Abstract

Title: Three Studies in Industrial Economics: Competition and Industry Structure

Author: Jan Keil

Chapter 1 reviews alternative theories of competition - the standard Neoclassical view, the contribution of the Chicago School as well as the two dynamic lines of thought which are part of Austrian economics and Classical Political Economy. The latter is presented as a consistent alternative to the other existing theories. Of special interest is the question if and how industry structure matters in these approaches, how profitability differentials are explained and what role market share concentration and mobility barriers play. Their predictions and implications for empirical research are compared. Ways to test and evaluate these different approaches are described.

Chapter 2 investigates econometrically how industry and micro level variables determine persistent differentials in the rate of return on assets in the U.S. The analysis is the first to use business segment data to explain long term profitability differentials. It presents new market concentration indicators that are superior to concentration ratios and allow to analyze an unpreceded amount of concentration and other data back to 1977. Critical concentration levels, non-linearities, interaction effects and previously ignored important control variables like industrial unionization are being considered. Concentration is found to have significant negative effects on profitability differentials. Barrier indicators are insignificant while market shares are positively correlated with long-run profitability. Concentration thus increases, not diminishes the degree of industrial competition. This is interpreted as evidence in support of Classical Political Economic competition theory.

Chapter 3 presents a costs of production based industry analytical study that aims at consistency with Classical Political Economic thought. It investigates how growth of renewable electricity in Germany forces conventional power plants to shift towards more more flexible operating regimes. The simulation of individual power plant load uses different current and future as well as alternative price and energy policy scenarios, four years of 15-minute interval data on system and renewable load as well as an unpreceded degree of detail on plant cost structures and technical characteristics. I find that the costs of electricity generation of cleaner, flexible thermal plants are positively effected by the transition. The competitiveness of inflexible baseload plants falls as they become more expensive than renewables. Lignite and nuclear power turns out to be unsuited to supplement renewable energy: a future exit reduces the average costs of electricity generation from conventional plants.

Three Studies in Industrial Economics:

Competition and Industry Structure

by

Jan Keil

May 2013

Submitted to The New School for Social Research of The New School in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Dissertation Committee:
Dr. Anwar Shaikh (chair)
Dr. Willi Semmler
Dr. William Milberg

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

For my parents.

For their love and support and for everything.

Acknowledgments

For comments and suggestions on my research as well as motivation and inspiration to conduct it, I am indebted especially to my dissertation committee chair, Anwar Shaikh as well as to the committee members William Milberg and Willi Semmler. For discussions, critique and support of the work that became part of my dissertation and/or for additional intellectual inspiration and motivation along the long path towards my doctoral degree I am also thankful to Josef Auer, Stephan Klasen, Christian Proano, Theresa Ghilarducci, Klaus and Johanna Heilos-Keil, Josef Esser, Bertram Schefold, John Kannankulam, Gregor Semieniuk, Siavash Radpour, Christian Schoder as well as the faculty and graduate student body at the Economics Department of the New School.

Contents

List of Tables

List of Figures

Chapter 1

Theories of Competition

This chapter reviews alternative theories of competition. The study works out differences in underlying assumptions, lists the most important theorems and predictions that can be derived from them and describes ways to test and evaluate these different approaches. A core focus lies on the questions how profit rate differentials are explained and how industry structure matters in these approaches. The impacts of market share concentration and of mobility barriers on competition and profitability differentials are of special interest. I am reviewing the standard Neoclassical view, the contribution of the Chicago School as well as the two dynamic lines of thought which are part of Austrian economics and Classical Political Economy.

Section 1 describes the static approaches, section 2 the dynamic ones and section 3 compares their predictions and implications for empirical research.

1.1 Static Competition

Neoclassical mainstream models of competition, oligopoly and monopoly rely on very minimalistic, rigid standard microeconomic or game theoretic frameworks and are usually mathematically highly formalized. These common baseline models of competition have several distinctive features. They are based on a perfectionist framework where the two models of perfect competition and of a perfect monopoly are held as the reference points against which deviations from the perfectionist states are compared.

Competition is thought of working instantaneously in such a way that it can be understood as a fixed state that can really exist at any given moment. Real, physical time is not part of the models which wholly abstract from time assuming simultaneous decision making or which in the case of game theory rely on fictional game time where players make their moves according to a certain sequence.

1.1.1 Immediate Competition

The basic starting point of most mainstream industrial organization models is the assumption of static oligopolistic competition - the analysis of a limited number of selling firms that exist in an isolated industry. Actors in this *isolated competition* within the industry generally differ from each other by defining different action spaces for the firms, different behaviors and expectations regarding their competitors' actions, different game rules and different degrees of heterogeneity among firms. Examples of such models are the ones originated by Cournot, Bertrand, Edgeworth, Stackelberg, dominant firm and coordination models. The outcomes lie always somewhere between the perfectly competitive outcome and the one of a monopoly output-price combination. Raising prices over marginal costs of production allows to obtain a certain amount of sustained excess profit on the firm and/or industrial level. In the strategic management business literature, Porter (1980) described this dimension where existing rival firms interact within an industry as one of the five forces of competition.

Prices can be raised above marginal costs when rivalry is undermined by collusion. The major variable that drives the possibility to collude and coordinate and that thus determines the degree of market power is the number of firms in the industry. In most cases, a rise in market share and/or concentration automatically leads to higher monopoly profits in the respective models (Cournot, Edgeworth models), unless a price war competition as in the Bertrand model is being pursued by firms. In game theoretic or coordinated behavior models, oligopolistic coordination through formal explicit or purely tacit collusion usually becomes easier and more likely when fewer firms are in the market. Due to these positive links between concentration and collusion as well as between coordination and excess profitability, these views are known as the *concentration doctrine*. The more imperfections - like firm or product heterogeneity, information incompleteness, non-marginality or instability of demand - are being introduced the more complicated collusion tends to become (Clarke, 1985, p. 39ff).

Some authors have refined the concentration doctrine and proposed that there is likely to be discontinuity and non-linearity in the concentration profitability relation. This is usually expressed in terms of *critical concentration levels* beyond which a positive effect of concentration on the ability to collude and to obtain excess profitability can be expected while there might be no effect below the value (see. e.g. Chamberlin (1929), Bain (1951, p. 323), Scherer (1980, p. 280), Ratnayake (1996)).

Another important augmentation to the basic model embeds the analyzed industry into an environment of other industries or the economy as a whole. Porter (1980) describes the bargaining power of suppliers and of buyers as the second and the third force of competition. The major determinant of the strength of suppliers and buyers is the size of the firms in these industries and the degree of market share concentration within these industries. This extension is still a static one, but embeds industry internal competition into the extended context of the environment of other industries. Correspondingly, one can speak of *embedded competition* and an *augmented concentration doctrine* according to which industrial concentration in an industry relative to the concentration up- and downstream of the value chains (it is placed into) is positively related to excess profitability. This implies that given everything else being equal, more excess profitability can be expected if an industry with a given degree of concentration supplies to and buys from sectors that are less concentrated.

1.1.2 Potential Competition / Threat of Industry Entry

A fundamental extension to standard models is to consider the competition that derives from the potential entry of outsiders into an industry (through backward or forward integration, from unrelated industries or through the formation of completely new companies). Porter (1980) described this potential which is opposed to actual competition as the fourth force of competition. The work of Bain (1956) has been generally considered to be seminal. It highlighted this idea and related the ability to set prices above unit costs to the existence of *entry barriers* (see also Clarke (1985, p. 71ff) and Porter (1980, p. 7ff)). Such barriers might be government-based, legal restrictions on the freedom of entry as well as cost advantages for established firms. The latter can be divided into scale - which requires obtaining a certain market share threshold or establishment size (minimum efficient scale), large capital volume requirements or a high (fixed) capital intensity; product differentiation - through advertising or research and development investments that result in product patents consumer loyalty, reputation; absolute cost advantages - though superior production technology that has been acquired over time and might be protected with patents, access to inputs of a superior quality and lower price. In addition to these objective conditions that form entry barriers, there are also 'endogenous entry barriers'. Potential entrants will form strategic considerations on how their entry will effect intra-industrial competition ex post (Caves and Porter, 1977). These expectations with regard to the behavior of existing insider firms can prevent entry when there is e.g. a history of predatory price competition in reaction to entrants, a high level of commitment to the industry (strategic role of a business unit, irreversible investments/sunk costs) or a significant fight back potential of incumbents (like a large excess capacity, financial reserves in the form of accumulated cash or open credit lines) (Porter (1980, p. 14), Clarke (1985, p. 87ff)).[1]

The *contestable market hypothesis* of Baumol (1982) represents a contribution to the barrier of mobility theories. He emphasized the decisiveness of net costs of entry, or of *barriers to exit* - the existence of sunk costs that represent a loss risk for potential entrants. If initial investments required for entry are fully recoverable and if entrants are able to produce under the same cost structure as incumbents, this threat of hit-and-run competition is enough to enforce marginal cost pricing, even in the case of a pure duopoly (Baumol (1982, p. 2); Clarke (1985, pp. 94-95)). On the other hand, the existence of exit barriers is expected to allow firms to price in a way that they can obtain monopoly profits in the long run. However, profits are limited by the net entrance costs faced by potential competitors. A practical problem for econometrical analyses is that irreversible investments are positively correlated with barriers to entry such as high capital requirements (Clarke, 1985, p. 95).

Barrier considerations generally imply that barriers to mobility can lead to excess profits. Viewed as an extension to the static models, the logic of the argument implies that concentration has mainly or even exclusively a significant positive effect on profitability when barriers are persistent. The conditions for surplus profits are the existence of few competitors that allow firms to coordinate as well as the occurrence of barriers to mobility which allow industry incumbents to capture these surpluses without attracting entrants that would undermine the incumbents ability to cooperate.

Barrier to mobility considerations are thus do not changing the nature of the argument of the concentration doctrine that is associated with plain neoclassical mainstream. Concentration still

---

[1] Of course, the degree of differentiation, the existence of absolute cost advantages or a high level of vertical integration are also a result of incumbents' actions and in this sense endogenous (Caves and Porter, 1977, pp. 245-247).

Another potential competitive threat that appears in the standard literature is not the entry of firms into an industry, but the development of substitute products by other producers or the search for substitutes by the buyers. Porter (1980, p. 23) calls it the fifth force of competition. As in the case of industry entry, advertising and R&D expenses are considered to have a potential of insulating against this competitive threat.

leads to excess profit, but it does so only if an additional condition - the existence of mobility hurdles - is fulfilled.

1.1.3 Chicago School

The Chicago School with Demsetz (1973) argued like the Austrian school (see below) against the concentration doctrine and for an endogeneity of industry structure. He highlighted with his *efficiency hypothesis* the possibility that concentration can be the result of single firms having superior efficiency[2] in producing and marketing products and thus enjoying stronger sales growth, obtaining larger market shares and leading to higher industrial concentration levels (Demsetz, 1973, p. 1). The main prediction is thus the existence of a strong significant positive relationship between market share and profitability. The often observed positive significant relation between concentration and profitability is a spurious one due to omitted variable bias since large market shares of firms in an industry imply higher industrial concentration as well as higher industry average profitability.

The approach relies on the ideas that markets produce naturally and automatically outcomes that are efficient and socially optimal as well as that the payoff of economic actors depends on their own contribution, skill and performance. The regulatory message is that policies aimed at deconcentration can lead to inefficiency if they destroy the incentives to obtain temporary monopolies through entrepreneurship or if concentrated market structures have production cost advantages due to economies of scale (Demsetz, 1973, pp. 3-4).

The possibility of a high level of persistence of excess profit is acknowledged and explained with difficulties to transfer knowledge that leads to higher productivity and that depends on the employees in their specific team and firm, by high costs of obtaining information and problems in duplicating techniques, or by limits to understand the reasons of competitive advantages due to the high organizational complexity of firms (Demsetz, 1973, pp. 2-3).

The approach stays closer to the stationary framework than the Austrian or Classical one and puts less emphasis on evolutionary processes and more on the absence of mobility barriers as a condition for competition. It reflects rather a different expectation about the occurrence of mobility barriers than a deviation from that theory's foundation. However, its predictions are similar to the ones of

---

[2] Proponents usually mention superiority in reducing costs e.g. due to superior managing skills. However, for the prediction that follows from different cost structures, it does not make a difference if advantages are alternatively due to plain economies of scale or luck in the investment decisions undertaken under true uncertainty.

these alternative theories.[3] According to Demsetz (1973, p. 4), there is evidence for the efficiency view if small firms do not earn higher profits when they are in more concentrated industries and when the profitability differential between small and large firms rises with the degree of concentration, since collusion is assumed to benefit all firms in the same way and monopoly profits to occur in proportion to firm sales and capital stocks. In an econometric analysis, testing this requires the use of firm and industry level data - market shares of firms and industrial concentration data in a single regression that allows to isolate industry concentration from individual market share effects. The Chicago School view highlights that the payoff to firms depends on their own efficiency. It is seen as rather independent from the actions of competitors and from industry structural dimensions.

1.2 Dynamic Theories of Competition

Even though the dimensions of potential competition introduce a dynamic aspect by including potential future competitive developments, they are still doing this within a static framework (Deeds and Hill, 1996, p. 440). Market power is assumed if excess profits can be observed at a given point of time and where deviations from the model of perfect competition lead potentially to such market power.

There are two approaches which promote a dynamic view on competition and which treat observed reality (here: profitability differentials) not as an imperfection of an idealistic model that theory aims to preserve through the integration of imperfections into perfectionist models. These dynamic theories rather rely on fundamentally different models that recognize observed reality as what it is: profit differentials are normal features of the economic system, part of the real competitive process how it is and not evidence of a lack of it. The approaches share the core ideas that the concepts of competition and economic equilibrium do not describe a given state respectively, but a process. This competitive process is viewed as being naturally disruptive and revolutionary, interactive and strategic, that there is never an observable equilibrium at any point of time (where all profit rates are equalized), that firms are not homogenous, that the outcomes of the competitive process can be unpredictable and unintended by the participants (Hayek, 2002, p. 10). One approach comes

---

[3] There are at least two other theoretical extensions to the mainstream framework that would lead to similar predictions where concentration is not or only very weakly positively related to profitability: 1. if there are costs involved in maintaining a monopoly or oligopoly that are not cost advantages themselves (like excess capacity utilization) (Semmler, 1981, p. 44); 2. if a lack of competitive pressure leads to X-inefficiencies - organizational inefficiencies that are reflected in higher costs, lower productivity and thus also lower profitability.

from the Austrian School and the other one originates from Classical Political Economy, while the former is significantly influenced by the latter (see Schumpeter (1947)).

1.2.1 (Neo-)Austrian School

The core idea of the evolutionary Austrian theory following Schumpeter (1947) is that competition is a dynamic process which is fundamentally violent and permanently creating new products, technologies and markets while erasing the fundaments of established old industries. *Creative destruction* leads to wholly competitive firm behavior - through its actual realization and as an ever present threat.[4] Monopolies can exist temporarily, serve as incentives for innovation and enable long-run planning. They can thus be understood better as being a part of the competitive process itself and not as constituting a fixed state that stands in opposition to competition. They are in general also virtually irrelevant over time.[5] Excess profit is the result of successful innovations of entrepreneurs that exploit temporary monopolies.

Potential monopolies are always limited by rival firms and technologies - old and new ones, those which exist and those which will emerge eventually - and very unlikely to persists in the long run if they really do behave monopolistically. Monopolists can also have access to superior technologies and other options which lead to lower prices and higher output than firms in a perfectly competitive structure. They are better able to set up research units, finance R&D expenditures and are thus themselves driving directly the competitive process.

Innovations, the utilization of new technologies allow firms to overcome mobility barriers like disadvantages that are emphasized by traditional theory and that result e.g. from a lack of experience in an industry or a lack of scale (Deeds and Hill, 1996, p. 441).[6]

The competitive process is driven by entrepreneurs that are active price and quality makers and no passive takers of given market parameters as in the mainstream model. They discover opportunities resulting from disequilibria (which might be caused by errors previously made by

---

[4] ‘ ...competition ... acts not only when in being, but also when it is merely an ever-present threat. It disciplines before it attacks.’ (Schumpeter, 1947, p. 44)

[5] ‘...competition... strikes not at the margins of the profits and the outputs of existing firms but at their foundations and their very lives. This kind of competition is as much more effective than the other as a bombardment is in comparison with forcing a door, and so much more important that it becomes a matter of comparative indifference whether competition in the ordinary sense functions more or less promptly;’ (Schumpeter, 1947, p. 44).

[6] A quoted example that fits this idea very well is the broad emergence of electric arc furnace technology in the steel industry of the 70s, which was relatively concentrated at that point of time and featured large scales and high capital intensity requirements. Utilized in ‘mini mills’, new competitors such as Nucor were able to enter the industry and produce steel from scrap at lower costs than established market leaders like U.S. Steel that relied on traditional coke blast furnace technology.

entrepreneurs, changed tastes, new technologies) and exploit them for profit under conditions of true uncertainty. This process tends systematically towards the allocative equilibrium. Competition leads dynamically to the discovery and in the long run to the spread of otherwise unknown and unused knowledge about supply and demand potentials (Kirzner (1997, pp. 62-73), Hayek (2002)). Conventional indicators of market power such as mergers, advertising or R&D expenditures are seen as being part of the competitive process and not as undermining it (Kirzner, 1997, p. 74).

Implications for empirical investigations are that cross sectional analyses of a single point of time cannot indicate monopoly power even if there are profitability differentials, but that a sufficient time period must be analyzed which is not dependent on business cycles and random shocks. One should expect that there is no significant and positive, persistent long term relation between market concentration or barriers to entry and excess profitability. Another implication is that if such a link is found, then it is likely that it is temporary and the result of successful innovations in the competitive process and due to productivity advantages (endogenous).

The version of Austrian theory supporting the resource-based view of the firm emphasizes the importance of the resources of the firm - mainly those that consist of knowledge, information, routines within the organization or 'software' (Deeds and Hill, 1996, p. 433).[7] As in the evolutionary approach, the recognition of firm heterogeneity is a precondition for the theory. The theory states that firm resources which can lead to sustainable competitive advantages and above average profits if they are not perfectly mobile between firms (Barney, 1991, p. 101), if they are 'valuable' (exploit opportunities, neutralize competitive threats for the firm), rare among current and potential competitors and hard to substitute with alternative resources (Barney, 1991, pp. 105, 111).

Imperfection in mobility and imitability can result if a resource depends on unique historical conditions, when the causal links between it and the competitive advantages are ambiguous (and thus not understood by outsiders and maybe even by the managers within the firm) or when it is based on social complexity that is hard for others to copy (Barney, 1991, pp. 107).

The implication is that persistent excess profits can be on ownership and utilization of superior resources which cannot be replicated, imitated or substituted. Differential profits are thus 'rents to efficiency', not monopoly (Barney, 1991, pp. 105, 116). Industry concentration, market share sizes and indicators of 'barriers' are endogenous and result from differences in performance and

---

[7] Plants, equipment, employees, locations ('hardware') are largely neglected due to the assumed tradability. Differences in them alone cannot lead to sustainable profitability differentials according to the theory (Deeds and Hill, 1996, pp. 433-434).

profitability (Deeds and Hill, 1996, p. 430).

An important notion is that firm heterogeneity implies that the more efficient firms do not have an incentive to collude with less efficient competitors, since competing on price allows them to increase profits through obtaining additional market shares (Deeds and Hill, 1996, p. 440). This is a thought similar to the classical idea that employing the lowest costs production technique in industrial price competition allows capitals to obtain surplus profits.

1.2.2 Classical Political Economy

While Austrian theorists often expect a relatively rapid convergence towards equilibrium and a quick erosion of profit differentials, Classical theory implies longer adjustment periods and higher degrees of persistence. There are two mechanisms that play a role in the equilibrating-disequilibrating competitive process and in the determination of profitability: a capital allocative one and an innovation based or innovative one.

Competition plays an important role in classical price theory. It is the mechanism that establishes natural prices or prices of production as centers of gravitation for the observable market prices (Marx (1991, p. 279); Smith (1991, pp. 61-63)). Competition among capitalists is the reason why this role is not played immediately by labor values (since they would not allow capitalists to obtain a uniform rate of profit under different organic compositions of capital). Prices of production modify market prices in such a way that they not only cover the costs of (re)production, but also allow to obtain a normal or average rate of profit on the total capital stock that is necessary for production. This 'capitalist law of exchange' is the long term mechanism that determines prices.

Prices of production are never actually observed and exist only as moving averages. The industrial rates of return can never be equal at a single point of time and constitute sets of disequilibria. Random events as well as demand and supply fluctuations lead to imbalances that raise market prices above or lower them below prices of production and lead to positive and negative profitability differentials. However, the concept of equilibrium through disequilibrium requires that competition works in the way envisioned by the classics and produces profitability equalizing tendencies. Without such an equilibrating force it becomes hard to argue that prices of production regulate observable market prices and that labor values are of any relevance (through their effect of determining prices of production). The profit rate equalizing mechanism of inter-industrial competition is thus an essential precondition for the classical value and price theory.

If a market price is persistently above its price of production, capital will flow in this inter-industrial competition out of industries with below average profit and into industries with above average profit. This reallocation of capital is the core equilibrating mechanism. It is a dynamic, gradual process and implies that supply is shrinking in the former and expanding in the latter industry relative to demand - through exit and entry of new capitalists or investment and disinvestment undertaken by existing ones. The reallocation can directly be undertaken from capitalists' own funds that are retained and reinvested or mediated through the credit system. It might only be reflected in growth rate differentials of the industry capital stocks versus their respective demand growth rates. No actual entry or exit of capitalist firms needs to actually occur. Due to downward sloping demand curves (Marx, 1991, p. 279), this 'allocative' competitive equilibrating mechanism will in turn lead to pressure on market prices to adjust toward prices of production and on industrial profit rate differentials to narrow and gravitate towards the average or normal rate (Marx (1991, pp. 297, 310); Marx (1997, p. 13); Ricardo (1996, pp. 61-64, 82-83); Smith (1991, pp. 51, 58, 98ff)). Before any capital stock adjustment, reallocation and physical investment or disinvestment takes place the adjustment of supply to demand first takes place through the alteration of the degree of utilization of the already existing capacity (Shaikh, 2008, p. 168).

Another aspect of competition in the classical framework is also dynamic and one that disequalizes (and re-equalizes) profit rates between firms within industries. Competition between firms within industries leads to technological change, mechanization, innovation, improvements in equipment and the organization of labor. This development of new methods of production by individual capitals allows to raise their productivity and lower their individual costs of production below the ones prevailing in the industry. Capitals that successfully innovate are thus able to obtain a surplus profit (a positive difference between them and the industry) by selling at or slightly below the price of production of the industry average that runs at higher costs ((Marx, 1991, p. 279, 345, 780, 783); Ricardo (1996, pp. 269-270); Smith (1991, pp. 63-64, 122)).[8] Since competition 1. creates arrays of different production conditions in each industry over time through these innovative dynamics and 2. enforces a single common selling price, a general disequalization of profit rates within industries is the result (Shaikh, 2008, p. 169).

However, this 'innovative' competitive process has also an equilibrating dimension which is em-

---

[8] This mechanism plays an important role in the classical theory of technological change and mechanization as well as in the theorem of the falling rate of profit.

phasized by the Austrians. Innovation tends to cancel out cost differences over time between firms within industries because new and more efficient modes of production will also be applied by competitors and become broadly used throughout the industry over time through imitation and copying. In each industry there are incentives for and pressures on existing and potential competitors to develop their own completely different, even newer and more effective methods. The long run effect is a turbulent equalization-through-disequalization process that will lower industry output prices ((Marx, 1991, pp. 783-784); Ricardo (1996, pp. 270, 273); Smith (1991, pp. 63-64, 122)).[9]

Another equilibrating effect of innovation is that it also enables capitals from outside the industry to enter and overcome potential barriers to entry. This is the case when new and lower cost production techniques are available for industry entrants. This potentially allows them to overcome the disadvantages associated with a lack of experience and scale when entering an industry (as illustrated in the mentioned electric arc furnace example). The ability to overcome conventional barriers to entry is thus judged as in the Austrian framework: new capitals are in a position where they can apply the newest, often most economic vintages. A difference to the Austrian view is that it is not clear that the equilibrating dimension of this innovative mechanism of competition is a rapid one. It might very well be persistent to a significant degree and let profit rate differentials only converge in the very long run.

The notion of the 'allocative' equilibrating mechanism shows another difference to the Austrian School (resource-based view) which focuses on immaterial assets like knowledge and information or social interaction and (corporate) cultural factors.[10] For the Classics, physical capital structures and fixed assets matter in many aspects of their analysis. Here it cannot be just assumed that perfect secondary markets for capital goods exist and lead to bids for these assets that leave them in each firm evaluated in such a way that the rates of return derived from these fixed assets just yield the average rate of profit (something that might be true for equity shares on the stock market). In the Classical view, fixed capital goods can also be extraordinarily devaluated, but in general need to be adjusted in their valuation in a way that reflects the consumption of their value during the production process. Certain fixed asset structures (including high capital intensity and high fixed capital shares)

---

[9] This means that it is not true that profit rates among capitals within one industry always exist without any tendency towards equalization. The equalizing mechanism is just a different one and it might work over shorter or longer periods. Of course, this does not mean that profit rates will ever be equalized at any point of time.

[10] There is no doubt that these can be important factors in a low-abstraction framework as in Classical Political Economy. However, a general problem is the lack of observability and thus testability of hypotheses based on such factors as well as the fact that these concepts involve usually a very high degree of vagueness and abstractness when described by Austrian economists.

can thus result in differentials in the rates of return and in higher degrees of persistence thereof if e.g. unexpected demand shocks create unfavorable conditions for sellers (Semmler, 1984, p. 37).

Shaikh (2008) stated the classical hypothesis more precisely in this respect: only the lowest-cost or best-practice methods that are generally available are subject to the equalizing process. These regulating capitals exclude past investments in older technologies that usually involve higher costs and sometimes lower ones (e.g. due to lower price levels at the time of investment) and those that involve non-reproduceable conditions of production. Technical change and persistent capital assets lead to arrays of production techniques and of capitals with different profit rates within each industry and within each firm. Accordingly, he proposed an approximation of profit rates on regulating capitals through profit on recent investment (which is estimated as the change of profit over lagged investment). The approximation of these incremental rates of profit is done in a similar way as a popular approximations of marginal Q in empirical investment theory econometrics[11]. My own previous research showed that they work very well in different applications with aggregate data, but due to high disturbances not in micro data as it is used here.

The adjustment process lasts for different periods in different industries due to unequal conditions of production where the main determination lies in different fixed capital turnover times and durabilities of capital (Marx, 1991, p. 311; Marx, 1997, p. 42; Semmler, 1981, pp. 41-42, Semmler, 1984, p. 36).[12] Slow fixed capital turnover times can act as entry and exit barriers that shield and protect profits in the boom and trap capitals or lock them into the industry in the bust. The overall effect is to slow down the entry and exit of capital. Profitability differentials are impacted in the short and medium run such that any adjustment of supply through capacity alterations will be slower and profit rate differentials will thus take more time to erode. This phenomenon is especially likely to occur where 1. a slow fixed capital turnover/high durability is combined with the already mentioned 2. high capital intensity (capital/output or capital/labor ratio) and 3. high fixed capital share in the sum of fixed and working capital.[13] However, this does not mean that there will be

---

[11] The formula looks similar to $q_t = \frac{M_t-(1-\delta_t)M_{t-1}}{I_t}$ where M is the market value of the firm, δ the rate of depreciation of the capital stock and I the investment or change of the capital stock that is assumed to have caused the change in the market value of the firm in the numerator. See e.g. Gugler, Mueller, and Yurtoglu (2004) or Eklund (2008).

[12] Ricardo highlighted for the determination of prices of production not only the division of the total capital stock into fixed and circulating capital (money for wages and working capital), but also the differences with respect to durability in each of the two capital types (Ricardo (1996, pp. 35-38); see also Smith (1991, pp. 224ff)).

[13] These differences in fixed capital structures have also important implications for industrial pricing patterns: high intensity will generate incentives to keep a higher percentage of unutilized capacity in order to be able to meet demand fluctuations with utilization instead of capacity adjustments. The result will be more fluctuations in capacity and less volatile price movements (since it is more likely that demand spikes will be met with quantity adjustments). Industries

any permanent differentials, but that the persistence of positive and negative differentials will be different (higher) and that the full (and in these cases: longer lasting) cycles of all industries considered have to be taken into account in order to avoid a flawed analysis. As plant life times can easily go beyond 20 and last even up to 50 years as in the electric power producing industry even a comparison of means of industrial rates of return is questionable.

The classics were explicit about the fact that differentials have to be expected as normal outcomes if the conditions which determine differences in risk are unequal (Ricardo (1996, pp. 62-63), Smith (1991, p. 117)). This implies that certain limited differentials are necessarily a competitive equilibrium outcome and that profit rates have to be adjusted for a risk dimension in empirical investigations.

The relation between market share and size on the one hand and profitability on the other hand is not immediately obvious in Classical Political Economy. It is clear that the introduction of new techniques leads tends to lead to cost advantages for individual capitals versus the rest of the industry. This enables firms to follow two distinct strategies. They could raise profitability by either 1. increasing the profit margin by selling at the industry average price and keeping sales constant, or by 2. keeping the profit margin constant by lowering their selling price below the one offered by competitors (in proportion to the newly achieved costs of production advantage) and raising the sales volume (due to price undercutting). Both strategies produce surplus profit for the individual capital. In reality it is likely that most or many firms in such a situation will pursue a combination of these two. Thus, a positive link between market share and profitability can be expected.

Classical technical change might in addition to that also create pressure to raise sales. This is the case because it involves the use of new methods that involve a higher amount of fixed capital that replaces labor, raises its productivity and/or economizes the use of working capital and intermediate inputs. This implies that there will be higher fixed and lower variable costs. One effect of it is that firms are required to produce and sell more output units in order to break even and that they will chose to produce under market price scenarios where firms with old techniques will chose to idle production (this is the situation where the market price is between the costs of production of the new and old production methods). The overall effect is that the successful introduction of new technologies leads to higher market shares and higher profitability and that these two measures

---

with low fixed capital requirements will more often have to respond with price adjustments to demand changes, since they operate closer to the capacity limit and will adjust capacity in more cases (which takes time).

should thus be positively associated in the short and medium run. On the other hand there might at the very same time be a negative relation between firm size and profit margins if higher productivity is associated with very aggressive price cutting and if higher sales and utilizations allow to obtain the same or a higher rate of profit with lower margins.

1.3 Predictions and Possible Tests

The econometrically testable hypotheses that are implied by the different approaches differ from each other significantly, as do the empirical approaches that are suited to test them. Due to the fact that the concepts and drivers of competition differ, alternative modeling methodologies and empirical investigations follow for industry analytical applications of the theories. The most important stylized predictions are illustrated in table 1.

Table 1: **Empirical predictions of competition theories**

	Importance of		Significance of	
	Industry	Firm	Concentration	Market share
Neoclassical Mainstream	+	-	+	-
Chicago School	-	+	-	+
Austrian Economics	-	+	-	+
Classical Polit. Econ.	+	+	-	+

"+": important & significant; "-": unimportant & insignificant

The most important prediction of the *neoclassical mainstream* is that indicators of industrial concentration should be significantly and positively related to profit differentials in an econometric analysis. This might only be observed above a certain critical concentration level or with a non-linear regression specification. Concentration ratios up- and downstream the value chain should have negative significant impacts on profitability. Alternatively, an industry's concentration relative to its suppliers and customers should be more significant than indicators of industrial concentration alone. While this extension of the concentration makes sense within the mainstream framework, this tweak has not yet been accounted for in econometric analyses that are representative for the economy. Strategic investments like advertising, research and development (R&D) spendings, as well as firm size indicators have significant positive effects on profitability. It might be the case that only interaction coefficients of concentration and barrier indicators should be positive and significant or that only regressions on a subsample of high barrier industries reveal significant positive relations between concentration-barrier interactions and profitability.

Industry structure is seen as the most important determinant of the fortunes of a company. This thought lead economists and business management academics conclude in the *structure-conduct-performance* theorem that the actions of firms follow automatically from industry structure and that the payoff depends again on these actions taken and indirectly on industry structure.

An application of static theory of competition thus implies an industry analysis which focuses on concentration and the existence and structure of mobility barriers in the respective industry under consideration. These industry structures in sectors connected along the value chain are also of interest. Industrial production and pricing is described best using oligopolistic models.

From the point of view of static competition theories it is possible to use cross sectional tests as the equilibrium state should exist at any moment of time. The sole detection of profitability differentials or the revelation of significant positive effects of concentration and barrier indicators on profitability are sufficient to indicate oligopolitist or monopolistic market power.

The alternative *Chicago School* approach implies different hypotheses. Market shares are predicted to have significant positive effects on profitability while concentration should have no significant effect once market shares are controlled for. The existence of significant exit barriers could theoretically move the prediction closer to the hypotheses of the standard theory. But the only actual major difference to the standard approach is that mobility barriers are seen usually to not exist on a large scale. In this view, industry structure is judged radically different. It is generally irrelevant while the success of a firm depends exclusively on its own capacities and efficiency. ‘Industry analysis’ becomes nothing but a set of individual evaluations of companies whose performances and activities are not effected by each other.

The need for the inclusion of market share as a control variable and a precondition to isolate the effect of concentration is the most important implication of the Chicago School approach for any empirical research. Any study that fails to include this variable needs to be rejected as unsuited to say anything about the effect of concentration.

Austrian economics predicts that excess profits might persist for some time, but will in general erode relatively rapidly. Market shares might be related positively to profitability throughout the medium run while concentration and barrier indicators will not have any effect. Markets (primary and secondary capital and managerial markets) work perfect or rapidly towards equilibria and the payoffs of economic actors depends largely on their own efficiency and performance. Thus, industry level dimensions are negligible and not have any effects while firm level factors are the only important ones. This does relate the Austrian somewhat to the Chicago point of view. Industry analysis should focus on industrial change and innovations introduced by new or existing economic actors that pursue entrepreneurial behavior.

Classic Political Economy states that profitability differentials need to be expected at any point of time as normal outcomes and elements of the competitive process itself. A finding of the existence of profit rate differentials at any point of time or even in an average over time does not imply a lack of competition. Unlike in the other approaches the dimensions industry and individual firm are both of joint importance for profitability: surplus profits are obtained from a divergence of individual from industry average costs. This implies that the *relative* costs of production, the relative efficiency of a firm versus its direct rivals are the decisive element.

In addition to that, the dimension industry matters because of differences in industrial fixed capital intensities and durabilities. They have the potential to influence pricing patterns throughout the business cycle and desired capacity utilizations or the speed of supply adjustments (in reaction to profitability differentials in a disequilibrium). It follows that findings of some impact of industrial factors on profit differentials do not suggest automatically a lack of competition as the exact nature of the relation might not be immediately apparent and opposed to mainstream stories. This idea must be applied for example to criticize the profit variance decomposition research following Schmalensee (1985) that often finds that some of business unit profit variance is explained by the dimension industry and that also aims at making statements about competitive intensity (see e.g. McGahan and Porter (2003, 1999)).

However, it is clear concentration and barriers to mobility have no positive effects on long-run profitability differentials. Market shares however might be significantly positively associated with profitability - at least in the medium run while there will be an (often slow) adjustment process in the long run.

From dynamic views on competition it is clear that cross sectional regressions which involve only observations of a single year are unsuited to detect competition or a lack thereof. Even time averages can be insufficient, especially from a Classical perspective as high fixed capital durabilities might imply very long periods of supply adjustment. Thus, only an approach that compares very long time averages or a profitability measure that utilizes time series methodology can detect a lack of competition from a classical point of view. Since it is also applicable to static theories, one can test different theories of competition with such a measure. An example is the persistence of profit methodology pioneered by (Mueller, 1977, 1986) that uses autoregressive models to obtain estimates of persistent long run profitability differentials which can then be used as the dependent variable in an industry structural regression.

The use of a classical dynamic framework in industry analysis calls for an identification and assessment of the drivers of any current and future change in 1. products; 2. conditions of production and 3. methods of production. Gradual and radical innovations and the commercial implementation of new technologies need to be assessed as well as changes in the composition of the industry average with respect to production techniques. In addition to that, the Classical perspective requires any industry analytical approach to focus on the costs of production. The different production methods that are generally available must be assessed with respect to cost and how innovation and industrial change effects them in the future. The average costs of production in the industry in relation to the the costs of individual companies are of paramount importance.

Chapter 2

New Econometric Evidence on Competition, Profitability and Industry Structure in the U.S.

The conclusions of the previous chapter were that cross sectional econometrics based on snapshot points of time or even on time averages are unsuited to test all mayor theories of competition - in particular the dynamic ones. In addition to that testable hypotheses were derived from the four different theories with respect to the importance of industry and firm level variables in general as well as with respect to the impacts of concentration and market share in particular. This chapter aims to test the four competition theories or the theorems derived from them in the last section of the previous chapter with a methodology that is appropriate for all approaches.

The persistence of profit literature following Mueller (1977, 1986) has provided an estimator of long-run persistence in rates of return on assets. It has become the standard in empirical research for investigating issues of long-run performance differences and is also suited to test dynamic theories of competition. However, its predecessor, the tradition of industry structural research on profitability differentials that usually utilized cross sectional regressions or short panels relied on sets of regressors and on regression equation specifications that described industry and company level characteristics in much more detail (see Ravenscraft, 1983, Webster, 1996 or Spanos, Zaralis, and Lioukas, 2004). Another research tradition that is an alternative to the persistence of profit approach decomposes the variance of profit differentials into segment, company and industry effects (see e.g. Schmalensee, 1985 or McGahan and Porter, 1997) and relies on the business segment as a micro unit which is broadly believed to be more suited than the company level. This paper combines the strengths of the different approaches in a persistence of profit analysis to shed light on the relation of industry concentration, market power and excess profitability and to test alternative theories of competition. It is the first study to use Compustat Segments data as the unit of analysis and it includes more than one sector of the economy (usually manufacturing) while relying at the same time not on an arbitrarily picked concentration ratio as an explanatory variable but on a concentration indicator that is superior to these and in its interpretation close to the Herfindahl Index. As a more serious attempt to avoid omitted variable problems, detailed data on industrial unionization, industry imports, exports and dummies for international and sectorial company diversification are included. I am testing for the impact of critical concentration levels and industrial world market integration on the regression results, consider interaction variables, non-linear specifications and pioneer approximations of concentration in industries up- and downstream the industrial value chains. The analysis updates the time frame by including years after 2000. It covers 2,052 segments in the 33-year period from 1976 to 2009.

The structure of the chapter is as follows: the next section reviews related empirical literature; thereafter the empirical analysis is presented and the results evaluated; the last section summarizes the chapter.

2.1 Studies on Competition and Profit Rate Differentials

There are three major traditions of econometric empirical investigation into the topics of profitability differentials and competition. All of them are related to and combined in the investigation presented here.

2.1.1 Persistence of Profit Analysis

The persistence of profit literature as pioneered by Mueller (1977, 1986) is the predominating econometric approach in investigations on profit differentials and competition today. This study here is included in the tradition. Most authors focus on modeling an autoregressive or an alternative time series process in usually firm level profit rate differentials and investigate their mean reverting properties. The individual rate of profit for segment i in period t, Π_{it} is assumed to be made up from an average 'competitive' return on capital, c, a permanent rent or surplus profit, r_i and a short-run rent, s_{it}, all of which can be written as a fraction of the average profit rate in the economy, $\overline{\Pi}$:

$$\Pi_{it} = (c + r_i + s_{it})\,\overline{\Pi}_t \tag{2.1.1}$$

or

$$\pi_{it} = \frac{\Pi_{it} - \overline{\Pi}_{it}}{\overline{\Pi}_{it}} = (c - 1) + r_i + s_{it} \tag{2.1.2}$$

where s_{it} is dependent on time and $(c - 1) + r_i$ is constant. When the latter term is written as a_i, the equation becomes

$$\pi_{it} = a_i + s_{it} \tag{2.1.3}$$

s_{it} is assumed to follow the path

$$s_{it} = \lambda_i s_{it-1} + \mu_{it} \tag{2.1.4}$$

where μ_{it} is a random error and the assumption $-1 < \lambda < 1$ implies convergence. Manipulating and substituting (2.1.3) into (2.1.4) results in

$$\pi_{it} = a_i + \lambda_i\,(\pi_{it-1} - a_i) + \mu_{it} \tag{2.1.5}$$

or

$$\pi_{it} = a_i\,(1 - \lambda_i) + \lambda_i \pi_{it-1} + \mu_{it} \tag{2.1.6}$$

and writing the constant term $((c - 1) + r_i)\,(1 - \lambda_i)$ or $a_i\,(1 - \lambda_i)$ as α_i leaves us with

$$\pi_{it} = \alpha_i + \lambda_i \pi_{it-1} + \mu_{it} \tag{2.1.7}$$

which can be estimated as an AR(1) process on profit rate differentials. Its unconditional mean

$$\hat{p}_i = \frac{\hat{\alpha}_i}{1 - \hat{\lambda}_i} \tag{2.1.8}$$

can be calculated from the estimated coefficients and is interpreted as the long-run, permanent or persistent profit rate differential that is not being eroded by competition. Most studies only run regression (2.1.7) and test for stationarity. Since the autoregressive coefficient $\hat{\lambda}_i$ measures the speed of erosion of temporary differentials, high values are often interpreted as indicators of low competitive pressure. This common interpretation is based on a misunderstanding, since $\left(1 - \hat{\lambda}_i\right)$ determines the speed of adjustment to the long-run equilibrium value that includes the competitive return (or zero in case of differences and ignoring risk) *and* the permanent rent or profit rate differential r_i. Thus, a fast short-run adjustment process (a small value of $\hat{\lambda}_i$) can be an adjustment to a mean value that might include a large and positive permanent above average rate of profit (a large α_i). On the other hand, a very high degree of persistency represented by a large $\hat{\lambda}_i$ can be perfectly consistent with high competition if the differential itself / the constant α_i is close or equal to zero (it might even be negative). Analyzing $\hat{\lambda}_i$ is thus not telling us anything valuable if we want to test predictions of alternative competition theorems. Instead, the appropriate variable to be analyzed for such an investigation is the estimated permanent profit rate differential $\hat{p}_i$, the differential that results after the adjustment process has taken place. Cable and Mueller (2008) also showed that the accuracy of $\hat{\lambda}_i$ is much more dependent on the length of the sample period than it is the case for $\hat{p}_i$. Studies typically use panels with at least five years of observations and some rely on up to 50 years but usually include only a few hundred firms. A common finding is a large profit differential persistence, which is usually being interpreted as an indicator of lacking competition. The major shortcoming is that competition can naturally give rise to permanent profit differentials to reflect different degrees of risk or as an outcome of the introduction of new techniques of production that imply heterogenous cost structures which might persist for a very long time.[1]

2.1.2 Two Stage Regressions (Profit Persistence plus Industry Structure-Profit Regression)

To assess competition one cannot just observe persistent profit differentials. They are only necessary, but not sufficient for a proof of market power. Instead, one needs to additionally explain the

---

[1] Innovations can throw companies into different profitability trajectories which might imply an even longer time horizon until competitors erode the corresponding excess profits.

determinants of permanent profit differentials and prove that those factors (like concentration or barriers) for which economic theory explains and predicts a causal competition harming effect are actually significantly associated with profitability differentials in the envisioned way. It cannot just be assumed that these factors are at work and that competition is being impeded since they might have no such effect and explanations for the differentials may have nothing to do with monopoly or market power. This can be done by combining the profit persistence analysis with an industry structural regression frame. There are only very few two-stage regression studies that accomplish this. The present study here is itself a part of that approach.

Analyses rely in the first stage on the profit persistence estimations of $\hat{\lambda}_i$ and / or $\hat{p}_i$. In a second stage, they use these regression results in an industry structural cross sectional regression

$$\hat{p}_i = \beta x_i + e_i \tag{2.1.9}$$

or

$$\hat{\lambda}_i = \beta x_i + e_i \tag{2.1.10}$$

where e_i is an error term, x_i a vector with structural and control variables such as concentration, asset size, advertising intensity or sales growth and β the corresponding coefficient vector.

Gschwandtner (2012) provides the most complete analysis on the subject until now. Using Compustat firm level data, her regression with the most observations and regressors was based on the 15-year period from 1984 to 1999 and included 1,099 firms. The regression on $\hat{p}_i$ that includes a concentration ratio, market share, advertising and R&D intensity, asset size, sales growth, exports and imports reveals an adjusted R^2 of 0.17, insignificant negative coefficients for concentration, market share and the standard deviation of the return on assets, positive insignificant ones for R&D and asset size as well as a positive and significant one for advertising. Unfortunately, the analysis only includes manufacturing firms and is based on company level data with all its shortcomings in this context (see below).

In another study, she uses 157 U.S. manufacturing companies from the same Compustat data base which had time series of 10 years or more between 1950 and 1999. Persistence parameters and the long run profit rate differentials were estimated. The second stage regression with $\hat{p}_i$ includes census industry data on growth (of the number of firms), industry size (number of firms, value of shipments) and concentration ratios as well as the firm level regressors such as market share and profit rate volatility. Regressions were run on subsets of the sample that divided it into groups of survivors and exiters which were active in 61 and 53 different industries respectively. The R^2 for two subsamples are -0.009 and 0.074 and the coefficients for concentration insignificant (positive in the first and negative in the second case). The negative correlation of profit rate volatility with $\hat{p}_i$ in one regression and the positive correlation of firm size with $\hat{\lambda}_i$ and with $\hat{p}_i$ in the other ones let her conclude that 'the competitive process is weak' and that 'a large firm ... operating in a concentrated industry ... is a persistent profitable company' (Gschwandtner, 2005, p. 804). This is a very brave statement given the fact that concentration is insignificant in all listed equations even at a 10% significance level. Also, observing a positive correlation between firm size and the persistence parameter does not tell us if large firms are persistently profitable above or *below* average.[2]

Acquaah (2003) uses Compustat firm level data on 119 firms from 1985 to 1997 and Fortune's AMAC database for his study. The second stage regression contains several industry structural variables including the four firm concentration ratio (CR4), capital intensity, firm size, R&D and advertising intensities. He finds insignificant values for concentration variables, capital, advertising and R&D intensities. However, CR4 is estimated using Compustat data, which is known to yield poor approximations and no substitute for census data (Ali, Klasa, and Yeung, 2009).

Yurtoglu (2004) tries to explain in the second-step regression of his analysis of 172 large manufacturing firms in Turkey from 1985 to 1998 not only $\hat{\lambda}_i$, but also $\hat{p}_i$ by using CR4 and a set of control variables. He finds the concentration ratio to be positive and significant. However, his second-step regression included only 100 observations and the industry data were originally on a highly aggregated three digit level[3] and does not include any mobility barrier indicators.

Kambhampati (1995)uses data on large public limited companies in India from 1970 to 1985. However undertakes an industry level analysis where industrial averages are calculated for the companies and grouped into 42 industries. Using industry aggregates allowed the author to include the structural variables of a concentration ratio (not mentioning which one), the advertising to sales ratio, a variable that is supposed to measure cost disadvantages of operating below the industry's minimum efficient scale, industry growth and an industrial policy dummy. He finds that the con-

---

[2] It could very well be that higher persistency across the industrial business cycle in fat as well as lean years could be due to the ability of producers to adjust only slowly if they operate at a high degree of capital intensity (which is likely to be correlated with size and not accounted for in her regression).

[3] The author disaggregated the data manually even though higher industry aggregates of concentration variables are not necessarily related to the equivalent variables at the lower aggregation level and usually diverge systematically from them.

centration ratio (and the advertising intensity) had positive coefficients and were insignificant at the 5% (and the 10%) confidence level in a second stage regression on $\hat{\lambda}_i$.

Unfortunately he did not try to explain $\hat{p}_i$. The industry level analysis does not enable to discriminate between market share and industry concentration to separate market power from efficiency effects. In addition to that, the industries used represent a level of aggregation that is very high and should be considered as inappropriate for analyzing industry structural factors in profit regressions, since even 6-digit detail levels can be too broad (see below).

Goddard, Liu, Molyneux, and Wilson (2011) present a paper with an interesting focus on a single industry. They study profit differential persistence in banking for 65 different countries using Bankscope data for 11,634 institutes and for a period from 1997 to 2007. It includes Herfindahl Indexes, macroeconomic as well as legal and regulatory variables where some of these are interpreted as barrier indicators. Unfortunately, the second stage regression is based on country-level $\hat{\lambda}_i$ which leaves only 59 observations and does not allow to separate market share from concentration effects. Another shortcoming is that the authors only use the persistency coefficient as a dependent variable, $\hat{\lambda}_i$, not the long term differential, $\hat{p}_i$. The regression that includes the Herfindahl Index has an adjusted R^2 of 0.142 and indicates a positive significant relation between concentration and legal barriers on the one hand and profit persistence on the other.

2.1.3 Industry Structure-Profit Regressions and Profit Variance Decompositions

The mayor weaknesses of most of these studies are that they analyze with few exceptions mainly the persistence of profit itself and only add a second step regression as an extension to that main investigation. Accordingly, they suffer from very small samples, an exclusive focus on manufacturing, the omission of important variables and the use of a unit of analysis that is unsuited for the evaluation of industry level variable ($\hat{\lambda}_i$). The presented results thus cannot be taken as conclusive. It is necessary to have a more careful regression analysis at the second stage and better data to improve the reliability of the regressions and rule out possible misinterpretations of the results. Some insights from the industry structural regressions seem to have been lost in the current analyses. The understanding of profit variance decomposition research that business segment level data is superior and an empirically important explanation for explaining profit variance has not been recognized so far.

The older tradition of industry structure-profit regression research aimed to explain returns on investment on industry or firm level with a broad sample of industry structural and/or firm level variables. Ravenscraft (1983) for example includes in his U.S. data cross sectional regressions indicators of concentration, market share, advertising, R&D and capital intensities, diversification, capacity utilization, industry growth, minimum efficient scale, exports, imports and various interaction effects, including interactions between concentration and market share. Ratnayake (1996) tests the existence of critical concentration levels that lead to higher profitability using dummy variables and a 2SLS simultaneous equation model. Scherer (1980, pp. 278-279) summarized his analysis of empirical studies by stating that 'there is a rather robust tendency for a positive association to emerge between seller concentration and profitability. ... this conclusion emerges unmistakably'. However, most studies involve significant weaknesses that suggest caution: 1. some regressions that were based on the industry level or that did not include market shares as a control variable in company level regressions measured the effect that market shares have on profitability through the concentration variable (Demsetz, 1973); 2. as studies were generally based on cross sectional regressions, they were only able to give statements about the extreme short-run, dependent on complications with respect to the business cycle stage of industries and thus unsuited from a dynamic view on competition; 3. the usual focus on the manufacturing sector limited the generality of results significantly; 4. with some exceptions, most studies relied on company level data from samples that are known to include firms which are active in more than one industry as well as in markets outside the U.S. or that changed their main line of business during the observation period. Segments or lines of business are thus the far more suitable unit of analysis (Scherer, 1980, p. 283); 5. the calculation and the theoretical effect of concentration measures can depend on imports into the industry from which the concentration ratios have been calculated (Cowling, Yusof, and Vernon, 2000), which is also true for barrier indicators or other industry structural variables. Such variables are bound to become inappropriate with a high world market integration. The last three points of critique are of course also applicable to the persistence of profit research with second stage regressions.

Following Schmalensee (1985), a significant amount of literature emerged that analyzes the variation in profit rate differentials and their dependence on industry, company, and segment (business unit) levels. The main goal is to evaluate the relative importance of these dimensions and search for evidence on industrial organization and strategic management theories like structure-conduct-performance or resource based models of the firm. Studies employ variance decomposition techniques like sequential analysis of variance or variance components analysis (coefficients of variance) that cannot be described in detail here. They usually rely on Compustat Segments Files data or other business unit data bases like the Federal Trade Commission Reports to include the segment dimension in the analysis. This is a mayor advantage relative to the other approaches discussed here which rely largely on industry or company level data. Results generally point to the industry dimension having a significant effect on segment level profitability while parent company and business segment levels are also (and often to a larger degree) able to explain the variation of segment profits. While the industry dimension is often assumed to matter for (short-run) profitability, the approach does not allow to explain why this is the case, which industry dimensions have an impact and if their respective effects are positive or negative. That a part of the variation of snapshot or time-average profitability differences can be explained by the industry dimension does not automatically imply market power or collusion just as the pure proof of the existence of profitability differentials is insufficient.[4]

Some studies combine the decomposition with the persistence of profit approach or try to incorporate long term effects (e.g. Cubbin and Geroski (1987); McGahan and Porter (2003, 1999)). Still, the nature of the approach does not allow to determine whether any of the common competition hazard suspects really do have any positive impact at all on persistent profitability differences. If such a link cannot be shown to exist, it becomes a matter of mere religious believe - not of scientific evidence and logic - that market power is present in an economy.

2.2 Empirical Analysis

In the previous chapter on theories of competition I argued that a test of theories of competition that also include dynamic ones requires the use of a dependent variable such as $\hat{p}_i$. I showed that the level of sophistication of competition theories requires the consideration of non-linearities, critical concentration levels, interaction effects between concentration and barrier indicators, the analysis of a longer period of time and the use of a broad set of control variables. The particular application of a 'persistence of profit plus industry structure-profit regression' that is presented in this chapter aims

---

[4] It can e.g. be very well the case that there are differences in industry specific business cycle stages - industrial demand and supply disparities - instead of structural factors like concentration which explain most of the industry dimension's impact.

to accomplish precisely that. It tests the hypotheses derived by alternative theories of competition that were discussed above.

2.2.1 Data

The unit of analysis used here, the company line of business or company business segment is defined as the aggregate of all activities of a company that are located within a NAICS industry and within the U.S. (in the following just called 'segment'). The business segment is a much more appropriate unit of analysis than the average stock market listed U.S. American company as they can e.g. be obtained from data bases like Compustat's Fundamentals. The profit variance decomposition literature showed that the segment dimension has an important impact on the variation of profit differentials. Companies often engage in several lines of businesses that are either part of different industries and / or placed in different geographic locations (in the U.S. and in foreign markets). On average in 2009, segments defined along the lines of industrial activities of Compustat companies (but including different locations) originated 13% of their sales from operating units abroad while sales from such units amounted to at least 56% for the 10% of segment groups which were operatively the most international ones. At the same time, 23% of all companies in the file were active in more than one and 11% in more than two industries. For companies like General Electric or United Technologies, one can identify up to 9 different operating segments respectively in the Compustat Segments file, none of which accounting for more than 30% of total company sales. The company level classification labels these two conglomerates as unclassifiable establishments in the case of the former and assigns a higher 4-digit industry in the case of the latter (where actually only 2 operating segments are a part of this higher aggregate). By assigning a main line of business to each company in the Compustat Fundamentals Files, all operating activity abroad and in 'non-core' industries is treated as being a part of the primary industrial engagement. Variables like market shares that are based on company level sales will thus be substantially distorted in many cases. For example, Chevron Corp's primary NAICS industry classification in 2000 was 'Petroleum Refineries' (NAICS 324110). However, its sales within that industry in the U.S. amounted to only $29,629 million while the company's activities in industry 'Crude Petroleum and Natural Gas Extraction' (NAICS 211111) in the U.S. amounted to $12,724 million and total foreign operations had a sales volume of $55,081 million. Chevron's true sales volume in the domestic petroleum refinery industry was thus just a third of the $97,434 million that would have been assumed using Compustat's North American

Fundamentals Files. Relying on Compustat Segments instead of Compustat Fundamentals data reduces the market share of this second largest U.S. oil company accordingly from an incorrect number of over 52% to less than 16% in the refinery business. Using the latter company level data base, the return on assets would have been recorded with 11.6% while in fact, historically much smaller profit margins in the refining than in the exploration industry implied a rate of return on assets of only 9.7% for Chevron's U.S. refining operations.

Concentration indicators are the explanatory variables in the focus of this study. The data used have been obtained from the Economic Census and include manufacturing and non-manufacturing industries. The focus on manufacturing in other studies is quite surprising since the sector accounts for less than 13% of value added in the U.S. and studies based on it can thus hardly claim to be representative. I included all sectors except the ones for which no data have been published (agriculture, mining, construction and management of companies and enterprises). The resulting data base with information on utilities, wholesale and retail trade, transportation and warehousing and other industries is more complete than other ones that include just manufacturing or add only easy-to-assemble non-manufacturing sectors for the years following 1997.[5] The file contains all concentration data that have ever been published by the Census for any industry in electronic form. 3,450 non-manufacturing industries could be added to a total count of 8,968 observations (which compares well to the Wooster College data file with 3,600 manufacturing industries). In the final regressions, U.S. industrial activities from non-manufacturing industries accounted for 35% of all U.S. industrial activities included, increasing the sample size significantly and allowing to obtain more general results.

Additional data used in this study are based on industrial unionization. They come from the Current Population Surveys and can be obtained from the NBER and for most recent years in a very well assembled form from Hirsch and Macpherson's Unionstats web page. With 52,837 observations collected on an annual basis, the final file covers a period from 1973 to 2010 and provides an important control variable in the industry structural profit regression equation. I also used input output tables to obtain information about the share of output that is going to households (to determine consumer good producing industries) or the role of imports and exports for the industries' output markets. Those data have also been used to obtain indicators for market concentration for suppliers of and

---

[5] see e.g. Gschwandtner (2012) who used the data base from the Wooster College, available at http://www3.wooster.edu/economics/archive/indconc.html

buyers from the industry. The appendix describes the different data files assembled more closely, the manipulations undertaken and the details about the categories and units of analysis used.

2.2.2 Variables

Most studies rely on a single n-firm concentration ratio, $CR^n = \sum_{i=1}^{n} MS_i$, as a dependent variable that is usually picked in a relatively arbitrary way - based on which ratio delivered regression results that were significant or that were believed to be the most appropriate ones. It is commonly known that the Hirschmann-Herfindahl Index (HHI[6]) is a by far more appropriate indicator for describing market concentration, since it consists of the squared market shares of all firms and not just the aggregate of the largest four (eight, 20 or 50):

$$HHI = \sum_{i=1}^{N} MS_i^2$$

where MS is the market share of firm i and N is either number 50 of the largest or the number of all firms in the industry. However, relying on this measure does neither allow to include non-manufacturing data, nor manufacturing data before 1982 since the index has been calculated only for manufacturing industries starting in 1982. Keeping in mind that the HHI is the ideal measure of concentration, I extracted the maximum amount of information on market concentration from the four concentration ratios that are published for every industry and survey year by calculating an index that can be understood as the floor of the Hirschmann-Herfindahl Index. This Hirschmann-Herfindahl Floor (HHF) is the minimum value which the true and unknown HHI will have, given the four known concentration ratios of that industry. This means that the calculation assumes perfecty equal market share distributions within the groups of firms for which published concentration ratios reveal the groups' aggregate market share:

$$HHF = 4\left(\frac{CR4}{4}\right)^2 + 4\left(\frac{CR8 - CR4}{4}\right)^2 + 12\left(\frac{CR20 - CR8}{12}\right)^2 + 30\left(\frac{CR50 - CR20}{30}\right)^2$$

where $CR4$, $CR8$, $CR20$ and $CR50$ are the concentration ratios for the largest four, eight, 20 and 50 companies in the industry. In the case of perfect equality in the distribution of market shares among the four largest companies, the top five to eight, the next largest 12 as well as the group from

---

[6] Industry level variables have abbreviations in capitalized letters and segment level variables in lower case letters.

number 30 to number 50 in the industry respectively, the HHF will be equal to the HHI. As soon as there is a less than equal distribution within these four groups, they will differ. Not reflecting this inequality, the HHF contains obviously less information about the distribution of market shares than the HHI, but also much more than any single concentration ratio. In the file that includes all concentration data, the HHI was on average 15.5% larger than the HHF. In 90% of all manufacturing industries, the difference was smaller than 33.4%. For example, the 'Iron and Steel Mills' industry (NAICS 331111) had a HHI value of 0.0786 in 2007 and a HHF value of 0.0762 which corresponded to concentration ratios of 52.1%, 67%, 84.4% and 95%. While most regressions here are based on HHF, a test on whether it leads to significantly different results was easily obtained by running the final regressions on a restricted sample that contains only manufacturing firms after 1982 - once with HHI and once with HHF as the concentration indicator (see table 23).

I also ran a set of regressions to investigate into the existence of critical concentration levels. One is the HHI value of 0.1 (CCL2) which Abbasoglu, Aysan, and Günes (2007, p. 110) mention as being the commonly accepted border line for moderately concentrated industries (while 0.18 would be the line of separation towards highly concentrated industries). It corresponds to the 70% CR8 prominently quoted by (Bain, 1951). Scherer (1980, p. 280) summarizes empirical studies which found support for critical concentration ratios at CR4 values between 45% and 59%. The lower bound is used as the second critical concentration level option CCL1 (below CCL2).[7] I used the percentiles of the distributions of the two concentration indicators in the sample of 3,240 manufacturing industry observations for which there were HHI data available. Those percentiles were applied on the distributions of the respective other five concentration indicators in order to find the equivalent values that correspond to the discussed critical concentration levels. The critical concentration levels expressed in the six different indicators with the respective percentiles are reported in the table below (rows two and three). For comparison, it includes also the medians (row one) and the higher concentration threshold considered by the Department of Justice (row four) that could not be investigated due to sample size problems. Only the HHF values are the ones being used in the subsequent analysis.

To test other forms of non-linearity I included as an additional regressor the squared HHF (HHF2)

---

[7] Unfortunately, higher threshold values could not be analyzed in a meaningful way since they would have reduced the sample size too much (see table). The Department of Justice considers e.g. industries to be moderately concentrated if the HHI is above 0.15 (see http://www.justice.gov/atr/public/guidelines/hhi.html, reviewed 7/16/2012) and Stigler (1963) mentions a 60% CR4 that is ranked slighly below this value in the percentiles of the respective distributions. The presented results do thus only hold for concentrated, not "highly concentrated" industries.

Table 2: **Median, critical concentration levels and "high concentration" with the corresponding indicator values**

	Percentiles	*OBS.*	*HHI*	*HHF*	*CR4*	*CR8*	*CR20*	*CR50*
Median	50	1026	0.046	0.040	34.5	49.0	68.0	83.8
Crit. Level 1	67	456	0.073	0.061	**45.0**	60.7	79.0	92.8
Crit. Level 2	77.7	236	**0.100**	0.079	52.8	69.0	86.5	96.5
"High Con."	88.5	63	0.150	0.107	63.0	78.8	93.4	99.0

Percentiles of the distributions of each concentration indicator used to match the indicators. HHI: Herfindahl Index; HHF: Herfindahl Floor; CR4-50: 4-50 firm ratios, OBS.: number of observations above the percentile usable in the final baseline regression model.

and ran auxiliary regressions with the whole regression equation in log terms (see table 22). Even though it has been highlighted by authors (see above) and represents a refinement, the impact of concentration in the buyer and seller industries has never been tested for in a representative way for the economy as a whole. In order to investigate that issue, I used input output data on the detailed aggregation level to calculate the average HHF based upstream (supplier) and downstream (buyer) concentration along the value chain of each industry, $HHFu$ and $HHFd$ as

$$HHFu_i = \sum_{j=1}^{N} HHF_j * \frac{I_{ij}}{I_i}$$

and

$$HHFd_i = \sum_{j=1}^{N} HHF_j * \frac{X_{ij}}{X_i}$$

where I_{ij} is the input of intermediate goods of industry j into industry i, I_i the sum of all intermediate inputs into industry i, X_{ij} the output of industry i that is sold to industry j and X_i the total output of industry i. While the quality of the measures is not perfect[8], most problems could be diminished by reducing the sample to industries whose sales share to government and private industries is neglectable and which do not have large sales shares going to and inputs coming from industries for which no concentration data were available. The regressions include the vertical value chain concentration $HHFv$, while $HHFu$, $HHFd$ and the relative concentration $HHFr$ have also been tried. $HHFv$ represents the concentration along the value chain the industry is placed in and

---

[8] There are industries 1. for which no concentration data are published; 2. that sell a significant amount of output to government and private sectors; 3. that are significantly impacted by imports of intermediate inputs and exports of industry output; 4. that rely heavily on fixed capital inputs (capital intensive) which are not included in regular input output data.

$HHFr$ the relative concentration on an industry versus the industries it interacts with as a customer and a supplier:

$$HHFv_i = \frac{HHFu_i + HHFd_i}{2}$$

and

$$HHFr_i = HHF_i - HHFv_i$$

The concentration data considered in the regression analysis are listed in the following table:

Table 3: **Concentration indicators used in the regression analysis**

Variable	*Description*
HHF	Herfindahl-Hirschmann Floor (lowest possible HHI, given CR4, CR8, CR20 and CR50)
*HHI**	Herfindahl-Hirschmann Index
CR4, CR8**	four- & eight-firm concentration ratios
HHFv	HHF of industries up- and downstream the value chain
*HHFr***	relative HHF concentration; HHF - HHFv
*HHFu***	HHF for industries upstream the value chain
*HHFd***	HHF for industries downstream the value chain
CCL1	1st critical concentration level dummy; HHF: 0.061
CCL2	2nd critical concentration level dummy; HHF: 0.079

** Used in auxiliary regressions (instead of HHF); *** Used in auxiliary regressions (instead of HHF and HHFv).*

There are several barrier indicators used as regressors in this study, including advertising and R&D intensities for each industry i, AD_i and RD_i which represent annual Compustat Segment File industry averages of advertising and R&D expenditure/sales ratios. As another barrier variable I included a minimum efficient scale indicator, MES_i. It equals the average fixed capital requirements per establishment

$$MES_i = \frac{IndFC_i}{IndEstabs_i} = \frac{IndSales_i}{IndEstabs_i} * (Capital/Sales)_i$$

where the total sales and the number of establishments in industry i, $IndSales_i$ and $IndEstabs_i$ were obtained from Census concentration data publications while the industry average fixed capital / sales ratio $(Capital/Sales)_i$ was calculated from Compustat Segments data as the median value in each industry year. As an alternative indicator I used the logarithm of the industry average capital stock, K_i, which was approximated from Compustat data on segments' total assets. It could be the case that each of these barrier indicators taken for themselves has a small impact while their combined effect might be significant. Considering both as strategic investments, I thus used $ADRD_i = AD_i + RD_i$ and $ADxRD_i = AD_i * RD_i$ as alternative specifications that either interpret these investments as perfect substitutes or (positive) complementaries. For another alternative specification for the profit equation I constructed the barrier index

$$BAR_i = AD_i * RD_i * MES_i$$

I also tried alternative versions of combining the three dimensions as $BAR2_i = (AD_i + RD_i) * MES_i$ and $BAR3_i = AD_i^{st} + RD_i^{st} + MES_i^{st}$ where st indicates a standardization of the variable.

Table 4: **Barrier indicators used in the regression analysis***

Variable	Description
AD	Advertising intensity; adv. expenditures / sales
RD	R&D intensity; R&D expenditures / sales
ADRD**	AD+RD (strategic investment; substitutes)
ADxRD**	AD*RD (strategic investment; complementaries)
MES	minimum efficient scale; average assets per establishment
K***	industry year median total assets
BAR****	Barrier Index; AD*RD*MES
BAR2****	Barrier Index; (AD+RD)*MES
BAR3****	Barrier Index; AD*RD*MES (each variable standardized)

** All variables are on the industry level; ** Used in auxiliary regressions (instead of AD and RD); *** Used in auxiliary regressions (instead of MES); **** Used in auxiliary regressions (instead of AD, RD and MES)*

As mentioned earlier, the static approach can be interpreted in such a way that interaction effects between concentration and mobility barriers should be positively related to profitability: both dimensions isolated from each other might not necessarily have that effect if the ability to exercise market power depends on the number of firms / the ability to coordinate as well as the protection from newcomers that threaten to emerge if excess profit occurs. Accordingly, I included the interaction between the HHF and the barrier indicators BAR and ADRD, HHFxBAR and HHFxADRD.

Usually, industry structural profit equations do not include unionization as a dependent variable even though it is an established stylized empirical fact in labor economics that there is a significant negative relation between unionization and profitability. A negative association could be due to a direct distributional reduction of profits through stronger bargaining for higher wages and larger shares of the value added that is up for distribution or due to productivity reducing and cost inflating improvements in work conditions (safety, hours per week, flexibility, work intensity). This might happen especially in industries where competition takes mainly place on a cost basis (Kaufman and Hotchkiss, 2006, pp. 648-652, Hirsch, 1991, Chapters 4 and 6).[9] Given the presumed importance of unionization, an omission of the variable in a profitability regression should lead to disturbed results. Some economists argue that unions often pursue strategies to organize large companies, companies with larger market shares within industries and industries with entry barriers and that they might be more successful in organizing these industries or in exercising bargaining power with a given industrial unionization degree (Kaufman *et al.*, 2006, pp. 637-638). A positive correlation between unionization and market share, concentration and barriers could lead to systematic distortion of the size-profitability and market share-profitability relationships. Thus, an appropriate indicator must be included. All regressions contain the percentage of worker unionization on the industrial level (UNION).

Another control variable included here is a dummy for membership in a diversified conglomerate (cong) that is equal to one if company sales in other industries accounted for at least 30% of total company sales and equal to zero otherwise. This could have a positive effect if corporations share technologies and involved R&D efforts or safe transaction costs. It could also have a negative one if different industrial activities have strategic importance for the rest of the company and will not be abandoned even if profitability conditions in the industry erode. I also included a variable that indicates how geographically diversified the industrial activities of companies are. It is measured as the share of the segment's total sales that are realized abroad (forshare). It could be the case that companies that are active internationally realize additional returns to scale or are able to exploit geographically specific comparative advantages abroad. To approximate risk, I included the standard deviation of each segment's differential to the economy wide annual average over the period of observation of that respective segment, sd. Capital intensity measures for the segment (capint) as well as for the industry level (CAPINT) were used where the latter one was estimated from Compustat as the median of all units within one year and industry. I calculated it as a capital/employees ratio,

---

[9] There are also arguments that unions increase productivity in certain industries by exercising a voice for employees, improving work satisfaction and in turn motivations while reducing employee turnover (and creating incentives to invest into company specific human capital). Additional control over the management might reduce negative effects that are associated with a lack thereof. Still, it could be that the negative distributive effect on profits dominates any positive productivity effect.

since a capital/sales ratio (CAPINT2 and capint2) would be the inverse of the demand indicator used and not yield any additional information. Total assets of a segment (k) were used as a segment level control variable for size. The potentially most important control variables for concentration measures are mktshare and gmktshare, the absolute level and the percentage change in the segments market share. Demand indicators used are the output/capital ratio as well as the growth rate of it on the micro level (xk, gxk) and on the industrial level (XK, gXK), calculated as industry-year median values of sales/total assets ratios from Compustat. I preferred to use the growth rate of real total industry sales, GROWTH, instead of gXK as obtained from Census concentration data files since it is a more representative reflection of the industry aggregate (gXK was applied instead only to test if it produced different results). There are different ideas on how to interpret sale/capital ratios. They could reflect productivity differences where more efficient firms produce the same amount of output from different capital input volumes. A higher ratio would thus imply less efficiency in the use of resources. On the other hand it can be seen as an indicator for capacity utilization. In that case it could reflect fight back potential of incumbents against potential newcomers - i.e., a barrier. In this more stationary view on competition, higher values would imply lower barriers and long term profits. Alternatively, differences in capacity utilization might just reveal different demand situations for industries or firms and should be positively associated with the ability to obtain higher profits.

Additional industry level control variables used and obtained from input-output tables include industry exports and imports as shares of total industry output (EXP and IMP). These variables could reflect the international competitiveness of industries at least to some degree in those industries whose production does not depend on unique geographical conditions (as with certain mining or agricultural industries). Higher exports and lower imports should in this case be associated with higher profitability. It is commonly known that high exports and imports undermine the applicability of concentration data which has been obtained from national data (Cowling *et al.*, 2000). To account for that problem, auxiliary regressions were included where samples are restricted to industries for which IMP and EXP values were below 5% respectively (table 18, regressions 1 and 2). Similarly, the sample in auxiliary regressions was restricted to consumer good industries if the share of industry output going into private household consumption exceeded 60% (table 18, regressions 3 and 4).

For all regressors - those obtained from Compustat and those from Economic Census, CPI and BEA data - averages were calculated across the runs of consecutive segment-year observations for each segment. All averages were calculated from at least seven and on average more than 13 segment

Table 5: **Control variables used in the regression analysis***

Variable	Description
mktshare	market share of the segment
gmktshare	% change in mktshare
UNION	% of workers that are union members, by industry
cong	diversified conglomerate dummy; equals 1 if company sales in other industries account for at least 30% of its total sales
forshare	world market integration; % of sales realized outside of the US by the company in that industry
sd	standard deviation of segment profit differentials
CAPINT	industrial capital intensity; median of the capital/employees ratios of the US industrial activities in each industry and year
CAPINT2*	industrial capital intensity; median of capital/sales
capint	segmental capital intensity; capital/employees ratio
capint2*	segmental capital intensity; capital/sales
k	total assets of the segment
XK	output/capital ratio; industry year median
gXK	growth rate of XK
xk	segment output/capital ratio
gxk	growth rate of xk
GROWTH	growth rate of total industry sales
IMP	imports of industry products in % of industry output
EXP	exports of industry products in % of industry output

** Used as alternatives to CAPINT and capint*

*of AD and RD); *** Used in auxiliary regressions (instead of AD, RD and MES)* years each.

The following table summarizes segment and industry variables on the segment average level that are used in the second stage regressions and that turn out to be important. It includes the number of segment averages as well as arithmetic means of variables by sectorial groups.

Table 6: **Summary of segment & industry data, sectorial aggregates***

	Num.	roi %	HHF	UNION %	IMP %	k $	capint $	cong %	forshare %	xk
Non-Manufacturing										
Utilities	25	8.5	2.77	20.4	0.0	1302	1411	48.0	0.0	0.74
Wh. Trade	77	9.1	2.13	3.9	0.0	273	319	19.5	3.2	2.04
Ret. Trade	110	9.2	3.64	2.7	0.0	452	158	6.4	2.2	2.11
FIRE	47	5.8	2.64	1.7	0.3	528	897	10.6	2.9	1.16
Prof. Serv.	80	2.5	1.45	1.2	0.6	288	267	12.5	5.7	1.51
Publishing	205	-26.2	3.06	1.4	0.4	257	315	4.4	18.6	1.02
Misc Info.	45	0.6	6.32	9.4	0.2	1004	901	20.0	1.5	0.56
Misc Serv.	45	4.8	3.04	5.3	0.0	268	234	8.9	4.6	1.36
Manufacturing										
Food	81	15.5	5.70	19.3	5.0	305	229	16.0	2.3	1.70
Chemical	79	10.5	4.98	11.4	7.7	237	333	8.9	9.4	1.37
Pharmac.	216	-30.5	5.45	5.8	20.2	204	610	1.4	6.1	0.66
Plastic	59	3.0	1.80	10.8	8.7	194	183	16.9	7.9	1.35
Machinery	83	3.2	5.00	15.0	20.9	154	200	7.2	15.4	1.38
Metal	52	11.0	4.47	20.1	15.6	291	228	17.3	6.2	1.41
Computer	131	-10.8	6.83	1.9	43.4	239	314	1.5	13.4	1.31
Communic	114	-0.1	6.32	9.7	21.6	107	237	3.5	8.2	1.17
Semicond	117	4.6	5.21	9.2	24.8	263	285	3.4	15.6	1.14
Electronic	172	-12.8	3.56	9.9	20.5	98	242	4.7	11.0	1.06
Electrical	60	7.1	4.82	12.4	23.1	110	209	5.0	9.3	1.35
Transport	39	6.2	6.90	24.9	13.6	305	157	7.7	11.1	1.56
Furniture	32	14.4	2.90	13.1	15.7	131	72	6.3	2.1	1.80
Other	232	0.9	3.31	11.2	15.4	164	214	5.6	7.4	1.30

** Sectorial groups largely represent 5, 4 or 3 digit 2007 NAICS industry classifications. Values are segment and industry data associated with segments of a total of 2,101 segment averages (Num.). In the mean these segment averages represent 13 segment years each (over 27,000 total observations. from 1973 to 2010). HHF was multiplied by 100. k is in millions of and capint in thousands of USD. cong is expressed as the percentage of industry segments classified as conglomerate members. roi is return on investment.*

One can see how important and numerically significant sectorial groups are added in this study to the manufacturing sectors usually analyzed. The table shows that the variables values are of reasonable magnitude and variation. The sample of segments in food processing industries is the most and the sample in pharmaceutical industries the least profitable. The most concentrated industries are part of the transportation as well as computer equipment producing sectorial groups while information services are also fairly concentrated. Segments in heavy and traditional manufacturing industries centered in the north east are highly unionized while trade and especially service sectors are at the lower end of the spectrum. Imports are essentially irrelevant for non-manufacturing and highest for the computer equipment producing sector. The utility sector tops the list when it comes to segmental conglomerate membership, firm size and capital intensity - with over $ 1.4 million in assets per employee. Within manufacturing, the pharmaceutical sector includes the segment sample that is most capital intensive while food processing and transportation equipment manufacturers are fairly large with respect to total assets. Trade and furniture manufacturing sectors have high output capital ratios. International sales are more important for segments in publishing, semiconductor and machinery producing sectors.

Table 7 presents the pairwise correlation coefficients between $\hat{p}_i$, $\hat{\lambda}_i$, $\hat{\alpha}_i$ and a sample of explanatory variables for all observations from the baseline regression in table 16. It shows that profit persistence, $\hat{\lambda}_i$, is negatively correlated with the average profitability differential of a spell of data, $\hat{\alpha}_i$. HHF and BAR are negatively and mktshare positively correlated with $\hat{p}_i$. Some control variable have correlations with $\hat{p}_i$ of a significant magnitude, including UNION, EXP, IMP, GROWTH (negative) and XK, k, cong and forshare (positive). HHF has higher correlations with GROWTH, XK (negative) and with HHFv, UNION, EXP, IMP and mktshare (positive). mktshare is correlated with EXP, GROWTH (negative) and with UNION, XK, k and forshare (positive).

Table 7: **Pairwise correlation coefficients matrix**

	$\hat{p}$	$\hat{\alpha}$	$\hat{\lambda}$	HHF	HHFv	BAR	UNION	EXP	IMP	CAPINT	GROWTH	XK	mktshare	k	cong	forshare
$\hat{p}$	1.00															
$\hat{\alpha}$	0.82	1.00														
$\hat{\lambda}$	-0.17	-0.32	1.00													
HHF	-0.08	-0.09	-0.02	1.00												
HHFv	-0.03	-0.04	-0.04	0.34	1.00											
BAR	-0.11	-0.10	-0.01	0.06	-0.10	1.00										
UNION	-0.01	-0.04	0.12	0.18	0.27	-0.08	1.00									
EXP	-0.09	-0.12	0.04	0.15	0.35	-0.02	0.06	1.00								
IMP	-0.11	-0.14	0.04	0.24	0.19	0.15	0.05	0.47	1.00							
CAPINT	0.03	0.09	-0.16	0.10	-0.02	0.24	-0.05	-0.06	-0.03	1.00						
GROWTH	-0.11	-0.15	0.12	-0.17	-0.22	0.05	-0.19	0.03	-0.25	-0.02	1.00					
XK	0.10	0.08	0.09	-0.14	-0.05	-0.34	0.13	-0.25	-0.22	-0.47	-0.12	1.00				
mktshare	0.18	0.20	0.03	0.16	-0.02	-0.08	0.18	-0.12	-0.04	-0.08	-0.10	0.23	1.00			
k	0.21	0.26	-0.09	0.06	-0.03	0.00	0.01	-0.13	-0.13	0.27	-0.02	-0.07	0.43	1.00		
cong	0.15	0.23	-0.17	-0.04	-0.01	-0.03	0.05	-0.10	-0.11	0.17	-0.10	0.03	-0.02	0.05	1.00	
forshare	0.08	0.09	-0.07	0.06	0.14	-0.05	-0.10	0.26	0.11	-0.04	0.01	-0.16	0.10	0.10	-0.18	1.00

Includes the pairwise correlation coefficients of all observations used in the second stage benchmark regression equation.

2.2.3 Regression Equations and Results

In the calculation of $\hat{p}_i$ from $\hat{\alpha}_i$ and $\hat{\lambda}_i$ all observations were excluded where the spells of consecutive segment observations counted less than seven, since values lower than that are unlikely to yield reliable AR(1) estimates. All segments with $\hat{\lambda}_i \geqq 0.95$ or $\hat{\lambda}_i < 0$ were dropped as the calculation of $\hat{p}_i$ becomes imprecise when $\hat{\lambda}_i$ is close to and not defined if $\hat{\lambda}_i$ is exactly equal to one (a random walk) or impossible to interpret economically if $\hat{\lambda}_i$ exceeds one (implying explosive behavior and $\hat{p}_i$) or turns negative.

Since the second stage regressions are based on a dependent variable that has been obtained from a previous estimation, White's heteroskedasticity consistent standard errors were used in order to adjust for heteroskedasticity that is likely to result from such a procedure. According to Lewis (2000), his method is superior to the usual approach to the problem that has been applied in other persistence of profit studies (e.g. Gschwandtner (2012); Yurtoglu (2004)) and that follows Saxonhouse (1976) in weighting the regression by the inverses of the sampling standard errors of the dependent variable.[10] A benefit of using White standard errors is that one could go ahead with OLS estimations and assume heteroskedasticity to be taken care of already. Thus, tests for forms of heteroskedasticity and questions about whether to change the functional form or apply FGLS as alternative responses to the problem were omitted. To reduce disturbances from outliers all observations beyond the 3rd and 97th percentiles were dropped for each micro level variable obtained from Compustat. The 1st and 99th were used for variables computed from aggregate data.

Each regression was tested for non-linearity using Ramsey RESET tests. For all equations, these tests failed to reject even at 10% the null that powers of predicted values of the dependent variable offer no additional explanatory value when added to the regression with the original predictors. It could thus be assumed that non-linear terms of the predictors would be insignificant and that the analyzed relationships were linear. Still, the squared value of HHF was included and log regressions were executed as alternatives.

Before running any regressions, I checked for multicollinearity by reviewing an extension of the correlation matrix reported in table 7. It did not indicate problems of extreme multicollinearity for any variable pair but HHF and HHF2 as well as BAR and HHFxBAR (correlations of 0.94 and 0.92),

---

[10] This latter approach is likely to lead to inefficient estimates and underestimated standard errors.

which is not surprising since one variable has been calculated from the other. The same is the case for industry averages variables calculated from Compustat segment level variables: XK and xk (0.73) and capint and CAPINT (0.55). Variance inflation factors were all below 10, with the exception of HHF and HHF2 (BAR and HHFxBAR, XK and xk, capint and CAPINT had values above five) which dropped to 1.5 when HHF2 was omitted. HHF2 was not included in the regression equations reported here due to insignificance in all cases and the inclusion of log linearized regressions included in table 22. All regression output tables are included in the appendix.

As the general specification of the baseline regression model (see table 16, regression 1) the choice was

$$\begin{aligned}\hat{p}_i = & HHF + HHFv + BAR + HHFxBAR + UNION \\ & + EXP + IMP + GROWTH + CAPINT + capint + mktshare \\ & + gmktshare + sd + cong + forshare + xk + XK + k + constant\end{aligned}$$

Departing from this specification, the number of regressors was reduced by stepwise eliminating insignificant ones at the 10% significance level (the one with the highest p value) until all regressors were significant. This technique has been applied in all subsequent regression equation variations. Result tables accordingly always show both the general and its specific reduced regression equation. The variables HHF, BAR and HHFxBAR were evaluated together and deleted if all of them had p-values above 0.1 since the elimination of the interaction term, or of one of the primary variables from which it was calculated, changes the interpretation of the remaining two variables. The resulting reduced equation for the baseline model (table 16, regression 2) is:

$$\begin{aligned}\hat{p}_i = & HHF + BAR + HHFxBAR + UNION + IMP + GROWTH \\ & + capint + mktshare + gmktshare + sd + cong + forshare + xk + XK \\ & + k + constant\end{aligned}$$

These first two regressions in table 16 reveal negative significant coefficients for HHF. BAR is significant and positive while the interaction term survives the elimination process as a highly significant negative variable. The regressors mktshare and gmktshare are positive and highly significant (as are cong, forshare, xk, and k). IMP, XK, GROWTH and UNION are significantly negative. Departing from the baseline regression specifications, the sample was restricted to segments with spells of consecutive observations lasting for at least 10 years, since a period of seven years might have been insufficient in the first stage regression. Starting from the full specification, a similar restricted equation as before resulted from the stepwise deletion process. The only differences are a slightly higher R^2, insignificance of GROWTH and a close insignificance of HHF at the 10% level (see regressions 3 and 4 are in table 16).

The problem with the use of barrier variables was that it was far from clear which the appropriate ones are that should be included in the regression model. Due to this uncertainty, I tried different specifications for the baseline model without interaction effects, using different sets of barrier indicators and combinations of AD, RD, MES, K, ADRD, ADxRD, BAR2 and BAR3 (not reported). Applying a stepwise deletion technique, MES got eliminated in the very first steps, while AD and RD also did not survive either. The only three variables that did survive as significant ones at the 10% level from any initial combination were ADRD, BAR and BAR2 (all having negative coefficients). BAR was the most significant one and resulted in the highest R^2 while ADRD was also highly significant. In both of these specifications the overall results were comparable (with respect to the set of surviving variables, their coefficient signs and significance levels). I thus went on with BAR as the barrier indicator in the baseline regression, but included an alternative pair of regressions in table 17 with ADRD and MES instead of BAR (regressions 1, 2). Unlike in the baseline model, the inclusion of the concentration barrier interaction term pushes the alternative barrier variable ADRD into insignificance. MES has a significant negative impact and HHF becomes insignificant. The table also includes versions of the baseline regression model (3, 4) that contain ad and rd as additional control variables but were omitted in other equations since their use would have reduced the samples by half. Ad and rd are both significant and have negative coefficients while the R^2 of the regression increases slightly.

Since imports and exports have the potential to invalidate concentration indicators and other industry structural parameters, the sample was restricted to industries where imports and exports accounted for less than 5% of the industry output (see table 18). As authors have found different values for concentration and barrier variables in the past when analyzing consumer good producing industries, I also tested for this by including only sectors where at least 60% of their output goes into private consumption (table 18). Differences to the baseline model are that in the case of the domestic industries regressions the R^2 drops and the variables GROWTH, BAR and HHFxBAR become insignificant in the restricted model while HHF and the market share indicators do not change. This implies that the impact of trade is not so important in the U.S. to dramatically disturb the regression results with respect to concentration in the sense that a positive relationships appear to be negative or insignificant. Import penetration is certainly not causing otherwise significant indicators of concentration and industry barriers to appear insignificant. In consumption good industries, HHF, BAR and HHFxBAR vanish. Trying different specifications of barrier variables (not shown) with AD, RD, MES, HHFxADRD as well as ADRD, HHFxADRD as alternatives for BAR and HHFxBAR also lead to eliminations and insignificance of all of these variables. The same holds when ad and rd are added. One could thus assume that consumer good industries are not revealing a dramatically different pattern with respect to the effect of industry barriers and concentration.

Table 19 displays the output for regressions whose samples are restricted using CCL1 and CCL2. While their application raises the R^2 of the regressions, HHF, BAR, HHFxBAR and mktshare become insignificant, not providing any evidence for critical concentration levels. A problem is here the extreme reduction of the sample size to 456 and 236 observations. Two other methods of testing for critical concentration levels were applied in tables 20 and 21. The first is equal to the baseline regression model but includes a dummy variable (CCL) that is zero if the HHF is below the critical value and one if it is greater. In the second one the multiplicative interaction between this dummy variable and the HHF (HHFxCCL) was included instead.

Table 22 shows the results of altering the specification to a logarithmic regression equation. The restricted regression eliminates more control variables and shows a clear insignificance for HHF. Table 23 reports the results of a restriction of the sample to the manufacturing sector. The second set of regressions has HHI substituted for HHF to see if the choice of HHI as the superior concentration indicator makes a difference. This is not the case with respect to the coefficient signs, significance levels of HHF and HHI or the R^2. BAR is insignificant in all regressions and the interaction variable HHIxBAR reaches a higher insignificance level in the equations that include HHI. These very similar results imply that HHF is a robust approximation of HHI and summarizes industrial concentration well.

The focus of the analysis lies clearly on $\hat{p}_i$. To investigate the determinants of $\hat{p}_i$ I however decomposed the measure into $\hat{\alpha}_i$ and $\hat{\lambda}_i$. Table 7 revealed a negative correlation between the two components. I ran regressions with both of them as dependent variables. Table 24 reports regressions with the left hand side variable being $\hat{\alpha}_i$ (regressions 3 and 4) and the average profit differential calculated as an average over the respective data spell (regressions 1 and 2). The results are very much in line with the ones from the regressions on $\hat{p}_i$. Compared to the baseline model in table 16, significance levels of variables and coefficient signs are equivalent while the R^2's are slightly higher. Table 25 reports the baseline regression model that was ran on $\hat{\lambda}_i$ as the dependent variable. The first two columns include the full sample. The middle includes all observations with positive and the last two columns segments with negative profitability differentials (as indicated by $\hat{\alpha}_i$) to see if there are asymmetrical patterns. The R^2 drops by two thirds in these type of equations below 12 % and most variables become insignificant. UNION, IMP, GROWTH and XK are all significant and positive in regressions on the whole and on the positive differential sample, meaning that they contribute to profitability persistence of positive differentials. This could e.g. indicate high degrees of persistence of favorable demand conditions which slow down the supply side and thus also the profit rate differential adjustment process. In the same two samples the variables CAPINT, cong and forshare are negative and significant which might reflect that segments which are part of conglomerate and international company structures are faster at supply adjustments which should ultimately lead to faster profitability convergence. HHFv is the only variable significant (with a negative sign) in all three samples which could mean that higher concentration along the value chain erodes excess profits quicker due to easier entry for larger firms. However, one needs to keep in mind that these described relations are very weak due to the low explanatory power of all six regressions in table 25.

2.2.4 Interpretation

As with most industry structure-persistent profit differential regressions, the explanatory value of the regression is fairly low and the R^2 shows that usually around one third and never more than 47.5% of the variation of $\hat{p}_i$ can be explained. Most of the variation has thus to be explained from unobservable factors for which the quantitative data are not available or observable. They might be of a more qualitative nature and cannot be accounted for in an econometric analysis or are related through different forms of interaction. Nevertheless, the explanatory values of the regressions are higher than the ones found in other studies (see above) that were based on much fewer observations and less detailed data. Given the problems associated with estimations of variables like MES as well as errors and imprecisions in the data, R^2 values between 30% and 45% are not surprising and should be seen as relatively satisfying.[11]

The following table summarizes significance levels and coefficient signs of a selection of variables in the regressions reported below.

Table 8: **Regression results summary***

Segment Variables	Significance		Coef. signs		**Industry Variables**	Significance		Coef. signs	
	at 10%	at 5%	+	-		at 10%	at 5%	+	-
xk	23	23	23	0	UNION	22	18	0	23
k	24	22	24	0	IMP	7	5	0	19
mktshare	23	21	23	0	BAR	10	2	15	7
cong	23	23	22	1	HHF	14	12	2	20
forshare	23	23	23	0	HHFxBAR	16	16	2	20

** Includes variables from the 24 regressions on* $\hat{p}$ *and* $\hat{\alpha}$ *in the appendix. From critical concentration level equations only results from the one based on HHFxCCL are included (high similarity). Numbers indicate in how many regressions a variable was significant or in how many it had the respective coefficient sign (and was not eliminated).*

In the majority of the reported regressions that are summarized in the table, concentration as measured by the HHF (or the HHI) is negatively and significantly related to persistent profit rate differentials on the segment level. In only two regressions the coefficient switches signs (but is also insignificant). HHFv is generally insignificant [12]. HHFxBAR is generally negative and significant in 16 out of the 24 regressions illustrated in the table.

These results of a negative significant and in some cases insignificant concentration profitability relationship represent empirical evidence against neoclassical mainstream and the concentration doctrine that expect stable and significant positive relations between concentration indexes and profitability from such regressions. Evidence points to exactly the opposite: the profitability of a firm is more likely to be lower if an industry has a higher market share concentration. This is especially true when mobility barriers are jointly present. Higher concentration leads to more, not less competition.

---

[11] Concentration data are e.g. published every five years. It is sure that some official industry classifications do not always group firms together into units that actually compete against each other with their products. NAICS industries are based on the use of similar production technologies, where it might make more sense to form groups based on the production of close substitutes. First Solar for example is clearly no direct competitor of Intel, but both are in the same industry group, semiconductor and related device manufacturing (NAICS 334413). The lowest level of aggregation is in many cases not detailed enough to group competitors. Some Compustat Segment data are also questionable as a seemingly huge degree of freedom allows companies to follow different reporting standards. Aggregates are often too high: Exxon has in 2009 activities in only two different industries listed, crude petroleum and natural gas extraction (NAICS 211111) and petroleum refineries (NAICS 324110), while its downstream operations make it one of the world's largest players in the chemical industry.

[12] HHFr, HHFu and HHFd did not yield more significant results.

The barrier variable BAR is insignificant in the majority of regressions. It tends to be positive in some but switches signs in one third of the estimations. An interesting pattern is that once the usually negative HHFxBAR is being dropped from the equation (not reported), the sign of BAR becomes negative too. Alternative barrier indicators such as AD, RD and ADRD are less significant (ad and rd are significant and negative). Thus, there is no clear evidence for any significant and positive effect of mobility barriers on profitability.

The variables mktshare, gmktshare and UNION all do not change their signs and achieve in nearly all regressions high significance levels (often at 1%). This yields support for the earlier mentioned suspicion that industrial unionization has important distributive and at least in some American industries negative productivity related impacts that need to be controlled. The performance of mktshare implies that there could be very well a positive market share-productivity linkage that explains the positive market share-profitability relation. These results confirm the findings from most other studies. This offers support for views that argue for a positive market share-competitiveness link (the two dynamic competition theories and the Chicago School). That innovating firms reduce costs and are likely to grow and make more profits is a possible explanation for the regression results. It could be stated within the dynamic frameworks that such a process reflects a trajectory or virtuous circle between innovation and profitability which can have long lasting effects. That excess profits do not die off after 10 years does not imply that there is a lack of competition. Just the other way around, there could be a behavior that is more, not less competitive than the one of rival companies or business segments. The positive and significant variable gmktshare supports the evidence about mktshare. It might capture the change of competitiveness of a business unit over time, since the general market demand trend is already captured by the growth of industry output, GROWTH.

The industry level control variable IMP reveals consistently a negative relation which however is not significant in several regressions. It could be a weak indicator for international competitiveness of U.S. businesses where a higher import penetration by foreign, potentially more competitive producers that attempt to enter or expand in the U.S. market might squeeze profits.

The firm level control variables cong, forshare, k and capint have positive significant coefficients in a relatively consistent way throughout the analysis, implying that conglomerate membership, international positioning, scale and fixed high capital intensity might allow segments to exploit comparative advantages that can be sustained for longer periods of time. In the case of cong and forshare, advantages could be due to shared technologies or costs between segments of a company or originated in the exploitation of factors of production that are not available within the U.S. (like lower wages abroad). The variable sd is mostly significant and has negative coefficients which is puzzling but also just might mean that the standard deviation of the profit rate differentials is a very bad indicator for risk. For example, in the case of a stable rising differential that is observed over the sequence of segment years, sd would be higher than for a segment with constant negative returns. It also does not say anything about the risk of bankruptcy or the rate of failure of firms within an industry, its financial risks (high leverage) or the potential loss of value involved with sunk costs in a bankruptcy (this could be reflected in variables like capint or k). The variable xk with its positive significant impact is likely to capture the demand situation and capacity utilization of the firm.

With significant positive market share coefficients and negative significant concentration-profitability links, evidence is not on the side of traditional static approaches and instead offering more support for dynamic competition theories and the efficiency hypothesis. However, it is not immediately clear from these latter views why concentration should be negatively related to permanent excess profits and why industry level demand indicators like GROWTH and XK are largely significant with negative signs as a rise of industry sales should imply a general improvement in demand which, if it has any effect, should support profits.[13] Other authors that previously found similarly negative and in some cases significantly negative concentration profitability relationships (Ravenscraft, 1983; Sass, 1975; Gale and Branch, 1982; Semmler, 1984) have not aimed at explaining their findings. Their results were generally weaker and much less significant throughout their sets of regressions reported. They were in most cases interpreted as evidence for the concentration doctrine link being weak (e.g. Ravenscraft, 1983). An exception is Semmler (1984) who sees clear empirical evidence against mainstream approaches and support for classical views on competition.

An explanation could be the following: according to the positive market share-profitability link as implied by dynamic theories, a successfully competing and innovating firm introduces new products that have an improved quality or are produced with new production techniques that reduce costs of production. This will have two effects: 1. an increased profitability of that firm due to temporary monopolies or individual cost differentials versus the industry average and 2. an increase in sales and firm size due to the appropriation of previously unserved markets or of market shares from

---

[13] The effect of XK is in any case evidence against the views that interpret industrial excess capacity as representing fight back potential to defend surplus profits against potential entrants.

competitors due to product quality advantages or lower prices. This means that the industry average costs of production, prices of production and average market prices for industry output will fall. Within industries, higher quality or lower cost products of one firm mean for this firm's competitors that they will lose market shares and/or profit margins. In order to sell inferior quality or higher cost products in such a situation, lower prices need to be offered to customers that reduce profit margins unless a drop in sales is being accepted by these competitors. Both, reduced sales at constant margins as well as narrowed margins at constant sales lead to lower profit rates on investment or on assets. The success of one competitor that increases its profitability lowers at the same time the profitability of the remaining industry members. If one or more firms in an industry are competing successfully and their market shares are growing, then the concentration ratio of that industry increases. Thus the logic behind the positive market share effect implies that concentration is an indicator for the competitiveness of firms within an industry. Keeping everything else constant, tougher direct competitors with lower cost structures produce lower average selling prices for the industry output that imply lower profits for any given firm. This determination of the industry price is a problem in a perfectionist neoclassical frame where the price depends on the marginal least efficient producer. In the Classical approach however, the production price of an industry product is given by the average conditions of production which are dependent on the exact relative proportions of the production techniques with different cost structures that exist in an industry. Techniques with average conditions of production determine the prices if they make up most of the industry output or if volumes from below and better than average conditions cancel each other out. If neither of that is the case, worse or better than average conditions will regulate the market price in that industry. As industry average cost structures and market prices are positively related, a negative market shares-cost structure link implies that high concentration is linked to low prices: in concentrated industries above average conditions of production dominate the industry cost average, determine the price of production and the market price. On these prices depends if there will be any excess profit for firms that are utilizing more productive techniques, how large the surplus will be and how big the losses for above average costs of production techniques are.

As this explanation requires that there are no instant or fast equilibria and no perfect capital good markets, it is not compatible with Chicago School views. This approach implies that superior efficiency is rewarded no matter which industry a company belongs to. Instead, incomes of economic actors must be in some degree dependent not only on their own actions, but also on factors beyond their immediate control and own will. The dimension industry needs to matter in some way.[14] In a Classical disequilibrium frame fixed assets are largely valued at costs of production while producers can be locked into their industry through fixed asset intensities and durabilities - exposing them to industrial cycles and changes in the competitiveness of rivals. This "industrial capital lock" that is a part of the classical story of surplus profit allows to explain that higher market concentration can be negatively associated with profitability.[15] It explains that the impacts of these structural factors on profitability can persists at least in the "medium" run which may last for decades in many industries. The impact of these barrier to mobility dimension explains why the interaction variable HHFxBAR is even more significant than HHF.

With this interpretation one might also attempt to explain why the GROWTH and XK variables are negatively associated with persistent profit differentials. They capture only the change in and levels of sales for the rest of the market, since gmktshare and xk are included and already capture the sales growth for the individual firm. If rising sales are associated with an increase in competitiveness of firms within an industry, higher GROWTH and XK values will imply improved cost structures and better product quality of competitors within that industry. Again, lower average costs and prices of production imply lower profits or higher losses for a given capitalist firm.

Of course this explanation does not imply that rising concentration always and necessarily leads to higher degrees of competition. One needs to consider that R^2 values do not exceed 47.5% and that concentration is insignificant in some regressions. It rather means that given the current institutional structure in the U.S. concentration does not automatically need to imply oligopolistic coordination or a more favorable environment for firms. Holding everything else constant high concentration is instead more often associated with greater competitive intensity. This needs to be expected and can consistently be explained if classical surplus profit and profitability differential theory is true. Moreover, a simple 'quantity theory of competition' is shown to be insufficient for judgements about

---

[14] It is also not clear how other approaches within the static tradition (that are also critical of the concentration doctrine and entry barrier hypotheses) would explain this outcome. Why for example should the costs to defend an oligopoly or the performance losses due to X-inefficiency be so high that they do not just erode excess profits but lead to profit levels that are even lower than the ones of 'competitive' industries?

[15] An alternative explanation could be that concentration is an indicator for the maturity of an industry. The observed relation results when the maturity reflects a life cycle stage where competition evolved in such a way that the more efficient (and profitable) companies grew over time more than the less efficient ones. This would result in more mature industries being dominated by fewer, larger, more competitive ones. If this industrial evolution leads to an abandonment of high cost technologies and a spread of low cost methods throughout the industry, remaining companies would end up to be more similar. Even determined by the least efficient producer, market prices will fall as this process develops. The result is a greater likelihood for a randomly picked member from a more concentrated industry to have returns that are below average. Thanks to my colleague Gregor Semieniuk for discussions and ideas on this point.

market power. Market power and threats to competition can still exist but their detection requires an analysis of qualitative and of industry specific factors that do not fit into an econometric framework which is applicable to the economy as a whole.

2.3 Summary and Conclusion

I presented a data set that allowed to analyze business segment level information as opposed to the usually used company data. It included the most extensive amount of information on concentration which has been used for any empirical study on profit differentials in the U.S. so far. Analyzing different specifications of regression equations, alternative samples and including new indicators, I found that market share is positively and concentration significantly negatively associated with persistent excess profit - especially when high concentration is combined with barrier indicators. There was no convincing evidence that non-linear, interactive specifications or restrictions that allow to focus on consumer goods industries or truly domestic sectors alter these observed results dramatically. I did not find any hints at critical concentration levels. Previously ignored control variables such as industrial unionization turn out to be important. I interpreted these results as providing evidence against the more traditional, stationary views on competition. They are seen as as strong support for the dynamic approaches that are critical of the concentration doctrine and promote the idea of positive market share and profitability-productivity linkages which underly a positive market share-profitability relation. I argued that the hypothesis of a positive market share-productivity correlation can imply in a classical context where ‘industry matters’ that there is also a negative concentration-profitability link like the one observed here. Classical political economic theory is thus able to explain these results which appear counterintuitive on first sight. The negative link results because higher concentration is correlated with having more productive rivals within an industry. Their lower costs of production (correlated with market share size) lowers the average price levels in an industry - i.e. increases the competitive pressure which reduces profitability of any given firm. The conclusion is that industry structure does matter - but in a very different way than commonly believed. The competition theory of classical political economic thought is better suited than the alternative approaches to explain economic reality of modern U.S. capitalism.

Chapter 3

Industrial Change and Costs of Production in the German Electric Power Industry

The first chapter has described classical competition theory as a consistent alternative to other theoretical approaches. Chapter two has shown how it is better able to explain empirical results than others when it comes to the determination of profitability differentials. This chapter pursues an industry analytical application, a case study that is methodologically inspired by classical ideas and that targets a greater consistence with the classical approach than other related studies that rely on mainstream analytical frameworks.

The electric power industry is experiencing dramatic changes. One new development that takes place on a world scale and that is not only restricted to the U.S. is the growth of electric energy production from renewable sources and the spectacular fall of energy generation costs from wind and photovoltaic (PV) sources. It is known that such a transformation in which these intermittent energy sources grow significantly brings a lot of challenges to the industry when it comes to the extension of transmission networks, the augmentation of electric energy storage or the introduction of smart grid technologies. Of paramount importance is the alteration of operating regimes of conventional power plants. They are forced to run under more flexible and irregular modes than they did in the past in such a way that even former baseload plants need to be operated in aggressive load following mode or have to act as temporary cold reserves in start-stop operation. This involves start-up and various cycling costs which change the economics of all electricity sources and add to the costs of an industrial transition towards renewable sources. This study analyzes how the relative cost competitiveness of plant technologies change within such a transformational process.

While several countries and some states in the U.S. pursue these paths towards a green economy, Germany is the country with the most ambitious plan. It is in a very advanced stage today as renewables contribute already more than 20 % of the total electricity generation. The country is suited very well as an analytical example because it does not benefit from exceptional climatic or geographic conditions that ease the transformation towards a green electric power industry as they are available in other countries: annual sun radiation is only moderate; onshore wind conditions are average with a very high population density; coast lines are short relative to the size of the country; hydroelectric energy accounts for a comparatively low share while additional potentials are virtually nonexistent. This industry analytical study thus relies on the German example.

The paper is inspired by a dynamic and classical economic view on competition which is, however, not much more than a loose guide for this research while the theory itself is neither discussed nor tested. The classical view is reflected in:

1. the focus of the study on costs of production. a) Model outcomes are used to assess different production techniques that form the array or mode-of-production hierarchy with respect to costs. This hierarchy exists in any industry, is expected by classical theory to be in the core focus of companies in their intra-industrial competition as it is the primary determinant of surplus profit. b) Variable costs of production are the prime variable within the simulation procedure that determines plant behavior;

2. the goal of the paper to analyze industrial change and its effect on the economics of production methods as these issues make up the core interest of any dynamic approach of competition as it can be found in classical or Austrian theory;

3. the critique of and the modeling differences to existing related studies that need to be judged as unsatisfying from a classical perspective. They assume perfect foresight, rely on marginalist calculation, concepts of perfect or imperfect competition and other ideas that are not part of the classical theoretical frame;

4. the emphasis on hard empirical and technical industry data of the supply side (plant types

> and technological characteristics) and the demand side (system load profiles) that does not abstract from or simplify issues related to capacity utilization or fixed capital stocks.

I am analyzing the alteration of operating behaviors, capacity utilization factors and levelized costs of electricity (LCOE) from the different conventional power plant types that are available today. Costs are compared between individual conventionals and between conventionals and renewables. Average electricity production costs of the sum of all conventionals are calculated. The assessments are undertaken by the means of a simulation model that is applied to past and future scenarios of the years 2010, 2025 and 2040. For each of these three time scenarios three different fuel and CO2 price configuration scenarios are used. Six more scenarios are analyzed that involve different capacity mixes and reflect a (more) nuclear intensive and a lignite free industry structure. This simulation utilizes 15-minute interval data for four years on total electricity consumption, load from wind and from PV installations in Germany - a total of 140,260 observations. Using such long time series of actual data constitutes a major difference to studies that simulated these load profiles entirely from scratch or used extremely short periods of just four weeks or less. The simulation models the load profiles of electric power generating technologies on the individual plant level according to institutional and technical determinants as well as cost-based commercial operating principles. A second important difference to related previous studies is that the model applies more realistic representations of the costs of start-ups and of load following operations as well as of technical limitations of power plants. It focuses exclusively on costs of production, demand and institutionally-technically determined load profiles while other papers model electricity generation by using simulated market prices (which require fairly strong assumptions) and models of perfect, Cournot or Stackelberg competition.

One aim of the paper is to prove that the effect of renewables to increase certain electricity generation costs is not as dramatic as many industry observers suggest. It also tries to confirm two previously found results (Keil, 2013). The first one is that it is not required to subsidize fossile fuels (at least in the long run). German utilities have demanded these for mainly natural gas fired plants and claimed that they are necessary incentives for the provision of sufficient back-up reserve capacity. The second one is that less flexible and/or more polluting conventional plants are less suited to play a constructive role in an industry that moves towards renewable energy than cleaner natural gas plants - the nuclear power industries and coal intensive utilities still claim that nuclear and lignite plants are useful in such a transition period.

3.1 Review and Critique of other Studies

Studies that model electric load from different technologies, companies or individual plants in the electric power market usually work with linear optimization models or with game theoretic models of imperfect competition. Siohansi (2011) applies a Stackelberg model to simulate the impact of increasing wind capacity and of storage on the ERCOT market. Wind generators act as Stackelberg leaders and conventional power producers submit supply functions that either maximize profits (for a group of 'strategic' generators) or reflect marginal costs ('competitive fringe' generators). Aims are to analyze how wind and storage effect the modeled electricity price and the assumed 'market power' of conventional generators. An optimization method that is methodologically closer than other ones to the simulation applied here is formally outlined in (Kuntz and Müsgens, 2011). It computes the optimal production schedule from a set of available technologies with respect to the minimization of variable and start-up costs and subject to the constraints that a perfectly inelastic exogenous demand profile is satisfied. The advantage from a classical perspective is the omission of questionable mainstream competition models and the reliance on technical cost of production determinants. However, this approach involves several methodological and technical details that need to be altered from a classical perspective (while it is very well suited from the point of view of mainstream neoclassical economics).

1. In their application of the presented algorithm, Kuntz and Müsgens (2011, p. 24) abstract from the *capacity limits* that exist in every industry. As fixed capital supply adjustments cannot be assumed to take place instantaneously and companies are at a given point of time and in the short run 'stuck' with the plant capacities they have, this needs to be accounted for in a model from a classical point of view;

2. Complementary to this issue is the need to assess the *average utilization* of plant capacities. Only this allows to analyze the full cost competitiveness of a technology (e.g. via LCOE calculations) and its relative cost competitiveness versus alternative methods of production. Not considering existing capacities or assuming full load operation of all used units (as this is done in several studies) make this impossible;

3. In their application, Kuntz and Müsgens (2011) abstract from the operation of plants under *partial load*, which might make sense if economic units (in this case: power plants) are assumed

to be marginally small or to be constantly operating at or close to normal capacity. Classical economics does generally not rely on such a significant abstraction from reality. Thus, realistic plant sizes need to be modeled while temporary patterns of operation below the desired rate of capacity utilization must be accounted for with respect to their cost effects. They are well documented in the power plant engineering and commercial operating literature and include a) additional maintenance or fuel costs due to deviations from the technically optimal output level and b) from changes in the utilization rate that produce thermal and other stress on the equipment as well as c) negative effects on total unit costs that result from a given amount of fixed costs being divided among a smaller amount of output units);

4. Straight linear optimization suffers from the inclusion of *start-up costs* that add a fixed component ('unit commitment problem'). This produces complications with respect to the computational effort and the interpretation of variables. Two popular solutions for that are assuming incrementally small plant units or approximating start-up costs by not including them as a fixed component but by assigning costs to any positive load gradients (Traber and Kemfert, 2011). These answers are not satisfying as they make it impossible to discriminate between the just mentioned additional costs that result from the often significantly important load following activity (partial load operation, ramping of plant output) on the one hand and start-stop operation on the other hand. They are computationally different and known to involve cost impacts that have a fundamentally different nature (Lefton and Hilleman, 2012; Kumar, Besuner, Lefton, Agan, and Hilleman, 2012). Careful modeling of these technical cost details of flexible operation is essential from a serious cost-of-production perspective or when the main goal is to assess the effects of increasing shares of renewable energy on the plant operating regimes (this is also the case in Traber and Kemfert, 2011);

5. Linear optimization models do rely on *perfect foresight*. This is questionable from a heterodox perspective and/or a model that is supposed to closely reflect processes as they actually take place in the real world: while utilities have a lot of information and knowledge of electricity demand patterns and can also access increasingly reliable weather and renewable energy output forecasts, predictability of these factors is still more than limited once the time horizon increases.

6. The approach to *optimize* costs on the aggregate level is methodologically questionable from

a classical point of view. The latter recognizes the fact that capitalist market economies and industries are shaped by the interactive, independent and rival behavior of capitalist firms. Their behavior should be simulated on the individual level and the outcomes may or may not lead to rational and optimal results in some sense for society and/or the capitalist members of the industry.

7. While most simulations assess the impact of rising wind energy shares, the German example shows that it becomes more important to evaluate the *effect of PV energy*. This is the case since its cost competitiveness is in many regions far below offshore wind power. Installed capacities already exceed the existing capacity of wind power in Germany, making PV the largest electricity generating technology by installed capacity today (PV in other markets also shaves off mayor shares of peak load).

8. Many studies (e.g. Traber and Kemfert, 2011) that also model the electricity market rely on profit maximizing criteria where the *price for electricity* is used as the major exogenous parameter in the simulation procedure. This is a weakness for analyzing counterfactual (future) scenarios as the data are either real data and as such the result of a fundamentally different and incomparable electricity market (of the past) or generated through an own model that has by nature a large degree of arbitrariness and uncertainty. The negative spot prices that emerge more frequently prove that new phenomena are emerging that were unthinkable even in the most recent past. The core principle applied in the study presented here is that flexible on-demand load is simulated depending on the relative values of electricity demand (total system load) and inflexible or intermittent (not on-demand generated) load. This is similar to Kuntz and Müsgens (2011). In cases where future scenarios (that are necessarily fictional) are being analyzed, this approach has advantages. It just assumes that the electricity market 'works'. The electricity market price moves in the immediate short run so that available capacity will be used in a way that the grid network does not break down: to satisfy the demand that is assumed to be perfectly inelastic while the price forces capacity to go offline if it is not needed and online if there is demand.[1]

9. Traber and Kemfert (2011) are also a prime example of modeling industries under the assump-

---

[1] As the alternative is a blackout which results in extreme costs for market participants, this seems to be a fair assumption. The negative spot prices are an example of how the price mechanism enforces a load balance in the immediate short run.

> tion of *perfect or imperfect competition* and the associated models (i.e. Cournot or Stackelberg competition games). Earlier research by economists including myself has shown that their underlying assumptions (like the one that higher market share concentration makes surplus profit possible) are empirically largely proven to be wrong. Instead, alternative dynamic theories of competition offer frameworks that describe reality in a better and more consistent way.

Any model is necessarily limited and can never account for all influential aspects of reality. Different models also involve trade-offs between different types of imprecisions and shortcomings. Keeping this in mind, the simulation procedure implemented here seeks to ease the mentioned problems of other approaches that constitute issues from a classical perspective.

3.2 Data

The basic data used in this study are 15-minutes interval data on actual load, output from wind turbines and from PV installations. They have been obtained from the four German network operators 50Hertz Transmission (2012), Amprion (2012), TenneT TSO (2012) and TransnetBW (2012).[2] Data on total system load were obtained from the European Network of Transmission System Operators for Electricity (ENTSOE, 2012). Data on wind output and total system load cover the four-year period starting in July 2008 while data on PV output cover two years starting in July 2010[3]. Files were merged and aggregate time series were created for all of the national aggregate of the four network zones. The resulting file contains a total of 140,260 Germany-wide observations on wind output as well as 70,180 on PV load.

Monthly data from the federal network agency (Bundesnetzagentur, 2012) on aggregate PV and end-of-year data from the German wind energy industry association (BWE, 2012) on aggregate wind power capacities installed were added to the file. Applying linear interpolation made it possible to obtain a very close approximation of actual capacity factors for all the 15-minutes intervals of the period for which load data were used.[4]

---

[2] All files have been downloaded on 10/09/2012.

[3] Since what matters is the utilization rate and the capacity estimate in each scenario analyzed I could use the PV load profile for the previous years (assuming that the first year PV profile was equal to the third year and the second one equal to the fourth year).

[4] Information on capacities has been reviewed on 10/25/2012. The interpolation of wind power between end of year data points was acceptable as wind turbines are added more smoothly than PV installations which tend to spike at the end of the year (or at other points of time). This is due to significant historical feed-in tariff reductions and very short time periods to plan and install systems.

Any significant occurrence of wind power curtailment due to high load from turbines is obviously diminishing the quality of wind load data and implies that the use of actual wind speed information might be superior.[5] While the scale of curtailment practice is large enough to justify this critique in areas like the zone of the Electric Reliability Council of Texas where 17.1 % of all power from wind (or 3,872 GWh) was curtailed in 2009 (Cusick, 2013), this phenomenon occurred relatively spoken on an extremely small scale in Germany during the period of observation used here. In 2010, curtailment losses from installations that are compensated according to the EEG renewable energy law (that covers the vast majority) amounted to only 0.16 % (127 GWh) with even lower values in the past Bundesnetzagentur (2011, : 28).[6] It follows that the disturbance of wind output data due to forced shut downs is very minor and can be safely neglected as a problem for the data used here.[7]

Wind farm operators often require as a minimum wind site data that cover three years in order to safely assess the expected capacity factors of projects. Even the aggregate wind power load profile is dominated by randomness over even a month where high or low load periods can last for several weeks. Load might also be alternating during other periods on a highly frequent basis (PV reveals also significant random patterns with respect to total daily energy production). Using only a few weeks and assuming them to be representative is thus a questionable simplification especially when it comes to wind power. A large data set that covers a period of four years of actual data allows to assume that weeks, months and years with extreme weather conditions or demand situations are averaged-out. All simulations presented here were applied on the whole four-year-period. From these results averages were calculated for an average single year.

3.3 Simulation Model

I am modeling the electricity market elements total system load, output from wind, PV, hydroelectric, biomass and waste burning power plants on the aggregate. I also model load from nuclear, lignite, hard coal, combined and single cycle gas turbine (CCGT and SCGT) sources on the individual power plant level. More basic versions of this method have been applied on similar data sets

---

[5] This problem is irrelevant for PV in the case of Germany today as most losses due to curtailment fall on wind power with 98.67 % in 2010 Bundesnetzagentur (2011, : 28).

[6] This number was roughly equivalent to the annual output of less than seven offshore wind turbines.

[7] It should also be noted that these curtailments do not perfectly correlate with wind output peaks on the aggregate level (even the aggregate of the respective network) as the output losses are due to local stress on parts of the distribution network (and not due to production exceeding demand). The far less than perfect correlation between wind condition on a local site and on the nationwide average is thus less systematic and disturbing than one might expect.

before to assess the market potential for electricity storage in Germany (Auer and Keil, 2012) or to analyze effects of the German energy transition in the electric power industry (Keil, 2013). The model is restricted to the national market and does not consider the effect of industry level imports and exports. All modeling involved in this study has been carried out using StataMP 11.2 and its native programming languages, ‘Stata’ and ‘Mata’. The simulations for the different scenarios took depending on the scenarios 40-100 minutes on computers with 2.66GHz quad-core processors and 8 GB of RAM.

The major interests of the analysis lie on 1. the alteration of capacity factors of conventional plants due to load following operation and forced shut downs; 2. the impact of their changing, more flexible operating behavior and the start-up and load following costs associated with it on their relative competitiveness. There are two modeling principles applied in this study on three consecutively applied stages: institutional-technical determination, cost-based commercial operation and a combination of both.

3.3.1 Modeling of renewables

Before conventional load can be modeled, ‘institutional-technical’ principles need to be applied for modeling renewable energy output. They produce load profiles from total consumption (L^T) on the one hand and from intermittent renewable electricity sources on the other hand. The latter one consists of wind (L^W) and PV (L^{PV}) output. This stage also models inflexible but relatively stable producing renewable load from the parts of hydroelectric, biomass and waste power capacities that do not produce in reaction to the conditions of the electricity market.

Modeling the listed technologies is institutionally and technically determined in the sense that they do not react in the very short term to load profiles from other plants or to relative cost criteria: electricity consumption is very inelastic on a minute-to-minute basis; wind and PV power is produced according to patterns determined by natural factors beyond the control of utilities; wind, PV as well as parts of hydroelectric, biomass and waste electricity are largely treated in a preferred way by law and even legally required to be used in the electric power grid before other sources. The outcomes of their operations thus represent fixed externally determined parameters for commercial operators which can only be modeled in a subsequent step.

3.3.1.1 Inflexible renewables

The load profiles that correspond to the total usage of electricity, output from wind and from PV installation are rescaled versions of the respective original real time data series described in the previous chapter. The rescaling of total load was undertaken in such a way that its annual sum corresponds exactly to the total net electricity consumption including transmission and excluding pumping losses in Germany that is assumed to occur in the respective scenario. The rescaling of PV and wind power was undertaken by using the 15-minute capacity factors and applying them to the total installed capacity that was assumed in each scenario.[8]

Hydroelectric power in Germany consists mainly of 'run-of-the-river' plants that deliver baseload power which can only be adjusted marginally. Such plants sum up to 3.8 GW (C^{Hb}) and are assumed to run at the average annual utilization rate of the last years of 55 % (U^{Hb}). On this average annual utilization I imposed the strong seasonal pattern where output is much higher in spring than in fall using monthly averages of the last years as published in(r2b energy consulting, 2012, : 19-20). It involved scaling 55% of capacity with differentials of monthly to average monthly full load hours for the middle of each month and applying linear interpolation to generate the load time series. Waste energy and biomass plants operate largely on a continuous basis, do not follow a significant seasonal pattern (r2b energy consulting, 2012, : 41) nor do they produce according to demand and supply parameters of the electricity market. Waste plants (C^{WA}) are thus modeled to produce base load constantly at a minimum utilization of 40 % of their total capacity installed (U^{WA}_{min}). Biomass plants usually produce heat and generate electricity only as a secondary product in a joint production process. 60 % of all plants (s^{Bb}) are assumed to produce electricity secondarily (C^B) and to constantly run at 75 % of their capacity (U^{Bb}_{const}).

This first stage of institutional-technical modeling results in new load data for the sample period (that lasts from $t = 1 \ldots T$). It is being used as an exogenous parameter in the next simulation stage and represents the sum of total renewable load that is not generated on demand for the needs of the electricity market, L_t^{Ri}:

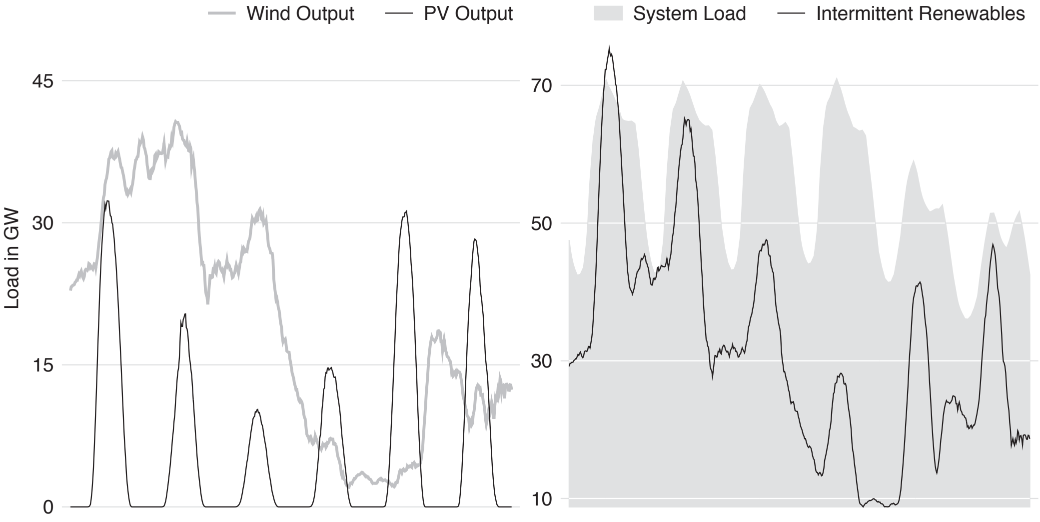
$$L_t^{Ri} = L_t^W + L_t^{PV} + C_t^{Hb}U^{Hb} + C_t^{WA}U_{min}^{WA} + C_t^B s^{Bb} U_{const}^{Bb},\ for\ t \in T$$

To illustrate this step of the procedure, I included graph 1. The left hand side shows the wind

---

[8] See Keil (2013) for details on predictions for PV and wind capacity and output in these scenarios.

Figure 1: **Intermittent load and electricity demand**



and PV output profiles. They are scaled from actual 2011 data (September, 00:00 6th to 00:00 12th) to a 2025 scenario that reflects government targets for total wind and PV capacities installed. The right hand side presents for the same interval the total system load scaled to correspond to the official annual consumption target, L^T. It also shows how the load profile L_t^{Ri} looks like that combines intermittent renewable sources (as well as inflexible, not-on-demand renewable output).

The right side of the graph shows that a large share of electricity can be satisfied by wind and PV power in 2025. It also reveals that there will be potential losses of renewable energy due to overproduction (the first plotted observations) and that one can expect shortages (the gap between the demand and renewable energy load profiles when the latter lies below the former) which represent potential production cycles for flexible conventional power plants.

3.3.1.2 Flexible renewables

A combination of institutional factors and commercial operating principles are applied to the simulation of flexible hydroelectric plant capacity, (C^{Hf}), biomass plants producing according to market demand, ($C^B\left[1-s^{Bb}\right]$), and the capacity of waste energy that is left unutilized ($C^{WA}\left[1-U_{min}^{WA}\right]$). They are modeled to produce load according to the demand and supply situation on the electricity market. As renewable sources they are assumed to obtain guaranteed fixed feed-in tariffs as a compensation and receive legal priority in electricity production. Hydroelectric power is the cheapest source of electricity that exists in Germany, highly flexible and does not involve any significant start-up costs. Biomass and waste plants are owned by independent operators that generally do not own any other conventional capacity. This means that given their preferred treatment they generate electric power before any conventional plant is started. Their characteristics also imply that it is not possible to clearly determine an order according to which plants are being ramped-up or restarted and put online within the group of renewables on-demand. Since modeling and analyzing their load is not the focus of this study it is valid to program the sum of their capacities as one unit (L_t^{Rf}).[9] Their output can take values within the space

$$L_t^{Rf} \begin{cases} \leqq C^{Hf} + C^{B}\left[1 - s^{Bb}\right] + C^{WA}\left[1 - U^{WA}_{min}\right] & = L^{Rf}_{max} \\ \geqq 0 & = L^{Rf}_{min} \end{cases} , \; for \; t \in T$$

implying that 'on-demand renewables' can neither produce load greater than the sum of their capacities (restriction 1) nor can they produce negative load (restriction 2). The actual load values are determined by the criteria

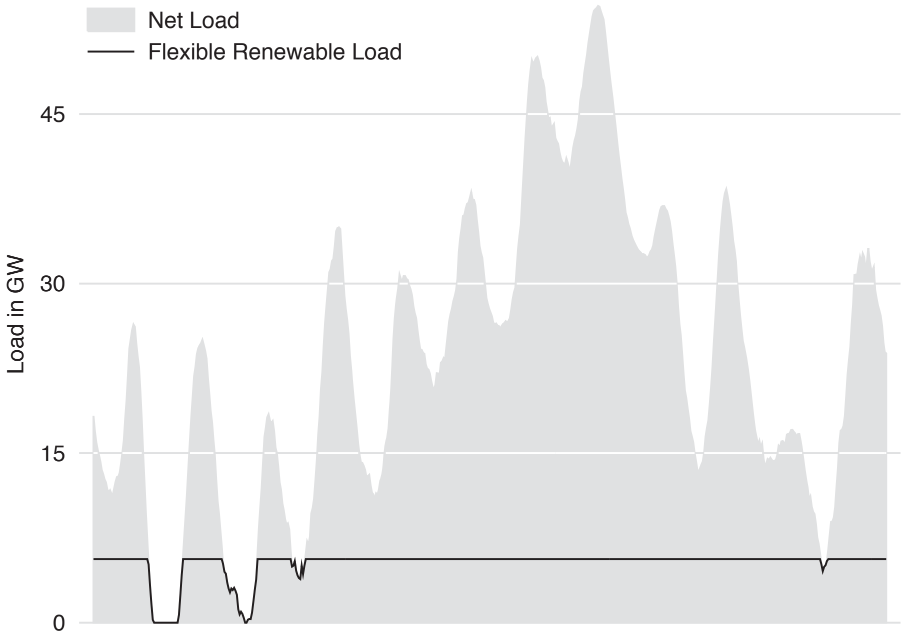
$$L_t^{Rf} \begin{cases} = L^{Rf}_{max} & if \; L_t^{T} - L_t^{Ri} \geqq L^{Rf}_{max} \\ = L^{Rf}_{min} & if \; L_t^{T} - L_t^{Ri} \leqq 0 \\ = L_t^{T} - L_t^{Ri} & if \; otherwise \end{cases} , \; for \; t \in T$$

meaning that they run at maximum capacity if the gap between electricity demand and 'institutional renewable output' exceeds or equals their maximum capacity (row 1); that they will not generate additional load if wind, PV and other fixed components produce more than what is being used (row 2); that they produce exactly the load gap between total demand and existing renewable output if this gap is positive and not exceeding the installed capacity (row 3). Maximum start-up or normal ramp rates (load gradients) were not considered here as installations are usually rather small and flexible. Hydroelectric plants are flexible due to their technology and can deliver 100 percent of capacity in less than 2 minutes during a start. From this modeling pre-stage of the simulation a new set of parameters results. They are used in the next and main modeling stage. Graph 2 illustrates the procedure of load generation according to commercial criteria just applied. The shaded area represents shortages ($L_t^T - L_t^{Ri}$) as they result from the first stage of modeling of renewables for exactly the data that were illustrated in the previous graph. The black line represents the output

---

[9] The load from this sum is divided among the three sources in proportion to their installed capacities of every 15-minute interval in the final evaluations of the model.

Figure 2: **Renewable energy plants producing for the market**



from capacities of flexible hydroelectric, biomass and waste plants that produce for the electricity market ('renewables on demand'). The graph reveals how shortages constitute potential production cycles for such (renewable) capacities that are still available. They are used to fill existing load gaps while leaving shortages that represent potential production cycles for conventional plants in the next modeling stage (in this graph five new cycles result).

3.3.2 Modeling of conventionals

The main simulation section models load according to 'cost-based commercial operation' principles. This program is an ***iterative cost-of-electricity based simulation of plant output*** (ICEsim). It works in principle similarly to the just illustrated modeling of renewable load that produces for the electricity market. This model has been applied to the rest of the existing industrial park of power plants - nuclear, lignite, hard coal, CCGT and SCGT plants.[10] Load from these plants was modeled on the individual power plant level where plants were neither allocated to companies nor broken down into blocks or turbines constituting a plant. Due to a lack of ready-to-use data and the

---

[10] They are in Germany mainly owned by the four large utilities, E.ON SE, RWE AG, EnBW AG and Vattenfall GmbH. A significant amount of plants is run by the German railroad company, Deutsche Bahn AG, by large energy manufacturing companies in energy intensive industries, by network operators and by public municipal utilities.

need to reduce simulation complexity, standard plant sizes of 3.5 GW for nuclear, 2.5 GW for lignite, 2 GW for hard coal, 1.5 GW for CCGT and 1 MW for SCGT were being assumed.[11] There is no differentiation between or individual identification of the plants in each technology class. The final evaluation summarizes average values of each technology. The number of plants in technology class j corresponds to $\overline{m}_j = \frac{C_j}{ps_j}$, which is the integer-rounded number of the sum of installed capacity in the German electricity industry (C_j) as predicted in the used benchmark scenarios, divided by the respective standard plant size (ps_j).

Cost-based modeling as applied here is carried out in four steps: 1. simulation of a round of potential production cycles for each plant type; 2. calculation of electricity generation costs in cent/kWh for each production cycle of the first plant output iteration for each technology; 3. evaluation to find the lowest cost technique for each cycle; 4. selection of the lowest cost technique and updating of the aggregate electricity production load profile. This *consecutive cumulation* is then being repeated with the next load profile.

3.3.2.1 Simulation of potential production cycles

ICEsim models load from the pool of the N plants existing in the industry ($N = \sum_{j=1}^{J} \overline{m}_j$). The ***potential load profile*** or output of plant n in production cycle k ($l^P_{n,k}$) depends on net load (the cumulative sum of load generated before $L^P_{n-1,k}$ minus the total consumption profile L^T) and the technical limits of the respective technique. If selected, the modeled load of plant n is then added to the aggregate output of the production cycle period $L^P_{n-1,k}$, updating the respective total load generation profile to $L^P_{n,k} = L^P_{n-1,k} + l^P_{n,k}$.[12]

Gaps between total system demand and renewable plus conventional load that has already been committed ('net load') constitute periods of potential temporary shortages (if not satisfied). These represent potential '***production cycles***' which are defined as the load that a plant generates between a start-up ($l^P_{n,t-1} = 0$ & $l^P_{n,t} > 0$) and its following shut down ($l^P_{n,t-1} > 0$ & $l^P_{n,t} = 0$). ICEsim

---

[11] The numbers represent sizes of large modern plants. Large sizes and small plant numbers allowed to reduce the computational effort significantly. The nuclear, lignite and hard coal plants sizes are reached when two large blocks constitute a plant. The CCGT size is in the range of a 3x1 shaft system that uses the largest turbine classes available while SCGT corresponds to a 3x design with three large turbines in one plant. The blocks or gas turbines can in reality be run more or less independently. However, modeling such units below the plant level is not expected to yield much better results while requiring much more computational resources.

[12] For the first conventional plants modeled, $L^P_{n-1,k}$ results from the two previous institutional modeling stages, $L^P_{n-1} = L^{Ri} + L^{Rf}$.

considers every potential cycle k ($k = 1 \ldots K$) individually for every available plant type. Plant start-ups are considered if there is a period ahead that exhibits a positive gap between load demand and committed supply, $L^T_{t+v} > L^P_{n-1,\, t+v}$ for at least one time period v_k. v_k is a 15-minute interval that is part of a production cycle which lasts for a total of V_k periods ($v_k = 1 \ldots V_k$) and is part of the period of observation analyzed here ($v_k \in T$ for each v).[13]

$l^P_{n,k}$ also depends technical features that are specific to the power plant such as plant size ps_j and limits to plant flexibility. Thermal power plants have plant and technology specific ***minimum utilization rates***. The absolute limits of l^P_{nk} are thus given by

$$l^P_{n,t} \begin{cases} \leqq & ps_j \\ \geqq & 0 \\ = & ps_j * U_{n,t} \geqq ps_j * U^{min}_j \; if\; d = 0 \end{cases} \quad , \; for\; t \in T \; and\; j \in J$$

where the third restriction enforces the actual utilization rate $U_{n,t}$ being not smaller than the minimum utilization rate U^{min}_j under normal operation mode ($d = 0$) - i.e. whenever the variable d indicates that the plant is neither in a start-up ($d = 1$), nor in a shut-down ($d = 2$) phase.

Flexibility is not only limited by these absolute load values, but also by maximum ramp rates or ***load gradients*** that describe plant behavior within such level boundaries. Each technology j has a technically determined maximum load gradient g^{max}_j when the plant is in normal operating mode. These ramp rates are expressed as the maximum percentage change of nominal plant capacity per time unit. Here they are assumed to be independent of preceding and succeeding gradients, of the level of plant capacity utilization ($U^{min}_j < U_{n,t} < 1$) and of whether output grows or contracts. In the model, the normal absolute maximum kW change per 15-minute interval within the absolute boundaries is given by

$$l^P_{n,t} \begin{cases} \leqq & l^P_{n,t-1} + g^{max}_j * ps_j \\ \geqq & l^P_{n,t-1} - g^{max}_j * ps_j \end{cases} \; , \; for\; t \in T,\; j \in J\; and\; if\; d = 0$$

The ***start-up*** process of a plant takes a minimum amount of time. This is captured with a

---

[13] The simulation is finished if $L^T_t = L^P_{N,\, t}$ or if no power plant is idle whenever $L^T_t > L^P_{N,\, t}$ for $t \in T$. This means that there is either no shortage or only shortages at points of time when all existing plants are already operating.

maximum possible positive load-gradient during the start-up process, g_j^{smax}, which is always lower than the restriction under a normal operating mode (negative gradients are assumed to not being altered). Cold starts are commonly assumed in related models and also in the one used here. The start-up is simplified by modeling its profile as a linear one where the same g_j^{smax} is applied to every positive ramp rate that falls into the interval of the fastest possible (cold) start. This minimum start-up phase lasts from the beginning of the production cycle $v_k = 1$ to $v_k = 1/g_j^{smax}$, which is indicated by $d = 1\, if\, v_k \leqq 1/g_j^{smax}$. Start-up ramp rate restrictions are thus given by

$$l_{n,t}^P \begin{cases} \leqq & l_{n,t-1}^P + g_j^{smax} * ps_j \\ \geqq & l_{n,t-1}^P - g_j^{max} * ps_j \end{cases} ,\ for\, t \in T,\ j \in J\, and\, if\, d = 1\, \&\, U_{n,t} \geqq U_j^{min}$$

The effect of minimum utilization is also assumed to impact the start-up process. Output cannot be reduced to follow load when the plant operates below this rate during a start-up process:

$$l_{n,t}^P \begin{cases} \leqq & l_{n,t-1}^P + g_j^{smax} * ps_j \\ \geqq & l_{n,t-1}^P * ps_j \end{cases} ,\ for\, t \in T,\ j \in J\, and\, if\, d = 1\, \&\, U_{n,t} < U_j^{min}$$

The enforcement of start-up times is also reflected in an additional restriction that permits a plant to only start-up if the operating time will be longer than its minimum start-up time ($l_{n,t}^P = 0\, if\, V_k < 1/g_j^{smax}\, for\, t \in v_k$).

Shut down periods are much shorter than start-ups and last for the minimum time it takes for a shut down of a plant operating at nominal capacity. The last 15-minute intervals of a production cycle are defined as such shut downs that are relevant for plants with normal maximum ramp rates below 100%: $d = 2\, if\, v \geqq V - 1/g_j^{max}$. During a shut-down period, output can fall below the minimum utilization and load cannot be increased once it falls below that minimum utilization rate even if net load rises again:

$$l_{n,t}^P \begin{cases} \leqq & l_{n,t-1}^P * ps_j \\ \geqq & l_{n,t-1}^P - g_j^{max} * ps_j \end{cases} ,\ for\, t \in T,\ j \in J\, and\, if\, d = 2\, \&\, U_{n,t} < U_j^{min}$$

Plants behave as if they were under normal operating mode otherwise:

$$l_{n,t}^{P} \begin{cases} \leqq & l_{n,t-1}^{P} + g_j^{max} * ps_j \\ \geqq & l_{n,t-1}^{P} - g_j^{max} * ps_j \end{cases} , \; for\, t \in T,\; j \in J\; and\, if\; d = 2\, \&\, U_{n,t} \geqq U_j^{min}$$

Within these capacity and load gradient limits, plant load was modeled to be exactly equal to net load or as close as possible - given the technical plant specific limitations.

3.3.2.2 Calculation of electricity generation costs

The *operating order* according to which the N existing conventional plants are being put online is determined by the availability of plants that are not started yet, by their technical ability to start and by the relative total variable costs in cent/kWh that are associated with the j different plant technologies in the k'th potential production cycle, $tc_{j,k}$.

Total variable costs of plant n, $tc_{n,j,k}$, consist of 1. variable operating and maintenance (O&M) costs, 2. start-up costs, 3. load following / cycling costs and 4. expected plant output for the production period.

Variable costs O&M of technology j in production cycle k in kWh terms are given by

$$OMC_{j,k} = \left(oomc_j + \frac{p_j^f + e_j p^e}{\eta_j} \right) * X_{n,k}^{gross}$$

$oomc_j$ represents 'other variable O&M costs', p_j^f the fuel price (per thermal kWh of fuel), e_j the emissions of CO2 (in grams per thermal kWh), p^e the price of CO2 emission rights and η_j the thermal plant efficiency. $X_{n,k}^{gross}$ is the total plant output occurring during a production cycle (including losses).

Abstracting from multiple independent turbines or separate 'blocks', plants are assumed to be only startable as whole units. Thus, total ***start-up costs*** depend on the plant capacity scaled sum of additional, start-up related maintenance and depreciation as well as operating and other start-up cost OSC_i, start-up fuel SF_i, p_j^f, e_j and p^e:

$$SC_j = ps_j * \left[OSC_j + SF_j \left(p_j^f + e_j p^e \right) \right]$$

Cycling costs for a plant n that is considered to be used in a potential production cycle k ($cc_{n,j,k}$) are reflected in a) the technology specific efficiency losses that result whenever plants operate under partial load ($U_{n,t} = \frac{l^P_{n,t}}{ps_j} < 1$); b) output losses due to technical restrictions on maximum load gradients and minimum utilization rates; c) deterioration and stress on the equipment that is associated with output fluctuations.

Efficiency losses due to producing at partial load result in increased fuel and emission costs. In the simulation the efficiency loss is assumed to increase linearly with $\left(1 - U^{gross}_{n,v}\right)$, which replicates the known fact that deep cycling is more expensive than a shallow load following operation. This phenomenon can be modeled by modifying the formula for omc_j. In the denominator one can incorporate the product of the percentage point loss in efficiency (that results from a one percentage point utilization gap between the actual and the nominal utilization, δ_j) and one minus the actual utilization ratio $U^{gross}_{n,v}$. Applying their product on the efficiency at nominal load results in the partial load efficiency.

If the load gradient $g^{gross}_{n,v}$ is unequal to zero, thermal stress on the equipment is caused which leads to *elevated maintenance and depreciation rates*. The rate γ_j at which this happens is technology specific and assumed to be constant throughout the spectrum of possible capacity utilizations (i.e. the same for every kW change in plant output). Being a simplification it still replicates the fact that deeper and more frequent cycling results in higher costs. Cycling costs of a production cycle that result from efficiency losses and elevated maintenance and depreciation can be expressed as

$$CC_{j,k} = \sum_{v=1}^{V} \left[\frac{p^f_j + e_j p^e}{\eta_j - (1 - U^{gross}_{n,v}) * \delta_j * \eta_j} * l^{P\,gross}_{n,v} + \gamma_j * g^{gross}_{n,v} * ps_j \right]$$

Output losses can result from cycling and start-stop operation. They imply that costs occur for the production of load which is not being used. This impact is reflected in the use of gross output (before such losses), utilization and load gradients in the $OMC_{j,k}$ and $CC_{j,k}$ formulas while net output (excluding such losses) is used as the denominator in the following $tc_{j,k}$ formula

$$X^{gross}_{n,k} = \left[\sum_{v=1}^{V} l^P_{nk}\right], \; X^{net}_{n,k} = \left[\sum_{v=1}^{V} l^P_{nk} - loss^P_{nk}\right]$$

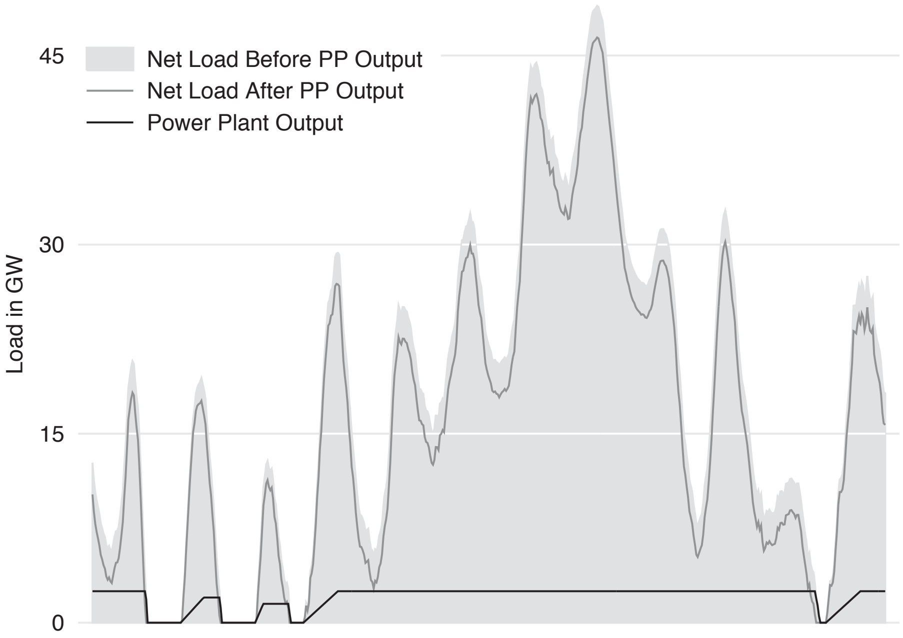
where $loss^P_{nk}$ is the difference between plant output and net load. It is equal to zero if the latter exceeds the former.[14]

For comparison plants' total variable, start-up and cycling costs need to be expressed in the form

---

[14] Using the sum of 15-minute load intervals is a reasonable approximation of electricity (whose exact value is equal to the integral of continuous load data throughout the respective period).

Figure 3: **Lowest cost output from conventional plants, iteration 1**



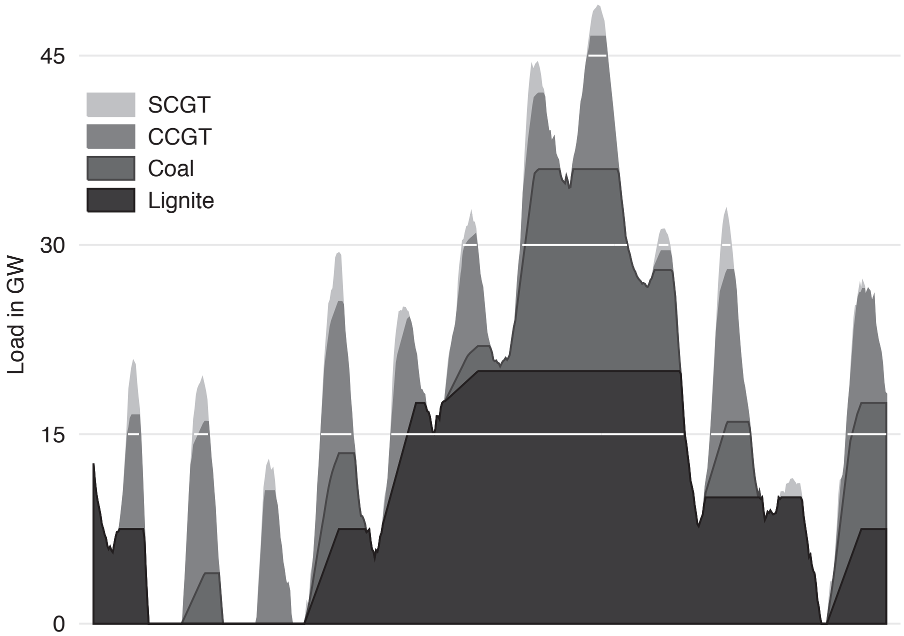
of $cent/kWh$ by weighting the sum of these costs with the net electricity output (in kWh) that can be expected during the production cycle. The sum of all variable costs that occur in a production cycle for a plant of a given technology amounts on a kWh basis to

$$tc_{j,k} = \frac{OMC_{j,k} + SC_j + CC_{j,k}}{X_{n,k}^{net}}$$

A plant of technology type j is chosen to be activated for production cycle k if exists, is not yet operating and its total variable costs $tc_{j,k}$ are lower than the ones of all other available plant types. Graph 3 shows for the same period as the other graphs how the first iterative round of power plant selection produces a load profile after the previous stage where renewables were modeled:

The shaded area represents net load before the first round of plant output is committed. The black line is the load profile of the lowest cost of production plant for the respective production period. In the case of the graph, the plant technology that delivers the lowest costs for the first period are lignite plants. For the shorter second interval hard coal plants are cheaper and will start. The third shortage lasts for just 6 hours and 15 minutes - making CCGT plants the most cost competitive ones. The remaining two longer shortages are answered by a lignite power plants. The slower start-ups of lignite versus hard coal and especially CCGT plants can easily bee seen at

Figure 4: **Output from conventional plants, after the last iteration**



the flatter incline of the load profile in the third and fourth production periods. The result of the simulation of conventional plants for the time period used in the illustration here are summarized in graph 4. It shows how baseload is being produced by plants with low variable and high start-up and cycling costs while peak differences between demand and renewables are served by more flexible coal and natural gas plants. The graph contains the load profiles of all conventional power plants that result at the end of the simulation. They are aggregated by technology class:

If the ICEsim model had no restrictions imposed on maximum capacities installed, it would automatically produce load curves that generate the lowest costs of production on the aggregate and ensure that no load gap would exist at the end of the simulation. However, enforcing limits creates potential problems that the result is not necessarily cost minimizing[15] and shortages are larger than they could be. Flexible plants can be producing at lower costs and start first. If the temporary shortage exceeds the sum of existing power plants with the same or higher degree of flexibility, less flexible plants can be forced to start after flexible ones. If they would have started earlier, inflexible plants' output involved the same amount of start-up costs while they caused a smaller sum of variable costs than their flexible counterparts (that would start later and produce

---

[15] As markets are no perfect institutions this point alone does not necessarily constitute a weakness. However, in this particular case here where utilities own many plants one can assume that they operate them in such a way that the type of puzzling result described in this paragraph should not occur.

less output in their production period). Thus, a different order would have produced lower costs of total conventional load per shortage. It could also happen that a larger final shortage results since inflexible plants do not produce for the minimum time required by the model. To avoid these problematic outcomes, a plant is programmed to break the direct relative cost comparing order and to start. This occurs whenever the sum of the nominal outputs of all unused plant types that are more flexible than the one considered is exceeded by the maximum value of net load of the temporary shortage (before that particular iteration). This seems to be a sensible assumption since several plant types are owned by the big utilities and the production decision of any given plant is also depending on its effect on the operation of other plants operated by a utility and on its total production costs.

3.4 Model inputs and evaluation method

3.4.1 Power plant technology and cost details

The details about costs of and technical limits to flexible operation of power plant types are summarized in the following table.

Table 9: **Power plants under flexible operation**

Cycling	Nuclear	Lignite	Coal	CCGT	SCGT
Min. utilization	60%	50%	35%	25%	20%
Max. ramp rate*	50%	35%	60%	100%	100%
Efficiency loss**	0.25%	0.2%	0.2%	0.25%	0.3%
O&M costs, ct/kW***	0.02	0.015	0.01	0.0025	0.0035
Start-up (cold)					
Max. ramp rate*	3%	4%	6%	17%	75%
Fuel, kWh/kW	17	9	4.5	2	0.5
Other costs, ct/kW	6	4	2.5	0.8	0.4
Maint.&Dep. costs, ct/kW	30	20	10	5	3
Total costs, ct/kW	40.7	31.1	17.7	9.3	4.7

** % of nominal capacity / 15-min.; ** % of nominal efficiency / %-point below 100% utilization; *** for each positive 1% load gradient*

The table shows how gas turbines are able to run at very low utilization rates while nuclear and lignite plants are much more suited for delivering constant baseload. The ramp rates give a similar picture. Only the numerically smaller effects of efficiency losses from partial load are higher for natural gas fired units. Additional O&M costs that can be associated to cycling activity are expressed in cent/kW per 1% load gradient. While this cost dimension is also rather small when its sum is divided by the total annual output, it still makes a difference for a production cycle as a deep cycle from maximum to minimum load and back results in additional costs. In cent/kW they amount (excluding costs due to efficiency losses) to 0.28 cent/kW for SCGT, 0.19 cent/kW for CCGT, 0.65 cent/kW for coal, 0.75 cent/kW for lignite and 0.8 cent/kW for nuclear plants. This illustrates how more flexible plant types can have advantages in short production cycles.

Start-up costs are much higher for the larger, less flexible nuclear and lignite units while hard coal takes an intermediate position. This is especially striking when it comes to start-up fuel consumption (which results in costs due to the assumed fuel prices and emission costs).[16] It is also present in maintenance and depreciation as well as additional operating and other costs (that include chemicals, water and auxiliary electricity) which are associated with plant start-stop operation. Since less flexible plants are also larger and start-up costs apply not only to each percentage point of capacity used but rather to the whole plant capacity, the advantage of flexible plants with respect to start-up costs can become even larger. The start-up time differentials that are expressed by the start-up ramp rates reinforce this picture. While SCGT can deliver full load after less than 30 minutes, nuclear plants take nearly 8.5 hours to produce at 100% of capacity. Start-up costs in cent/kWh for a production cycle will be elevated through this as the fixed start-up costs are divided upon fewer output units. In order to avoid too much computational flexibility cold starts were assumed as a standard.[17]

The data come mainly from a review of literature on the topic. It largely represents average values from the information sample collected. The most important sources on costs are Kumar, Besuner, Lefton, Agan, and Hilleman (2012) and Lefton and Hilleman (2012) with their studies on start-up and cycling costs. For further technical information on power plant characteristics shown in the table I consulted publications from Areva (2010), Auer and Keil (2012), Dena (2005, 2010), Hundt, Barth, Sun, Wissel, and Voß (2009), IEA (2010), Kehlhofer (1997), Kumar, Besuner, Lefton, Agan, and Hilleman (2012), Lefton and Hilleman (2012), RWE AG (2009, 2011), Umweltbundesamt (2011) and Vuorinen (2007).

While certain plant types have cost disadvantages when it comes to flexible operation, they are

---

[16] The fuel quantities in the table are reduced in the simulation by 60 % due to the fact that they include the fuel that produces load in addition to the fuel that is used for the pure start-up of a unit and that does not produce any electrical output.

[17] While this is an obvious simplification given the huge differences in start-up times and costs between cold and hot starts other authors have pursued the same strategy in their simulations.

also the ones which involve lower variable costs at base load operation. Variable O&M costs at nominal load, investment costs, fixed costs and other important parameters used in the evaluation of the simulation results are summarized in the following table.

Table 10: **Power plant characteristics**

	Nuclear	Lignite	Coal	CCGT	SCGT
Investment cost (EUR/kW)	4,000	1,400	1,200	650	400
Construction time (years)	6	4	4	3	2
Lifetime (years)	50	50	50	40	40
Efficiency	36%	43%	45%	60%	40%
Variable costs					
Emissions (t CO2/MWh th)	0	0.41	0.34	0.2	0.2
O&M costs (EUR/MWh)	6	1.6	1.2	0.4	0.3
Fuel costs* (EUR/MWh th)	2.5	6	11	25	25
Total variable (ct/kWh)	1.4	2.1	3	4.4	6.6
Fixed costs					
Staff (people/MW)	0.12	0.1	0.1	0.06	0.08
Staff costs (EUR/person)	90,000	90,000	90,000	90,000	90,000
Outage insurance (EUR/kW)	8	8	8	8	8
Total fixed (EUR/kW)	18.8	17	17	13.4	15.2

**Includes transportation costs. Emission costs of 6 EUR/t CO2 were assumed.*

The table shows how flexible technologies are characterized by much higher variable O&M costs - SCGT run at 4.7 times as much on constant nominal capacity than nuclear plants. Each more flexible technology is inferior to baseload plants on a cost basis whenever long periods of output at full capacity can be realized. All maintenance costs shown as variable O&M costs in this table are independent of start-stop and cycling and partial load operations. They are especially high for nuclear plants due to high maintenance requirements associated with elevated risks and higher costs of plant components. Since coal and lignite plants burn less clean, higher costs result for emission removal and cleaning. For hydrocarbon fired plants the most important factor are fuel costs which are especially high for gas fired plants. Hard coal costs much more than lignite which is mined in open pit mines on a constant basis and hauled automatically with conveyor belts to pit-mouth plants. Parameters used in this table are based on information from a large German utility, own research on capital expenditures and construction time in the news and the previously listed literature on technical information in table 9 (for plant life and plant efficiencies in table 10).

3.4.2 Scenarios

The three baseline scenarios analyzed with the simulations are similar to the ones used in two other studies Auer and Keil (2012); Keil (2013). They reflect three different points of time: 2010, 15 years from now and 30 years from now. Conventional thermal capacities are the 2010 gross values as published by the Bundesministerium für Wirtschaft und Technologie. The 2025 and 2040 values are derived from the official lists of power plants and of plants under construction of the Bundesnetzagentur as well as from the mentioned plant life assumptions.[18] The plants representing the future capacities are thus the ones that already exist and that live throughout that time period as well as plants currently under construction. Nuclear power will be phased out by 2025. I assumed that 30% of all gas fired plants were SCGT since information for such units was only available for that aggregate. On the renewable side, hydroelectric plants exclude pumped-storage hydroelectricity. Future values assume a moderate augmentation in capacity (as in the case of waste energy) due to turbine replacement. Biomass development is aligned to government targets. Net electricity consumption excluding transmission and including pumping losses was used for total system load. Future values represent government target values that envision reductions of 12.5% and 20% from the 2010 level.

Table 11: **Baseline industry scenarios used - capacities in GW**

Conventionals	2010	2025	2040
Nuclear	21.5	0	0
Lignite	22.7	19.9	12.8
Hard coal	30.2	27.3	12.3
CCGT	16.7	11.4	8.1
SCGT	7.1	4.9	3.5
Renewables intermit.			
Wind	27.3	49.5	65
PV	17	57	63.5
Renewables other			
Hydroelectric, inflexible	4.4	4.8	5.2
Hydroelectric, flexible	1.1	1.2	1.3
Biomass	4.8	7.5	10
Waste	1.65	1.85	2.05
Total consumption (TWh)	562	492	450

Rounding was applied to use mentioned standard large plant sizes. ICEsim modeled the behavior

---

[18] The data can be accessed under http://www.bundesnetzagentur.de/cln_1911/DE/Sachgebiete/ElektrizitaetGas/Sonderthemen/Kraft and have been reviewed on February 19th, 2013.

of 6 nuclear, 9 lignite, 15 coal, 11 CCGT and 7 SCGT plants in 2010. These numbers amounted in 2025 (2040) to 0 (0), 8 (5), 14 (6), 8 (5) and 5 (4) plants respectively.

For PV and wind modeling of the individual time paths was applied for: 1. new capacity added; 2. old capacity retired; 3. capacity replaced; 4. rates of technical improvements that yield higher capacity factors (see also Keil, 2013). Values for ***PV*** result from a growth path of capacity that is significantly slower than the current pace of around 7.5 GW per year. It equals the average value of the current target corridor of the federal government (3 GW), starting from 30 GWp installed in August 2012. Once 52 GW of installed capacity are reached[19] capacity is assumed to grow by 1 GW per year until 2025 and by 500 MW per year after 2025.[20]

Future scenarios for ***wind*** assume that starting from 30 GW installed in 2012, offshore installations come online according to the current plan with 17.5 GW by 2025 and that there will be annual additions amounting up to 200 MW after 25 GW are installed in 2030. This will lead to 27.7 GW in 2040 including offshore repowering effects. In the scenario applied here in the study, new onshore installations are assumed to grow by only 50 MW annually (a significant slowdown from the recent pace). However, the analysis of data on past annual installments indicates that there will be 6.1 GW of capacity that are going to be older than 25 years by 2025 and that have to be retired or replaced. Assuming that 60% of all turbines will be repowered with an increased capacity of 200% per turbine (the required minimum to obtain a special repowering bonus), will yield 7.31 GW in new onshore capacity (a net addition of 1.2 GW). Under the same assumptions, 24.1 GW will go offline between 2025 and 2040 and repowering will produce 28.9 GW of new turbines (a net plus of 4.8 GW). Equivalently, the first offshore turbines will retire in 2039: by 2040 3.5 GW will have to go offline and 4.2 GW of new turbines will replace them. The result of this offshore intensive growth in new installation and of (mainly) onshore repowering effects is an increase of wind power capacity from 30 GW in 2012 to 49.5 GW by 2025 and 65 GW by 2040. Due to continuing technical improvements and more importantly, a rise in the share of offshore wind power, average capacity factors of cumulated total capacity installed are increasing from 18.5% to 28% in 2025 and 32% in

---

[19] The subsidy for PV is being discontinued for new installations once 52 GW of PV capacity are installed according to the newest revision of the German Renewable Energy Act.

[20] Furthermore, I assume that installations will produce power for 30 years and that after that period 30% of PV installations will not be replace while 70% will be repowered. Repowering is predicted to increase capacity by 35% since rising efficiencies allow to install more kWp per square meter. It is assumed that the average capacity factor of newly installed modules increases every year by 0.2 percentage points due to continuing trends of technical improvements (such as rising inverter efficiencies). At the same time, annual output is assumed to fall by 0.5% per year for modules installed in the past due to technical module efficiency degression. This implies moderate overall increases of average utilizations from 10.5% today to 11% in 2025 and 12.5% in 2040. Total annual output from PV thus amounts to 52.7 TWh in 2025 and to 66.7 TWh in 2040.

2040.[21] Such advancements are assumed to raise annual wind outputs to 121.5 TWh in 2025 and 182.5 TWh in 2040.[22]

For each scenario a ***transformation of the original load profiles*** of electricity from PV and wind installations had to be conducted. In the case of PV it was rescaled by applying the future capacity estimates C^F on the capacity utilization rates U for PV in the file and by inflating the load profile with the predicted efficiency improvement from its current (e^C) to its future value (e^F):

$$L_{PV}^{scenario} = C^F * U * \left[1 + \frac{e^F - e^C}{e^C}\right]$$

For wind, the reshaping process is more complicated as capacity additions take place in offshore and low wind locations whose installations have very different distributions of capacity utilization rates and differently shaped load profiles. Technological developments also change these shapes as they mainly produce higher capacity factors under partial load at low wind speeds. There is no apparent and unique way how to deal with this. However, it is clear that the actual empirical load profile needs to be rescaled by multiplying predicted scenario capacities with real utilization rates ($C^F * U$). One also needs to apply a filter or a function on the time series that reshapes or adds to the load curve in such a way that lower utilization rates are elevated more than high rates (the function applied here is $C^F * [(1-U) - (1-U)^z]$). It is also clear that the overall effect of this reshaping transformation needs to increase aggregate annual wind load such that the predicted annual capacity utilization results. The parameter z was adjusted precisely that this condition was fulfilled. A higher value leads to a smoother time series where lower original output values would be scaled by a greater factor than higher ones.[23] The whole transformation is summarized by the formula

$$L_W^{scenario} = C^F * U + C^F * [(1-U) - (1-U)^z]$$

The drivers behind the transformations are illustrated for the 2040 scenario in figure 5. On the left side it contains for an arbitrarily chosen interval the original wind output data, its rescaled

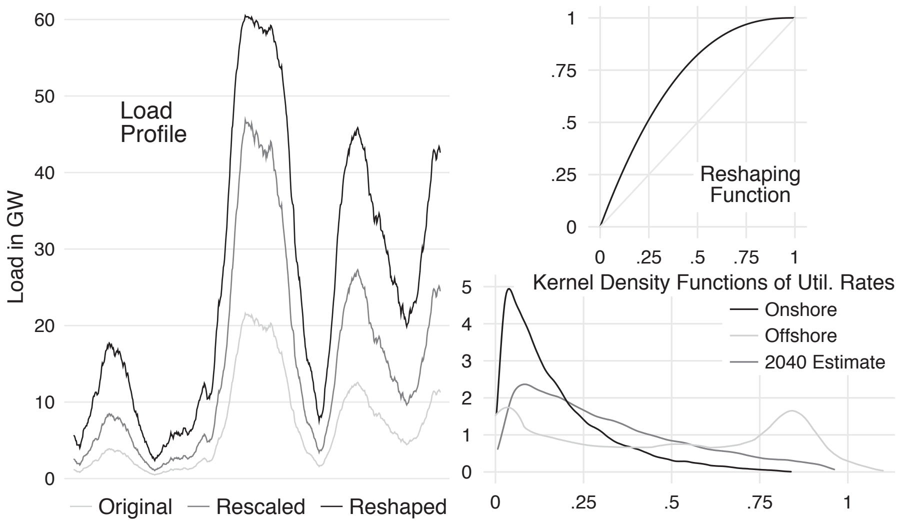
---

[21] Utilizations of new on- and offshore turbines are assumed to increase rather slowly by 0.1 percentage points annually.

[22] In order to stay with the scenarios as close as possible to the official targets of the government, these estimations are overly conservative as it is likely that technical improvements (mainly greater hub heights and larger rotor diameters) take place at a faster pace. Actual repowering could also have a greater impact as some installations should go offline before they reach an age of 25 years and new turbines often increase capacity by much more than 200%. Of course, it is easily possible that more than 60% of all retired turbines could be replaced.

[23] Values of 1.88 were chosen for wind in 2025 and 2.28 for wind in 2040.

Figure 5: **Rescaling and reshaping of wind output (2040 scenario)**



version $(C^F * U)$ and the effect of reshaping these rescaled time series with the reshaping function $(C^F * [(1 - U) - (1 - U)^z])$. It is drawn on the upper right hand side. The lower right hand side graph shows how the capacity factors are distributed in the original onshore and offshore wind output series as well as in the series produced for the 2040 scenario (after reshaping and rescaling).[24]

While this remodeling approach is of course not perfect, it should be considered superior to any 'solution' that either simulates load profile from scratch without the use of real and empirical data or that does not consider the effect of technical change and the change of the capacity mix behind the load profiles.

The baseline scenario includes current fuel (gas €25/kWh, coal €11/kWh) and CO2 emission prices (€6/tCO2) as of late 2012. To evaluate the impact of different prices, I used two alternative scenarios which include values at the end of the range that could be observed over the last three years. In the first alternative 'gas favorable' scenario, CO2 and coal prices are at heights of €16 and €13 while gas prices are as low as €22. In the second alternative 'coal favorable' scenario, CO2 and coal prices are at lows of €4 and €9 while gas prices are at €28.

In order to evaluate the impact of different energy policies I ran the model on alternative scenarios for the all three points of time. The nuclear future scenarios assume that nuclear power plants are retired after 60 years (which is unrealistic but done for the sake of analyzing the effect of a nuclear

---

[24] Epanechnikov kernel density estimates were used. Offshore data contains only offshore wind power for the TenneT TSO and covers only a single year. The load is dominated by just two wind farms. Onshore data covers all installations countrywide and for all four years.

energy based future). This would yield 20.5 and 14.3 GW of nuclear capacity in 2025 and 2040 respectively. However, if nuclear plants were not phased out this would have two effects: a change in the mix (technologies' shares in total conventional capacity) and a change in total capacity. To evaluate how suited a nuclear based energy policy is and to assess nuclear energy's impact on costs of production (that are highly dependent on utilization rates) reasonable one should isolate the first effect. I thus used the nuclear capacity estimates, calculated the shares in capacity and multiplied these with the total conventional capacity excluding nuclear plants. This changed the plant count for the 2025 (2040) scenario in such a way that it produced 4 (3) nuclear, 6 (4) lignite, 10 (4) coal, 6 (4) CCGT and 4 (3) SCGT plants. Due to the fact that nuclear plants exist in 2010, a different method was applied by using a more nuclear intensive scenario than in the 2010 benchmark version. Nuclear capacity was increased by two thirds of installed 2010 nuclear capacity while the capacity of all other technologies was reduced by exactly these increases. Reductions were chosen to let relative proportions (as taken from the benchmark scenario) stayed constant. Plant numbers fell from 9 to 7 for lignite, from 15 to 11 for coal, from 11 to 8 for CCGT and from 7 to 5 for SCGT while nuclear plants increased from 6 to 10.

The second alternative scenario is cleaner than the baseline forecasts. It analyzes the hypothetical replacement of lignite capacity by hard coal and natural gas fired units. These technologies replace lignite in those scenarios in proportion to the relative capacity sizes modeled in the baseline scenario in the respective year. The plant counts of hard coal, CCGT and SCGT rise in the 2025 (2040) scenario in this lignite-free future from 14 to 19 (6 to 9), 8 to 12 (5 to 8) and 5 to 9 (4 to 6) respectively. Benchmark fuel prices were used for the lignite-free and nuclear energy based scenarios. For the 2010 'lignite free' scenario the shares of all technologies (including nuclear) in total capacity without lignite were multiplied with total conventional capacity installed including lignite. This changed the plant count in the 2010 scenario from 6 to 8 for nuclear, from 15 to 20 for coal, from 11 to 14 for CCGT and from 7 to 9 for SCGT plants.

3.4.3 Levelized costs of electricity modeling

The LCOE model of power generating units is the standard way for evaluating electricity generation costs.[25] The model discounts all costs and all output in kWh that occur during the lifecycle of a plant and divides the latter by the former. The alternative approach is to calculate the costs/kWh

---

[25] See for example (Branker, Pathak, and Pearce, 2011; Kost and Schlegl, 2010).

by dividing all electricity generated by all costs occurring during one single year. This calculation has the disadvantage that it requires modeling depreciation (where one needs to decide on the economically correct method). It also means that the technically determined evolution of cost and output throughout the plant life time cannot be expressed in a simple way. PV modules are for example experiencing a normal annual output degradation that leads to production being roughly 20 % lower after 20 years of operation; O&M costs of wind turbines are low in the first 5 years of operation and rise to a level that is more than 40 % higher after that. Thus, a LCOE model was used in this study to estimate total costs of electricity generation. Its basic version is

$$LCOE = \frac{\sum_{t=0}^{T} \frac{C_t}{(1+r)^t}}{\sum_{t=0}^{T} \frac{E_t}{(1+r)^t}}$$

where T is the maximum economic life time of the plant or installation, C_t the sum of investment, total variable and fixed costs that occur in year t and E_t the electricity produced in year t. The LCOE inputs total variable costs (that include cycling costs) and utilization rates for conventional plants come from the simulation model. All other information was summarized in tables 9 and 10. For PV and wind, the exactly identical model inputs were used as described in Keil (2013).

3.5 Results and evaluation

In the interpretation one needs to be cautions as variables such as the aggregate annual average utilization factors necessarily differ from the actually observed parameters due to several reasons. 1. The model applied here assumes plants to be all current state-of-the-art types and to be all identical while differences in technology and variable costs are present in reality. 2. It abstracts from scheduled maintenance and the start-stop operations and shut down times associated with it. This abstraction raises annual capacity factors of baseload plants artificially in the simulation. 3. The model does not include unscheduled and unintended plant failures. 4. It abstracts from local grid network capacity limits that might enforce start-ups and shut-downs of plants located in certain network zones. 5. ICEsim does not account for plant ownership and the effects of fuel supply or electricity provision contracts at different prices and for different durations. These are likely to result in the use of higher cost plants even when the capacity of lower cost plants is available. Small municipal utilities have e.g. only a few or even just one plant which is smaller and more flexible (natural gas or hard coal fired). When large utilities would start the lowest-cost plant available in their portfolio, a small utility might prefer to use own (higher cost) plants and accept a lower profit margin on their own plants instead of obtaining electricity from other utilities with cheaper available plant capacity on the wholesale spot market.[26] Ignoring these effects will yield lower capacity factors in the model for hard coal and even more so for natural gas fired plants while inflexible plants have higher factors. 6. The simulation ignores the provision of network stability services like primary and secondary reserves. These services can be reflected in plants operating not always at full load or not starting-up even if the demand situation and the competitiveness would allow for it. While including these factors is beyond the scope of this analysis, the cost-of-production dimension applied here should however be considered as the most important long run driver in a predominantly capitalist industry as the German electric utility sector. The results are thus expected to be of sufficient value.

The most important scenario results are summarized in the three tables of this section. The results for benchmark scenario prices are in both tables in bold text while coal and gas favorable price patterns are included in gray ink above and below the benchmark scenario row respectively for each plant technology. Table 12 includes annual output of all plants by technology, the average annual aggregate utilizations of each plant type and its average annual amount of start-ups. Start-ups are a prime mover of cycling (including start-up) costs and utilization rates determine the amount of fixed costs (capital, fixed O&M) and cycling costs per output unit (e.g. kWh).

One can see how the utilization rates for baseload plants decline significantly. For lignite, the drop is especially significant from 2025 to 2040. Annual average output falls constantly and steeply. Alterations in prices have only limited effects. Interestingly, lignite is threatened more by the ‘coal favorable’ prices even though this price scenario involves low CO2 prices which lower costs of lignite plants. This is so because hard coal is more suited to provide baseload power than CCGT units and is thus a more immediate competitor of lignite technology. Utilization rates and output volumes for hard coal plants under benchmark prices stay constant until 2025. Thereafter, utilizations rise while aggregate output falls - reflecting the fact that the natural retirement of capacities is faster than the shrinkage of the market for coal plant output. The benchmark scenario is already relatively favorable for coal. Thus ‘coal favorable’ prices do not improve the situations much while ‘gas favorable’ prices imply a great danger for coal fired units. Capacity factors of gas fired units are heading upwards in

---

[26] One very obvious reason can be of political nature. Small public utilities are likely to follow goals that are alien to capitalist utilities such as supporting regional employment and self-sufficiency.

Table 12: **Output, utilizations and start-ups***

	Output (TWh)			**Utilization (%)**			**Annual Starts**		
	2010	2025	2040	2010	2025	2040	2010	2025	2040
Nuclear	184.0			100			2		
	184.0			**100**			**2**		
	184.0			100			2		
Lignite	147.3	112.6	44.5	75	64	41	72	122	265
	154.0	**119.4**	**45.9**	**78**	**68**	**42**	**113**	**163**	**277**
	147.4	114.5	44.8	75	65	45	72	130	267
Hard coal	80.4	83.2	39.0	30.6	33.9	39.0	366	420	540
	67.1	**63.0**	**36.2**	**26**	**26**	**34**	**264**	**231**	**486**
	12.2	54.6	36.0	4.6	22.3	34.2	29	176	478
CCGT	7.6	13.8	16.4	5.2	13.1	24.9	184	427	646
	14.1	**26.4**	**17.7**	**10**	**25**	**27**	**265**	**656**	**680**
	73.3	38.7	18.7	50.7	36.8	28.5	528	732	689
SCGT	2.6	3.0	4.9	4.3	6.9	14	311	575	668
	2.7	**3.8**	**5.1**	**5**	**9**	**14**	**318**	**625**	**680**
	5.0	4.7	5.2	8.1	10.8	14.7	377	677	686

** All values are annual averages for the average year and net of any losses. The first line for each plant type reflects the 'coal favorable', the second the benchmark and the third the 'gas favorable' price scenario.*

the two future benchmark scenarios. However, prices will change in the alternative pricing scenarios - annual output and capacity factors for CCGT and SCGT plants will go up in the future (versus the 2010 benchmark scenario). SCGT technology is the only technology whose output grows both from 2010 to 2025 as well as from 2025 to 2040 in the benchmark (and every other price) scenario. The data for SCGT illustrate how the technology that is made to serve very short peaks is really delivering a different product than all other plant types. Its market potential and its economics depend almost exclusively on the occurrence of these potential production cycles. Thus, even unfavorable fuel prices do not present a major threat to lose market shares to CCGT or hard coal.[27] 'Gas favorable' prices show how (on an international scale) even moderate gas price changes turn CCGT plants into baseload units that can replace all existing hard coal units and render them as wholly unnecessary. This illustrates the potential threat to coal plant operators from hydraulic fracturing in Europe or large scale future natural gas exports from the US where natural gas eliminates bituminous thermal coal.

Start-ups increase for all plant types for both future benchmark scenarios. The only exception is hard coal in 2025 where its starts decrease as it replaces baseload nuclear plants and such plants

---

[27] SCGT units are competing more with flexible hydroelectric and electricity storage plants. Since this could not be modeled in detail in this study, at least some negative impact of higher fuel prices for SCGT plants is likely to occur.

run for longer periods. The increase of starts is (with this exception) relatively dramatic. Gas fired units more than double their starts from less than one per day in 2010 to more than twice a day in 2040. Coal also starts more than once a day in 2040 where even lignite plants are forced to start 277 times in 365 days - nearly operating like peak load plants operate today.

The cost components of the baseline scenarios that correspond to the operating regimes discussed in the preceding paragraph are summarized in table 13. The costs labeled 'variable' are variable operating costs under nominal load that exclude start-stop or cycling operation costs. They are constant over time due to identical prices and technologies assumed. One can see that the 'cycling' cost component which includes load following/partial load as well as start-up costs can be of a significant magnitude whenever many start-ups and lower utilization rates occur.[28] The most significant development takes place for lignite and SCGT plants. The former's cycling costs double until 2025 and quadruple until 2040 and the latter's fall by nearly 50 %. The generally much higher cycling cost values of flexible plants are rooted in the much more frequent plant starts and the shorter operating

Table 13: **Cost factors, in cent/kWh***

	Variable	**Cycling**			**Fixed**			**Total**		
		2010	2025	2040	2010	2025	2040	2010	2025	2040
Nuclear	1.4	0			4.31			5.67		
	1.4	**0**			**4.31**			**5.71**		
	1.4	0			4.31			5.67		
Lignite	1.9	0	0.29	0.98	2.04	2.38	3.76	4.12	4.6	6.67
	2.1	**0.23**	**0.37**	**0.94**	**2.04**	**2.24**	**3.73**	**4.28**	**4.71**	**6.77**
	3.1	0.17	0.34	1.03	2.0	2.34	3.73	5.28	5.75	7.9
Coal	2.4	1.23	1.28	1.50	4.37	3.94	3.61	8.02	7.64	7.53
	3.0	**1.14**	**0.99**	**1.55**	**6.24**	**5.20**	**3.89**	**9.38**	**9.19**	**8.44**
	4.2	0.78	0.97	1.71	29.1	6.0	3.91	34.07	11.19	9.84
CCGT	4.8	2.72	2.5	1.98	14.84	5.89	3.1	22.4	13.23	9.92
	4.4	**2.05**	**2.04**	**1.88**	**7.90**	**3.07**	**2.86**	**14.4**	**9.41**	**9.14**
	4.3	0.77	1.46	1.78	1.52	2.1	2.7	6.53	7.8	8.72
SCGT	7.2	4.41	5.05	2.88	12.59	7.84	3.87	24.23	20.12	13.98
	6.6	**4.25**	**4.35**	**2.79**	**12.0**	**6.37**	**3.76**	**22.9**	**17.3**	**13.2**
	6.3	2.8	3.7	2.75	6.68	5.0	3.68	15.8	15.0	12.8

'Variable': var. operating costs at normal capacity; 'Cycling': start-up & cycling costs; 'Fixed': fixed operating & capital costs. The first line for each plant reflects the 'coal favorable', the second the benchmark and the third the 'gas favorable' price scenario.

---

[28] Partial load and load alteration cycling costs usually amount to roughly 10 % of the 'cycling' costs in the table while start-up costs account for around 90 %.

times that lead to lower utilization factors. Rising utilizations for the two natural gas technologies are the drivers behind the fall of this cost component per output unit. ‘Fixed’ cost components include fixed O&M as well as capital costs. Since they are also expressed in cent/kWh terms, the determinant behind their change is the average annual utilization rate. Consequently, fixed costs for all plant types but lignite fall - for natural gas plants dramatically.

The overall effect is summarized by the ‘total’ LCOE which are the sum of ‘variable’, ‘cycling’ and ‘fixed’ costs. The differentials in these total costs of production that exist between inflexible and flexible plants under the late 2012 pricing scenario are shrinking dramatically and constantly under the type of industrial change to which the German electricity industry is exposed. Lignite power becomes 50 % more expensive. At the same time the advantage of hard coal versus natural gas plants declines as costs of the former stay constant until 2025 and decline by only 1 cent until 2040. SCGT plants are bound to become much more competitive in each of the six future scenarios reviewed. Still the hierarchy where lignite is cheapest, followed by coal, CCGT and SCGT does not change under benchmark and ‘coal favorable’ price scenarios. However, one needs to keep in mind that the value of peaking power is much higher and flexible plants are able to sell their output for higher prices than inflexible plants. Given the rise in intermittent renewable load and the increase in overproduction and shortages that are hard to fill by conventional plants, one can expect much lower average selling prices of baseload power and significantly higher ones for flexible power plants. The recent surge in the share of coal and lignite in power generation thus seems to be only of temporary nature and a phenomenon of the early stage of the German energy transition that will last for at least four decades. If prices for CO2, coal and natural gas move within the ranges observed over the last three years into directions that are favorable to natural gas, it is possible that CCGT render coal as uncompetitive on a full cost basis. The core insight of the simulation is that CCGT as well as SCGT units will become more competitive and more compatible with an industry that has a large and growing share of intermittent renewable energy. However, the scenarios also show that there are threats to the commercial viability of natural gas plant operators. In order to save the plant capacity that will be essential in future, policy should intervene. A capacity market where operators are payed for keeping capacity reserves available might make sense in the immediate future. However, a systematic and gradual phase-out period should be implemented at the same time since the fundamentals of flexible plants (costs, utilization rates) are about to improve significantly in all price scenarios as time goes on. One viable option might be to index possible capacity subsidies to fuel input costs. Interestingly, policy might also be forced to support hard coal instead of gas fired plants if the fundamentals of input prices were about to change persistently as they did in the US.

A change in the cost-of-production hierarchy only occurs when renewables are included in the list. This is illustrated by figure 6. It includes the current LCOE from renewable sources as of late 2012/early 2013 and the projected future values for 2025 as described in Keil (2013). According to the 2025 cost prediction for renewables, onshore and offshore wind will both be cheaper than all flexible conventional sources (including coal) in 2025. Even PV power will be cheaper than coal and gas fired units. The '2040' scenario includes 2025 cost predictions for renewables and not 2040 values. This is due to too much uncertainty while it is nevertheless certain that technological improvements will further drive down renewable energy costs. This 2040 cost comparison shows that unrealistically expensive (at 2025 costs) onshore and offshore wind will become the cheapest energy sources. They generate electricity even in a more cost competitive way than lignite. Utility companies will thus have incentives to replace conventional with renewable capacity on a pure cost basis.[29]

The described two alternative energy policy scenarios that include nuclear power and that exclude lignite plants are used to further evaluate how suited baseload plant are to support the energy transition. The next table summarizes the most important outcomes.

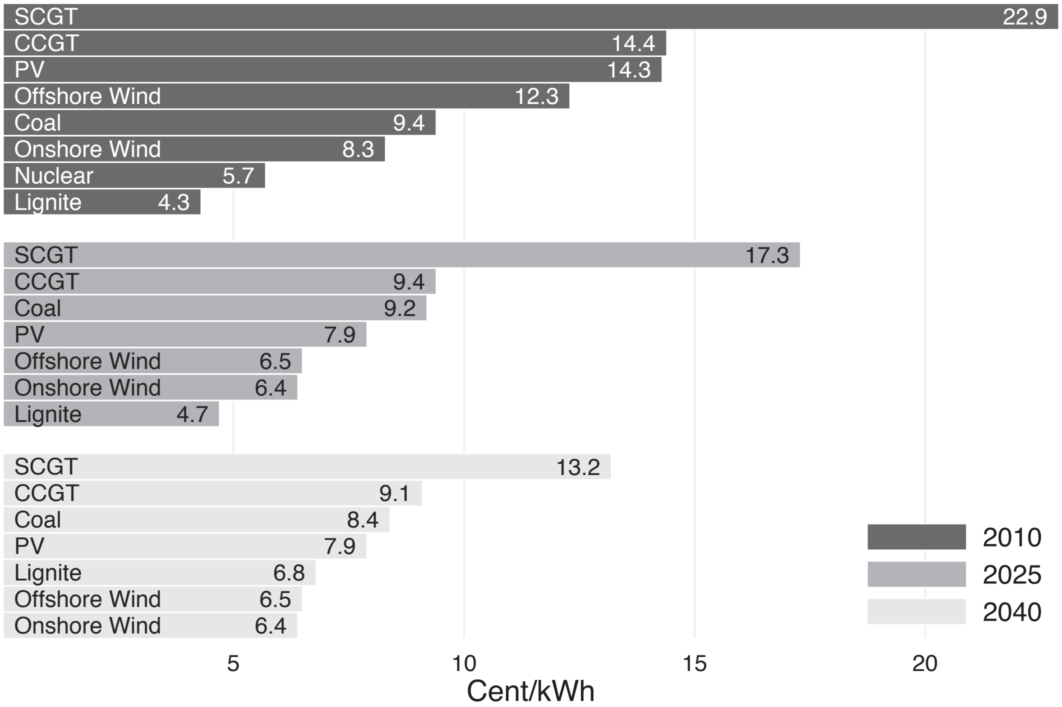
In the lignite free scenario coal delivers electricity for the longest lasting production cycles, resulting in fewer starts, much higher utilizations and thus significantly lower costs. The effect on the economics of CCGT plants is negative in 2010 and 2025, but becomes positive (with a moderate size impact) in 2040. Since a part of lignite capacity is replaced by SCGT, this augmentation in capacity produces fewer annual starts and utilization rates per plant - as well as higher costs. The overall effect of replacing lignite with coal, CCGT and SCGT plants would inflate the average costs of electricity generation of all conventional plant output in 2010 from 6.17 cent/kWh to 6.52 cent/kWh (see following table). In 2025, there would be no effect on the costs of 6.84 cent/kWh. In 2040, average conventional costs of production would actually decrease from 8.06 to 7.72 cent/kWh. It is interesting to note that the benchmark scenario numbers represent only a moderate future increase in electricity generation costs.

The nuclear future scenarios and the more nuclear intensive 2010 scenario is moderately favorable

---

[29] The market value of intermittent renewable output is of course also lower than that of conventional plants that deliver on demand.

Figure 6: **Levelized costs of electricity - today and in future**



Conventionals: baseline scenario LCOE. Renewables: 2025 and 2040 values are predictions for 2025. 2012 values were taken for the 2010 scenario. Calculations for renewables are from Keil (2012).

Table 14: **Outcomes for alternative policy scenarios***

		Utilization, %			**Start-ups**			**Tot. Cost**, ct/kWh		
		2010	2025	2040	2010	2025	2040	2010	2025	2040
Nuclear	N	96.9	73.1	41.1	30	185	334	5.89	7.66	13.1
	B	**100**			**2**			**5.71**		
	L	99.5			8			5.74		
Lignite	N	33.9	40.6	34.3	146	177	447	7.27	6.53	8.55
	B	**78**	**68**	**42**	**113**	**163**	**277**	**4.28**	**4.71**	**6.77**
Coal	N	28.8	25.3	28.3	328	369	671	8.89	9.88	10.3
	B	**26**	**26**	**34**	**264**	**231**	**486**	**9.38**	**9.19**	**8.44**
	L	45.9	53	68	247	207	287	6.5	5.95	6.84
CCGT	N	13.6	27.5	24.8	375	773	740	12.1	9.3	9.72
	B	**10**	**25**	**27**	**265**	**656**	**680**	**14.4**	**9.41**	**9.14**
	L	7.6	20	29	201	471	539	16.54	9.95	8.46
SCGT	N	6.1	9.8	14.9	439	730	755	19.8	16.5	13.2
	B	**5**	**9**	**14**	**318**	**625**	**680**	**22.9**	**17.3**	**13.2**
	L	3.4	5	13	247	353	538	26.83	22.6	13.3

**N: 'nuclear scenario'; B: benchmark scenario; L: 'lignite free scenario'.*

on a cost basis for SCGT plants - indicating that this technology is a good or at least necessary supplement to nuclear power (if no flexible hydroelectric capacity is available). However, the impacts on CCGT and coal plants in the future tend to be negative (in 2010 positive). This is because all flexible conventional plants are forced to start-up even more often than in the already cycling intensive benchmark future. This elevates their high cycling costs even more. Lignite utilization rates drop dramatically, start-ups spike up and costs are driven to 200 % of the value found in the 2010 benchmark scenario. This reflects the rivalry of the two baseload technologies. The costs of nuclear power itself go up to 7.66 and 13.1 cent/kWh in 2025 and 2040. Only the highly flexible SCGT output would be higher (by 0.1 cent) than the less flexible nuclear power in 2040. This demonstrates how unsuited nuclear technology is as a complement for large amounts of renewable energy. This clearly disproves opposite claims by the nuclear industry. The total average costs of electricity from conventional sources would also be elevated significantly in every scenario to 6.76, 8.15 and 10.83 cent/kWh in 2010, 2025 and 2040.

Table 15: **Average costs of production (conventionals), cent/kWh**

Policy Scenario	2010	2025	2040
Nuclear intensive	6.76	8.15	10.83
Benchmark	6.17	6.84	8.06
Lignite-free	6.52	6.84	7.72

The total average costs of production of all conventional plants are summarized in table 15. It shows how the energy mix in 2010 of the really existing benchmark scenario involves the lowest costs. Increasing or decreasing the amount of plants suited for baseload power would lead to cost spikes. The costs of production in the future scenarios are positively effected by substituting inflexible for flexible plants, while going into the other direction produces negative results. The policy recommendation from this is that Germany and other countries that target higher shares of intermittent renewable power should rapidly phase out nuclear power. Once renewables reach a share of 40 %, lignite plants (and old coal plants that are also less flexible) should be retired and replaced by cleaner burning, flexible new coal or natural gas fired units. This shows that under the condition of a growing renewable energy share there is on a pure commercial cost basis no long run environment-cost/competitiveness trade-off. Dirty and more risky electricity generation becomes more expensive while clean(er) conventional energy helps to reduce the costs of electricity generation and thus to potentially lower electricity prices, increase national wealth and industrial competitiveness. The simulation also shows that conventional energy costs do not need to increase as dramatically due to renewables as some industry observers suggest. If the conventional mix is being adjusted in an optimal way, additional costs of conventional power due to flexibility do not need to exceed an increase of 25 % or 1.55 cents/kWh.

Bibliography

50Hertz Transmission (2012): “http://www.50hertz.com/de/Netzkennzahlen.htm. Reviewed 10/09/2012,” .

Abbasoglu, O. F., A. F. Aysan, and A. Günes (2007): “Concentration, Competition, Efficiency and Profitability of the Turkish Banking Sector in the Post-Crisis Period,” *Banks and Bank Systems*, 2(3).

Acquaah, M. (2003): “Corporate Management, Industry Competition and the Sustainability of Firm Abnormal Profitability,” *Journal of Management and Governance*, 7(1), 57–85.

Ali, A., S. Klasa, and E. Yeung (2009): “The Limitations of Industry Concentration Measures Constructed with Compustat Data: Implications for Finance Research,” *The Review of Financial Studies*, 22(10), 3839–3871.

Amprion (2012): “http://www.amprion.net/netzkennzahlen. Reviewed 10/09/2012,” .

Areva (2010): “Regenerative Energieträger und Kernkraftwerke – passt das zusammen?,” *argumente - Energie Umwelt Gesellschaft*.

Auer, J., and J. Keil (2012): “State-of-the-Art Electricity Storage Systems, Indispensable Elements of the Energy Revolution,” Current issues, Deutsche Bank Research, Frankfurt / Main, Germany.

Bain, J. S. (1951): “Relation of Profit Rate to Industry Concentration: American Manufacturing, 1936-1940,” *The Quarterly Journal of Economics*, 65(3), 293–324.

Bain, J. S. (1956): *Barriers to New Competition*. Harvard University Press, Cambridge, Massachusetts.

Barney, J. (1991): "Firm Resources and Sustained Competitive Advantage," *Journal of Management*, 17(1), 99–120.

Baumol, W. J. (1982): "Contestable Markets: An Uprising in the Theory of Industry Structure," *The American Economic Review*, 72(1), 1–15.

Branker, K., M. J. Pathak, and J. M. Pearce (2011): "A Review of Solar Photovoltaic Levelized Cost of Electricity," *Renewable & Sustainable Energy Reviews*, 15, 4470–4482.

Bundesnetzagentur (2011): "Monitoringbericht," Discussion paper, Bonn, Germany.

——— (2012): "http://www.bundesnetzagentur.de/cln_1912/DE/ Sachgebiete/ElektrizitaetGas/ErneuerbareEnergienGesetz/. Reviewed 10/25/2012," .

BWE (2012): "http://www.wind-energie.de/infocenter/statistiken. Reviewed 01/25/2012," Discussion paper, Bundesverband WindEnergie.

Cable, J. R., and D. C. Mueller (2008): "Testing for Persistence of Profits' Differences Across Firms," *International Journal of the Economics of Business*, 15(2), 201–228.

Caves, R. E., and M. E. Porter (1977): "From Entry Barriers to Mobility Barriers: Conjectural Decisions and Contrived Deterrence to New Competition," *The Quarterly Journal of Economics*, 91(2), 241–262.

Chamberlin, E. H. (1929): "Duopoly: Value Where Sellers Are Few," *Quarterly Journal of Economics*, 43, 63–100.

Clarke, R. G. (1985): *Industrial Economics*. Oxford: Basil Blackwell.

Cowling, K., F. M. Yusof, and G. Vernon (2000): "Declining Concentration in UK Manufacturing? A problem of measurement," *International Review of Applied Economics*, 14(1), 45–54.

Cubbin, J., and P. Geroski (1987): "The Convergence of Profits in the Long Run: Inter-Firm and Inter-Industry Comparisons," *The Journal of Industrial Economics*, 35(4), 427–442.

Cusick, D. (2013): "Technology helps wind industry overcome 'start-stop' hurdles. http://www.governorswindenergycoalition.org/?p=4448. Reviewed 01/25/2012," .

DEEDS, D. L., AND C. W. L. HILL (1996): "The Importance of Industry Structure for the Determination of Firm Profitability: A Neo-Austrian Perspective," *Journal of Management Studies*, 33(4), 429–451.

DEMSETZ, H. (1973): "Industry Structure, Market Rivalry, and Public Policy," *Journal of Law and Economics*, 16(1), 1–9.

DENA (2005): "Energiewirtschaftliche Planung für die Netzintegration von Windenergie in Deutschland an Land und Offshore bis zum Jahr 2020," .

——— (2010): "Integration erneuerbarer Energien in die deutsche Stromversorgung im Zeitraum 2015 – 2020 mit Ausblick 2025.," .

EKLUND, J. (2008): "Q-Theory of Investment and Earnings Retentions - Evidence from Scandinavia," *CESIS Electronic Working Paper Series, Paper No. 157, The Royal Institute of technology, Centre of Excellence for Science and Innovation Studies (CESIS).*

ENTSOE (2012): "European Network of Transmission System Operators for Electricity. https://www.entsoe.eu/data/data-portal/consumption/. Reviewed 09/28/2012," .

GALE, B. T., AND B. S. BRANCH (1982): "Concentration versus Market Share: Which Determines Performance and Why Does it Matter?," *The Antitrust Bulletin*, 27, 83–105.

GODDARD, J., H. LIU, P. MOLYNEUX, AND J. O. WILSON (2011): "The Persistence of Bank Profit," *Journal of Banking and Finance*, 35, 2881–2890.

GSCHWANDTNER, A. (2005): "Profit Persistence in the 'Very' Long Run: Evidence from Survivors and Exiters," *Applied Economics*, 37, 793–806.

——— (2012): "Evolution of Profit Persistence in the USA: Evidence from three Periods," *The Manchester School*, 80(2), 172–209.

GUGLER, K., D. C. B. MUELLER, AND B. YURTOGLU (2004): "Marginal q, Tobin's q, Cash Flow, and Investment," *Southern Economic Journal*, 70(3).

HAYEK, F. (2002): "Competition as a Discovery Procedure," *The Quarterly Journal of Austrian Economics*, 5(3), 9–23.

HIRSCH, B. T. (1991): *Labor Unions and the Economic Performance of Firms.* W. E. Upjohn Institute, Kalamazoo, Michigan.

HIRSCH, B. T., AND D. A. MACPHERSON (2003): "Union Membership and Coverage Database from the Current Population Survey: Note," *Industrial and Labor Relations Review*, 56(2), 349–354.

HUNDT, M., R. BARTH, N. SUN, S. WISSEL, AND A. VOSS (2009): "Verträglichkeit von erneuerbaren Energien und Kernenergie im Erzeugungsportfolio; Technische und ökonomische Aspekte," .

IEA (2010): "Power Generation from Coal - Measuring and Reporting Efficiency, Performance and CO2 Emissions," .

KAMBHAMPATI, U. S. (1995): "The Persistence of Profit Differentials in Indian Industry," *Applied Economics*, 27, 353–361.

KAUFMAN, B., AND J. HOTCHKISS (2006): *The Economics of Labor Markets.* Thomson South-Western, Mason, Ohio.

KEHLHOFER, R. (1997): *Combined-Cycle Gas and Steam Turbine Power Plants.* PennWell Publishing Company, Tulsa, Oklahoma.

KEIL, J. (2013): "The German Energy Transition – Issues and Perspectives (upcoming)," American Institute for Contemporary German Studies - John Hopkins University.

KIRZNER, I. M. (1997): "Entrepreneurial Discovery and the Competitive Market Process: An Austrian Approach," *Journal of Economic Literature*, 35(1), 60–85.

KOST, C., AND T. SCHLEGL (2010): "Studie Stromgestehungskosten Erneuerbare Energien," Discussion paper, Fraunhofer Institut für Solare Energiesysteme ISE.

KUMAR, N., P. M. BESUNER, S. A. LEFTON, D. D. AGAN, AND D. D. HILLEMAN (2012): "Power Plant Cycling Costs," Discussion paper, Intertek APTECH, Sunnyvale, CA and Houston, TX.

KUNTZ, L., AND F. MÜSGENS (2011): "Modelling start-up costs of multiple technologies in electricity markets," *Mathematical Methods of Operations Research*, 66, 21–32.

LEFTON, S. A., AND D. HILLEMAN (2012): "PE, Intertek-Aptech. Make Your Plant Ready for Cycling Operations. http://www.powermag.com/coal/3885.html. Reviewed 10/15/2012," .

Lewis, J. B. (2000): "Estimating Regression Models in which the Dependent Variable Is Based on Estimates with Application to Testing Key's Racial Threat Hypothesis," .

Marx, K. (1991): *Capital Volume III*. Penguin Books, New York, New York.

——— (1997): *Wages, Price, and Profit*. Foreign Languages Press, Peking, China.

McGahan, and M. E. Porter (1997): "How Much Does Industry Matter, Really?," *Strategic Management Journal*, 18(Special Issue: Organizational and Competitive Interactions), 15–30.

McGahan, A. M., and M. E. Porter (1999): "The Persistence of Shocks to Profitability," *The Review of Economics and Statistics*, 81(1), 143–153.

McGahan, A. M., and M. E. Porter (2003): "The Emergence and Sustainability of Abnormal Profits," *Strategic Organization*, 1(1), 79–108.

Mueller, D. C. (1977): "The Persistence of Profits Above the Norm," *Economica*, 44, 369–80.

Mueller, D. C. (1986): *Profits in the Long Run*. Cambridge University Press, New York.

Porter, M. E. (1980): *Competitive Strategy, Techniques for Analyzing Industries and Competitors*. The Free Press, New York, New York.

r2b energy consulting (2012): "Jahresprognose zur deutschlandweiten Stromerzeugung aus EEG geförderten Kraftwerken für das Kalenderjahr 2013," .

Ratnayake, R. (1996): "Industry Concentration–Profitability Relationship and Competition Policy: is there a Critical Concentration Level?," *Applied Economics Letters*, 3, 611–614.

Ravenscraft, D. J. (1983): "Structure-Profit Relationships at the Line of Business and Industry Level," *Review of Economics and Statistics*, 65, 22–31.

Ricardo, D. (1996): *Principles of Political Economy and Taxation*, Great Mind Series. Prometheus Books, Amherst, New York.

RWE AG (2009): "The Need for Smart Megawatts Power Generation in Europe – Facts & Trends," Discussion paper, Essen, Germany.

——— (2011): "FACTS & FIGURES," Discussion paper, Essen, Germany.

SASS, P. (1975): *Die Untersuchung der Profitraten-Unterschiede zwischen den westdeutschen Industriebranchen nach dem 2. Weltkrieg*. Mohr, Tuebingen, Germany.

SAXONHOUSE, G. R. (1976): “Estimated Parameters as Dependent Variables,” *American Economic Review*, 66, 178–183.

SCHERER, F. (1980): *Industrial Market Structure and Economic Performance*. Rand McNally College Publishing Company, Chicago, Illinois, 2 edn.

SCHMALENSEE, R. (1985): “Do Markets Differ Much?,” *The American Economic Review*, 75(3), 341–351.

SCHUMPETER, J. A. (1947): “The Dynamics of Competition and Monopoly (Excerpt from Capitalism, Socialism and Democracy),” pp. 40–67.

SEMMLER, W. (1981): “Competition, Monopoly, and Differentials of Profit Rates: Theoretical Considerations and Empirical Evidence,” *Review of Radical Political Economics*, 13(4), 39–52.

——— (1984): *Competition, Monopoly, and Differential Profit Rates. On the Relevance of the Classical and Marxian Theories of Production Prices for Modern Industrial and Corporate Pricing*. Columbia University Press, New York, NY.

SHAIKH, A. (2008): “Competition and Industrial Rates of Return,” *Issues in Finance and Industry, Essays in Honour of Ajit Singh*, pp. 167–194.

SIOHANSI, R. (2011): “Increasing the Value of Wind with Energy Storage,” *The Energy Journal*, 32(2), 1–29.

SMITH, A. (1991): *Wealth of Nations*, Great Mind Series. Prometheus Books, Amherst, New York.

SPANOS, Y. E., G. ZARALIS, AND S. LIOUKAS (2004): “Strategy and Industry Effects on Profitability: Evidence from Greece,” *Strategic Management Journal*, 25(2), 139–165.

STIGLER, G. J. (1963): *Capital and Rates of Return in Manufacturing Industries*. Princeton University PRess, Princeton, NJ.

TENNET TSO (2012): “http://www.tennettso.de/site/Transparenz/veroeffent lichungen/netzkennzahlen. Reviewed 10/09/2012,” .

Traber, T., and C. Kemfert (2011): “Gone with the Wind? - Electricity Market Prices and Incentives to Invest in Thermal Power Plants under Increasing Wind Energy Supply,” *Energy Economics*, 33(2), 249–256.

TransnetBW (2012): “http://www.transnetbw.de/kennzahlen. Reviewed 10/09/2012,” .

Umweltbundesamt (2011): “CO2-Emissionsminderung durch Ausbau, informationstechnische Vernetzung und Netzoptimierung von Anlagen dezentraler, fluktuierender und erneuerbarer Energienutzung in Deutschland,” .

Vuorinen, A. (2007): “Fundamentals of Power Plants. Planning of Optimal Power Systems.,” .

Webster, A. L. (1996): “The Concentration-Profitability Relationship in American Industry: A Varying Parameters Model,” *Journal of Applied Business Research*, 12(4), 44–50.

Yurtoglu, B. B. (2004): “Persistence of Firm-Level Profitability in Turkey,” *Applied Economics*, 36(6), 615–625.

Appendix

The Compustat Segments File was downloaded from WRDS on April/03/2011 and the variables net sales (SALE), total assets (AT), operating profit (OPS), capital expenditures (CAPX), employment (EMP) as well as research and development expenditures (RD) were extracted. The file could not be used in its raw state due to coding practices, errors and modifications that were necessary for this present analysis. The dummy values for ‘insignificant’ figures (-0.008) were decoded to zero, those for ‘not available’ and ‘not meaningful’ and ‘combined’ figures (-0.001, -0.007, -0.004) to missing values. Entries were decoded as missing if an operating segment reported negative values for sales, total assets, employment and R&D expenditures. For the industry classification of segments, I used the primary NAICS segment industry classification codes (SNAICS1) wherever possible and the primary segment SIC code (SSIC1) where SNAICS1 data were missing (data from older classification systems were translated later). To avoid duplication of information, I deleted all non-trading companies which had trading divisions (their ‘Ticker Symbol’ variable SMBL ended with ‘.CM’ and was created by S&P) as well as all subsidiaries of publicly traded companies (their symbol consisted of numbers that are preceded by the parent company’s ticker symbol). In both cases there was a significant amount of duplication. I also excluded all subsidiaries of companies with no publicly traded common stock (their symbol was a four digit number followed by the letter A). In some cases data have been restated (indicated by the source year variable SRCYR), resulting in more than one observation for each segment-year unit. Since reporting systems change, one should not use different segments from different source years for the same year. I used those restatements with the fewest missing values for the variables CAPX, EMP and RD, the highest amount of segments per company-year-restatement year or in the case of identical numbers, the most recent restatement. Some firms in the file divided their segments into operating and non-operating ones. It was not directly obvious for every single company whether it reports data for non-operating activities separately in an own segment. In cases where this has been done it was not necessarily clear which segments were operating and which non-operating. In most cases where the variables SNAICS1 or SSIC1 and SGEOTP (the geographic segment type variable that indicates domestic and non-domestic) had missing values, the segment names (variable SNAME) implied either: 1. non-operating segments or shared functions (‘Corporate’, ‘Headquarters’, ‘Holding’, ‘Common Expenses’, ‘Central Functions’, ‘Administrative’, ‘Group Operations’, ‘General Costs’, ‘Overhead’, ‘Shared Services’, ‘Non-Operating’ etc.); 2. minor or not allocable activities (‘Other’, ‘Miscellaneous’, ‘Non-Core’, ‘Unassignable’, ‘Not-Allocated’); or

3. accounting categories that cannot be assigned to specific segments (‘Elimination’, ‘Adjustment’, ‘Reconcilitation’, ‘Intersegment’, ‘Intangibles’, ‘Discontinued’, ‘Sold Businesses’, ‘Liquidating’). I assumed that all observations with missing values for SNAICS1 and SSIC1 were such cases that could not be allocated along industrial lines and that all observations with missing SGEOTP information were cases that could not be allocated along geographical lines. In some cases there was lacking information and/or errors regarding geographical information in the OPSEG and GEOSEG segments. I modified the geographical codes for foreign and domestic classifications by e.g. recoding segments whose segment names variable was equal to ‘United States’ or ‘USA’ and which were incorporated in the U.S. as domestic segments or by coding every segment as foreign if the segment name contained e.g. ‘Foreign’. Foreign incorporated companies were treated equivalently. Companies which were incorporated in offshore locations like the Cayman Islands or the Bermudas were recoded as U.S. incorporated companies if their headquarters were located in the U.S. (the variable indicating this was obtained from Compustat’s North American Fundamentals File). The business segment unit of analysis in this paper includes all segments within an industry in the U.S. and thus differs from the different Compustat segments which are also different from each other depending on the segment type (OPSEG, GEOSSEG etc.) as well as on the different reporting options within the OPSEG segment type. My unit makes sense for analyses in areas like competition theory or industrial organization that investigate factors like the effects of firms size and market concentration. Merging segments in the same country and industry makes sense as they often share necessary expenditures, jointly purchase of inputs or play coordinated roles in the overriding corporate strategy when it comes to pricing, differentiation, investment or output decisions (it is usually also ownership that is considered as the central dimension for any analysis of market power and collusion). Some companies in the data set report operating and non-operating segments while others only report operating ones or are made up of a single segment only. In cases where non-operating-segment-information was reported separately, most companies did not allocate these to the different operating segments along industrial and/or geographical lines. To ensure comparability of these different cases, non-operating segments were allocated across operating segments in proportion to their sale volumes. This has been applied to the company level aggregate of all segments without geographical and industrial information - usually ‘corporate’ or ‘elimination’. The company level aggregate of all segments with missing NAICS information (headquarters etc. assigned to the U.S.) was multiplied with the respective share of the industry in all domestic company sales and allocated across the different U.S.

industrial activities of the firm. Entries with NAICS but no geographical information were aggregated for each NAICS industry and allocated to the operating company segments that were active in these industries. In cases where there was no single operating segment at all that was identified as a U.S. segment and that had an assigned industry classification code, variable values were coded as missing values and not considered any further. These allocations of all segments that represent some headquarter, corporate or other non-operating activity between operating segments produced a higher degree of comparability and makes economically sense as these activities and their associated costs are likely to be to a large degree necessary for any segment business activity to function at all and should thus be included in the unit of analysis. The allocation of non-operating segments to the operating ones implied that I had to calculate domestic and industrial shares of U.S. industrial activities in their company aggregate (that includes all industries and countries). I calculated the U.S. shares in total sales[30] for each of the group of operating segments within an industry classification (domestic for U.S. incorporated, U.S. for foreign incorporated companies). This was only possible when the geographical information was in the form of OPSEG data and located on the NAICS industrial segment level; when the company was only active in a single industry (geographical information being in OPSEG or GEOSEG form); or when none of the company operations was reported to be located outside of the U.S. (I also included segments whose industrial-U.S. unit accounted for at least 90% of total company sales). All other cases were dropped from the file. In the case of no foreign or no geographical information at all, companies were assumed to be wholly domestic. In the case of companies that only reported foreign and missing geographical information, segments with missing values for the geographical variable were assumed to be domestic ones (which seemed to be a common reporting pattern in some cases).[31] The method of calculating U.S. shares of industrial activities of foreign incorporated companies was equivalent to this calculation of domestic shares for U.S. incorporated companies. The industrial shares were calculated from OPSEG and BUSSEG information. Non-operating segments were ignored in the calculations. I used sales data and only NAICS-U.S. units for which there was non-missing, non-zero AT, SALE and OPS data in at least

---

[30] It was a common pattern in many cases that companies reported a detailed breakdown of profit, investment and other data on industrial level in the OPSEG or BUSSEG reporting types only and had sales as the variable to describe the geographical division of business activities. The 'GEOSEG' segment type divides the company along geographical lines only, while the 'BUSSEG' type divides it along the different industrial activities. 'OPSEG' allocates business activities of a company along either industrial or geographical dimensions or a combination of both (information from the state segment type 'STSEG' has been deleted and ignored). Companies either report for a given year only along one segment type or under BUSSEG and GEOSEG or under OPSEG and GEOSEG.

[31] All of these geographical share calculations ignored non-operating segments and segments that were lacking geographical information (I only used those where the SGEOTP variable indicated either 2=domestic or 3=foreign).

one operating segment in the NAICS industry respectively. The resulting version of the data set is superior to 'solutions' which focus on the U.S. company level as the unit of analysis instead and that do not consider geographical divisions or activities in different industries at all. It is a major improvement over the use of segments as reported in the file without dealing with the problems of non-operating segments, of more than one company segment being active in the same industry or of geographical divisions of segment sales (when using BUSSEG information).

The concentration file assembled here includes all official concentration data that have ever been published electronically by the U.S. Census Bureau at the U.S. Department of Commerce, including files which are not available online. The Economic Census usually collects concentration data every five years and includes the share in the total value of industry sales of the biggest 4, 8, 20 and 50 companies. The concentration ratios are available on different levels of aggregation, from 2 to 6 digit-NACIS codes. The files also include the HHI of the largest 50 firms for manufacturing industries from 1982 onwards as well as the total number of establishments in the industry and the sale volumes for all publication years and industries. Non-manufacturing data have not been published in electronic form before 1987, but from then on in five year intervals until 2007. Manufacturing concentration data (without the HHI) reach back to 1935.[32] In the case of non-manufacturing concentration data of the years 1997 and 2002, unofficial NAICS codes with 7 and 8 digits have been used in some cases. There were also cases where additional variables broke industries down according to e.g. different types of seller groups (like merchant and other seller types in the wholesale trade) or according to their tax status (taxable, tax exempt) in the service sector. I ignored these subgroups since the aggregate that included all categories was usually available. For cases where it was not, I used the non-taxable subgroup (in case of the division along this line) or the largest group in all other cases and also in cases with missing explanations for the subgroups. These approximations were, however, not considered in the econometrics here. The non-manufacturing files for 1992 and 1987 contained key variables that had to be combined with the broader trade which the industries were a part of and then be translated with concordance tables (some not being available online) in order to arrive at official industry classifications. As with the other complications mentioned, cases where there was no clear one-on-one translation into an official SIC industry possible were excluded in the present analysis. There were also instances where single concentration ratios were missing or inconsistent

---

[32] No data have ever been collected by the Census for construction, mining and agriculture industries as well as for management of companies and enterprises.

(e.g. when the market share of the largest 8 companies was larger than the one of the largest 20). All of these observations were excluded. Missing values were filled through linear interpolation that treated the four ratios as equal distances (e.g. a missing CR8 being interpolated as 20% if the CR4 and CR20 were 10% and 30% respectively). If at least one ratio was missing, industries were excluded in the regressions. The final file contains concentration data of 8,971 industry-years. For 7.513 of those, none of the described data quality uncertainties occurred.[33]

The unionization data were obtained from the Unionstats webpage for the years 1983 to 2010[34] and from the raw CPS files for the years 1973 to 1981[35]. The raw CPS data required data manipulations to bring them into a usable format and to match them to the Unionstats data files. They contained data on the number of workers surveyed in total, the number of union members and workers covered by union contracts as well as the total number of employees within the industry by 3-digit Census Industry Classification (CIC). This most detailed level available was used for the higher digit industries that were part of each 3-digit industry group. Since some of the older surveys included only a few observations for some industries, I calculated centered moving averages.[36] The old codes 1970 CIC, 1980 CIC and 1990 CIC had to be translated into 2000 CIC, and in a second step into official 1997 NAICS classifications.[37] All industry years where less than 25 observations were available in the original year or with less than 50 in the whole calculation (including the observations of the years that were used in the application of moving averages) were deleted (a total of 4,228 observations). Missing values were estimated through linear interpolation (generating 1,481 observations). The final file contains all published data on industry level unionization in the U.S. that have ever been produced by official sources with a total of 52,837 final industry year data points.

The variables HHFv, HHFd, HHFu, HHFd, IMP, EXP described above as well as the share of

---

[33] The three most recent years, 2007, 2002 and 1997, have more than 1,500 observations each, 1992 899, 1987 870 and the years in the more distant past less than 400 each (after 1960: less than 200 per year). In the case of NAICS codes, 1,758 observations are on the 6-digit, 1,782 on the 5-digit, 791 on the 4-digit, 256 on the 3-digit and 32 on the 2-digit level. In the case of SIC codes, 3,868 are on the 4-digit, 421 on the 3-digit and 63 on the 2-digit level. There are 3,240 manufacturing industry observations that included the HHI.

[34] The files that are already industry aggregates are available at http://www.unionstats.com/. The data base is being maintained by Barry T. Hirsch and David A. Macpherson. Further details can be obtained from Hirsch and Macpherson (2003).

[35] The employee level survey files are provided by the NBER and available at http://www.nber.org/data/cps_may.html. Additional information was provided by the BLS and was reviewed at http://www.bls.census.gov/cps/bindcd.htm and http://www.bls.census.gov/cps/bocccd.htm. No information on unionization has been collected in the 1982 survey.

[36] For those industry years with less than 100 observations, I calculated a three year centered moving average where weights of 0.5 were being used for t-1 and t+1 (and weight of 1 for t). For cases where less than 50 workers were surveyed, I applied a five year centered moving average with weights of 0.25 for t-2 and t+2 (the other weights as before).

[37] The years 1973 - 1981 are in 1970 CIC; 1983 - 1991 in 1980 CIC; 1992 - 2002 in 1990 CIC 2003; and thereafter in 2000 CIC. All unpublished translation tables were provided by a person from the IPUMS Project.

industry output that goes to private consumption (used to define and analyze a subsample) were obtained from detailed input output 'use tables' after redefinition. They were published by the BEA in the six benchmark years between 1977 and 2002. Direct instead of total (which include indirect) inputs were used. I had to apply bridge tables to translate input output classifications into the official classification systems.

Table 16: **Regression results**

	Baseline model		Long spells [>10 obs]	
	(1)	(2)	(3)	(4)
HHF	-.45024	-.46188	-.36541	-.36798
	(0.051)	(0.040)	(0.122)	(0.109)
HHFv	-.06176		-.10566	
	(0.873)		(0.800)	
BAR	4.1e-05	4.1e-05	3.2e-05	3.3e-05
	(0.078)	(0.079)	(0.223)	(0.214)
HHFxBAR	-.00186	-.00186	-.00179	-.00181
	(0.001)	(0.001)	(0.005)	(0.005)
UNION	-.00492	-.00496	-.00404	-.00403
	(0.000)	(0.000)	(0.000)	(0.000)
EXP	-.00449		-.01388	
	(0.946)		(0.851)	
IMP	-.11215	-.11367	-.11648	-.11281
	(0.003)	(0.002)	(0.011)	(0.006)
GROWTH	-.0031	-.00308	-.00076	
	(0.018)	(0.017)	(0.547)	
CAPINT	.04765	.04794	.04801	.05162
	(0.089)	(0.089)	(0.108)	(0.094)
capint	.05502	.05511	.06057	.06172
	(0.000)	(0.000)	(0.000)	(0.000)
mktshare	.89367	.89992	.89713	.91403
	(0.000)	(0.000)	(0.000)	(0.000)
gmktshare	.00226	.00226	.00249	.00256
	(0.000)	(0.000)	(0.000)	(0.000)
sd	-.5916	-.59158	-.62746	-.62369
	(0.000)	(0.000)	(0.000)	(0.000)
cong	.16302	.16298	.12265	.12384
	(0.000)	(0.000)	(0.000)	(0.000)
forshare	.20653	.20504	.18085	.17863
	(0.000)	(0.000)	(0.000)	(0.000)
xk	.09521	.09526	.0817	.08189
	(0.000)	(0.000)	(0.000)	(0.000)
XK	-.10756	-.10737	-.09456	-.09212
	(0.000)	(0.000)	(0.000)	(0.000)
k	3.8e-05	3.8e-05	2.1e-05	2.1e-05
	(0.000)	(0.000)	(0.010)	(0.010)
N	2055	2055	1473	1473
R^2	0.302	0.302	0.369	0.369
adj. R^2	0.296	0.297	0.362	0.363

p-values in parentheses

Table 17: **Regression results for auxilliary specifications**

	ADRD [not BAR]		With rd & ad	
	(1)	(2)	(3)	(4)
HHF	-.00318	.00163	-.57361	-.60663
	(0.991)	(0.996)	(0.050)	(0.025)
HHFv	.02012		-.26423	
	(0.959)		(0.590)	
BAR			3.3e-05	3.5e-05
			(0.377)	(0.346)
ADRD	.07558	.07592		
	(0.317)	(0.317)		
MES	-2.6e-07	-2.6e-07		
	(0.008)	(0.007)		
HHFxBAR			-.00145	-.00139
			(0.147)	(0.153)
HHFxADRD	-4.2512	-4.2103		
	(0.009)	(0.011)		
UNION	-.005	-.00491	-.00347	-.00375
	(0.000)	(0.000)	(0.003)	(0.001)
EXP	.05289		-.07271	
	(0.435)		(0.419)	
IMP	-.11037	-.09657	-.09031	-.09721
	(0.005)	(0.009)	(0.071)	(0.032)
GROWTH	-.00287	-.00278	-.00074	
	(0.031)	(0.035)	(0.575)	
CAPINT	.07081	.06925	.27528	.26838
	(0.031)	(0.035)	(0.018)	(0.017)
capint	.06708	.06671	.09002	.09099
	(0.000)	(0.000)	(0.010)	(0.007)
mktshare	.84969	.83453	.88823	.945
	(0.000)	(0.000)	(0.000)	(0.000)
gmktshare	.00227	.00227	.00198	.00205
	(0.000)	(0.000)	(0.000)	(0.000)
sd	-.59051	-.58996	-.38999	-.38708
	(0.000)	(0.000)	(0.000)	(0.000)
cong	.15514	.15539	.13922	.14312
	(0.000)	(0.000)	(0.000)	(0.000)
forshare	.18558	.19186	.15891	.15473
	(0.000)	(0.000)	(0.000)	(0.000)
xk	.0947	.09428	.05275	.05322
	(0.000)	(0.000)	(0.002)	(0.002)
XK	-.11389	-.1149	-.06475	-.05854
	(0.000)	(0.000)	(0.003)	(0.007)
k	3.3e-05	3.3e-05	3.1e-05	2.9e-05
	(0.000)	(0.000)	(0.023)	(0.026)
ad			-.54103	-.50989
			(0.009)	(0.012)
rd			-.09558	-.09477
			(0.000)	(0.000)
N	1952	1952	1011	1011
R^2	0.303	0.303	0.347	0.346
adj. R^2	0.296	0.296	0.334	0.335

p-values in parentheses

Table 18: **Regression results for comestic and consumer industries**

	Domestic		Consumption	
	(1)	(2)	(3)	(4)
HHF	-.69883	-.6669	-.05719	
	(0.039)	(0.071)	(0.874)	
HHFv	.12359		.32836	
	(0.945)		(0.878)	
BAR	-1.3e-05	-8.0e-06	1.9e-05	
	(0.722)	(0.815)	(0.609)	
HHFxBAR	.001	.0008	-.00091	
	(0.381)	(0.472)	(0.348)	
UNION	-.00634	-.00558	-.00478	-.00514
	(0.000)	(0.000)	(0.010)	(0.007)
EXP			-.28157	
			(0.536)	
IMP			-.11991	-.14397
			(0.198)	(0.060)
GROWTH	-.00204		-.00858	-.00977
	(0.303)		(0.000)	(0.000)
CAPINT	.0289		.126	
	(0.373)		(0.281)	
capint	.0404	.04882	.00748	
	(0.011)	(0.001)	(0.808)	
mktshare	.8419	.84167	.92923	.91254
	(0.003)	(0.002)	(0.021)	(0.007)
gmktshare	.00189	.00212	.00214	.00214
	(0.000)	(0.000)	(0.000)	(0.000)
sd	-.34244	-.3319	-.65675	-.65604
	(0.004)	(0.006)	(0.000)	(0.000)
cong	.11356	.11865	.1438	.15985
	(0.000)	(0.000)	(0.001)	(0.000)
forshare	.1639	.157	.3131	.30897
	(0.039)	(0.049)	(0.001)	(0.001)
xk	.04698	.04899	.11409	.11255
	(0.014)	(0.011)	(0.000)	(0.000)
XK	-.08772	-.09459	-.12089	-.12229
	(0.001)	(0.001)	(0.001)	(0.000)
k	2.0e-05	2.1e-05	4.9e-05	5.4e-05
	(0.058)	(0.045)	(0.049)	(0.015)
N	504	504	435	435
R^2	0.242	0.239	0.437	0.433
adj. R^2	0.217	0.219	0.413	0.418

p-values in parentheses

Table 19: **Regression results for critical concentration levels**
(restricted to only moderately concentrated industries)

	CCL 1		CCL 2	
	(1)	(2)	(3)	(4)
HHF	-.22543		.18117	
	(0.736)		(0.867)	
HHFv	.14564		.46532	
	(0.848)		(0.660)	
BAR	4.5e-05		.00026	
	(0.726)		(0.231)	
HHFxBAR	-.0014		-.0035	
	(0.381)		(0.164)	
UNION	-.00128		-.00343	
	(0.462)		(0.173)	
EXP	.19327		.15166	
	(0.107)		(0.456)	
IMP	-.10357		-.12939	
	(0.247)		(0.368)	
GROWTH	-.00593	-.00576	-.00477	-.00527
	(0.003)	(0.004)	(0.147)	(0.061)
CAPINT	.12574	.22527	.09518	
	(0.111)	(0.004)	(0.448)	
capint	.09764	.09307	.19492	.21835
	(0.004)	(0.006)	(0.052)	(0.011)
mktshare	.37257		.27493	
	(0.226)		(0.592)	
gmktshare	.00277	.00272	.00298	.00282
	(0.000)	(0.000)	(0.000)	(0.000)
sd	-.82054	-.82013	-1.0349	-.99011
	(0.000)	(0.000)	(0.000)	(0.000)
cong	.29376	.30528	.31247	.32956
	(0.000)	(0.000)	(0.000)	(0.000)
forshare	.31466	.3852	.33086	.36302
	(0.000)	(0.000)	(0.003)	(0.001)
xk	.1526	.11858	.16844	.18264
	(0.000)	(0.000)	(0.008)	(0.002)
XK	-.1005		-.11753	-.13466
	(0.076)		(0.136)	(0.059)
k	4.8e-05	5.1e-05	2.8e-05	6.1e-05
	(0.021)	(0.002)	(0.409)	(0.027)
N	456	456	236	236
R^2	0.463	0.447	0.475	0.461
adj. R^2	0.441	0.436	0.431	0.439

p-values in parentheses

Table 20: **Regression results for critical concentration levels**
(CCL dummy variable "CCL")

	CCL 1		CCL 2	
	(1)	(2)	(3)	(4)
CCL	.00047		.00118	
	(0.980)		(0.959)	
HHF	-.45585	-.46188	-.46027	-.46188
	(0.159)	(0.040)	(0.116)	(0.040)
HHFv	-.06144		-.06168	
	(0.874)		(0.873)	
BAR	4.1e-05	4.1e-05	4.1e-05	4.1e-05
	(0.077)	(0.079)	(0.077)	(0.079)
HHFxBAR	-.00187	-.00186	-.00186	-.00186
	(0.001)	(0.001)	(0.001)	(0.001)
UNION	-.00492	-.00496	-.00492	-.00496
	(0.000)	(0.000)	(0.000)	(0.000)
EXP	-.0047		-.00429	
	(0.945)		(0.949)	
IMP	-.11197	-.11367	-.1119	-.11367
	(0.003)	(0.002)	(0.003)	(0.002)
GROWTH	-.0031	-.00308	-.00309	-.00308
	(0.018)	(0.017)	(0.018)	(0.017)
CAPINT	.04764	.04794	.04769	.04794
	(0.089)	(0.089)	(0.089)	(0.089)
capint	.05502	.05511	.05507	.05511
	(0.000)	(0.000)	(0.000)	(0.000)
mktshare	.89345	.89992	.89348	.89992
	(0.000)	(0.000)	(0.000)	(0.000)
gmktshare	.00226	.00226	.00226	.00226
	(0.000)	(0.000)	(0.000)	(0.000)
sd	-.59155	-.59158	-.59163	-.59158
	(0.000)	(0.000)	(0.000)	(0.000)
cong	.16303	.16298	.163	.16298
	(0.000)	(0.000)	(0.000)	(0.000)
forshare	.20655	.20504	.20652	.20504
	(0.000)	(0.000)	(0.000)	(0.000)
xk	.0952	.09526	.09522	.09526
	(0.000)	(0.000)	(0.000)	(0.000)
XK	-.10756	-.10737	-.10761	-.10737
	(0.000)	(0.000)	(0.000)	(0.000)
k	3.8e-05	3.8e-05	3.8e-05	3.8e-05
	(0.000)	(0.000)	(0.000)	(0.000)
N	2055	2055	2055	2055
R^2	0.302	0.302	0.302	0.302
adj. R^2	0.296	0.297	0.296	0.297

p-values in parentheses

Table 21: **Regression results for critical concentration levels**
(CCL dummy-HHF interaction variable "HHFxCCL")

	CCL 1		CCL 2	
	(1)	(2)	(3)	(4)
HHFxCCL	.11413		.06791	
	(0.665)		(0.785)	
HHF	-.57911	-.46188	-.51023	-.46188
	(0.122)	(0.040)	(0.099)	(0.040)
HHFv	-.05415		-.06044	
	(0.889)		(0.876)	
BAR	4.3e-05	4.1e-05	4.1e-05	4.1e-05
	(0.067)	(0.079)	(0.076)	(0.079)
HHFxBAR	-.0019	-.00186	-.00187	-.00186
	(0.001)	(0.001)	(0.001)	(0.001)
UNION	-.00491	-.00496	-.00492	-.00496
	(0.000)	(0.000)	(0.000)	(0.000)
EXP	-.00647		-.00271	
	(0.923)		(0.968)	
IMP	-.10865	-.11367	-.11087	-.11367
	(0.004)	(0.002)	(0.003)	(0.002)
GROWTH	-.00307	-.00308	-.00308	-.00308
	(0.019)	(0.017)	(0.019)	(0.017)
CAPINT	.04742	.04794	.04782	.04794
	(0.092)	(0.089)	(0.089)	(0.089)
capint	.05518	.05511	.05527	.05511
	(0.000)	(0.000)	(0.000)	(0.000)
mktshare	.88942	.89992	.89265	.89992
	(0.000)	(0.000)	(0.000)	(0.000)
gmktshare	.00226	.00226	.00226	.00226
	(0.000)	(0.000)	(0.000)	(0.000)
sd	-.59107	-.59158	-.5919	-.59158
	(0.000)	(0.000)	(0.000)	(0.000)
cong	.16301	.16298	.16285	.16298
	(0.000)	(0.000)	(0.000)	(0.000)
forshare	.20667	.20504	.20643	.20504
	(0.000)	(0.000)	(0.000)	(0.000)
xk	.09516	.09526	.09526	.09526
	(0.000)	(0.000)	(0.000)	(0.000)
XK	-.10792	-.10737	-.10792	-.10737
	(0.000)	(0.000)	(0.000)	(0.000)
k	3.8e-05	3.8e-05	3.8e-05	3.8e-05
	(0.000)	(0.000)	(0.000)	(0.000)
N	2055	2055	2055	2055
R^2	0.303	0.302	0.303	0.302
adj. R^2	0.296	0.297	0.296	0.297

p-values in parentheses

Table 22: **Regression results for a logarithmic specification**

	(1)		(2)	
ln_HHF	8.6e-05	(0.918)		
ln_HHFv	-2.2e-05	(0.938)		
ln_BAR	.00033	(0.693)		
ln_HHFxBAR	-.00045	(0.593)		
ln_UNION	-.00022	(0.222)		
EXP	.00057	(0.611)		
IMP	-.00062	(0.435)		
ln_GROWTH	-.00046	(0.272)		
ln_CAPINT	.00093	(0.184)	.00105	(0.002)
capint	.00108	(0.028)		
ln_mktshare	.00014	(0.180)		
ln_gmktshare	.00279	(0.000)	.00292	(0.000)
ln_xk	.0018	(0.000)		
ln_XK	-.00072	(0.308)		
ln_forshare	3.0e-05	(0.680)		
ln_sd	-.00055	(0.075)		
cong	-.00043	(0.838)		
ln_k	.00068	(0.000)	.00076	(0.000)
ln_capint			-.00144	(0.000)
N	907		897	
R^2	0.215		0.298	
adj. R^2	0.199		0.295	

p-values in parentheses

Table 23: **Regression results for manufacturing**

	HHF		HHI	
	(1)	(2)	(3)	(4)
HHF	-.6644	-.64791		
	(0.012)	(0.014)		
HHI			-.30677	-.29369
			(0.040)	(0.048)
HHFxBAR	-.00137	-.00142		
	(0.020)	(0.015)		
HHIxBAR			-.00108	-.00111
			(0.003)	(0.002)
BAR	1.4e-06	-2.1e-07	-4.3e-06	-6.5e-06
	(0.956)	(0.993)	(0.826)	(0.743)
UNION	-.00171	-.00175	-.00197	-.00195
	(0.093)	(0.088)	(0.051)	(0.055)
EXP	.08523		.08949	
	(0.218)		(0.196)	
IMP	-.00199		-.01113	
	(0.973)		(0.849)	
GROWTH	-.00523	-.00511	-.00529	-.00511
	(0.001)	(0.002)	(0.001)	(0.002)
CAPINT	.29051	.2904	.26546	.26615
	(0.001)	(0.001)	(0.002)	(0.002)
mktshare	.43854	.41627	.37951	.35451
	(0.035)	(0.045)	(0.063)	(0.081)
gmktshare	.00254	.00254	.00255	.00255
	(0.000)	(0.000)	(0.000)	(0.000)
sd	-.6989	-.69848	-.69967	-.69972
	(0.000)	(0.000)	(0.000)	(0.000)
cong	.19344	.19271	.19424	.19404
	(0.000)	(0.000)	(0.000)	(0.000)
forshare	.21253	.22303	.21478	.22526
	(0.000)	(0.000)	(0.000)	(0.000)
xk	.1345	.13417	.13438	.13392
	(0.000)	(0.000)	(0.000)	(0.000)
XK	-.1087	-.11407	-.1108	-.11551
	(0.000)	(0.000)	(0.000)	(0.000)
k	6.0e-05	6.0e-05	6.0e-05	6.0e-05
	(0.000)	(0.000)	(0.000)	(0.000)
capint	.06759	.06721	.06883	.06802
	(0.016)	(0.015)	(0.014)	(0.014)
N	1411	1411	1411	1411
R^2	0.344	0.344	0.344	0.343
adj. R^2	0.336	0.337	0.336	0.336

p-values in parentheses

Table 24: **Regression results on average profitability differentials**

	av. RoI diff.		$\hat{\alpha}_i$	
	(1)	(2)	(3)	(4)
HHF	-.16757	-.25701	-.28393	-.33296
	(0.304)	(0.095)	(0.036)	(0.017)
HHFv	-.0788		-.0826	
	(0.781)		(0.749)	
BAR	8.6e-06		3.2e-05	3.3e-05
	(0.697)		(0.056)	(0.067)
HHFxBAR	-.00104	-.00085	-.00127	-.0013
	(0.035)	(0.000)	(0.001)	(0.001)
UNION	-.00358	-.00409	-.00316	-.00358
	(0.000)	(0.000)	(0.000)	(0.000)
EXP	-.03516		-.00122	
	(0.532)		(0.977)	
IMP	-.11659	-.13998	-.09894	-.11616
	(0.000)	(0.000)	(0.000)	(0.000)
GROWTH	-.00358	-.00638	-.00278	-.00525
	(0.000)	(0.000)	(0.000)	(0.000)
CAPINT	.04852	.04846	.05228	.05131
	(0.025)	(0.022)	(0.016)	(0.015)
capint	.04274	.0239	.03366	.01784
	(0.000)	(0.029)	(0.000)	(0.053)
mktshare	.6552	.6455	.57258	.55929
	(0.000)	(0.000)	(0.000)	(0.000)
gmktshare	.00152		.00139	
	(0.000)		(0.000)	
sd	-.53649	-.58961	-.29053	-.33144
	(0.000)	(0.000)	(0.000)	(0.000)
cong	.14391	.13975	.13888	.13439
	(0.000)	(0.000)	(0.000)	(0.000)
forshare	.21209	.18104	.15346	.12991
	(0.000)	(0.000)	(0.000)	(0.000)
xk	.07615	.07457	.06281	.0617
	(0.000)	(0.000)	(0.000)	(0.000)
XK	-.06425	-.0654	-.06806	-.06932
	(0.000)	(0.000)	(0.000)	(0.000)
k	4.4e-05	5.6e-05	3.4e-05	4.4e-05
	(0.000)	(0.000)	(0.000)	(0.000)
N	1997	1997	2055	2055
R^2	0.418	0.385	0.348	0.305
adj. R^2	0.412	0.381	0.342	0.300

p-values in parentheses

Table 25: **Regression results on profitability differential persistence**

	all		pos. diff.		neg. diff.	
	(1)	(2)	(3)	(4)	(5)	(6)
HHF	-.15538		-.27862		.38839	
	(0.443)		(0.188)		(0.490)	
HHFv	-.75091	-.76307	-.53271	-.62628	-2.1553	-1.9434
	(0.037)	(0.019)	(0.167)	(0.076)	(0.028)	(0.030)
BAR	-1.8e-05		-3.5e-05		8.7e-05	
	(0.406)		(0.144)		(0.122)	
HHFxBAR	.00054		.00081		-.00104	
	(0.265)		(0.194)		(0.272)	
UNION	.00493	.00502	.00555	.00557	.00233	
	(0.000)	(0.000)	(0.000)	(0.000)	(0.251)	
EXP	.07063		.02304		.38106	.31342
	(0.326)		(0.762)		(0.058)	(0.077)
IMP	.10428	.12001	.13117	.13905	-.05753	
	(0.007)	(0.001)	(0.002)	(0.000)	(0.548)	
GROWTH	.0056	.00569	.00694	.00656	.00168	
	(0.000)	(0.000)	(0.000)	(0.000)	(0.342)	
CAPINT	-.08769	-.09554	-.09174	-.11248	-.03907	
	(0.025)	(0.006)	(0.027)	(0.003)	(0.808)	
capint	-.00418		.00156		-.06361	
	(0.746)		(0.908)		(0.155)	
mktshare	.36004		.28153		1.786	2.5886
	(0.090)		(0.200)		(0.016)	(0.000)
gmktshare	5.2e-05		.00035		-5.6e-05	
	(0.780)		(0.145)		(0.847)	
sd	.01475		-.00743		-.00572	
	(0.683)		(0.884)		(0.919)	
cong	-.12511	-.12653	-.12732	-.12971	-.02754	
	(0.000)	(0.000)	(0.000)	(0.000)	(0.646)	
forshare	-.12512	-.11538	-.11775	-.12464	-.10051	
	(0.000)	(0.001)	(0.002)	(0.001)	(0.340)	
xk	.00335		-.0005		.0351	.04138
	(0.752)		(0.966)		(0.177)	(0.018)
XK	.02231	.02558	.02112	.02377	.01326	
	(0.165)	(0.029)	(0.222)	(0.063)	(0.757)	
k	-2.5e-05	-1.9e-05	-2.3e-05		.00018	
	(0.045)	(0.082)	(0.065)		(0.152)	
N	2055	2055	1664	1664	391	391
R^2	0.098	0.096	0.114	0.109	0.085	0.060
adj. R^2	0.090	0.092	0.104	0.105	0.041	0.050

p-values in parentheses