices and dividends are indeed cointegrated implying no evidence for an explosive bubble component in stock prices.

Backward-looking tests for a structural break

A natural empirical approach to identify speculative bubbles is to apply statistical tests for a structural change from a random walk to an explosive regime. Such tests were originally proposed by Phillips, Wu, and Yu (2011) (henceforth: PWY) and further developed by Phillips and Yu (2011), Homm and Breitung (2012), and Phillips, Shi, and Yu (2013).

The simplest version of the test is based on a first order autoregressive process

$$y_t = \rho y_{t-1} + \epsilon_t \ , \tag{6}$$

where $E(\epsilon_t) = 0$ and $E(\epsilon_t^2)$ for $t = 1, \ldots, T$. The null hypothesis is that y_t follows a simple random walk, i.e.

$$H_0: \quad \rho = 1. \tag{7}$$

Under the alternative hypothesis y_t starts as a random walk but changes at time

$T^* = \tau^* T$ into an explosive regime:

$$H_1: \quad y_t = \begin{cases} y_{t-1} + \epsilon_t & , \text{ for } t = 1, \ldots, \tau^* T \\ \rho y_{t-1} + \epsilon_t & , \text{ for } t = \tau^* T + 1, \ldots, T \quad \text{with } \rho > 1 \end{cases} . \tag{8}$$

Assume for the moment that the true break fraction (or relative break date) $\tau^* = T^*/T$ is known. Then it is straightforward to test the hypothesis of a structural break at period T by a Chow test for the parameter $\phi = \rho - 1$ which is zero before T^* and positive after T^*. This gives rise to the regression function

$$\Delta y_t \;=\; \delta\, y_{t-1} D_t(\tau^*) + \varepsilon_t, \tag{9}$$

where $D_t(\tau^*)$ is a dummy variable which is zero for $t < \tau^* T$ and changes to one for $t \geq \tau^* T$. Correspondingly, the null hypothesis of interest $H_0 : \delta = 0$ is tested against the (one-sided) alternative $H_1 : \delta > 0$ by using an ordinary t-statistic:

$$\mathrm{DFC}_{\tau^*} = \frac{\sum\limits_{t=\tau^* T+1}^{T} \Delta y_t y_{t-1}}{\widetilde{\sigma}_{\tau^*} \sqrt{\sum\limits_{t=\tau^* T1}^{T} y_{t-1}^2}} ,$$

where

$$\widetilde{\sigma}^2_{\tau^*} = \frac{1}{T-2} \sum_{t=2}^{T} \left(\Delta y_t - \widehat{\delta}_{\tau^*} y_{t-1} D_t(\tau^*) \right)^2$$

and $\widehat{\delta}_{\tau^*}$ denotes the OLS estimator of δ in (9). This test is labeled as Chow-type Dickey-Fuller (DFC) test in Homm and Breitung (2012).

An important problem is that in practice the absolute (T^*) or relative (τ^*) break date is usually unknown. Following Andrews (1993) a test with unknown break date can be constructed by trying out all possible break dates τ^* within a reasonable interval $\tau^* \in [\tau_0, 1 - \tau_0]$, where τ_0 is often set to 0.10 or 0.15. The reason for leaving out break dates at the beginning and end of the example is that the regression (9) performs poorly if the number of observations in one of the two regimes is too small. The test statistic for an unknown break date is the maximum of all $T(1 - 2\tau_0)$ DFC statistics, labelled as supDFC. The limiting distribution of this test statistic can be expressed as a supremum of a function of Brownian motions (cf. Homm and Breitung (2012)).

It is important to note that under the alternative it is assumed that the bubble runs up to the end of the sample, that is, the price series does not switch back from the bubble regime into the random walk regime. As indicated by the some Monte Carlo simulations presented in Homm and Breitung (2012) the power of the test may deteriorate dramatically if the sample includes the collapse of the bubble. Phillips, Shi, and Yu (2013) suggest a search strategy for three and more regimes which is computationally demanding and may imply a severe loss in power. A simpler solution is to adapt the sample such that it focuses on the upswing of stock prices only. Obviously this strategy involves the risk of data mining which may bias the test result towards finding a bubble.

Real time monitoring

Classical test procedures like the sequential DFC statistic are designed to detect speculative bubbles within a given historical data set. Accordingly, these tests can be employed to answer the question whether a bubble occurs in a particular time span (e.g. the dot.com bubble in the late 1990s). From a practical point of view it is often more interesting to analyze whether some asset class is *currently* characterized by a speculative bubble. To this end a sequence of statistics is constructed that summarizes the accumulated evidence for a bubble at each point in time. If the statistic exceeds some threshold, we are able to conclude that with a prespecified probability the price series has entered an explosive regime. An important advantage of this approach is that the test statistic is not affected by a later collapse of the bubble.

Assume that, when the monitoring starts, a training sample of n observations is available and that the null hypothesis of no structural break holds for the training sample. Then, in each period $n+1, n+2, \dots,$ a new observation arrives. Following Chu, Stinchcombe, and White (1996) we consider two different statistics (detectors):

$$\text{CUSUM:} \qquad S_r = \frac{1}{\widehat{\sigma}_r} \sum_{j=n+1}^{r} (y_j - y_{j-1}) = \frac{1}{\widehat{\sigma}_r}(y_r - y_n) \quad (r > n) \tag{10}$$

$$\text{FLUC:} \qquad \mathrm{DF}_r = \frac{1}{n+r}(\widehat{\rho}_r - 1)/\widehat{\sigma}_{\rho_r} \tag{11}$$

where $\widehat{\sigma}_r^2$ is some consistent estimator of the residual variance, $\widehat{\rho}_r$ denotes the OLS estimate of the autoregressive coefficient (including a constant) based on the subsample $1, \dots, n, n+1, \dots, r$ and $\widehat{\sigma}_{\rho_r}$ denotes the associated standard error. Assume that we enter the bubble regime at some period $T^* > n$. Then for $r > T^*$ the estimate $\widehat{\rho}_r$ tends to become large and eventually exceed the critical level c_α. Similarly, within the bubble regime the $r - n$ step ahead forecast error $y_r - y_n$ based on a random walk forecast tends to large positive values. Phillips, Wu, and Yu (2011) derived a fixed critical value for the FLUC detector, which implies that at a significance level of 0.05 a bubble is detected if $\mathrm{DF}_r > 1.468$. For the CUSUM detector Chu, Stinchcombe, and White (1996) developed a time dependent critical value $c_\alpha(r) = \sqrt{b_\alpha + \log(r/n)}$, where for the one-sided bubble test Homm and Breitung (2012) obtained $b_{0.05} = 4.6$. In their Monte Carlo simulations, Homm and Breitung (2012) found that the FLUC detector tends to outperform the CUSUM detector.

Break Date Estimation

An important practical problem is to date stamp the bubble start. The sequential Dickey-Fuller test (which is equivalent to the FLUC monitoring approach) proposed by Phillips, Wu, and Yu (2011) can straightforwardly be used as an estimation procedure for the bubble start. The estimate for the starting date of the bubble is the smallest value r such that DF_r is greater than the right-hand critical value of the asymptotic distribution of the standard Dickey-Fuller t-statistic. In order to obtain a consistent estimator of the break date the critical value is specified as a function of the sample size T. Specifically, Phillips, Wu, and Yu (2011) proposed to apply time dependent critical values for estimating the emergence and crash of the bubble computed as $c_r = \log[\log(r)]/100$ for $r = n+1, \ldots, T$. Homm and Breitung (2012) pointed out that this approach involves a substantial delay, which can be avoided by using the ML estimator proposed by Bai (1994). The ML estimator results in minimizing the sum of squared residuals which is equivalent with the maximum of the DFC statistic. In order to date stamp the end of a bubble episode Breitung and Kruse (2013) consider tests for a change from a bubble regime into a random walk regime. Accordingly such test procedures reverse the null hypothesis and alternative of the Chow test considered in Section . To this end we run the usual Chow for the null hypothesis $H_0 : \rho_t \geq 1$ versus the alternative $H_1 : \rho_t = 1$ by running the regression

$$y_t = \varrho y_{t-1} + \phi y_{t-1} D_t(\tau) + e_t \quad t = 2, 3, \ldots, T, \tag{12}$$

where $D_t(\tau)$ is defined as in Section (). Unfortunately, under the null hypothesis of an explosive process the t-statistic for the hypothesis $\phi = 0$ possesses a degenerate limiting distribution as it converges to zero in probability. To sidestep this difficulty Breitung and Kruse (2013) develop a simple modification that yields a test statistic with a standard normal limiting distribution. This modified test statistic allows us to test whether the autoregressive coefficient switches to a lower value within some range of possible break dates in the interval $\tau \in [\tau_0, 1 - \tau_0]$. Again, the break date τ^* that minimizes the residual sum of squares in the (12) is the ML estimator for the end date of the bubble. This estimator avoids the build-in time lag of the date stamping procedure proposed by Phillips, Wu, and Yu (2011) and Phillips, Shi, and Yu (2013).

An empirical example

In this section I illustrate the econometric methods for analyzing the well known "dot.com bubble" in the late 1990s.[2] A visual inspection of the (real) monthly NASDAQ Composite price index and dividends already suggest that the accelerating upswing starting with 1995 cannot be explained by fundamental factors alone (see Figure 1). Indeed, as shown in Figure 2, the sequential Dickey-Fuller statistic (the FLUC detector for the monitoring procedure) rejects the null hypothesis of no speculative bubble in the NASDAQ index from June 1995 up to May 2001. Thus according to Phillips, Wu, and Yu (2011) the bubble runs for years. Applying the sequential Chow statistic considered in Section we obtain the test statistic *supDFC*= 3.0467, which is significant according to the 0.05 critical value of 1.296. In contrast, the *supDFC* statistic is insignificant for the dividend series. Furthermore, the maximum of the DFC statistic is obtained for February 1995, four months before the sequence of Dickey-Fuller statistic becomes significant. This demonstrates the delay of the latter approach when date stamping the bubble.

Applying the Chow test for a collapse of the bubble we are able to reject the hypothesis that the bubble continuous at March 2000, more than one year before

[2]The empirical results presented in this section are taken from Homm and Breitung (2012) and Breitung and Kruse (2013).

the sequential Chow statistic becomes insignificant. Note that during these 14 months the price index looses 50 percent of its value! This demonstrates that for a practical implementation, a timely signal for the collapse of the bubble is essential. An investor who was perfectly riding the bubble in the NASDAC Composite index receives an annual return of 28 percent per year (on top of the dividends).

Conclusion

In recent years various econometric tools for analyzing speculative bubbles were developed. Among these methods, tests for a switch from a fundamental to a bubble regime (and vice verse) are particularly promising. These tests are easy to apply and have been shown to provide reliable and powerful evidence for or against a bubble. An obvious drawback of these tests is that they do not incorporate information on the fundamental value of the price series but merely summarize the evidence for an explosive path of the price series. By analyzing the difference between the actual and fundamental prices we should be able to design more powerful tests for persistent and accelerating mispricing of the asset which is due to a speculative excess.

References

Andrews, D. W. K. (1993): "Tests for parameter instability and structural change with unknown change point," *Econometrica*, 61, 821 – 856.

Bachelier, L. (1900): *Théorie de la Spéculation, Annales Scientifiques de l'Ecole Normale Supérieure, 3rd. Ser. 17, 21-88.* Translated in: The Random Character of Stock Market Prices, edited by Paul Cootner (1964), Cambridge.

Bai, J. (1994): "Least squares estimation of a shift in linear processes," *Journal of Time Series Analysis*, 15, 453–472.

Blanchard, O., and M. Watson (1982): *Bubbles, Rational Expectations and Financial Markets*pp. 295–315. In P. Wachtel (ed.), Crisis in the Economic and Financial Structure, Lexington: Lexington Books.

Breitung, J., and R. Kruse (2013): "When Bubbles Burst: Econometric Tests Based on Structural Breaks.," *Statistical Papers*, 54, 911 – 930.

Brock, W. (1982): "Asset Prices in a Production Economy," In McCall, J., (ed.), The Economics of Information and Uncertainty. Chicago University Press, 1-43.

Campbell, J., A. Lo, and A. MacKinlay (1997): *The econometrics of financial markets.* Princeton Univ. Press, Princeton, NJ.

Campbell, J., and R. Shiller (1987): "Cointegration and Tests of Present Value Models," *Journal of Political Economy*, 95, 1062 – 1088.

Chu, C.-S. J., M. Stinchcombe, and H. White (1996): "Monitoring Structural Change.," *Econometrica*, 64, 1045 – 1065.

Diba, B., and H. Grossman (1988): "Explosive rational bubbles in stock prices?.," *American Economic Review*, 78, 520 – 530.

Engle, R., and C. Granger (1987): "Co-integration and Error Correction: Representation, Estimation, and Testing.," *Econometrica*, 55, 251 – 276.

Flood, R., and R. Hodrick (1990): "On Testing for Speculative Bubbles.," *Journal of Economic Perspectives*, 4, 85 – 101.

Homm, U., and J. Breitung (2012): "Testing for Speculative Bubbles in Stock Markets: A Comparison of Alternative Methods," *Journal of Financial Econometrics*, 10(1), 198 – 231.

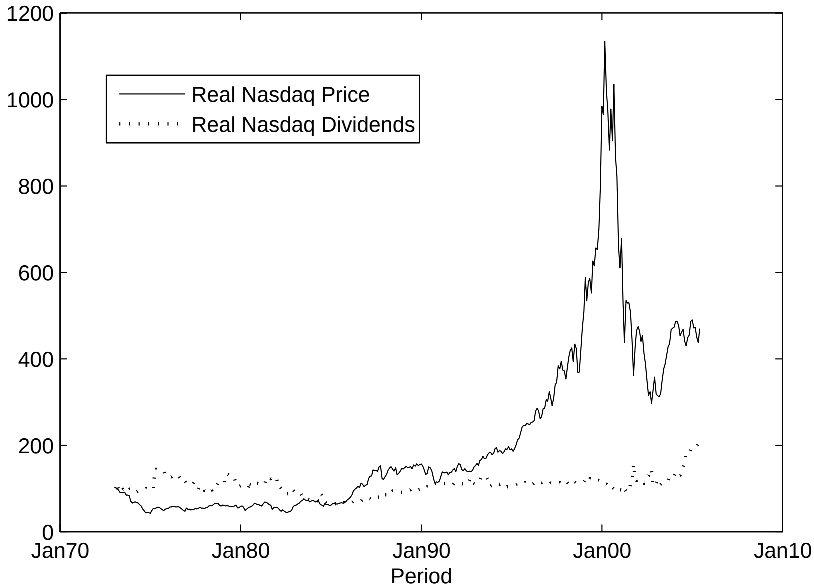
Phillips, P., S.-P. Shi, and J. Yu (2013): "Testing for Multiple Bubbles, Working Paper, Research Collection School of Economics. Paper 1302," .

Phillips, P., Y. Wu, and J. Yu (2011): "Explosive Behavior in the 1990s Nasdaq: When Did Exuberance Escalate Asset Values?.," *International Economic Review*, 52, 201 – 226.

Phillips, P., and J. Yu (2011): "Dating the Timeline of Financial Bubbles During the Subprime Crisis.," *Quantitative Economics*, 2, 455 – 491.

Appendix

Figures



Note: Both series were normalized to 100 at the beginning of the period.

Figure 1: Real Nasdaq Price and Dividends (normalized)

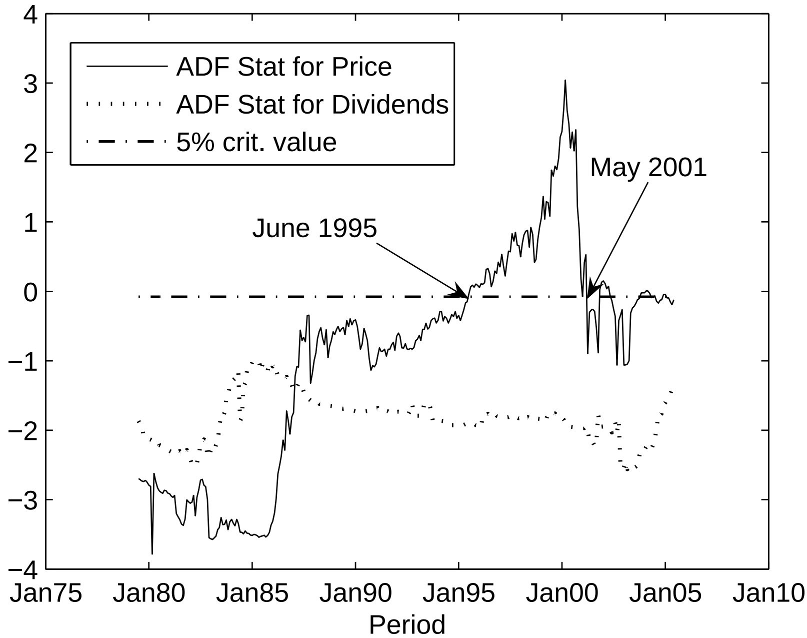


Figure 2: DF_r statistic for real Nasdaq price index and dividends

The Bonn Journal of Economics