Modelling industry dynamics in agriculture – an equilibrium approach

Dissertation

zur Erlangung des akademischen Grades

Doctor of Philosophy (Ph.D.)

eingereicht an der Lebenswissenschaftlichen Fakultät
der Humboldt-Universität zu Berlin von

Stefan Kersting

Präsidentin der Humboldt-Universität zu Berlin:
Prof. Dr.-Ing. Dr. Sabine Kunst

Dekan der Lebenswissenschaftlichen Fakultät:
Prof. Dr. Bernhard Grimm

Gutachter/innen:

1. Prof. Dr. Martin Odening
2. Prof. Dr. Silke Hüttel
3. Prof. Dr. Thomas Heckelei

Tag der mündlichen Prüfung: 11.11.2016

Zusammenfassung

Die Struktur der Landwirtschaft hat sich in vielen westlichen Ländern grundlegend geändert. In Westdeutschland, zum Beispiel, hat die Anzahl der Betriebe deutlich abgenommen, während die durchschnittliche Betriebsgröße von $7,4$ ha im Jahr 1950 auf $42,9$ ha im Jahr 2013 stieg. Um diese Entwicklung zu verstehen, müssen insbesondere die begrenzten Produktionskapazitäten berücksichtigt werden, die hinsichtlich der verfügbaren landwirtschaftlichen Nutzfläche oder der ehemaligen Milchquote existieren. Solche Kapazitätsbeschränkungen erzeugen einen direkten Zusammenhang zwischen dem Wachsen und Schrumpfen von Betrieben, da die Expansionsmöglichkeiten eines einzelnen Betriebes von frei werdenden Kapazitäten anderer Betriebe abhängen. Aus diesem Grund wird hier zunächst ein dynamisch stochastisches Gleichgewichtsmodell entwickelt, das auf mikroökonomischen Entscheidungen basiert und Markteintritt/-austritt für den Fall ausgeschöpfter Produktionskapazitäten darstellt. Der Kapazitätspreis ist dabei markträumend in dem Sinne, dass die Anzahl neuer Produktionseinheiten mit den aufgegebenen übereinstimmen muss. Die Industrie konvergiert unter diesen Bedingungen gegen ein stationäres Gleichgewicht, in dem Ein- und Austritt vorkommen, die Größenverteilung jedoch unverändert bleibt. Die Anpassung an diesen Zustand wird u.a. beschleunigt, wenn die fixen Eintrittskosten sinken. Eine Version des Modells wird dann auf den westdeutschen Milchsektor angewendet, um den Einfluss einer Quote auf den Strukturwandel zu untersuchen. Verglichen mit dem Szenario freien Wettbewerbs hemmt eine nicht-handelbare Quote die Anpassungsprozesse, während eine handelbare Quote Marktaustritte fördern und zu einer höheren durchschnittlichen Produktivität führen kann. Die Phase nach einer Quotenabschaffung ist geprägt von enormen Kapazitätsausweitungen und stark fallenden Preisen, falls die fixen Eintrittskosten niedrig sind. Diese Entwicklung ist weniger stark ausgebildet für höhere Eintrittskosten.

Schlagwörter: begrenzte Produktionskapazität; Markteintritt/-austritt; dynamisch stochastisches Gleichgewicht; Strukturwandel

Abstract

The agricultural industry in Western countries has undergone a substantial structural change. In West Germany, for instance, the number of farms declined notably during the last decades while the average farm size increased from 7.4 ha in 1950 to 42.9 ha in 2013. A factor that needs to be considered when explaining this development is the limited sectoral production capacity, which exists in terms of agricultural land or the former milk quota regime. This limited capacity generates a direct interrelation between farm growth and farm shrinkage as a single farm's investment option depends crucially on the possible release of production capacity by competitors. The contribution of this thesis is twofold: First, a dynamic stochastic equilibrium model is developed that accounts for microeconomic decision-making and represents an industry operating at an upper capacity limit. The capacity price is determined endogenously such that it offsets the mass of entering and exiting firms in an equilibrium. It is proven that the industry tends to a stationary equilibrium in the long-run, in which entry and exit still occur but the firm size distribution remains constant. Moreover, the adjustment speed to this steady state increases if either the discount factor or the fixed entry costs decrease. Second, the impact of a production quota on industry dynamics is assessed with regard to the former milk quota regime. After calibrating the model to the West German dairy sector, the quota constrained industry performance is compared to a quota free setup. While a non-tradable quota slows down the adjustment processes within the industry, a tradable production quota can stimulate firm turnover and lead to a higher average productivity level. If the entry costs are rather low, the transition phase after a quota removal is characterised by an enormous expansion of production capacity and a considerable output price drop. This development is less pronounced though for higher entry costs.

Keywords: limited production capacity; firm entry and exit; dynamic stochastic equilibrium; industry dynamics

Contents

List of figures v

List of tables vi

1 Introduction 1

2 Review of relevant literature 5

3 Industry dynamics under limited capacity supply 11

3.1 The model 12

3.2 Equilibrium definitions 18

3.3 Existence of a stationary equilibrium 20

3.4 Convergence to the steady state 27

3.5 Existence of a converging dynamic equilibrium 36

3.6 Altering entry distributions 42

3.7 Conclusive remarks 53

4 Impact of the milk quota 55

4.1 The formal model 57

4.2 Equilibrium definitions 63

4.3 Existence of an applied stationary equilibrium 65

4.4 Existence of a finite dynamic equilibrium 74

4.5 Dynamics in the Western German dairy industry 80

4.5.1 Stylised facts: agricultural policy and development of the dairy farm size distribution 80

4.5.2 Model calibration 82

4.5.3 Equilibrium calculation 88

4.5.4 Findings 92

4.6 Final remarks 98

5 Conclusions **100**

A Appendix **104**

A.1 General mathematical concepts and theorems 104

A.2 Lemmas and proofs used in Chapter 3 107

A.3 Lemmas and proofs used in Chapter 4 111

References **113**

List of Figures

1.1 Agricultural change in West Germany 2

3.1 Stochastic productivity distributions 14

3.2 Illustration of the exit condition 33

3.3 Impact of the discount factor and entry costs on the exit condition . . 34

3.4 Converging sequence of exit-points in a dynamic equilibrium 39

3.5 Converging sequence of exit-points and capacity-values under changing entry distributions 52

4.1 Dairy farm size distribution in Western Germany 1960-2010 81

4.2 Kernel density estimate for the farm size distribution in 2000 and 2008 82

4.3 Histogram for the number of dairy cows and milk output in 2003 . . . 83

4.4 QQ-Plot for the milk output in 2003 85

4.5 Firm size distribution in a stationary equilibrium 94

A.1 Farm gate price for milk in Germany 112

List of Tables

4.1 Utilised functional forms and estimated parameters 89

4.2 Stationary equilibrium outcome 93

4.3 Dynamic equilibrium outcome for entry costs $c_e = 0$ € 95

4.4 Dynamic equilibrium outcome for entry costs $c_e = 10000$ € 96

4.5 Dynamic equilibrium outcome for entry costs $c_e = 20000$ € 97

1. Introduction

Changes in the composition of an industry, such as number and size of firms, is a fundamental phenomenon in market economies. Such changes stem from firms' adjustment to an altering economic environment and are induced, for instance, by price changes, policy changes, or technological progress (Dunne et al., 2013). While the agricultural literature usually calls these adjustment processes within an industry 'structural change', the economic literature refers to this phenomenon as 'industry dynamics'. In this thesis, both expressions will be used synonymously describing changes in a sector's composition that are provoked by market entries and exits, growth and shrinkage of firms,[1] changes in the production structure, or the adoption of new key technologies (Caves, 1998). Understanding such adjustment processes is of great interest because they determine a sector's competitiveness (Jorgenson and Timmer, 2011). Moreover, the industry's evolution has consequences for distributional issues, regional development, rural employment, and other policies (Piet et al., 2012).

In the last decades, the agricultural production in Western European countries has fundamentally changed as basic food production has been replaced by complex (bio-) technological production systems. As a consequence thereof, the number of farms has significantly declined. In Western Germany, for example, the number of farms decreased from nearly 1.8 million in 1950 to just $260,100$ in 2013 (see left side of Figure 1.1). On the other hand, the average farm size increased from 7.4 ha to 42.9 ha in the same time period (see right side of Figure 1.1). These two figures alone illustrate that a considerable structural change has taken place in the agricultural industry.

A peculiarity of agricultural sectors, which needs specific attention, is that important production factors are short in supply. This means their availability is restricted. Prominent examples of such production factors include agricultural land

[1]Farms and firms will be used interchangeably throughout the thesis.

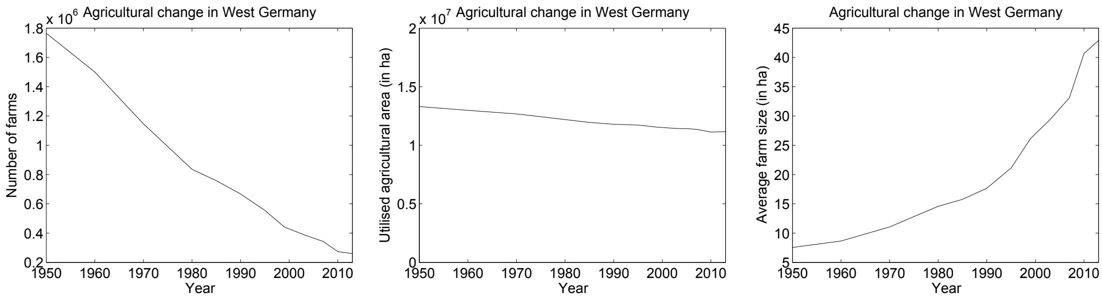

Figure 1.1: Agricultural change in West Germany reflected by the number of farms (left), total utilised land area (middle), and the average farm size (right).
Source: Bundesministerium für Ernährung und Landwirtschaft (BMELV)

and the milk quota, which affected the European dairy production from 1984 until 2015. Both factors represent some kind of production capacity that is limited to the sector level. While the available arable land constitutes a natural capacity constraint, the milk quota served as a policy instrument that was designed to avoid excess production and generate price support for milk producers.

The aforementioned shortage of production factors causes a strong interdependence of farms' decisions within a region (e.g. Chavas, 2001). That is, farms usually cannot grow in size unless other farms exit since only the capacity of ceasing firms provides new factor supply like land or other limited inputs (cf. Balmann et al., 2006). Hence, the price for such scarce production capacities strongly depends on the exit/shrinking rate determining the amount of free capacity, as well as on the expansion activities of other firms (e.g. Weiss, 1999; Zepeda, 1995; Richards and Jeffrey, 1997). The impact of this scarcity on firms' decision-making and the resulting competition for such a production factor may further increase if some firms expect to benefit from economies of size. Given that limited production capacity as such represents a valuable asset, a firm's liquidation value may increase under binding capacity constraints.

The complex relationships between farms make a coherent analysis of structural change difficult. In fact, decisions have to be made in a dynamic, stochastic environment as farms face both aggregate and firm-specific uncertainty. While price fluctuations are a common phenomenon in agricultural markets that affect all competitors,[2] a solitary farm's production process is also subject to idiosyncratic risk. Firm-specific productivity shocks can stem, for instance, from bad weather conditions or a suboptimal use of input factors. Due to the complete usage of production

[2]The development of the farm gate milk price in Germany, which is displayed in Figure A.1, serves as an example of such price fluctuations.

capacity, the dynamic evolution of industry structure plays an important role as well. That means, farms must anticipate the potential release of production capacity by competitors when they evaluate their own investment/disinvestment options. These might be some reasons why other studies dealing with structural change in agriculture analyse either farm growth or farm exit, but only little attention is spent on the interrelation between both. The illustrations above indicate, however, that it is necessary to consider farm growth and farm exit simultaneously.

The objective of this thesis is to investigate how farms' entry/exit decisions are mutually affected by the limited sectoral production capacity, and to find out what this means for the evolution of industry structure. Two major research questions are tackled in this context: First, how does a capacity constrained industry like the agricultural sector evolve in the long-run, and second, how does such an industry perform in comparison to an industry with unlimited capacity supply? A response to the latter question is important from a policy perspective as it targets the impact that possible production quota, like the European milk quota, has on the industry dynamics. It is frequently hypothesised that the introduction of a production quota slows down structural change and hinders efficient adjustment processes (Colman, 2000). But is this also true if quotas can be traded and thus increase the liquidation value of exiting firms (Barichello, 1995)?

With regard to the first point, I develop a dynamic stochastic equilibrium model that accounts for microeconomic decision-making and represents an industry operating at an upper capacity limit. The theoretic framework is based upon the seminal work by Jovanovic (1982) and Hopenhayn (1992a), who first analysed simultaneous entry and exit of firms in a dynamic, stochastic model. In contrast to their setup however, in which capacity constraints do not play a role and new firms have hypothetically free access to the industry, I assume that the total production capacity is distributed among all active firms, and potential entrants are compelled to acquire production capacity from exiting firms. The capacity price is determined endogenously such that it balances the number of demanded units with the released ones. Due to the dynamic nature of the model, the capacity value depends particularly on the firms' expected profits and, thus, on the prospective composition of the industry. Here, changes in the industry structure are not only induced by firms' entry/exit decisions, but also driven by a stochastic component. Potential differences, as for instance in firm size or production efficiency, are captured by a firm-specific productivity parameter that follows a Markov process. I will show that this framework possesses a dynamic equilibrium, in which the industry tends to a uniquely determined steady state in the long-run. This steady state is characterised by the fact that essential parameters, such as the firm size distribution, remain constant over time. Moreover, the industry's adjustment speed to this steady state is particularly defined by the size of the discount factor and the fixed entry costs.

The European Union has recently abolished a milk quota regime that was effective for the last 30 years. This offers a perfect opportunity to answer the second research question and analyse the impact that the introduction/abolition of a production quota has on the industry dynamics. To this end, I apply an advanced version of the introduced model to the Western German dairy sector, and compare the equilibrium outcome under different quota schemes (tradable, non-tradable) to a scenario without production quota. While the non-tradable quota slows down the adjustment processes within an industry, a tradable production quota can even stimulate firm turnover and lead to a higher average productivity level. In light of the recent milk quota abolition, I further visualise the transitional phase from a quota constrained to a quota free industry. The period right after the quota removal is characterised by an enormous expansion of production capacity and a considerable output price drop if the sunk entry costs are low. But, many of those newly created production units withdraw from the industry shortly thereafter, and the quota-free steady state is attained. I find that this development is less pronounced for higher entry costs.

The remainder of this thesis is structured as follows. After discussing the relevant literature about industry dynamics in Chapter 2, the theoretical framework is developed in Chapter 3 and the convergence to a steady state is proven. The basic model is modified and applied to the Western German dairy sector in Chapter 4. Chapter 5 concludes.

2. Review of relevant literature

The analysis of industry dynamics is a well-established topic in agricultural economics and much effort has been spent on modelling farm adjustment processes, both theoretically and empirically (Balmann et al., 2006). However, most existing models consider either single farms (eg. Foltz, 2004), adopt an aggregated view of the sector (Wolf and Sumner, 2001), or compare cross-country differences (Adamopoulos and Restuccia, 2014). Individual farm models account for dynamic stochastic adjustment behaviour but only under the assumption of exogenous price processes, and thus rational expectations equilibria on product and factor markets cannot be ensured with these models. In contrast, sectoral equilibrium models leave little room for a micro-economic foundation of decision-making (Féménia and Gohin, 2011; Verikios and Zhang, 2013). Only few modelling approaches take the interdependencies of growing and shrinking farms into account when describing the adjustment processes within a sector; these include multi-agent models, real options models and game theoretic models.

Multi-agent models (MAM) consist of sets of rules defining how agents make individual decisions, and how they interact with each other and their environment. Agent-based models have been quite successful in explaining stylised empirical facts such as the path dependency of systems, which is not well accounted for by existing representative-agent equilibrium models (Balmann et al., 2013). An advantage of this modelling framework is its flexibility. The agents' heterogeneity with regard to their production capacities and constraints can easily be implemented. The usefulness of MAM, however, is not unchallenged; they have been criticised because the outcome of complex dynamic simulations is difficult to interpret and generalise (Leombruni and Richiardi, 2005). Moreover, it is practically impossible to implement the concept of a rational expectations equilibrium.

Real option models have been developed to derive optimal investment and disinvestment strategies for firms facing uncertainty and sunk costs while having some managerial flexibility with regard to the timing of (dis)investments. Sunk costs usually reflect an investment's (partial) irreversibility, which is a common feature of many investments in reality, and commits a firm to its decision once it has been made. Therefore, the possibility of postponing investment and waiting for other market conditions has a positive value (option value) that must be considered when analysing optimal investment/disinvestment strategies under uncertainty. In this regard, Dixit (1989) shows that the optimal investment (disinvestment) trigger price exceeds (falls below) its counterpart that is determined with the classical net present value. But, these results are derived from a single firm's point of view, and they are based on the assumption of an exogenous price process (geometric Brownian motion). The impact from joint entries and exits on the aggregate output, and thus on the price process, are not explicitly taken into account in this basic model.

Leahy (1993) proves, however, that the optimal (dis)investment strategies derived for individual firms are also valid in a competitive environment with free market entry and homogeneous firms. As all firms are considered identical in this framework, they must have the same strategies in a competitive equilibrium. This implies immediately that all firms invest (disinvest) when the price reaches the upper (lower) trigger threshold. Due to this conform behaviour the increased (decreased) aggregate output will prevent the price from increasing (decreasing) any further. The emerging endogenous price process is thus bounded between these reflecting barriers. Although the value of (dis)investment is reduced under this bounded price process, the trigger thresholds coincide with the traditional ones that are valid for a solitary firm. A recent application of this framework to the agricultural context can be found in the study of Feil and Musshoff (2013). These authors utilise a real options model to evaluate agricultural policy schemes in a dynamic stochastic context. By applying heuristic solution procedures the authors are able to relax simplifying assumptions of previous models that strive for closed-form solutions of the dynamic equilibrium in a sector.

Many real option models have in common that they account for aggregate uncertainty (e.g. through demand shocks) and presume firms to be homogeneous. But, it is frequently observed that firms differ in their cost structure, their efficiency, or their strategic position. As this kind of heterogeneity is expected to have severe implications on industry dynamics, an incorporation of firm-specific differences seems to be necessary. Caballero and Pindyck (1996), for instance, extend the basic models of Dixit (1989) and Leahy (1993) by adding firm-specific uncertainty. They find, however, that idiosyncratic shocks, which affect only an individual firm, have less impact on the willingness to invest than aggregate shocks. Novy-Marx (2007) investigates the investment strategies of heterogeneous firms in a competitive, uncertain environment. The author can show that a heterogeneous industry structure reduces the competitive effect on the option value and the timing of investment. This means, firms have a higher option value and delay investment in the sense that their trigger price is higher. In general, the real options framework is an appropriate tool to analyse investment/disinvestment (or entry/exit) decisions in a competitive industry. However, the focus is more on the optimal timing than on the direct interdependency between investment and disinvestment, which is characteristic for the agricultural industry.

Game theoretic approaches have been applied to analyse the relationship between the dynamics of market structure and competition. While early research studies in industrial organisation used to assume a one-dimensional causality between market structure and firm behaviour (structure-conduct-performance paradigm), it is widely accepted nowadays that the relationship between both is more complex. Sutton (1991) already emphasised that the market structure is shaped by firm entry and exit, which in turn depend on the firms' expectations of future market structure and nature of competition. This interdependency is also confirmed in empirical studies by Syverson (2004) and Dunne et al. (2013), who investigate determinants defining the market structure in the U.S. concrete industry, and in the sector of American dentists and chiropractors respectively.

The seminal work by Ericson and Pakes (1995) defines a dynamic stochastic game to describe the development of an oligopolistic market structure with heterogeneous firms. Such models are particularly useful for explaining the emergence of asymmetric industry structures. Besanko and Doraszelski (2004) model a dynamic capacity accumulation game, for instance, with ex-ante identical firms and reversible investment. The authors can show that an asymmetric industry structure arises under price (Bertrand) competition while firms stay symmetric if they compete in quantities (Cournot competition). Hanazono and Yang (2009) consider a dynamic entry/exit game, in which firms learn about their relative cost positions. Their equilibrium findings confirm two frequent empirical observations: first, entry occurs gradually over time with lower cost firms entering earlier than higher cost firms, and second, exiting firms are those that entered the industry later.

Game theoretic models dealing with industry dynamics often neglect the possibility of an upper capacity limit but treat capacity as a freely available production factor. An exception to this is the paper by Esö et al. (2010). These authors model a two-stage game, in which ex-ante identical firms compete for limited production capacity in an upstream market before facing Cournot competition in the downstream market. They find that the industry remains symmetric if the available capacity is sufficiently scarce. Surprisingly, an increased capacity gives rise to an asymmetric industry structure with one large (capacity-hoarding) firm and a fringe of smaller (capacity-constrained) firms. This outcome is somewhat contrary to the result of Besanko and Doraszelski (2004), who illustrate that firms rather stay symmetric under Cournot competition and potentially unlimited capacity supply. Summing up, we can conclude that game theoretic models are capable of modelling the growth and shrinkage of firms in a given market with endogenous supply or constrained capacities. However, they are difficult to handle, particularly if firms are heterogeneous and there are more than two firms within the market.[1]

Dynamic stochastic equilibrium models can sometimes provide a framework to overcome this shortcoming. Authors like Jovanovic (1982) or Hopenhayn (1992a) model a heterogeneous industry by a continuum of firms, when they analyse simultaneous entry and exit of firms in a dynamic, stochastic environment. Idiosyncratic uncertainty, which every firm faces in its production process, thus cancels out at the aggregate level. As a consequence, the evolution of the industry follows deterministic paths, and changes in the industry structure can be pursued in greater detail.

Jovanovic (1982), for instance, sets up a selection model in which a single firm's productivity is repeatedly drawn from an individual distribution function. Although this individual distribution is unknown to each firm when entering the industry, firms can gather more and more information about their true distribution by observing a new productivity level every period (Bayesian learning process). As the firms' uncertainty about their own productivity distribution continuously resolves this way, relatively efficient firms stay in the industry while inefficient ones cease production and leave the industry. In the long-run, however, the industry tends to a state that does not contain entry and exit anymore. Next to this interesting result, Jovanovic's great achievement was to illustrate that an equilibrium in the proposed framework is equivalent to the optimum of a social welfare function when firms are atomistic and the industry is perfectly competitive. This finding is essential for many dynamic stochastic models since it provides a convenient method to prove the existence of an infinite-dimensional equilibrium.

Because stochastic models with an infinite time horizon are rather difficult to

[1] These are just few examples of studies that use a game theoretic approach for the analysis of industry dynamics. More details can be found in Doraszelski and Pakes (2007), for instance.

handle though, Hopenhayn (1992a) introduced the important concept of a stationary equilibrium. A stationary equilibrium can be considered as a particular dynamic equilibrium, in which significant parameters do not change over time. Although the overall firm size distribution does not alter in a stationary equilibrium, it may still exhibit entry and exit of individual firms. Hopenhayn (1992a) proves the existence of a stationary equilibrium for the case that the industry is perfectly competitive, and the firm's productivity follows a Markov process. In this state, firms leave the industry as soon as their productivity shock falls below a reservation rule. The released space is then occupied by new firms, who enter the industry to the same extent. Together with the productivity shocks and the respective production decisions, entry and exit thus determine the firm size and profit distribution within that industry. Furthermore, Hopenhayn's findings reveal that the size distribution is stochastically increasing with age, meaning that larger firms have a higher probability of survival.[2]

In general, a stationary equilibrium can be understood as the steady state of a dynamic system. The analysis of such a steady state offers a huge simplification to complex dynamic models as the dimension reduces from infinite sequences to single parameter values. A further advantage of the stationary equilibrium concept is that comparative statics can be carried out quite easily. The effect that a possible parameter change or an introduction of policy instruments would have on the industry can be evaluated by comparing the respective stationary equilibria.

Many authors thus apply the concept of stationary equilibria when dealing with firm turnover and industry dynamics. Asplund and Nocke (2006), for instance, extend the simultaneous entry/exit model from Hopenhayn (1992a) to imperfectly competitive markets. By comparing the resulting stationary equilibria they conclude that the turnover rate is increasing in market size. Melitz (2003) adapts the steady state idea to assess the impact of international trade on industry structure. He finds that the most productive firms enter the export market while less productive firms keep producing for the domestic market or leave the industry completely. From a more global perspective this will lead to an international re-allocation towards the more productive firms. The core model assumption that the patterns of entry and exit are systematically related to productivity differences among firms is confirmed by Fariñas and Ruano (2005). The authors show that sunk costs are one source of persistent heterogeneity in productivity. This means lower productivity becomes more likely in markets with high and sunk entry costs. Gomes (2001) and Miao (2005) transfer the stationary equilibrium concept to a slightly different framework

[2]See also Bento (2014), who takes up the idea of Hopenhayn in modelling entry costs.

as they investigate how liquidity constraints and financing costs affect the firms' investment decisions.

While a lot of research papers deal with the existence or the properties of a stationary equilibrium, only little effort is spent on the question whether this steady state is really attained in the long-run. Das and Das (1997) assert, for example, that the evolution of industry structure does not necessarily follow monotone paths in a dynamic equilibrium, and the adjustment of mature industries to the steady state is rather an assumption than a trivial outcome of the industry dynamics. In fact, industries that have been undergoing a consolidation process and are still displaying changes in the firm size distribution have certainly not achieved a steady state yet. Moreover, it is rather unlikely that a steady state will persist forever in real world industries because firms must continuously react to demand shocks or volatile prices. Even without this kind of aggregate uncertainty it is rather difficult to rule out cycles or fluctuations in the industry structure. Hopenhayn and Prescott (1992) show convergence for stochastically monotone operators. But, this assumption is generally not fulfilled in a dynamic entry/exit model. Does the industry tend to a stationary equilibrium then anyhow? This question is answered in the next chapter.

3. Industry dynamics under limited capacity supply

It is well documented that also in industries with unlimited capacity supply firm entry and exit can be positively correlated to each other (cf. Cabral, 1997; Caves, 1998). This correlation is even more pronounced though, when the available production capacity is limited to the sector level, and potential entrants have to compete with incumbents for that scarce factor. Our intention is to incorporate such a sectoral capacity limit into the dynamic, stochastic, entry/exit framework proposed by Hopenhayn (1992a). For this reason, we assume that the total production capacity is distributed among all active firms, and capacity units can be traded dircetly between incumbents and potential entrants. Before producing any output new firms have to build up production capacity. As they must acquire this capacity from withdrawing firms, the potential entrants create a positive demand for capacity. This increases the incumbents' liquidation value, on the other hand, who can now sell their capacity units to entrants after ceasing production. The emerging capacity price should then offset the mass of entering firms with the mass of exiting ones in an equilibrium. We will show that the industry possesses an infinite-dimensional dynamic equilibrium in this setup, which converges to a uniquely determined steady state in the long-run. We will illustrate especially that convergence takes place irrespective of the current industry structure. The convergence rate depends on the size of the discount factor as well as the size of the fixed entry costs. We find that a declining discount factor increases the adjustment speed, and the same applies to the fixed entry costs. The fixed entry costs have to be paid by new firms in addition to the capacity acquisition costs. In contrast to the capacity costs, however, those costs are sunk after entering the industry.

Compared to the models of Jovanovic (1982) and Hopenhayn (1992a), we do not consider perfect competition in particular but presume a separable structure of the period profits instead. Although this is restrictive to some extent, such a separable structure may also emerge in imperfectly competitive markets or under product differentiation. In this case, however, an infinite-dimensional equilibrium is no longer equivalent to the maximum of a social welfare function, and a different approach is required to prove the existence of an equilibrium. As a consequence, we show the existence of an equilibrium directly by utilising Schauder's Fixed Point Theorem.[1]

The remainder of this chapter is structured as follows. After presenting the theoretical framework in Section 3.1, we give the formal definition of a dynamic equilibrium and the stationary equilibrium as its steady state in Section 3.2. Section 3.3 covers the existence and uniqueness of such a stationary equilibrium, while necessary conditions implying convergence of the firm size distribution are derived in Section 3.4. The existence of an infinite-dimensional dynamic equilibrium converging to a steady state is then proven in Section 3.5. In Section 3.6, we show that the convergence can be sustained even for altering entry distributions. Section 3.7 concludes.

3.1 The model

We set up a dynamic stochastic equilibrium model accounting for entry and exit under limited sectoral production capacity. Competition takes place in discrete time $t \in \mathbb{N}_0$, and the industry is composed of a continuum of firms that are distinguished according to their productivity level $\varphi \in [a, b]$. This stochastic parameter φ is supposed to follow a Markov process with continuous state space $S \equiv [a, b]$ and conditional cumulative distribution function $F(\varphi'|\varphi) = \text{Prob}(\varphi_{t+1} \leq \varphi'|\varphi_t = \varphi)$. New firms are assigned with a productivity value that is drawn from the initial distribution ν with cumulative distribution function (cdf) G. We impose the following assumptions:

(i) the conditional cdf $F(\varphi'|\varphi) = \text{Prob}(\varphi_{t+1} \leq \varphi'|\varphi_t = \varphi)$ is continuous with respect to φ' and φ,

(ii) F is strictly decreasing in φ (stochastic dominance),

(iii) $F(\varphi'|b) \leq G(\varphi') \leq F(\varphi'|a)$ for all $\varphi' \in S$,

[1]A formal definition of Schauder's Fixed Point Theorem can be found in the Appendix A.1.

(iv) F is continuously differentiable with respect to φ' and there exists a constant $m_F > 0$ such that

$$\frac{dF(\varphi'|\varphi)}{d\varphi'} \leq m_F, \quad \forall \varphi', \varphi \in [a,b], \tag{3.1}$$

(v) G is a continuously differentiable distribution function satisfying

$$\frac{dG(\varphi)}{d\varphi} \leq m_G, \quad \forall \varphi \in [a,b] \quad \text{and a constant } m_G > 0. \tag{3.2}$$

If the conditional cdf is characterised by a density function $p : [a,b] \times [a,b] \to \mathbb{R}_+$, then

$$F(\varphi'|\varphi) = \int_a^{\varphi'} p(\varphi, x)dx \tag{3.3}$$

and assumption (iv) would simply imply $p(\varphi, x) \leq m_F, \ \forall x, \varphi$. We denote the probability kernel that is related to F and defined on the Borel sets $\mathcal{B}(S)$ by:

$$P(\varphi, A) := \int_A \mathbf{1}(\varphi')dF(\varphi'|\varphi) \quad \text{for } A \in \mathcal{B}(S). \tag{3.4}$$

The following assumption should be met by the kernel P and the measure ν:

(CON) $\exists \varepsilon > 0$ such that for any $A \in \mathcal{B}(S)$ one of the following conditions is satisfied:
EITHER $\nu(A) \geq \varepsilon$ and $P(\varphi, A) \geq \varepsilon, \ \forall \varphi \in [a,b]$
OR $\nu(A^c) \geq \varepsilon$ and $P(\varphi, A^c) \geq \varepsilon, \ \forall \varphi \in [a,b]$.

The set A^c is supposed to be the complement to the set A. Figure 3.1 illustrates that condition (CON) basically requires all corresponding density functions (pdf) to overlap. The distribution of productivity levels across firms in period $t \in \mathbb{N}_0$ is described by the probability measure $\mu_t : \mathcal{B}(S) \to [0,1]$. That is, for all sets $A \in \mathcal{B}(S)$ the value $\mu_t(A)$ describes the share of firms having a productivity level φ in A.

A new productivity level is revealed to all active firms at the beginning of each period. Being aware of the productivity level, every firm needs to choose an optimal output quantity or output price for its product. As in Asplund and Nocke (2006) the exact type of competition at this production stage is not specified any further. Yet, we assume that a single-period equilibrium exists for any given industry structure μ, and the equilibrium profits of a firm with productivity φ are separable like

$$\pi(\varphi, \mu) = g(\varphi)h(\mu) = g(\varphi)h\left(\int_a^b q(\varphi)\, d\mu(\varphi)\right). \tag{3.5}$$

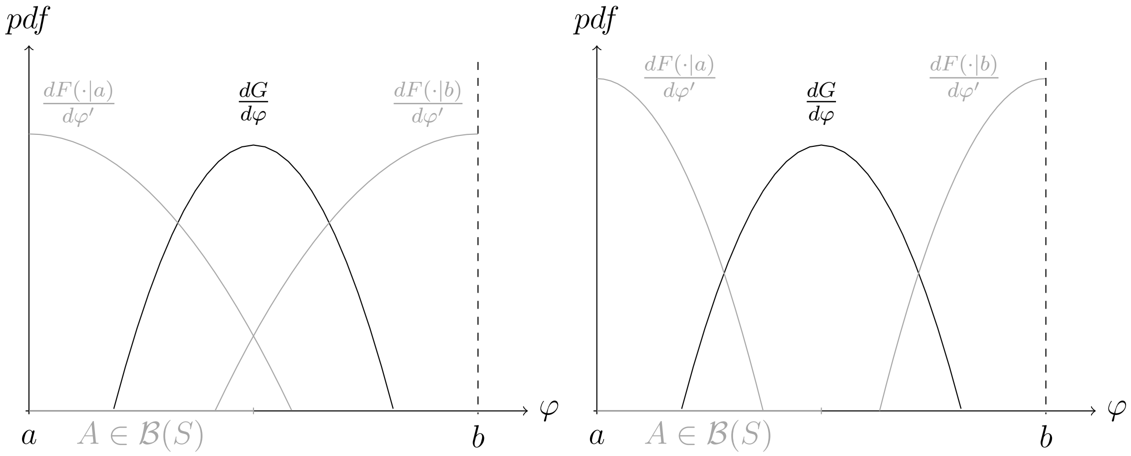

Figure 3.1: Assumption (CON) is fulfilled on the left side, but not on the right side

This structure of period profits may indeed arise under perfect competition with homogeneous goods (see the numerical example in Hopenhayn, 1992b) or in the case of monopolistic competition with differentiated products (cf. Melitz, 2003).

Both functions $g, h \in C^1(\mathbb{R})$ are supposed to be continuously differentiable and non-negative, such that the period profits $\pi(\varphi, \mu) \geq 0$ are always larger or equal to zero. Since firms with a higher productivity level are assumed to make higher profits, we further request g to be strictly increasing with respect to φ. In the single-period equilibrium a firm's profit will not just depend on its own productivity parameter but also on the distribution of productivity levels across other firms. A high share of productive firms should either intensify competition or lead to a higher aggregate output, thus having a negative effect on the single firm's profit. We account for this by introducing the strictly decreasing function h and the aggregate variable

$$Q := \int_a^b q(\varphi)\, d\mu(\varphi) \tag{3.6}$$

that reflects the total industry production output emerging under the industry structure μ in a single-period equilibrium . The function $q : S \to \mathbb{R}_+$ is assumed to be non-decreasing such that $\mu_1 \overset{FOSD}{\geq} \mu_2$ implies $Q_1 \geq Q_2$ and $\pi(\varphi, \mu_1) \leq \pi(\varphi, \mu_2)$ for a fixed productivity level φ.[2] Furthermore, the industry output is bounded between $Q_{\min} = q(a)$ and $Q_{\max} = q(b)$ for any probability measure μ on $(S, \mathcal{B}(S))$.

As mentioned in the introduction, the total production capacity is supposed to be limited to the sector level and completely distributed among established firms.

[2]The expression $\mu_1 \overset{FOSD}{\geq} \mu_2$ means that the measure μ_1 (first order) stochastically dominates μ_2, i.e. $\int f\, d\mu_1 \geq \int f\, d\mu_2$ for any non-decreasing function f.

Each firm can, thus, be considered as a marginal production unit that is equipped with a marginal capacity unit. At the end of each period incumbents have the option to cease production and leave the industry. If they decide to do so, they release their production capacity, which then becomes available to new firms. Due to the boundedness of production capacity entrants are forced to buy capacity units from exiting firms before setting up production. This creates a positive demand for capacity and induces firm turnover. The capacity is traded between both groups at a common price $y_t \geq 0$, which exactly offsets the mass of exiting firms with the mass of entrants in an equilibrium.

All firms are assumed to discount future profits with a constant factor $0 \leq \beta < 1$. If $w = \{\mu_t, y_t\}_{t\in\mathbb{N}}$ is an infinite sequence of probability measures and (bounded) capacity prices, the value of an active firm with realised productivity φ at time t can be defined recursively by

$$v_t(\varphi, w) = \pi(\varphi, \mu_t) + \beta \max\left\{ y_t, \int_a^b v_{t+1}(\varphi', w)\, dF(\varphi'|\varphi) \right\}. \tag{3.7}$$

According to (3.6) each probability measure μ_t on $(S, \mathcal{B}(S))$ is associated to an aggregate industry output Q_t. If $z = \{Q_t, y_t\}_{t\in\mathbb{N}}$ denotes the sequence of output and capacity values corresponding to the sequence w, an equivalent formulation of the value function is given by

$$v(\varphi, z_t) = \pi(\varphi, Q_t) + \beta \max\left\{ y_t, \int_a^b v(\varphi', z_{t+1})\, dF(\varphi'|\varphi) \right\}, \tag{3.8}$$

where $z_t := \{Q_j, y_j\}_{j=t}^{\infty}$ denotes the sub-sequence starting at time t. Hence, the value function consists of the current profits plus the discounted earnings that a firm may generate in prospective periods. The latter part inside the curly brackets is the continuation value designating the expected future profits of the firm if it keeps producing. In contrast to this, the price of production capacity y_t represents the exit value, i.e. the value that an incumbent firm receives when it withdraws from production by the end of the current period t.

Theorem 3.1. *(a) A unique, continuous, and bounded solution $v(\varphi, z)$ to equation* (3.8) *exists. (b) The value function is strictly increasing in φ. (c) It is monotone with respect to the sequences in z.*

Proof. **ad (a).** Let $f(\varphi, z)$ be a continuous, bounded function. We define the operator T on the space of continuous, bounded functions by

$$Tf(\varphi, z_t) := \pi(\varphi, Q_t) + \beta \, \max \left\{ y_t, \int_a^b f(\varphi', z_{t+1}) \, dF(\varphi'|\varphi) \right\}. \tag{3.9}$$

The value function v thus constitutes a fixed point of the operator T. To show the existence of such a fixed point, satisfying the properties in (a), we apply the sufficient conditions of Blackwell (1965).

First, we illustrate that the operator T maps continuous, bounded functions into other continuous, bounded functions. If $f(\varphi, z)$ is continuous in both arguments, then $\widetilde{f}(\varphi, z) := \int_a^b f(\varphi', z) \, dF(\varphi'|\varphi)$ is also continuous with respect to φ and z (see Stokey et al. (1989), Lemma 9.5 for a proof of this). The transition from $z_t \mapsto z_{t+1}$ is a continuous mapping on the space of bounded sequences that can be characterised by $z_{t+1} = \Gamma(z_t)$. As the period profits $\pi(\varphi, Q_t)$ are continuous by assumption, the function

$$Tf(\varphi, z_t) = \pi(\varphi, Q_t) + \beta \, \max \left\{ y_t, \widetilde{f}(\varphi, \Gamma(z_t)) \right\} \tag{3.10}$$

must be continuous as well. The boundedness of this function follows immediately from the boundedness of z_t and the compact support $\varphi \in S = [a, b]$.

Second, T is a monotone operator. If $f_1 \geq f_2$ for all possible combinations (φ, z), the inequality

$$\int_a^b f_1(\varphi', \Gamma(z)) \, dF(\varphi'|\varphi) \geq \int_a^b f_2(\varphi', \Gamma(z)) \, dF(\varphi'|\varphi) \tag{3.11}$$

holds and implies $Tf_1(\varphi, z) \geq Tf_2(\varphi, z)$ for all (φ, z).

Third, for any constant $c \geq 0$ we have $T(f + c)(\varphi, z) \leq Tf(\varphi, z) + \beta \, c$. According to Theorem 5 in Blackwell (1965), T is a contraction operator with modulus $\beta < 1$ on the space of bounded, continuous functions. Banach's Fixed Point Theorem (see Appendix A.1), hence, asserts the existence of a unique, bounded, and continuous solution v to equation (3.8).

ad (b). Due to the contractive behaviour of the operator T, the value function is given by $v = \lim_{n \to \infty} T^n f$, for any continuous and bounded function f. Furthermore, the operator T preserves (weak) monotonicity meaning that: $f(\varphi, z)$ non-decreasing in φ implies that $Tf(\varphi, z)$ is non-decreasing in φ. Thus, the limiting function v must also possess this property. As $\pi(\varphi, Q)$ is assumed to be strictly increasing with respect to φ, however, the value function

$$v(\varphi, z_t) = \pi(\varphi, Q_t) + \beta \; \max \left\{ y_t, \int_a^b v(\varphi', z_{t+1}) \, dF(\varphi'|\varphi) \right\} \tag{3.12}$$

must be strictly increasing as well.

ad (c). Monotonicity of v with respect to z follows by the same argument as in (b). In this case, monotonicity means that the value function increases with respect to every single y_t contained in the sequence z, and decreases with respect to every Q_t. If z, z' are two identical sequences that differ just in the element $y_k < y'_k$ (for some $k \in \mathbb{N}$), this implies $v(\varphi, z_t) \leq v(\varphi, z'_t)$. Analogously, for $Q_k < Q'_k$ and everything else kept identical we get $v(\varphi, z_t) \geq v(\varphi, z'_t)$ □

An incumbent stays in the industry as long as its continuation value exceeds the exit value. The exit point x_t characterises the critical productivity threshold for staying in the industry:

$$x_t := \inf \left\{ \varphi \in S : \int_a^b v(\varphi', z_{t+1}) \, dF(\varphi'|\varphi) \geq y_t \right\}. \tag{3.13}$$

Firms with a productivity level $\varphi_t \geq x_t$ keep producing while all firms with $\varphi_t < x_t$ drop out of the industry. If the set in (3.13) is empty, we define $x_t := b$.

At the end of each period new firms have the possibility to enter the industry. The expected profits of an entrant are given by

$$v^e(z_{t+1}) := \int_a^b v(\varphi, z_{t+1}) \, dG(\varphi). \tag{3.14}$$

New firms will enter the industry as long as their expected profits cover the entry costs. The entry costs are composed of a fixed part $c_e \geq 0$, which is sunk afterwards, and the variable part y_t representing the costs for acquiring production capacity from exiting firms. Free access to the industry implies that the condition $v^e(z_{t+1}) \leq c_e + y_t$ must be satisfied in an equilibrium, and hold with equality if positive entry occurs. We denote the mass of entering firms by $M_t \geq 0$.[3]

Recall that the variable part of the entry costs will be determined endogenously by supply and demand for production capacity. In our framework demand equals the

[3]In this context entry can be understood as the creation of a new marginal production unit. This can either be created by a really new firm, or an established firm that tries to extend its production capacity. As the entry option is supposed to be the same for all firms, we do not differentiate between those two groups but refer to both as 'entering' or 'new firms'.

mass of firms willing to enter the industry, and supply is generated by the number of firms leaving the industry. In an equilibrium the capacity price y_t must be chosen such that

$$M_t = \mu_t([a, x_t)), \tag{3.15}$$

i.e. the mass of entrants must coincide with the mass of firms leaving the industry.

The industry has been defined as a continuum of firms. Because of this uncountable number of firms we do not have to deal with uncertainty on the aggregate level. Hence, the evolution of industry structure follows deterministic paths, and is completely specified by the entry/exit behaviour of firms.[4] It can be described recursively:

$$\mu_{t+1}([a, \varphi')) = \int_{\varphi \geq x_t} F(\varphi'|\varphi)\, d\mu_t(\varphi) + M_t\, G(\varphi'). \tag{3.16}$$

3.2 Equilibrium definitions

All firms are assumed to have perfect information on the current industry structure and their competitors' entry/exit strategies. This allows them to foresee future capacity prices $\{y_t\}$ and the aggregate industry output $\{Q_t\}$. According to the definition, each firm bases its optimal entry/exit policy on these sequences. If the resulting entry/exit decisions, which are characterised by $\{M_t\}$ and $\{x_t\}$, give in turn rise to the anticipated output and capacity sequences, then the industry is in an equilibrium. We define:

Definition 3.1 (Dynamic Equilibrium)**.** Let a continuous probability distribution μ_0 on the state space S be given. A dynamic equilibrium consists of bounded sequences $\{\mu_t^*\}$, $\{Q_t^*\}$, $\{y_t^*\}$, $\{x_t^*\}$, $\{M_t^*\}$ such that for every $t \in \mathbb{N}_0$:

(i) the aggregate industry output Q_t^* is represented by (3.6),
and the period profits are given as $\pi(\varphi, Q_t^*) = g(\varphi)\, h(Q_t^*)$,
(ii) the exit point x_t^* satisfies (3.13),
(iii) $v^e(z_{t+1}) \leq c_e + y_t^*$ with equality if $M_t^* > 0$,
(iv) the capacity market clears, i.e. entry equals exit $M_t^* = \mu_t^*([a, x_t^*))$,
(v) μ_t^* follows the industry dynamics in (3.16).

[4] A deterministic development of the industry structure is justified by the law of large numbers. Evidence can be found in Judd (1985) or Feldman and Gilles (1985)

We will show in Section 3.5 that a dynamic equilibrium exists for any distribution μ_0. As the proof requires some preparation, however, we assume this being true for the moment. It is evident that different starting distributions μ_0 will provoke different adjustment processes and, thus, lead to different equilibria. This begs the question whether the industry tends to a uniquely determined steady state in either case, or whether fluctuations will still occur in the long-run. We define a stationary equilibrium as the steady state of the dynamic system laid out above.

Definition 3.2 (Stationary Equilibrium)**.** A stationary equilibrium consists of a vector $(\mu^*, Q^*, y^*, x^*, M^*)$ such that the constant sequences, constructed with these values, satisfy the conditions (i)–(v) of a dynamic equilibrium. The following conditions must hold in particular:

(i) $Q^* = \int_S q(\varphi)\, d\mu^*(\varphi)$ and $\pi(\varphi, Q^*) = g(\varphi)\, h(Q^*)$,

(ii) $\int_S v(\varphi, z^*)\, dF(\varphi|x^*) = y^*$,

(iii) $\int_S v(\varphi, z^*)\, dG(\varphi) = c_e + y^*$,

(iv) $M^* = \mu^*([a, x^*))$,

(v) $\mu^*([a, \varphi')) = \int_{\varphi \geq x^*} F(\varphi'|\varphi)\, d\mu^*(\varphi) + M^* G(\varphi')$.

The stationary equilibrium can be understood as a particular dynamic equilibrium starting from the invariant distribution $\mu_0 = \mu^*$. It may still exhibit entry and exit of firms. But, once the industry has achieved this state, the significant parameters will stay constant over time.

An equilibrium without entry and exit of firms can arise if the fixed entry costs c_e are too high to be covered by the expected profits. This implies $M_t = 0$ (for all $t \in \mathbb{N}$), and the convergence to a steady state follows by exactly the same argument as in Hopenhayn and Prescott (1992). In such a no-entry equilibrium multiple capacity price sequences $\{y_t\}_{t\in\mathbb{N}}$ may furthermore satisfy the conditions (ii),(iii). We will rule out this situation, however, later on. Since our study deals with the impact that scarce production capacity has on entry/exit of firms, we assume that new firms are really pushing into the industry and create a positive demand for capacity units.

3.3 Existence of a stationary equilibrium

The properties of the value function can be translated one-to-one to the continuation value. It is trivial that the continuation value

$$v^c(\varphi, z) := \int_a^b v(\varphi', z)\, dF(\varphi'|\varphi) \tag{3.17}$$

is continuous and (weakly) monotone with respect to φ, z. We have shown in Theorem 3.1 that the value function $v(\varphi, z)$ strictly increases with respect to the firm-specific productivity level φ. As $F(\varphi'|\varphi_1) > F(\varphi'|\varphi_2)$ for $\varphi_1 < \varphi_2$ by assumption, this implies $v^c(\varphi_1, z) < v^c(\varphi_2, z)$ for any positive, bounded sequence of aggregate output and capacity prices z (cf. Lemma A.8). Hence, the continuation value is strictly increasing in φ as well.

These features imply that the exit point, which has been defined as the critical productivity threshold for staying in the industry, is a continuous function $x_t = \gamma(z_t)$ of the output/capacity sequence. If $x_t \in (a, b)$, then it must be the unique solution to the equation

$$\int_a^b v(\varphi', z_{t+1})\, dF(\varphi'|x_t) = y_t. \tag{3.18}$$

According to the definition, the stationary equilibrium resembles a dynamic equilibrium with constant output/capacity sequences. If $\bar{z} = \{Q, y\}_{t\in\mathbb{N}}$ is such a constant sequence, the continuation value will not depend on the time t anymore and can be rephrased as:

$$v^c(\varphi, \bar{z}) = \int_a^b \pi(\varphi, Q)\, dF(\varphi'|\varphi) + \beta \left(\int_a^{\gamma(\bar{z})} y\, dF(\varphi'|\varphi) + \int_{\gamma(\bar{z})}^b v^c(\varphi', \bar{z})\, dF(\varphi'|\varphi) \right). \tag{3.19}$$

We will focus on constant sequences $\bar{z}$ in the following. The next Lemma shows how the continuation value reacts to changes in the parameters Q and y:

Lemma 3.2. *The continuation value $v^c(\varphi, \bar{z})$ is almost everywhere differentiable with respect to Q and y. The gradients are bounded by:*

(i) $\quad \dfrac{g_F\, h'_{\min}}{1-\beta} \le \dfrac{\partial v^c}{\partial Q}(\varphi, \bar{z}) \le 0$

(ii) $\quad 0 \le \dfrac{\partial v^c}{\partial y}(\varphi, \bar{z}) \le \beta$

with $g_F = \int_a^b g(\varphi)\, dF(\varphi|b)$ *and* $h'_{\min} \leq 0$ *being the minimal value of the derivative* $h'(Q)$ *on* $[Q_{\min}, Q_{\max}]$.

Proof. The differentiability follows by the monotonicity and continuity of v^c (see Kolmogorov and Fomin (2012), Chapter 9, Theorem 6, for instance). Hence, it suffices to show the inequalities in (i) and (ii).

ad (i). The second inequality is clear. To show the first inequality, let us fix a productivity level $\bar{\varphi}$. An increase in the aggregate industry output Q has the highest effect on the continuation value if the probability of exit is equal to zero and the firm stays in the industry forever (eg. for $y = 0$). In this case $\gamma(\bar{z}) = a$ and the continuation value is equal to

$$v^c(\bar{\varphi}, \bar{z}) = \sum_{t=0}^{\infty} \beta^t \int_a^b \pi(\varphi, Q)\, d\bar{\mu}_t(\varphi). \tag{3.20}$$

Here, $\bar{\mu}_t$ denotes the distribution of the firm-specific productivity that a firm with current productivity level $\varphi_0 = \bar{\varphi}$ will have at time t. Differentiating the expression in (3.20) with respect to Q yields:

$$\begin{aligned}
\frac{\partial v^c}{\partial Q}(\varphi, \bar{z}) &\geq \sum_{t=0}^{\infty} \beta^t \int_a^b \frac{\partial \pi}{\partial Q}(\varphi, Q)\, \bar{\mu}_t(\varphi) \\
&= h'(Q) \sum_{t=0}^{\infty} \beta^t \int_a^b g(\varphi)\, \bar{\mu}_t(\varphi) \\
&\geq h'_{\min} \sum_{t=0}^{\infty} \beta^t \int_a^b g(\varphi)\, dF(\varphi|b) \\
&= \frac{g_F\, h'_{\min}}{1-\beta}.
\end{aligned}$$

Note that the exchange of summation and differentiation in the first line is justified by Dini's Theorem on uniform convergence.

ad (ii). The continuation value is clearly increasing with respect to the capacity value. Therefore, the first inequality holds. A marginal change of y has the highest impact on v^c if $\gamma(\bar{z}) = b$ and the firm leaves the industry almost surely by the end of the next period (eg. for $y \geq \frac{\pi(b,Q)}{1-\beta}$). The continuation value is then given by:

$$v^c(\varphi, \bar{z}) = \int_a^b \pi(\varphi', Q)\, dF(\varphi'|\varphi) + \beta y. \tag{3.21}$$

Taking the derivative with respect to y implies that $\frac{\partial v^c}{\partial y}(\varphi, \bar{z}) \leq \beta$. $\square$

It is trivial that the same properties apply to the expected profits v^e of a new firm. To see why this is true, we just need to replace the conditional cdf $F(.|\varphi)$ by the distribution G in the previous proof.

In a stationary equilibrium the expected value v^e complies exactly with the total price that firms pay for entering the industry. The equation $v^e(Q, y) - y = c_e$, which must be satisfied according to Definition 3.2, thus determines the capacity price y^* implicitly. If the firms' incentive to enter the industry is high enough, the capacity price can be characterised as a function $y = \kappa(Q)$ of the aggregate industry output.

Lemma 3.3. *If $v^e(Q_{\max}, 0) \geq c_e$, then there exists a continuous, decreasing function $\kappa : [Q_{\min}, Q_{\max}] \to \mathbb{R}_+$ such that the equality $v^e(Q, \kappa(Q)) - \kappa(Q) = c_e$ holds for all $Q \in [Q_{\min}, Q_{\max}]$. The function is bounded by*

$$0 \leq \kappa(Q) \leq \sum_{t=0}^{\infty} \beta^t \int_a^b \pi(\varphi, Q_{\min})\, dF(\varphi|b) = \frac{g_F\, h(Q_{\min})}{1 - \beta}. \tag{3.22}$$

Proof. For convenience, we first define the function $H(Q, y) := v^e(Q, y) - y$. If we fix an aggregate output level $\bar{Q} \in [Q_{\min}, Q_{\max}]$, we must show that a unique capacity price $\bar{y}$ exists, which satisfies the equation $H(\bar{Q}, \bar{y}) = c_e$.

Existence. The function v^e is decreasing with respect to Q. Therefore, the inequality $v^e(\bar{Q}, y) - y \geq c_e$ is clearly satisfied for $y = 0$. On the other hand, every capacity value $y \geq \frac{g_F\, h(\bar{Q})}{1-\beta}$ would make the firm leave the industry immediately and, thus, lead to

$$\begin{aligned} v^e(\bar{Q}, y) - y &= \int_a^b \pi(\varphi, \bar{Q})\, dG(\varphi) + \beta\, y - y \\ &\leq g_F\, h(\bar{Q}) - (1 - \beta)\, y \\ &\leq 0 \leq c_e. \end{aligned}$$

Due to the continuity of $v^e(Q, y)$ in y there must be a value $\bar{y} \in \left[0, \frac{g_F\, h(\bar{Q})}{1-\beta}\right]$ such that $v^e(\bar{Q}, \bar{y}) - \bar{y} = c_e$ (Intermediate Value Theorem). This fact implies also the boundedness statement in (3.22).

Uniqueness. Referring to Lemma 3.2, we have $\frac{\partial v^e}{\partial y}(Q, y) \leq \beta$. Hence, the function $H(Q, y)$ must be strictly decreasing with respect to y. This implies, however, that the value $\bar{y}$ is uniquely determined. Since the value $\bar{Q}$ has been arbitrarily chosen, this defines a function $y = \kappa(Q)$, which is subject to $v^e(Q, \kappa(Q)) - \kappa(Q) = c_e$.

The continuity and monotonicity of $\kappa(Q)$ follow immediately from the continuity and (strict) monotonicity of the function $H(Q, y)$.

Continuity. Let $Q_n \to Q$ be a converging sequence of aggregate output values. We define $y_n := \kappa(Q_n)$ and $y := \lim_{n\to\infty} y_n$. By construction $H(Q_n, y_n) = c_e$ holds for all $n \in \mathbb{N}$. The continuity of H implies $H(Q, y) = \lim_{n\to\infty} H(Q_n, y_n) = c_e$. Hence, y is a solution to the equation $v^e(Q, y) - y = c_e$. Due to the uniqueness of solutions, however, we must have $\kappa(Q) = y$ and $\kappa(Q_n) \to \kappa(Q)$ for $n \to \infty$.

Monotonicity. Assume that $Q_1 < Q_2$. As v^e decreases with respect to the aggregate output, this yields $c_e = H(Q_1, \kappa(Q_1)) \geq H(Q_2, \kappa(Q_1))$. But, then we must have $\kappa(Q_1) \geq \kappa(Q_2)$. □

A crucial point in the dynamic equilibrium is that the mass of entering firms may not exceed the amount of exiting ones (see condition (iv)). Consequently, the total mass of the industry $\mu_t(S)$ must not vary over time. If the starting distribution μ_0 constitues a probability measure, then the measures $\{\mu_t\}_{t\in\mathbb{N}}$, which are derived from μ_0 via the industry dynamics (3.16), will form probability measures as well.

We define the linear operator $T_x : \mathcal{M}^1(S, \mathcal{B}) \to \mathcal{M}^1(S, \mathcal{B})$ on the space of probability measures by:

$$\begin{aligned} T_x\mu(A) &= \int_{\varphi\geq x} P(\varphi, A)d\mu(\varphi) + \int_{\varphi<x} \nu(A)d\mu(\varphi) \\ &= \int_x^b P(\varphi, A)d\mu(\varphi) + \mu([a, x))\,\nu(A) \end{aligned} \tag{3.23}$$

for $x \in [a, b]$ and $A \in \mathcal{B}(S)$. Under the assumption that $M_t = \mu_t([a, x_t))$ holds, equation (3.16) can thus be rewritten as

$$\mu_{t+1} = T_{x_t}\mu_t. \tag{3.24}$$

Notice here that T_x is not a monotone operator, i.e. it does not preserve first order stochastic dominance in the sense: $\mu_1 \overset{FOSD}{\geq} \mu_2 \quad \not\Rightarrow \quad T_x\mu_1 \overset{FOSD}{\geq} T_x\mu_2$. This is namely the reason why the convergence argument from Hopenhayn and Prescott (1992) does not apply in our case. The assumption (CON), however, which has been imposed on $F(.|\varphi)$ and G, implies the following important result.

Theorem 3.4. *The operator T_x is contractive with respect to the total variation norm $\|.\|_{TV}$. This means, there is an $\varepsilon > 0$ such that for any two probability measures $\mu_1, \mu_2 \in \mathcal{M}^1(S, \mathcal{B})$:*

$$\|T_x\mu_1 - T_x\mu_2\|_{TV} \leq (1-\varepsilon)\|\mu_1 - \mu_2\|_{TV}. \tag{3.25}$$

Proof. Let $x \in (a, b)$ be given and fixed. Assume that (CON) is fulfilled and $A \in \mathcal{B}(S)$. First, define the probability kernel

$$\hat{P}_x(\varphi, A) := \begin{cases} \nu(A) & \text{if } \varphi \in [a, x) \\ P(\varphi, A) & \text{if } \varphi \in [x, b] \end{cases} \tag{3.26}$$

This kernel satisfies the following properties:

(i) $\hat{P}_x(\varphi, A) \geq \varepsilon,\ \forall \varphi \in [a, b]$ **OR** $\hat{P}_x(\varphi, A^c) \geq \varepsilon,\ \forall \varphi \in [a, b]$,

(ii) $T_x\mu(A) = \int_a^b \hat{P}_x(\varphi, A) d\mu_(\varphi)$.

For two probability measures μ_1, μ_2 there are finite measures $\lambda, \alpha_1, \alpha_2$ such that α_1 and α_2 are mutually singular, and $\mu_k = \lambda + \alpha_k$ for $k \in \{1, 2\}$ (see Lemma A.6). We can conclude:

$$\begin{aligned} \|T_x\mu_1 - T_x\mu_2\|_{TV} &= \|T_x\alpha_1 - T_x\alpha_2\|_{TV} \\ &\overset{(ii)}{=} 2 \sup_{A \in \mathcal{B}(S)} \left| \int_a^b \hat{P}_x(\varphi, A) d\alpha_1(\varphi) - \int_a^b \hat{P}_x(\varphi, A) d\alpha_2(\varphi) \right| \end{aligned}$$

We fix $A, A^c \in \mathcal{B}(S)$ and suppose without loss of generality that the inequality $\hat{P}_x(\varphi, A) \geq \varepsilon$ holds for all $\varphi \in [a, b]$.[5]

It is clear that $\varepsilon \leq \hat{P}_x(\varphi, A) \leq 1$ for all $\varphi \in S$. Since $\lambda(S) + \alpha_k(S) = 1$, we get that $\alpha_1(S) = 1 - \lambda(S) = \alpha_2(S)$. For this reason the inequality

$$\left| \int_a^b \hat{P}_x(\varphi, A) d\alpha_1(\varphi) - \int_a^b \hat{P}_x(\varphi, A) d\alpha_2(\varphi) \right| \leq (1-\varepsilon)\, \alpha_1([a, b]) \tag{3.27}$$

must be satisfied. Recall that α_1, α_2 are mutually singular measures, and therefore $\|\alpha_1 - \alpha_2\|_{TV} = \alpha_1([a, b]) + \alpha_2([a, b]) = 2\alpha_1([a, b])$. Using this and (3.27) we obtain:

$$\|T_x\mu_1 - T_x\mu_2\|_{TV} = 2 \sup_{A \in \mathcal{B}(S)} \left| \int_a^b \hat{P}_x(\varphi, A) d\alpha_1(\varphi) - \int_a^b \hat{P}_x(\varphi, A) d\alpha_2(\varphi) \right|$$

[5]If $\hat{P}_x(\varphi, A^c) \geq \varepsilon$ is true instead, we can just replace $\hat{P}_x(\varphi, A)$ by $1 - \hat{P}_x(\varphi, A^c)$ and get the same results.

$$\le (1-\varepsilon)\ \|\alpha_1 - \alpha_2\|_{TV}$$
$$= (1-\varepsilon)\ \|\mu_1 - \mu_2\|_{TV}$$

But, this had to be shown. □

A major implication of the previous theorem is the existence of a stationary distribution. In this context, a stationary distribution resembles a fixed point of the mapping T_x.

Theorem 3.5. *(a) The operator T_x has a unique fixed point, i.e. $\exists$ exactly one probability measure $\lambda \in \mathcal{M}^1(S, \mathcal{B})$ such that the equality $\lambda = T_x\lambda$ holds. (b) λ is a continuous probability distribution.*

Proof. **ad(a).** The space of probability measures $\mathcal{M}^1(S, \mathcal{B})$ equipped with the total variation norm is a complete metric space.[6] The statement thus follows from Theorem 3.4 and Banach's contraction mapping theorem.

ad(b). This follows from the fact that $F(\cdot|\varphi)$ and G represent continuous distributions themselves. □

To show the existence of a stationary equilibrium, we first define the mapping $T_f : [a, b] \to \mathcal{M}^1(S, \mathcal{B})$. The image of $x \in [a, b]$ under T_f is supposed to be the invariant probability measure of the operator T_x, i.e.

$$T_f(x) = \mu, \quad \text{such that } \mu = T_x\mu. \tag{3.28}$$

This mapping is well defined as, according to Theorem 3.5, each operator T_x has exactly one fixed point.

Lemma 3.6. *The mapping T_f is continuous.*

Proof. Let $\{x_n\}_{n\in\mathbb{N}} \subseteq S$ be a sequence and $x_n \to x$ for $n \to \infty$. We have to show that:

$$\|\underbrace{T_f(x_n)}_{\mu_n} - \underbrace{T_f(x)}_{\mu}\|_{TV} \to 0, \quad \text{if } n \to \infty. \tag{3.29}$$

For each set $A \in \mathcal{B}(S)$ we have

$$\|\mu_n - \mu\|_{TV} = \|T_{x_n}\mu_n - T_x\mu\|_{TV}$$

[6]For a proof of this fact see for example Stokey et al. (1989) p.343

$$
\begin{aligned}
&= \|T_{x_n}\mu_n - T_{x_n}\mu + T_{x_n}\mu - T_x\mu\|_{TV} \\
&\leq \|T_{x_n}\mu_n - T_{x_n}\mu\|_{TV} + \|T_{x_n}\mu - T_x\mu\|_{TV} \\
&\leq (1-\varepsilon)\,\|\mu_n - \mu\|_{TV} + 2\,\mu([x_n, x]).
\end{aligned}
$$

This is equivalent to:

$$\|\mu_n - \mu\|_{TV} \leq \frac{2}{\varepsilon}\,\mu([x_n, x]).$$

As shown in Theorem 3.5 the invariant distribution μ must be continuous. Therefore, the term on the right hand side tends to zero if $n \to \infty$, and the statement is proven. □

A stationary equilibrium exists in the dynamic framework if an exit-point x^* and its corresponding invariant distribution $\mu^* = T_f(x^*)$ satisfy the entry/exit conditions of Definition 3.2 at the same time. According to Lemma 3.3, the function $y = \kappa(Q)$ defines an capacity value for every output level such that the entry equality of a stationary equilibrium is fulfilled. Taking this into consideration, the existence of a stationary equilibrium reduces to the exit condition

$$x^* = \inf\left\{\varphi \in S : \int_a^b v(\varphi', Q^*, \kappa(Q^*))\,dF(\varphi'|\varphi) \geq \kappa(Q^*)\right\}. \tag{3.30}$$

This is essentially a fixed point problem since the aggregate output $Q^* = \displaystyle\int_a^b q\,dT_f(x^*)$ as well as $\kappa(Q^*)$ are also functions of x^*.

By assumption, the conditional distribution $F(.|b)$ stochastically dominates G. For all $x^* \in S$, we thus have

$$
\begin{aligned}
\int_a^b v(\varphi', Q^*, \kappa(Q^*))\,dF(\varphi'|b) &\geq \int_a^b v(\varphi, Q^*, \kappa(Q^*))\,dG(\varphi) - c_e \\
&= v^e(Q^*, \kappa(Q^*)) - c_e \\
&= \kappa(Q^*).
\end{aligned}
$$

As a consequence of this, the set in (3.30) cannot be empty, and the infimum will be situated inside the interval $[a, b]$. The continuity of Q^* and $\kappa(Q^*)$ with respect to x^* implies, moreover, that the infimum denotes a continuous function itself. But, then we can conclude the existence of a fixed point x^* – and thus the existence of a dynamic equilibrium – from the Intermediate Value Theorem. This proves the following theorem:

Theorem 3.7 (Existence of a stationary equilibrium)**.** *The dynamic, stochastic framework laid out in Section 3.1 possesses a stationary equilibrium.*

A stationary equilibrium without entry/exit will emerge whenever the continuation value of a firm with productivity level $\varphi = a$ is larger than the expected profits of entrants. This would be the case for $v^c(a, Q^*, \kappa(Q^*)) \geq \kappa(Q^*)$. A necessary condition for positive entry and exit is $a < x^* < b$. This would, for instance, be induced by the assumption

$$v^e(Q, 0) - c_e > v^c(a, Q, 0) \tag{3.31}$$

$$\Leftrightarrow \quad \sum_{t=0}^{\infty} \beta^t \int_a^b \pi(\varphi, Q)\, d\bar{\mu}_t^e(\varphi) - c_e > \sum_{t=0}^{\infty} \beta^t \int_a^b \pi(\varphi, Q)\, d\bar{\mu}_t^a(\varphi). \tag{3.32}$$

Here, $\bar{\mu}_t^a$ denotes the distribution of productivity levels that an incumbent firms with current productivity $\varphi = a$ will have at time t. By the same token, $\bar{\mu}_t^e$ displays the productivity distribution of a new firm t periods after its entry. In this situation, the firms' prospective productivity distributions do not involve any exit probability. As the assumed capacity value equals zero, firms neglect any exit possibility and will stay in the industry almost surely. Due to the separability of the period profits, the condition (3.32) can be generalized by

$$h(Q_{\max}) \sum_{t=0}^{\infty} \beta^t \left[\int_a^b g(\varphi)\, d\bar{\mu}_t^e(\varphi) - \int_a^b g(\varphi)\, d\bar{\mu}_t^a(\varphi) \right] > c_e. \tag{3.33}$$

So far, we have just shown the existence of a stationary equilibrium. The uniqueness of this steady state will follow by the convergence argument that we will demonstrate in the subsequent section.

3.4 Convergence to the steady state

It is clear that the industry converges to a steady state if $\beta = 0$ and $c_e = 0$. In this situation firms act myopically and just account for the very next period when making their entry/exit decisions. The exit rule simplifies to

$$x_t = \inf \left\{ \varphi \in S : \int_a^b \pi(\varphi', Q_{t+1})\, dF(\varphi'|\varphi) \geq \int_a^b \pi(\varphi', Q_{t+1})\, dG(\varphi') \right\} \tag{3.34}$$

$$= \inf \left\{ \varphi \in S : \int_a^b g(\varphi')\, dF(\varphi'|\varphi) \geq \int_a^b g(\varphi')\, dG(\varphi') \right\}, \tag{3.35}$$

such that the exit-point does not depend on the aggregate output Q_{t+1}, and accordingly on the industry structure μ_{t+1} anymore. The infimum in (3.35) is the same $x_t = x^*$ for all time periods, thus giving rise to a constant exit sequence in the equilibrium. For any starting distribution μ_0, the resulting industry structure in period t is consequently given by $\mu_t = T^t_{x^*}\mu_0$.[7] Due to the contraction property of T_{x^*}, which was shown in Theorem 3.4, the sequence $\{\mu_t\}_{t\in\mathbb{N}}$ converges to a stationary distribution $\lambda = T_{x^*}\lambda$ in the total variation norm. The convergence of those measures then entails the convergence of all other relevant parameters $\{Q_t\}$, $\{y_t\}$, $\{M_t\}$ for $t \to \infty$.

If either $\beta \neq 0$ or $c_e \neq 0$, however, convergence to a steady state is not straightforward. In this case, the dynamic development of the industry structure will likely provoke a fluctuating sequence of exit-points $\{x_t\}_{t\in\mathbb{N}}$. This leads to alternating operators T_{x_t} such that we cannot infer the convergence to a stationary distribution from the contractive behaviour of the mapping T_{x_t} any longer. Keeping in mind that

$$T_{x_1}\mu(A) - T_{x_2}\mu(A) = \int_{x_1}^{x_2} \Big\{P(\varphi, A) - \nu(A)\Big\}\, d\mu(\varphi) \tag{3.36}$$

for all sets $A \in \mathcal{B}(S)$ and $x_1 < x_2$, we rather have the inequality:

$$\begin{aligned} \|T_{x_1}\mu_1 - T_{x_2}\mu_2\|_{TV} &= \|T_{x_1}\mu_1 - T_{x_1}\mu_2 + T_{x_1}\mu_2 - T_{x_2}\mu_2\|_{TV} \\ &\leq \|T_{x_1}\mu_1 - T_{x_1}\mu_2\|_{TV} + \|T_{x_1}\mu_2 - T_{x_2}\mu_2\|_{TV} \\ &\leq (1-\varepsilon)\|\mu_1 - \mu_2\|_{TV} + 2\,\mu_2([x_1, x_2]). \end{aligned} \tag{3.37}$$

Because of the last term, the difference between $T_{x_1}\mu_1$ and $T_{x_2}\mu_2$ will depend on the distance of the two exit-points x_1, x_2 as well. Therefore, the inequality (3.37) does not necessarily constitute a contraction, and the simple convergence argument from the myopic scenario ($\beta = 0$) cannot be transferred to this case.

Generally speaking, the inequality above holds for all arbitrarily chosen values x_1 and x_2. In our analysis, however, these values are rather considered to be equilibrium solutions representing the firm's optimal exit-policy for current industry structures μ_1 and μ_2. We conjecture therefore that the distance between two exit-points tends to zero, i.e. $|x_1 - x_2| \to 0$, whenever $\|\mu_1 - \mu_2\|_{TV} \to 0$. This brings us to the following method of resolution. We will illustrate that for any value $(1-\varepsilon) < \theta < 1$ the inequality

$$\|T_{x_1}\mu_1 - T_{x_2}\mu_2\|_{TV} \leq \theta\, \|\mu_1 - \mu_2\|_{TV} \tag{3.38}$$

[7] Here, $T^t_{x^*}$ denotes the $t-$fold composition of the operator T_{x^*}.

can be justified if x_1, x_2 are the exit-points referring to μ_1, μ_2 in an equilibrium, and both the entry costs c_e and the discount factor β are below some upper boundaries $\bar{c} > 0$ and $\bar{\beta} > 0$. This would still imply contractive behaviour of the industry structure and, thus, result in convergence of $\{\mu_t\}_{t\in\mathbb{N}}$ to a stationary distribution.

For a fixed (continuous) probability measure $\lambda \in \mathcal{M}^1(S, \mathcal{B})$ and $\eta > 0$ we define the distance function

$$d_\lambda(\eta) := \sup_{\mu:\|\mu-\lambda\|_{TV}\leq\eta} |x_\mu - x_\lambda| \tag{3.39}$$

indicating the maximum possible difference between the two corresponding exit-points x_μ and x_λ in an equilibrium.[8] It follows immediately from the definition that d_λ is an upward sloping function and $\lim_{\eta\to 0} d_\lambda(\eta) = 0$. Furthermore, the shape of the whole function will particularly be affected by the height of the entry costs and the discount factor. We have discussed above why $d_\lambda \equiv 0$ for $c_e = 0$ and $\beta = 0$.

For convenience let $\lambda = T_{x_\lambda}\lambda$ henceforth denote the invariant distribution belonging to the stationary equilibrium, and let x_λ be the solution of the exit-rule. The assumptions imposed on $F(\cdot|\varphi)$ and G imply that λ is a continuous probability distribution, and its cdf F_λ satisfies $\frac{dF_\lambda(\varphi)}{d\varphi} \leq m_\lambda$ for some positive value $m_\lambda = \max\{m_F, m_G\}$. We can conclude from this that:

$$0 \leq \lambda([x_1, x_2]) \leq m_\lambda \cdot |x_1 - x_2|. \tag{3.40}$$

According to (3.37), the contraction property (3.38) is satisfied as long as the inequality

$$(1-\varepsilon)\,\|\mu - \lambda\|_{TV} + 2\,\lambda([x_\mu, x_\lambda]) \leq \theta\,\|\mu - \lambda\|_{TV} \tag{3.41}$$

holds for any probability measure $\mu \in \mathcal{M}^1(S, \mathcal{B})$. If $\|\mu - \lambda\|_{TV} = \eta$, it suffices to show that

$$\begin{aligned} & (1-\varepsilon)\eta + 2\,m_\lambda d_\lambda(\eta) \leq \theta\,\eta \\ \Leftrightarrow \qquad & \frac{d_\lambda(\eta)}{\eta} \leq \frac{\theta - (1-\varepsilon)}{2\,m_\lambda} \end{aligned} \tag{3.42}$$

for all $\eta \in [0, 2]$.

In the next step, we will determine a range for the possible exit solutions x_μ. For this purpose, let the probability measure μ describe the current industry structure at time t and $\|\mu - \lambda\|_{TV} = \eta$. Depending on the total variation η, we will construct an interval I that is always centered around the stationary exit-point x_λ and certainly

[8] The supremum in (3.39) is taken over all continuous probability measures $\mu \in \mathcal{M}^1(S, \mathcal{B})$.

contains the exit solution x_μ. As will become clear later on, the length of this interval I will not just depend on the parameter η but also on the size of c_e and β. If $z_t^* = \{Q_j^*, y_j^*\}_{j=t}^\infty$ denotes the sequence that μ entails in an equilibrium,[9] the exit-point x_μ must be subject to

$$\begin{aligned} x_\mu &= \inf\left\{\varphi \in S : \int_a^b v(\varphi', z_{t+1}^*)\, dF(\varphi'|\varphi) \geq y_{t+1}^*\right\} \\ &= \inf\left\{\varphi \in S : v^c(\varphi, z_{t+1}^*) - v^e(z_{t+1}^*) + c_e \geq 0\right\}. \end{aligned} \tag{3.43}$$

If we introduce the function $u(\varphi, z_t) = \max\{y_{t+1}, v^c(\varphi, z_{t+1})\}$, the exit-point can be redefined by

$$\begin{aligned} x_\mu = \inf\Bigg\{\varphi \in S : &\int_a^b \pi(\varphi', Q_{t+1}^*)\, dF(\varphi'|\varphi) + \beta \int_a^b u(\varphi', z_{t+1}^*)\, dF(\varphi'|\varphi) \\ &- \int_a^b \pi(\varphi', Q_{t+1}^*)\, dG(\varphi') - \beta \int_a^b u(\varphi', z_{t+1}^*)\, dG(\varphi') + c_e \geq 0\Bigg\}. \end{aligned} \tag{3.44}$$

Using the separability of period profits, this is equivalent to

$$\begin{aligned} x_\mu = \inf\Bigg\{\varphi \in S : &\int_a^b g(\varphi')\, dF(\varphi'|\varphi) + \frac{\beta}{h(Q_{t+1}^*)} \int_a^b u(\varphi', z_{t+1}^*)\, dF(\varphi'|\varphi) \\ &- \int_a^b g(\varphi')\, dG(\varphi') - \frac{1}{h(Q_{t+1}^*)}\left(\beta \int_a^b u(\varphi', z_{t+1}^*)\, dG(\varphi') - c_e\right) \geq 0\Bigg\}. \end{aligned} \tag{3.45}$$

To simplify the notation, we define the function inside the curly brackets as $s_\mu(\varphi)$, such that $x_\mu = \inf\left\{\varphi \in S : s_\mu(\varphi) \geq 0\right\}$.

Let δ_a and δ_b denote the Dirac measures in a and b. These are defined for any set $A \in \mathcal{B}(S)$ by

$$\delta_a(A) = \begin{cases} 1 & \text{if } a \in A \\ 0 & \text{if } a \notin A \end{cases} \qquad \text{and} \quad \delta_b(A) = \begin{cases} 1 & \text{if } b \in A \\ 0 & \text{if } b \notin A. \end{cases} \tag{3.46}$$

[9] Recall that we implicitly presume the existence of a converging dynamic equilibrium at this stage. The proof requires some preparation and will be presented in Section 3.5.

It is proven in Lemma A.9 that for any probability measure μ being subject to $\|\mu - \lambda\|_{TV} \leq \eta$ the following inequalities hold:

$$\mu + \frac{\eta}{2}\,\delta_b \overset{FOSD}{\geq} \lambda + \frac{\eta}{2}\,\delta_a \quad \text{and} \quad \mu + \frac{\eta}{2}\,\delta_a \overset{FOSD}{\leq} \lambda + \frac{\eta}{2}\,\delta_b.$$

This allows us to specify boundaries for the aggregate output $Q_\mu = \int_a^b q(\varphi)\, d\mu(\varphi)$ that materialises under the industry structure μ. The lower/upper boundaries depend solely on the stationary distribution λ and the variational distance η between both distributions:

$$\underbrace{\int_a^b q(\varphi)\, d\lambda(\varphi) - \frac{\eta}{2}(Q_{\max} - Q_{\min})}_{=:Q^\lambda_{\min}(\eta)} \leq Q_\mu \leq \underbrace{\int_a^b q(\varphi)\, d\lambda(\varphi) + \frac{\eta}{2}(Q_{\max} - Q_{\min})}_{=:Q^\lambda_{\max}(\eta)} \tag{3.47}$$

For the sake of convenience, the lower boundary is henceforth denoted by $Q^\lambda_{\min}(\eta)$ and the upper boundary by $Q^\lambda_{\max}(\eta)$ respectively. The aggregate output in the subsequent period, which can be interpreted as a function of the exit-point

$$Q_\mu(x) = \int_a^b q(\varphi)\, dT_x\mu(\varphi), \tag{3.48}$$

will also be situated in the interval $\left[Q^\lambda_{\min}(\eta), Q^\lambda_{\max}(\eta)\right]$ when the exit-point x_μ is close enough to the stationary solution x_λ, and the contraction inequality (3.41) is fulfilled.

Minimum and maximum capacity values depending on the variational difference η can be found in a similar way. If we presume a constant industry output over time, we can define $y^\lambda_{\min}(\eta) = \kappa\left(Q^\lambda_{\max}(\eta)\right)$ and $y^\lambda_{\max}(\eta) = \kappa\left(Q^\lambda_{\min}(\eta)\right)$ by means of the function κ, which is the solution to the entry equality $v^e(Q, \kappa(Q)) - \kappa(Q) = c_e$. For the constant output/capacity sequences $\bar{z}_{\min}(\eta) = \left\{Q^\lambda_{\min}(\eta), y^\lambda_{\max}(\eta)\right\}_{t\in\mathbb{N}}$ and $\bar{z}_{\max}(\eta) = \left\{Q^\lambda_{\max}(\eta), y^\lambda_{\min}(\eta)\right\}_{t\in\mathbb{N}}$ we define moreover

$$\begin{aligned} s_L(\eta, \varphi) := &\int_a^b g(\varphi')\, dF(\varphi'|\varphi) + \frac{1}{h\left(Q^\lambda_{\max}(\eta)\right)} \left(\beta \int_a^b u\left(\varphi', \bar{z}_{\min}(\eta)\right) dF(\varphi'|\varphi) + c_e\right) \\ &- \int_a^b g(\varphi')\, dG(\varphi') - \frac{\beta}{h\left(Q^\lambda_{\min}(\eta)\right)} \int_a^b u\left(\varphi', \bar{z}_{\max}(\eta)\right) dG(\varphi') \end{aligned} \tag{3.49}$$

and

$$s_U(\eta,\varphi) := \int_a^b g(\varphi')\, dF(\varphi'|\varphi) + \frac{1}{h\left(Q^\lambda_{\min}(\eta)\right)} \left(\beta \int_a^b u\left(\varphi', \bar{z}_{\max}(\eta)\right) dF(\varphi'|\varphi) + c_e\right)$$
$$- \int_a^b g(\varphi')\, dG(\varphi') - \frac{\beta}{h\left(Q^\lambda_{\max}(\eta)\right)} \int_a^b u\left(\varphi', \bar{z}_{\min}(\eta)\right) dG(\varphi'). \qquad (3.50)$$

When the output/capacity sequences referring to the distribution μ in an equilibrium are subject to $\left\{Q^*_j\right\}_{j=t}^\infty \subset \left[Q^\lambda_{\min}(\eta), Q^\lambda_{\max}(\eta)\right]$ and $\left\{y^*_j\right\}_{j=t}^\infty \subset \left[y^\lambda_{\min}(\eta), y^\lambda_{\max}(\eta)\right]$, the monotonicity of $h(\cdot)$ and $u(\cdot)$ implies

$$s_L(\eta,\varphi) \geq s_\mu(\varphi) \geq s_U(\eta,\varphi) \quad \text{for all} \quad \varphi \in S. \qquad (3.51)$$

In this case, the intersection points of the functions s_L, s_U with the origin constitute lower and upper boundaries to the exit solution x_μ (see also Figure 3.2 for an illustration). Regarding the definition of x_μ in (3.45), the boundaries are formally defined by:

$$x_L(\eta) = \inf\Big\{\varphi \in S : s_L(\eta,\varphi) \geq 0\Big\}, \qquad (3.52)$$
$$x_U(\eta) = \inf\Big\{\varphi \in S : s_U(\eta,\varphi) \geq 0\Big\}. \qquad (3.53)$$

The way, in which s_L and s_U have been designed, means that we can take $x_\mu \in [x_L(\eta), x_U(\eta)]$ for granted. The monotonicity of s_L, s_U with respect to η implies moreover that the intervals containing the exit solution are nested, i.e. if $\{\eta_j\}_{j\in\mathbb{N}}$ is a positive, decreasing sequence, the intervals $I_j = [x_L(\eta_j), x_U(\eta_j)]$ satisfy $I_{j+1} \subseteq I_j$. It follows immediately from the continuity of s_L, s_U that the length of these intervals tends to zero if $\eta_j \to 0$. Since the exit-point $x_\lambda = x_L(0) = x_U(0)$ lies within all intervals I_j, an upper bound for the distance function can be estimated by

$$d_\lambda(\eta) \leq \max\Big\{\left|x_L(\eta) - x_L(0)\right|, \left|x_U(\eta) - x_U(0)\right|\Big\}. \qquad (3.54)$$

Lemma 3.8. *The functions $x_L(\eta)$ and $x_U(\eta)$ are continuous with respect to the distance paramter $\eta \in [0,2]$.*

Proof. We just show the continuity of x_L here as the continuity of x_U follows by exactly the same argument. Furthermore, the monotonicity and continuity properties

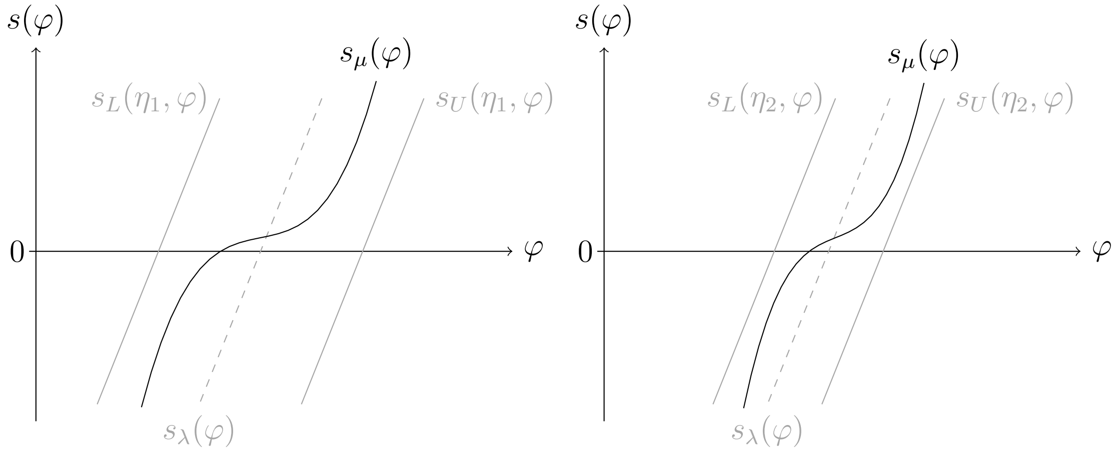

Figure 3.2: Illustration of the exit condition for different distances $\eta_1 > \eta_2$

of the value function can directly be transferred to the function s_L. Hence, $s_L(\eta, \varphi)$ is continuous in both parameters and strictly increasing with respect to φ.

The statement is proven by contradiction. We assume that x_L is not continuous in a point $\bar{\eta} \in [0, 2]$. Then, there is a sequence $\{\eta_m\}_{m\in\mathbb{N}}$ such that $\eta_m \to \bar{\eta}$ for $m \to \infty$ but

$$\lim_{m\to\infty} x_L(\eta_m) = \bar{x} \neq x_L(\bar{\eta}). \tag{3.55}$$

By the defintion of x_L, we must have $s_L(\eta_m, x_L(\eta_m)) \geq 0$ for all $m \in \mathbb{N}$. Since s_L is continuous this implies

$$\lim_{m\to\infty} s_L(\eta_m, x_L(\eta_m)) = s_L(\bar{\eta}, \bar{x}) \geq 0.$$

Two scenarios are possible. If $s_L(\bar{\eta}, \bar{x}) = 0$, we can conclude from the infimum property in (3.52) that $s_L(\bar{\eta}, x_L(\bar{\eta})) = 0$. This, however, is a contradiction to the strict monotonicity of $s_L(\eta, \varphi)$ with respect to φ.

If $s_L(\bar{\eta}, \bar{x}) > 0$, the continuity of s_L implies that $s_L(\eta_m, x_L(\eta_m)) > 0$ for all $m \geq m_0$. Due to x_L's infimum property on S, the equality $x_L(\eta_m) = a$ applies for all $m \geq m_0$ as well. But, then we have $\bar{x} = a = x_L(\bar{\eta})$, which contradicts (3.55). $\square$

The length of the interval I_j does not only depend on the distance parameter η_j. As Figure 3.3 indicates, the slope of the functions $s(\varphi)$ is also affected by the size of the entry costs and the discount factor. In fact, s_L and s_U are situated closer to each other if either the entry costs or the discount factor decline. Remember that we have already argued at the beginning of this section why $s_L(\eta_j, \varphi) \equiv s_U(\eta_j, \varphi)$ for $c_e = \beta = 0$, and the interval I_j just contains the stationary exit-point x_λ in this case.

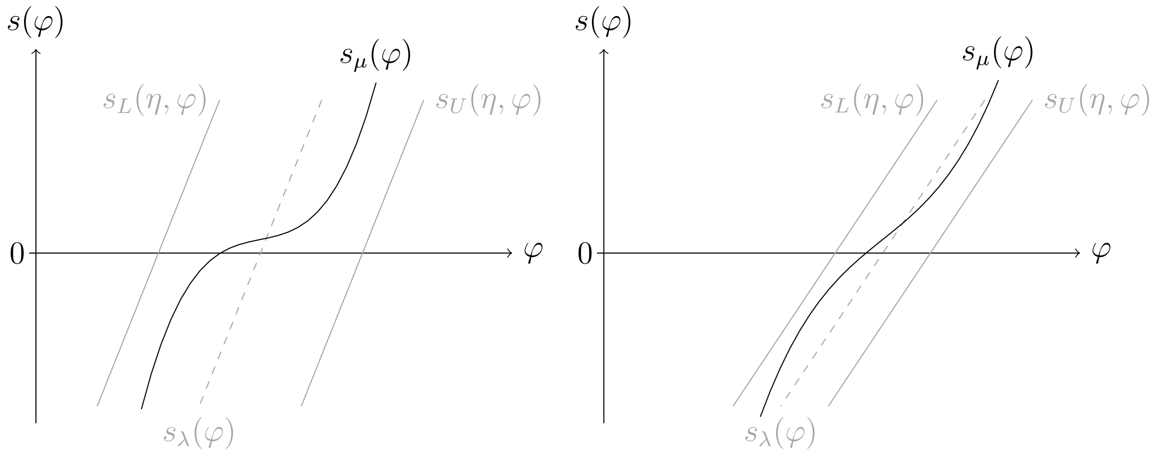

Figure 3.3: Illustration of the exit condition for a fixed distance $\eta > 0$ but declining discount factor β and entry costs c_e

Our goal is now to specify a range for the values c_e and β, such that the contraction condition in (3.42) is indeed satisfied for any $\eta \in [0, 2]$. Because of the boundedness in (3.54), we do not have to analyse the expression $\frac{d_\lambda(\eta)}{\eta}$ but can consider x_L's differential quotient and show that

$$\left| \frac{x_L(\eta) - x_L(0)}{\eta} \right| \leq \frac{\theta - (1 - \varepsilon)}{2\, m_\lambda}. \tag{3.56}$$

The same applies, of course, to the function x_U. In the following, we will exemplary focus on x_L however.

The crucial question is how x_L reacts to changes in the distance parameter. Because the function essentially represents a solution to $s_L(\eta, x_L(\eta)) = 0$, the first derivative can be computed by means of the chain rule[10]

$$x'_L(\bar{\eta}) = -\frac{\frac{\partial s_L}{\partial \eta}\big(\bar{\eta}, x_L(\bar{\eta})\big)}{\frac{\partial s_L}{\partial \varphi}\big(\bar{\eta}, x_L(\bar{\eta})\big)}. \tag{3.57}$$

If the absolute value of this term can be bounded for all $\bar{\eta} \in [0, 2]$ by the constant value $\frac{\theta - (1-\varepsilon)}{2\, m_\lambda}$ then so will be the differential quotient in equation (3.56).[11]

According to Lemma A.8, the function $\int_a^b g(\varphi')\, dF(\varphi'|\varphi)$ is strictly increasing in the productivity parameter φ. Hence, the derivative of this expression with respect to φ must be larger or equal to zero. It is necessary, however, to rule out any saddle

[10] If $s_L(\eta, \varphi) > 0$ for all $\varphi \in S$, then $x_L(\eta) = a$ and $x'_L(\eta) = 0$.

[11] This is a direct implication of the Mean Value Theorem, which states that $\frac{x_L(\eta) - x_L(0)}{\eta} = x'_L(\bar{\eta})$ for some $\bar{\eta} \in (0, \eta)$.

point here and presume that

$$\frac{d}{d\varphi}\left(\int_a^b g(\varphi')\, dF(\varphi'|\bar{\varphi})\right) \geq g_F' > 0, \quad \forall\, \bar{\varphi} \in S. \tag{3.58}$$

As a consequence of this, the estimate $\frac{\partial s_L}{\partial \varphi}(\bar{\eta}, \bar{\varphi}) \geq g_F'$ surely holds, too. The partial derivative of s_L with respect to η can be bounded by

$$0 \leq \frac{\partial s_L}{\partial \eta}(\bar{\eta}, \bar{\varphi}) \leq -\frac{h'_{\min}\,(Q_{\max} - Q_{\min})}{h^2(Q_{\max})}\left[\frac{\beta\,(2-\beta)\, g_F\, h(Q_{\min})}{(1-\beta)^2} + \frac{c_e}{2}\right]. \tag{3.59}$$

A proof of this inequality is presented in Lemma A.12. If both parts are combined, it follows that[12]

$$\frac{h'_{\min}\,(Q_{\max} - Q_{\min})}{g_F'\, h^2(Q_{\max})}\left[\frac{\beta\,(2-\beta)\, g_F\, h(Q_{\min})}{(1-\beta)^2} + \frac{c_e}{2}\right] \leq x_L'(\eta) \leq 0. \tag{3.60}$$

Hence, the contraction inequality (3.56) holds whenever

$$\begin{aligned} & \frac{|h'_{\min}|\,(Q_{\max} - Q_{\min})}{g_F'\, h^2(Q_{\max})}\left[\frac{\beta\,(2-\beta)\, g_F\, h(Q_{\min})}{(1-\beta)^2} + \frac{c_e}{2}\right] \leq \frac{\theta - (1-\varepsilon)}{2\, m_\lambda} \\ \Leftrightarrow \quad & \frac{\beta\,(2-\beta)\, g_F\, h(Q_{\min})}{(1-\beta)^2} + \frac{c_e}{2} \leq \frac{(\theta - (1-\varepsilon))\, g_F'\, h^2(Q_{\max})}{2\, m_\lambda\, |h'_{\min}|\,(Q_{\max} - Q_{\min})}. \end{aligned} \tag{3.61}$$

The same boundary can be derived for the function $x_U'(\eta)$. Thus, the condition above implicitly defines a positive range of discount factors and entry costs supporting contractive behaviour of the industry.[13] This brings us to the following convergence statement.

Theorem 3.9 (Convergence to a stationary distribution)**.** *Let a continuous starting distribution $\mu_0 \in \mathcal{M}^1(S, \mathcal{B}(S))$ be given, and assume that $\|\mu_t - \lambda\|_{TV} =: \eta_t$. If the discount factor β as well as the entry costs c_e satisfy the condition* (3.61)*, then the sequence $\mu_{t+1} = T_{x_L(\eta_t)}\mu_t$ converges in the total variation norm to the stationary distribution λ. The same applies to the sequence $\mu_{t+1} = T_{x_U(\eta_t)}\mu_t$.*

Proof. The way, in which we have constructed the boundary condition for β and c_e,

[12]Notice here that $h'_{\min} < 0$ is the only negative term. All others terms are supposed to be non-negative.

[13]Condition (3.61) does not necessarily specify the highest possible threshold for β and c_e. The estimate can, for instance, be improved if g_F' is modified and more (higher order) terms are taken into consideration.

implies that the contraction condition (3.42) is fulfilled. It follows that

$$\begin{aligned}
\|\mu_{t+1} - \lambda\|_{TV} &= \|T_{x_L(\eta_t)}\mu_t - T_{x_\lambda}\lambda\|_{TV} \\
&\leq \theta\, \|\mu_t - \lambda\|_{TV} \\
&= \theta\, \|T_{x_L(\eta_{t-1})}\mu_{t-1} - T_{x_\lambda}\lambda\|_{TV} \\
&\leq \theta^2\, \|\mu_{t-1} - \lambda\|_{TV} \\
&\vdots \\
&\leq \theta^{t+1}\, \|\mu_0 - \lambda\|_{TV} \quad \to 0 \quad \text{for} \quad t \to \infty \quad \text{since} \quad \theta < 1.
\end{aligned}$$

The proof for $\mu_{t+1} = T_{x_U(\eta_t)}\mu_t$ is carried out in exactly the same way. □

The convergence theorem implies some interesting findings. Notice, first of all, that the convergence rate is displayed by the parameter θ. Referring to condition (3.61), declining discount factors or entry costs validate smaller values for θ and, thus, increase the speed of convergence. The highest possible convergence rate $(1-\varepsilon)$ is certainly achieved for the myopic scenario with no fixed entry costs (i.e. for the case$\beta = c_e = 0$). The proof also demonstrates why the stationary industry equilibrium must be unique. If there were two different equilibria with stationary distributions $\lambda_1 \neq \lambda_2$, this would lead to the contradiction

$$\|\lambda_1 - \lambda_2\|_{TV} = \|T_{x_{\lambda_1}}\lambda_1 - T_{x_{\lambda_2}}\lambda_2\|_{TV} \leq \theta\, \|\lambda_1 - \lambda_2\|_{TV} < \|\lambda_1 - \lambda_2\|_{TV}.$$

We can, moreover, conclude from Theorem 3.9 that the sequence $\mu_{t+1} = T_{x_t}\mu_t$ will also converge to the stationary distribution λ for any exit-point sequence being subject to $x_t \in [x_L(\eta_t), x_U(\eta_t)]$. The existence of such a sequence representing the equilibrium exit solutions is shown in the next section.

3.5 Existence of a converging dynamic equilibrium

In this section we will finally prove the existence of a dynamic equilibrium that converges to a steady state in the long-run. The key to the proof is Schauder's Fixed Point Theorem, which generalises Brouwer's Fixed Point Theorem to infinite-dimensional Banach spaces. The theorem ensures that every continuous mapping of a compact, convex set into itself has a fixed point. Hence, we will proceed as follows: First, we define a continuous mapping $\tau : \ell^\infty \times \ell^\infty \to \ell^\infty \times \ell^\infty$ on the space of bounded sequences such that a fixed point of this mapping coincides with the dynamic equilibrium. Second, we use the contraction property derived in the previous section to specify a compact, convex subset $N \subset \ell^\infty \times \ell^\infty$, which is mapped into itself by τ, i.e. $\tau(N) \subseteq N$. According to Schauder's Theorem, this mapping possesses a fixed point, and therefore a dynamic equilibrium exists in the assumed framework.

Let a continuous starting distribution $\mu_0 \in \mathcal{M}^1(S, \mathcal{B}(S))$ as well as a sequence of exit-points and capacity values $\{x_t, y_t\}_{t\in\mathbb{N}_0}$ be given. This means that $x_t \in [a, b]$ and $y_t \geq 0$ hold in particular for all $t \in \mathbb{N}_0$. We construct τ such that it maps this sequence into another sequence of exit-points and capacity values. According to the industry dynamics, the presumed exit sequence provokes a certain evolution of industry structures $\mu_{t+1} = T_{x_t}\mu_t$. Each distribution is connected to an aggregate industry output via the equality

$$Q_t = \int_a^b q(\varphi)\, d\mu_t(\varphi). \tag{3.62}$$

For the resulting output sequence $\{Q_t\}_{t\in\mathbb{N}_0}$ and the given capacity sequence $\{y_t\}_{t\in\mathbb{N}_0}$ we can determine new capacity values $\widetilde{y}_t$ by

$$\widetilde{y}_t = \int_a^b v(\varphi, z_{t+1})\, dG(\varphi) - c_e \tag{3.63}$$

$$= v^e(z_{t+1}) - c_e. \tag{3.64}$$

As before, the variable $z_t := \{Q_j, y_j\}_{j=t}^\infty$ is supposed to denote the combined output-capacity sequence starting from time t. The calculated capacity value $\widetilde{y}_t$ is indeed non-negative as long as the assumption $v^e(Q_{\max}, 0) \geq 0$ from Lemma 3.3 is met. These capacity values are then used to compute new exit-points by the rule

$$\widetilde{x}_t := \inf\left\{\varphi \in S : \int_a^b v(\varphi', z_{t+1})\, dF(\varphi'|\varphi) \geq \widetilde{y}_t\right\} \tag{3.65}$$

$$= \inf\left\{\varphi \in S : v^c(\varphi, z_{t+1}) \geq \widetilde{y}_t\right\}. \tag{3.66}$$

Both variables are calculated for each time period $t \in \mathbb{N}_0$ such that this yields a new sequence $\{\widetilde{x}_t, \widetilde{y}_t\}_{t\in\mathbb{N}_0}$ of exit-points and capacity prices. The mapping τ on the space of bounded sequences is, thus, defined by $\tau : \{x_t, y_t\}_{t\in\mathbb{N}_0} \mapsto \{\widetilde{x}_t, \widetilde{y}_t\}_{t\in\mathbb{N}_0}$. It is evident that this is a continuous mapping, and every fixed point of τ represents a dynamic equilibrium in the sense of Definition 3.1.

The challenge is now to specify a compact $N = N_1 \times N_2 \subset \ell^\infty \times \ell^\infty$ that is mapped into itself by τ. We will define compact sets $N_1 \subset \ell^\infty$ for the exit sequences and $N_2 \subset \ell^\infty$ for the capacity sequences separately. Because the space of bounded sequences is infinite dimensional, the simple cartesian product of compact intervals $N_1 = [a,b] \times [a,b] \times \dots$ does not form a compact set in this space. But, if sequences $\{a_t\}_{t\in\mathbb{N}_0}$ and $\{b_t\}_{t\in\mathbb{N}_0}$ exist such that $|a_t - b_t| \to 0$ for $t \to \infty$ then the cartesian product

$$N_1 = \prod_{t=0}^{\infty} [a_t, b_t]$$

is a compact, convex set in ℓ^∞ (see Lemma A.13 for a proof of this fact).

Such bounding sequences for the exit-points can be found by using the convergence property from Theorem 3.9. If μ_0 is the given starting distribution and $\|\mu_0 - \lambda\|_{TV} = \bar\eta$, we set $a_t := x_L(\theta^t\,\bar\eta)$ and $b_t := x_U(\theta^t\,\bar\eta)$. As the term $\theta^t\,\bar\eta \to 0$ for $t \to \infty$, this generates monotone sequences $\{a_t\}, \{b_t\}$ converging to the stationary exit-point x_λ. Hence, these sequences satisfy $|a_t - b_t| \to 0$ in particular, and we can define a compact, convex subset for the exit sequences by

$$N_1 := \Big\{ \{x_t\}_{t\in\mathbb{N}_0} \in \ell^\infty : \; a_t \le x_t \le b_t, \; \forall\, t \in \mathbb{N}_0 \Big\}. \tag{3.67}$$

The shape of the set N_1 is depicted in Figure 3.4. Bounding sequences for the capacity values are constructed in similar fashion. If we define values $c_t := \kappa\left(Q^\lambda_{\max}(\theta^t\,\bar\eta)\right)$ and $d_t := \kappa\left(Q^\lambda_{\min}(\theta^t\,\bar\eta)\right)$, then both sequences $\{c_t\}, \{d_t\}$ tend to the stationary capacity value y_λ in the limit, and a compact, convex subset for the capacity sequences is given by

$$N_2 := \Big\{ \{y_t\}_{t\in\mathbb{N}_0} \in \ell^\infty : \; c_t \le y_t \le d_t, \; \forall\, t \in \mathbb{N}_0 \Big\}. \tag{3.68}$$

As a final step it remains to be shown that the image of any combined sequence $\{x_t, y_t\}_{t\in\mathbb{N}_0} \in N_1 \times N_2$ under the mapping τ is also an element of the subset $N_1 \times N_2$. We will see that this is indeed the case, whenever the discount factor β and the entry costs c_e satisfy the inequality (3.61). Under this assumption the exit sequence $\{x_t\}$ implies an evolution of the industry structure $\{\mu_t\}$ that is subject to $\|\mu_t - \lambda\|_{TV} \le \theta^t\bar\eta$ for every time period t. Moreover, this involves an aggregate output sequence $\{Q_t\}$ that is bounded by $Q^\lambda_{\min}(\theta^t\,\bar\eta) \le Q_t \le Q^\lambda_{\max}(\theta^t\,\bar\eta)$. But, if $z_t := \{Q_j, y_j\}_{j=t}^\infty$ denotes

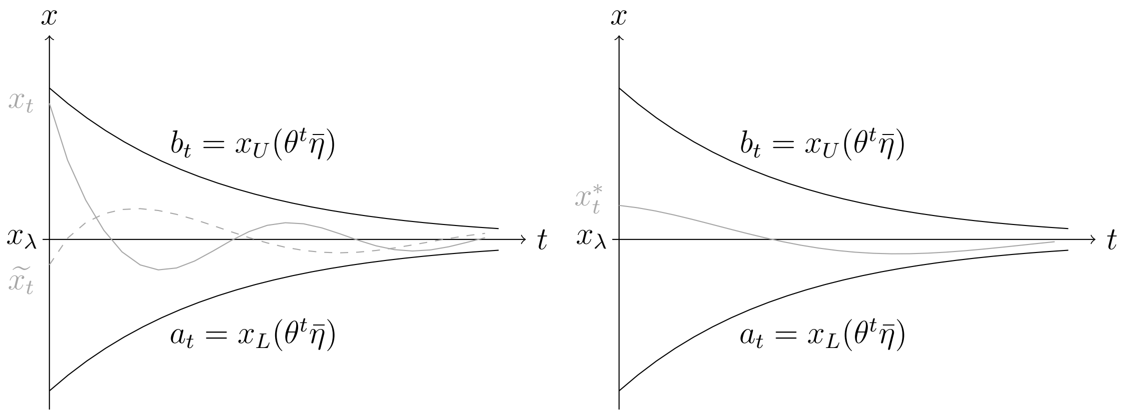

Figure 3.4: Illustration of the mapping τ with respect to the exit sequence (left side) and its fixed point (right side).

the derived output-capacity sequence, this means nothing else than

$$\begin{aligned} c_t &= v^e(Q^\lambda_{\max}(\theta^t\,\bar{\eta}), c_t) - c_e \\ &\le v^e(z_{t+1}) - c_e \\ &= \widetilde{y}_t \\ &\le v^e(Q^\lambda_{\min}(\theta^t\,\bar{\eta}), d_t) - c_e \\ &= d_t \end{aligned}$$

and $\{\widetilde{y}_t\}_{t\in\mathbb{N}_0} \in N_2$.

The mapping τ defines the exit-point as $\widetilde{x}_t = \inf\big\{\varphi \in S : v^c(\varphi, z_{t+1}) - \widetilde{y}_t \ge 0\big\}$. We can conclude from the boundedness of z_t that

$$s_L(\theta^t\bar{\eta}, \varphi) \ge v^c(\varphi, z_{t+1}) - \widetilde{y}_t \ge s_U(\theta^t\bar{\eta}, \varphi)$$

holds for all productivity levels $\varphi \in S$. This implies

$$\begin{aligned} a_t &= \inf\big\{\varphi \in S : s_L(\theta^t\bar{\eta}, \varphi) \ge 0\big\} \\ &\le \inf\big\{\varphi \in S : v^c(\varphi, z_{t+1}) - \widetilde{y}_t \ge 0\big\} \\ &= \widetilde{x}_t \\ &\le \inf\big\{\varphi \in S : s_U(\theta^t\bar{\eta}, \varphi) \ge 0\big\} \\ &= b_t \end{aligned}$$

for all times t and consequently $\{\widetilde{x}_t\}_{t\in\mathbb{N}_0} \in N_1$. Hence, the sequence $\{\widetilde{x}_t, \widetilde{y}_t\}$ is an element of $N_1 \times N_2$, and according to Schauder's Fixed Point Theorem the mapping τ must have a fixed point. This allows us to make the following significant statement.

Theorem 3.10 (Existence of a converging dynamic equilibrium)**.** *Let the discount factor β and the entry costs c_e satisfy condition* (3.61)*. For any continuous starting distribution $\mu_0 \in \mathcal{M}^1(S, \mathcal{B}(S))$ there exists a dynamic equilibrium that converges to the uniquely determined steady state in the long-run.*

Unfortunately, Schauder's Fixed Point Theorem just guarantees the existence of a dynamic equilibrium. The uniqueness of this equilibrium is not necessarily the case, and we cannot fully exclude the existence of a non-converging equilibrium. The advantage, however, is that we are not restricted to perfect competition. Many other studies dealing with dynamic stochastic equilibria assume perfect competition, and prove the existence of an equilibrium by utilising the equivalence of an equilibrium to the optimum of a social welfare function and showing that this function possesses a unique optimum. In our framework, a dynamic equilibrium may, instead, even exist for monopolistic or oligopolistic competition.

To demonstrate that the convergence to a steady state is not trivial, and violations of the model assumptions may lead to a diverging industry evolution, we first give an example featuring the existence but not the convergence to a stationary equilibrium. In a way, the following example can be interpreted as a simplified, discrete version of the general model.

Example 3.1. In contrast to the continuous framework introduced in Section 3.1, we presume that the firm-specific productivity φ just takes on three possible values, and we define this discrete state space as row vector $s = (s_1, s_2, s_3)$. The productivity process is given by the matrix

$$P = (p_{ij}) = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \tag{3.69}$$

containing the transition probabilities $p_{ij} = \text{Prob}(\varphi_{t+1} = s_j | \varphi_t = s_i)$. The entry distribution is supposed to be $\nu = (0, 0, 1)$. This means new firms almost surely receive a productivity level $\varphi = s_3$ when they enter the industry.

The structure of the industry at time t is reflected by the vector $\mu^t = (\mu_1^t, \mu_2^t, \mu_3^t)$, in which μ_j^t denotes the mass of firms with productivity $\varphi_t = s_j$. The total industry mass is thus given by $\mu^t(s) = \sum_{j=1}^{3} \mu_j^t$. In accordance to the general model, we assume that $\mu(s) = 1$ defines the capacity constrained size of the industry, and the entry-mass M must be equal to the mass of exiting firms. For any exit-point

$s_{i-1} < x \leq s_i$, we define the matrix

$$P_x = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \\ p_{i1} & \cdots & p_{i3} \\ \vdots & & \vdots \\ p_{31} & \cdots & p_{33} \end{bmatrix} \tag{3.70}$$

and obtain the industry structure in period $t+1$ by the equation

$$\mu^{t+1} = \mu^t \cdot P_x + M \cdot \nu. \tag{3.71}$$

The period profits are supposed to satisfy $\pi(\varphi, \mu) = g(\varphi)\,(\mu(s))^{-1}$, with g being subject to $0 < g(s_1) < g(s_2) < g(s_3)$. For simplification, we assume that firms act myopically ($\beta = 0$), and just account for the subsequent period when deciding about entry and exit. The expected value of potential entrants is thus

$$v^e(\mu) = \nu \cdot \pi(s, \mu) = g(s_3)\,(\mu(s))^{-1}. \tag{3.72}$$

Apart from the fixed entry costs $g(s_3) - g(s_2) < c_e < g(s_3) - g(s_1)$, new firms must purchase production capacity at a price y. The firms' continuation value, which consists only of the expected profits in the upcoming period, is represented by the vector

$$v^c(\mu) = P \cdot \pi(s, \mu) = \begin{bmatrix} g(s_1) \\ g(s_1) \\ g(s_2) \end{bmatrix} (\mu(s))^{-1}. \tag{3.73}$$

An active firm remains in the industry as long as its continuation value is larger than the capacity price, i.e. $v^c(\mu) \geq y$.

The framework presented above possesses a stationary equilibrium. For the exit-point $x^* = s_3$ and entry-mass $M^* = \frac{1}{2}$, a stationary distribution is given by the vector $\mu^* = \left(0, \frac{1}{2}, \frac{1}{2}\right)$. Under this industry structure, entrants are willing to pay a capacity price $y^* = g(s_3) - c_e$. Hence, all firms with current productivity level $\varphi = s_1$ or $\varphi = s_2$ are better off ceasing production and leave the industry. This completes the necessary conditions for a stationary equilibrium.

If a starting distribution $\mu^0 = (0, \alpha, 1-\alpha)$ with $\alpha \neq \frac{1}{2}$ is presumed though, an alternating dynamic equilibrium emerges. The exit-points $x_t^* = s_3$, as well as the capacity values $y_t^* = g(s_3) - c_e$, stay constant over time. But, the sequence of entry-masses alternates as $\{M_t^*\} = \{\alpha; 1-\alpha; \alpha; 1-\alpha; \alpha; ...\}$. Consequently, the industry structure switches from $\mu^{2t} = (0, \alpha, 1-\alpha)$ in even periods to $\mu^{2t+1} = (0, 1-\alpha, \alpha)$ in uneven periods. We infer from this that the industry will never attain a steady state in this setup.

The particular setup in Example 3.1 does not involve convergence to a steady state. One reason for the alternating development is that the condition (CON) of the general model is violated. However, even slight modifications of the framework can change this outcome and lead to an adjustment of the industry to the stationary equilibrium. One modification supporting this behaviour is described in the next example.

Example 3.2. Referring to the model in Example 3.1, a different entry distribution can lead to a converging dynamic equilibrium. If we presume the entry distribution $\nu = \left(0, \frac{1}{2}, \frac{1}{2}\right)$, the expected value changes to $v^e(\mu) = \frac{1}{2}\left[g(s_2) + g(s_3)\right](\mu(s))^{-1}$. In case the fixed entry costs are bounded by

$$\frac{1}{2}\left[g(s_2) + g(s_3)\right] - g(s_2) < c_e < \frac{1}{2}\left[g(s_2) + g(s_3)\right] - g(s_1),$$

a stationary equilibrium is given by the capacity value $y^* = \frac{1}{2}\left[g(s_2)+g(s_3)\right] - c_e$, the exit-point $x^* = s_3$, the entry-mass $M^* = \frac{2}{3}$, and the industry structure $\mu^* = \left(0, \frac{2}{3}, \frac{1}{3}\right)$.

The dynamic equilibrium starting from the distribution $\mu^0 = (0, \alpha, 1-\alpha)$ is characterised by the entry-mass $M_{t-1}^* = \alpha\left(-\frac{1}{2}\right)^t + \sum_{j=0}^{t-1}(-1)^j\left(\frac{1}{2}\right)^j$ and the exit-point $x_{t-1}^* = s_3$. This yields the industry structure

$$\mu^t = \left(0, \alpha\left(-\frac{1}{2}\right)^t + \sum_{j=0}^{t-1}(-1)^j\left(\frac{1}{2}\right)^j, -\alpha\left(-\frac{1}{2}\right)^t - \sum_{j=1}^{t-1}(-1)^j\left(\frac{1}{2}\right)^j\right)$$

in period t. As time tends to infinity, this vector converges to the stationary distribution $\mu^* = \left(0, \frac{2}{3}, \frac{1}{3}\right)$, and the industry adjusts to the steady state.

3.6 Altering entry distributions

In our analysis we have presumed that the productivity distribution of entering firms stays constant over time. This has been a necessary condition for the existence of a stationary equilibrium. However, the entry distribution ν may randomly alter at some point in time due to technological progress. This would change the setup of the model in such a way that the expected value of new firms needs to be modified, and the original steady state can no longer be sustained. If the alternate distribution is still subject to the convergence criteria though, the industry will adjust to a new stationary equilibrium.

It is easy to see that the emerging stationary equilibrium is continuous with respect to the underlying entry distribution ν. That is minor changes in the entry distribution have only little impact on the steady state. Unfortunately, monotonicity cannot be guaranteed in this context. This means, we cannot infer from $\nu_1 \overset{FOSD}{\geq} \nu_2$ that the same applies to the invariant distributions λ_1 and λ_2 arising in the respective stationary equilibria.

The aforementioned stochastic dominance of ν_1 versus the entry distribution ν_2 goes along with a higher willingness of new firms to enter the industry. In general, this should also facilitate a higher capacity price in the stationary equilibrium although exceptions to this finding do exist. If the consumers' demand is very elastic to aggregate output changes, the function h has a steep downward slope, and the emerging capacity price under the distribution ν_1 may be smaller than under ν_2.

It is also possible that the entry distribution constantly changes over time and forms a certain sequence $\{\nu_t\}_{t\in\mathbb{N}}$. We will show that the industry still tends to a stationary equilibrium as long as the predetermined sequence $\{\nu_t\}_{t\in\mathbb{N}}$ is monotone and converging. Before we prove this statement, we have to make some slight modifications regarding the model though. Notice, first of all, that an altering entry distribution changes the industry dynamics significantly. If the sequence $\{G_t\}_{t\in\mathbb{N}}$ represents the corresponding cumulative distribution functions, the industry structure is given by

$$\mu_{t+1}([a,\varphi')) = \int_{\varphi\geq x_t} F(\varphi'|\varphi)\, d\mu_t(\varphi) + M_t\, G_t(\varphi'). \tag{3.74}$$

Hence, the operator T_x, which has been defined for a fixed, predetermined entry distribution in (3.16), is no longer capable of reflecting the industry development. We need to introduce a new operator $T_{\nu,x} : \mathcal{M}^1(S,\mathcal{B}) \to \mathcal{M}^1(S,\mathcal{B})$ that does not just account for different exit-values but also different entry distributions by

$$T_{\nu,x}\, \mu(A) = \int_{\varphi\geq x} P(\varphi, A)\, d\mu(\varphi) + \mu([a,x))\, \nu(A). \tag{3.75}$$

If $M_t = \mu_t([a, x_t))$, the industry dynamics can be rephrased by $\mu_{t+1} = T_{\nu_t, x_t}\, \mu_t$. But, the contraction inequality in equation (3.37) is not valid anymore. Since we deal with different entry distributions, we must adjust this inequality as follows:

$$\begin{aligned}
&\|T_{\nu_1,x_1}\,\mu_1 - T_{\nu_2,x_2}\,\mu_2\|_{TV} \\
&= \|T_{\nu_1,x_1}\,\mu_1 - T_{\nu_2,x_1}\,\mu_1 + T_{\nu_2,x_1}\,\mu_1 - T_{\nu_2,x_1}\,\mu_2 + T_{\nu_2,x_1}\,\mu_2 - T_{\nu_2,x_2}\,\mu_2\|_{TV} \\
&\le \|T_{\nu_2,x_1}\,\mu_1 - T_{\nu_2,x_1}\,\mu_2\|_{TV} + \|T_{\nu_2,x_1}\,\mu_2 - T_{\nu_2,x_2}\,\mu_2\|_{TV} + \|T_{\nu_1,x_1}\,\mu_1 - T_{\nu_2,x_1}\,\mu_1\|_{TV} \\
&\le (1-\varepsilon)\|\mu_1 - \mu_2\|_{TV} + 2\,\mu_2([x_1, x_2]) + \|\nu_1 - \nu_2\|_{TV}.
\end{aligned} \tag{3.76}$$

The last term indicates that the difference between two consecutive industry structures particularly depends on the difference between the two entry distributions ν_1 and ν_2. Consequently, a converging dynamic equilibrium can only exist if the difference between those entry distributions tends to zero in the long-run. We will point out in following that convergence to a steady state can indeed be sustained if the sequence $\{\nu_t\}_{t\in\mathbb{N}}$ is stochastically increasing (i.e. $\nu_{t+1} \overset{FOSD}{\geq} \nu_t$ for all t) and the limiting distribution ν satisfies the necessary criteria from the previous sections. For this purpose, we initially act on the assumption that such a converging dynamic equilibrium exists. Based on this assumption, we start out to derive some key properties that help us prove the existence in a final step.

Notice first of all that switching entry distributions do not only control the evolution of industry structure but also the expected value of new firms. This means, the expectation in (3.14) is taken with respect to the different cdfs G_t now, and the expected value $v_t^e(\bar{z})$ varies with the time t even if a constant output/capacity sequence $\bar{z} = \{Q, y\}$ is given. As the expected value reflects the firms' willingness to pay for production capacity, the emerging capacity values will be affected, too. Hence, the function $\kappa(Q)$, which determines the capacity value for a constant output sequence and has been specified as solution to

$$\int_a^b v(\varphi, Q, \kappa(Q))\, dG(\varphi) - \kappa(Q) = c_e, \tag{3.77}$$

needs to be modified. Depending on the different entry distributions G_t and the constant aggregate output level Q, we define the functions $\kappa_t(Q)$ as solution to

$$\int_a^b v(\varphi, Q, \kappa_t(Q))\, dG_t(\varphi) - \kappa_t(Q) = c_e. \tag{3.78}$$

In principle, the functions κ_t possess the same qualities as the ones we derived for κ in Lemma 3.3. Since the distributions $\{\nu_t\}$ are stochastically increasing, the functions κ_t form an increasing sequence that converges to the function κ. According to Dini's Theorem, the convergence is even uniform on the compact interval $[Q_{\min}, Q_{\max}]$. This implies particularly that

$$\sup_{Q\in[Q_{\min},Q_{\max}]} |\kappa_t(Q) - \kappa(Q)| = \|\kappa_t - \kappa\|_{\sup} \to 0 \tag{3.79}$$

for $t \to \infty$.

Just like the expected value of entrants, the exit-points will also be affected by the altering entry distributions ν_t. Hence, we must adjust the function $s_\mu(\varphi)$ and replace the distribution G by G_t in equation (3.45).[14] The resulting function is denoted by $s^t_\mu(\varphi)$. Accordingly, the functions s_L, s_U, which we introduced in dependency of G and the stationary distribution λ, can no longer be used to specify lower and upper boundaries to the exit solution $x^t_\mu = \inf\left\{\varphi \in S : s^t_\mu(\varphi) \geq 0\right\}$. Instead, we define the constant sequences

$$\begin{aligned}
\bar{z}^t_{\min}(\eta) &= \left\{Q^\lambda_{\min}(\eta)\ ,\ y^\lambda_{\max}(\eta) + \|\kappa_t - \kappa\|_{\sup}\right\}_{j\in\mathbb{N}} \\
\bar{z}^t_{\max}(\eta) &= \left\{Q^\lambda_{\max}(\eta)\ ,\ y^\lambda_{\min}(\eta) - \|\kappa_t - \kappa\|_{\sup}\right\}_{j\in\mathbb{N}}
\end{aligned}$$

and based on this the functions

$$\begin{aligned}
s^t_L(\eta,\varphi) :=& \int_a^b g(\varphi')\, dF(\varphi'|\varphi) + \frac{1}{h\left(Q^\lambda_{\max}(\eta)\right)} \left(\beta \int_a^b u\left(\varphi', \bar{z}^t_{\min}(\eta)\right) dF(\varphi'|\varphi) + c_e\right) \\
&- \int_a^b g(\varphi')\, dG_t(\varphi') - \frac{\beta}{h\left(Q^\lambda_{\min}(\eta)\right)} \int_a^b u\left(\varphi', \bar{z}^t_{\max}(\eta)\right) dG_t(\varphi')
\end{aligned} \tag{3.80}$$

and

$$\begin{aligned}
s^t_U(\eta,\varphi) :=& \int_a^b g(\varphi')\, dF(\varphi'|\varphi) + \frac{1}{h\left(Q^\lambda_{\min}(\eta)\right)} \left(\beta \int_a^b u\left(\varphi', \bar{z}^t_{\max}(\eta)\right) dF(\varphi'|\varphi) + c_e\right) \\
&- \int_a^b g(\varphi')\, dG(\varphi') - \frac{\beta}{h\left(Q^\lambda_{\max}(\eta)\right)} \int_a^b u\left(\varphi', \bar{z}^t_{\min}(\eta)\right) dG(\varphi').
\end{aligned} \tag{3.81}$$

Both s^t_L and s^t_U depend on the distance $\eta = \|\mu - \lambda\|_{TV}$, which reflects the difference between the industry structure μ at time t and the stationary distribution λ referring to the steady state under the entry distribution G.

[14] Irrespective of the timing t, the distribution μ always entails the same equilibrium sequence if the entry distribution G stays constant over time. This is not the case for altering entry distributions though. The dynamic equilibrium differs whether μ occurs at time t or t'.

If the output/capacity sequences entailing μ in a dynamic equilibrium are subject to

$$\begin{aligned} \{Q_j^*\}_{j=t}^{\infty} &\subset \left[Q_{\min}^{\lambda}(\eta)\ ,\ Q_{\max}^{\lambda}(\eta)\right] \\ \text{and}\quad \{y_j^*\}_{j=t}^{\infty} &\subset \left[y_{\min}^{\lambda}(\eta) - \|\kappa_t - \kappa\|_{\sup}\ ,\ y_{\max}^{\lambda}(\eta) + \|\kappa_t - \kappa\|_{\sup}\right], \end{aligned} \tag{3.82}$$

we can infer from h's and u's monotonicity that

$$s_L^t(\eta, \varphi) \geq s_\mu^t(\varphi) \geq s_U^t(\eta, \varphi) \tag{3.83}$$

holds for all $\varphi \in S$. But, this implies immediately that the exit-point x_μ^t is bounded by

$$x_L^t(\eta) = \inf\Big\{\varphi \in S : s_L^t(\eta, \varphi) \geq 0\Big\} \tag{3.84}$$

$$\text{and}\quad x_U^t(\eta) = \inf\Big\{\varphi \in S : s_U^t(\eta, \varphi) \geq 0\Big\}, \tag{3.85}$$

i.e. the intersection of s_L^t and s_U^t with the origin. From the convergence $\nu_t \to \nu$ we can further deduce:

Lemma 3.11. *Let $\{\nu_t\}_{t\in\mathbb{N}}$ be a stochastically increasing sequence of probability measures (i.e. $\nu_{t+1} \overset{FOSD}{\geq} \nu_t$ for all t) and $\|\nu_t - \nu\|_{TV} \to 0$ for $t \to \infty$. For any $\delta > 0$ there exists a $t_0 \in \mathbb{N}$ such that $\|x_L^t - x_L\|_{\sup} \leq \delta$ for all $t \geq t_0$. The same applies to the functions x_U^t and x_U.*

Proof. Recall that G_t and G mark the cdfs, which belong to the probability measures ν_t and ν. Because g, u are bounded, continuous, and increasing functions, the presumed strong convergence $\|\nu_t - \nu\|_{TV} \to 0$ involves also pointwise convergence of the functions $s_L^t(\eta, \varphi) \to s_U(\eta, \varphi)$. In fact, the convergence is even uniform. This follows immediately from Dini's Theorem since $\{s_L^t(\eta, \varphi)\}_{t\in\mathbb{N}}$ constitutes a monotone sequence and (η, φ) is an element of a compact set.

The uniform convergence means that $\|s_L^t - s_L\|_{\sup} \leq \varepsilon$ for an arbitrary $\varepsilon > 0$ and sufficiently large t. This is equivalent to

$$s_L(\eta, \varphi) - \varepsilon \leq s_L^t(\eta, \varphi) \leq s_L(\eta, \varphi) + \varepsilon \tag{3.86}$$

for all (η, φ). We define the function on the left side as $\hat{s}_L(\eta, \varphi, \varepsilon) := s_L(\eta, \varphi) - \varepsilon$, and we denote the intersection with the origin by $\hat{x}_L(\eta, \varepsilon) = \inf\{\varphi \in S : \hat{s}_L(\eta, \varphi, \varepsilon) \geq 0\}$. Since $\hat{s}_L$ is continuous and strictly increasing with respect to φ, the function $\hat{x}_L$ must be continuous as well. This implies $\hat{x}_L(\eta, \varepsilon) \to \hat{x}_L(\eta, 0) = x_L(\eta)$ for $\varepsilon \to 0$.

If $\{\varepsilon_m\}_{m\in\mathbb{N}}$ is a decreasing sequence that converges to zero, then $\{\hat{x}_L(\eta, \varepsilon_m)\}_{m\in\mathbb{N}}$ is also a monotone sequence. For any arbitrary value $\delta > 0$, we can thus refer to Dini's Theorem to find a $m_0 \in \mathbb{N}$ such that $|\hat{x}_L(\eta, \varepsilon_m) - x_L(\eta)| < \delta$ holds for all $\eta \in [0, 2]$ and $m \geq m_0$. In the same way, we can define the function $\check{s}_L(\eta, \varphi, \varepsilon) := s_L(\eta, \varphi) + \varepsilon$ as well as the intersection point $\check{x}_L(\eta, \varepsilon) = \inf\{\varphi \in S : \check{s}_L(\eta, \varphi, \varepsilon) \geq 0\}$, and find a $m_1 \in \mathbb{N}$ such that $|\check{x}_L(\eta, \varepsilon_m) - x_L(\eta)| < \delta$ holds for all $\eta \in [0, 2]$ and $m \geq m_1$.

Let $t_0 \in \mathbb{N}$ be chosen such that $\|s_L^{t_0} - s_L\|_{\sup} \leq \min\{\varepsilon_{m_0}, \varepsilon_{m_1}\}$. According to the inequality (3.86), we must have

$$\check{x}_L(\eta, \varepsilon) \leq x_L^t(\eta) \leq \hat{x}_L(\eta, \varepsilon). \tag{3.87}$$

This implies

$$x_L(\eta) - \delta \leq x_L^t(\eta) \leq x_L(\eta) + \delta \tag{3.88}$$

for all $t \geq t_0$, and the Lemma is proven. □

Recall that x_λ characterises the exit-point, which belongs to the stationary equilibrium under the entry distribution ν. The industry converges to this steady state even in the case of altering entry distributions if the expression in (3.76) represents a contraction. This applies, when a value $\bar{\theta} \in [\theta, 1)$ exists such that

$$\begin{aligned} \|\mu_{t+1} - \lambda\|_{TV} &\leq (1 - \varepsilon)\|\mu_t - \lambda\|_{TV} + 2\,\lambda([x_\mu^t, x_\lambda]) + \|\nu_t - \nu\|_{TV} \\ &\leq \bar{\theta}\,\|\mu_t - \lambda\|_{TV}. \end{aligned} \tag{3.89}$$

Whether the latter inequality holds or not depends particularly on the difference between the entry distributions ν_t, ν and the distance between the exit-points x_μ^t, x_λ. While the term $\|\nu_t - \nu\|_{TV} \to 0$ tends to zero by assumption, the term $\lambda([x_\mu^t, x_\lambda])$ can be bounded by

$$\begin{aligned} \lambda([x_\mu^t, x_\lambda]) &\leq m_\lambda \cdot |x_\mu^t - x_\lambda| \\ &\leq m_\lambda \cdot \max\{|x_L^t(\eta) - x_\lambda|, |x_U^t(\eta) - x_\lambda|\} \end{aligned} \tag{3.90}$$

if the output/capacity sequences entailing the distribution μ_t in an equilibrium satisfy condition (3.82). This brings us directly to the next result, which is essentially an analogy to Theorem 3.9.

Lemma 3.12. *Let $\{\nu_t\}_{t\in\mathbb{N}}$ be a stochastically increasing sequence of probability measures (i.e. $\nu_{t+1} \overset{FOSD}{\geq} \nu_t$ for all t) and $\|\nu_t - \nu\|_{TV} \to 0$ for $t \to \infty$. Furthermore,* *let β and c_e be subject to the convergence criteria in* (3.61)*. Based on the starting distribution μ_0 and the distances $\eta_t := \|\mu_t - \lambda\|_{TV}$, we calculate the distributions $\mu_{t+1} = T_{\nu_t, x_L^t(\eta_t)}\mu_t$. For any $\bar{\eta} > 0$, we can then find a $t_0 \in \mathbb{N}$ such that $\|\mu_t - \lambda\|_{TV} \leq \bar{\eta}$ for all $t \geq t_0$.*

Proof. The Lemma is proven by induction. In the first step, we specify a $t_0 \in \mathbb{N}$ such that $\|\mu_{t_0} - \lambda\|_{TV} \leq \bar{\eta}$. Afterwards, we show that $\|\mu_t - \lambda\|_{TV} \leq \bar{\eta}$ implies $\|\mu_{t+1} - \lambda\|_{TV} \leq \bar{\eta}$ for any $t \geq t_0$.

Step I. For any $\delta > 0$ there is a number $t_0 \in \mathbb{N}$ such that both $\|\nu_t - \nu\|_{TV} \leq \delta$ and $|x_L^t(\eta) - x_L(\eta)| \leq \delta$ hold for all $\eta > 0$ and $t \geq t_0$. This follows directly from Lemma 3.11, and we can further conclude that

$$\begin{aligned}\lambda([x_L^t(\eta), x_\lambda]) &\leq m_\lambda \cdot \left|x_L^t(\eta) - x_\lambda\right| \\ &= m_\lambda \cdot \left|x_L^t(\eta) - x_L(\eta) + x_L(\eta) - x_L(0)\right| \\ &\leq m_\lambda \cdot \big(\, |x_L(\eta) - x_L(0)| + \delta \,\big).\end{aligned}$$

The contraction condition in (3.89) is thus satisfied if

$$\begin{aligned} & (1-\varepsilon)\,\eta + 2m_\lambda\big(\,|x_L(\eta) - x_L(0)| + \delta\,\big) + \delta \;\leq\; \bar{\theta}\,\eta \\ \Leftrightarrow\quad & (1-\varepsilon) + 2m_\lambda \left|\frac{x_L(\eta) - x_L(0)}{\eta}\right| + \frac{(2m_\lambda + 1)\delta}{\eta} \;\leq\; \bar{\theta}.\end{aligned}$$

As β and c_e are subject to the convergence criteria in (3.61), we have

$$\left|\frac{x_L(\eta) - x_L(0)}{\eta}\right| \leq \frac{\theta - (1-\varepsilon)}{2m_\lambda}.$$

If we apply this property to the contraction condition above, we get

$$\begin{aligned} & \theta + \frac{(2m_\lambda + 1)\delta}{\eta} \;\leq\; \bar{\theta} \\ \Leftrightarrow\quad & \frac{(2m_\lambda + 1)\delta}{\bar{\theta} - \theta} \;\leq\; \eta \\ \Leftrightarrow\quad & \delta \;\leq\; \frac{(\bar{\theta} - \theta)\eta}{2m_\lambda + 1}.\end{aligned}$$

Let t_0 be chosen such that δ is small enough to satisfy the inequality above for the presumed distance $\bar{\eta}$. The contraction inequality in (3.89) thus holds for all $t \geq t_0$ and distributions μ_t that are subject to $\|\mu_t - \lambda\|_{TV} = \eta_t \geq \bar{\eta} \geq \frac{(2m_\lambda+1)\delta}{\bar{\theta}-\theta}$. As the maximum possible difference between two probability distributions is always bounded by $\|\mu - \lambda\|_{TV} \leq 2$, we can find a $t_1 \in \mathbb{N}$ such that

$$\bar{\theta}^{t_1} \|\mu_{t_0} - \lambda\|_{TV} \leq 2\, \bar{\theta}^{t_1} \leq \bar{\eta}. \tag{3.91}$$

But, this implies $\|\mu_t - \lambda\|_{TV} \leq \bar{\eta}$ for some $t_0 \leq t \leq t_0 + t_1$.

Step II. Now, we show that the previous result gives rise to $\|\mu_{t+1} - \lambda\|_{TV} \leq \bar{\eta}$. Because $t \geq t_0$, we have

$$\begin{aligned}
\|\mu_{t+1} - \lambda\|_{TV} &\leq (1-\varepsilon)\|\mu_t - \lambda\|_{TV} + 2\lambda([x_L^t(\eta_t), x_\lambda]) + \|\nu_t - \nu\|_{TV} \\
&\leq (1-\varepsilon)\,\eta_t + 2m_\lambda \eta_t \left(\frac{\theta - (1-\varepsilon)}{2m_\lambda} + \frac{\delta}{\eta_t} \right) + \delta \\
&= \theta\eta_t + 2m_\lambda\delta + \delta \\
&\leq \theta\bar{\eta} + (\bar{\theta} - \theta)\bar{\eta} \\
&= \bar{\theta}\bar{\eta} \\
&\leq \bar{\eta}.
\end{aligned}$$

This concludes the proof. □

It is evident that the statement of Lemma 3.12 can also be deduced for the sequence $\mu_{t+1} = T_{\nu_t, x_U^t(\eta_t)}\mu_t$. Based on this, we can now show the existence of a converging dynamic equilibrium for altering entry distributions. The procedure resembles the one from Section 3.5. That means we construct a mapping $\tau : \ell^\infty \times \ell^\infty \to \ell^\infty \times \ell^\infty$ such that a fixed point of this mapping coincides with the dynamic equilibrium. The existence of a fixed point is proven by means of Schauder's Fixed Point Theorem once again.

Theorem 3.13 (Existence of a converging dynamic equilibrium for altering entry distributions)**.** *Let $\{\nu_t\}_{t\in\mathbb{N}}$ be a stochastically increasing sequence of continuous probability measures (i.e. $\nu_{t+1} \overset{FOSD}{\geq} \nu_t$ for all t) and $\|\nu_t - \nu\|_{TV} \to 0$ for $t \to \infty$. Furthermore, let β and c_e be subject to the convergence criteria in* (3.61)*. For any continuous starting distribution $\mu_0 \in \mathcal{M}^1(S, \mathcal{B}(S))$ there exists a dynamic equilibrium that converges to the uniquely determined steady state in the long-run.*

Proof. We begin to construct the mapping τ on the space of bounded sequences. For any sequence $\{x_t, y_t\}_{t\in\mathbb{N}_0}$ of exit-points and capacity values we derive the probability distributions $\mu_{t+1} = T_{\nu_t, x_t}\mu_t$. The resulting aggregate industry output $\{Q_t\}_{t\in\mathbb{N}_0}$ and the assumed sequence $\{y_t\}_{t\in\mathbb{N}_0}$ allow us to calculate new capacity values by

$$\widetilde{y}_t = \int_a^b v(\varphi, z_{t+1})\, dG_t(\varphi) - c_e. \tag{3.92}$$

Here, $z_t = \{Q_j, y_j\}_{j=t}^{\infty}$ denotes the output/capacity sequence starting at time t. A new exit-point is defined as

$$\widetilde{x}_t := \inf\left\{\varphi \in S : \int_a^b v(\varphi', z_{t+1})\, dF(\varphi'|\varphi) \geq \widetilde{y}_t\right\}. \tag{3.93}$$

It is clear that the mapping $\tau : \{x_t, y_t\}_{t\in\mathbb{N}_0} \to \{\widetilde{x}_t, \widetilde{y}_t\}_{t\in\mathbb{N}_0}$ is continuous, and each fixed point represents a dynamic equilibrium under altering entry distributions. To apply Schauder's Fixed Point Theorem, which guarantees the existence of a fixed point in this case, we must specify a compact, convex set $N = N_1 \times N_2 \subset \ell^\infty \times \ell^\infty$ that is mapped into itself by τ.

A compact, convex set N_1 for the exit-sequences can be constructed by means of the functions x_L^t and x_U^t. Let $\{\eta_m\}_{m\in\mathbb{N}_0}$ be a positive, decreasing sequence that converges to zero if m tends to infinity. According to Lemma 3.11, we can find a $t_0 \in \mathbb{N}$ such that both

$$\begin{aligned} \|\nu_t - \nu\|_{TV} &\leq \frac{(\bar{\theta}-\theta)\eta_0}{2m_\lambda + 1} \quad \text{and} \\ \max\left\{\|x_L^t - x_L\|_{\sup}, \|x_U^t - x_U\|_{\sup}\right\} &\leq \frac{(\bar{\theta}-\theta)\eta_0}{2m_\lambda + 1} \end{aligned} \tag{3.94}$$

hold for all $t \geq t_0$. As the proof of Lemma 3.12 illustrates, the contraction condition is met in this situation, and we can determine a $\bar{t}_0 \in \mathbb{N}_0$ that is subject to $2\,\bar{\theta}^{\bar{t}_0} \leq \eta_0$. The procedure is repeated for η_1. This means, we first look for a number $t_1 \in \mathbb{N}_0$ that satisfies the inequalities in (3.94) for η_1. Afterwards, we calculate the smallest value $\bar{t}_1 \in \mathbb{N}_0$ that implies $\eta_0\,\bar{\theta}^{\bar{t}_1} \leq \eta_1$. If the described procedure is carried out for every η_m, we get the maximum possible distance values

$$\bar{\eta}_t = \begin{cases} 2 & \text{if} \quad 0 \leq t \leq t_0 \\ 2\,\bar{\theta}^{t-t_0} & \text{if} \quad t_0 < t < t_0 + \bar{t}_0 \\ \eta_m & \text{if} \quad t_m + \bar{t}_m \leq t < t_{m+1} \\ \eta_m\,\bar{\theta}^{t-t_{m+1}} & \text{if} \quad t_{m+1} \leq t < t_{m+1} + \bar{t}_{m+1}. \end{cases}$$

Based on this sequence $\{\bar{\eta}_t\}$, we can define lower boundaries for the set N_1 by

$$a_t = \begin{cases} a & \text{if} \quad 0 \le t \le t_0 \\ x_L^t(\bar{\eta}_t) & \text{if} \quad t_0 < t \end{cases}$$

and upper boundaries by

$$b_t = \begin{cases} b & \text{if} \quad 0 \le t \le t_0 \\ x_U^t(\bar{\eta}_t) & \text{if} \quad t_0 < t. \end{cases}$$

Since η_m converges to zero by assumption, we can conclude that $\bar{\eta}_t \to 0$ for $t \to \infty$. Hence, both $a_t \to x_\lambda$ and $b_t \to x_\lambda$ for $t \to \infty$. A compact, convex set containing the possible exit sequences is thus given by

$$N_1 := \Big\{ \{x_t\}_{t \in \mathbb{N}_0} \in \ell^\infty : \, a_t \le x_t \le b_t, \, \forall \, t \in \mathbb{N}_0 \Big\}.$$

Lower and upper boundaries for the capacity sequences are

$$c_t = \begin{cases} \kappa(Q_{\max}) - \|\kappa_t - \kappa\|_{\sup} & \text{if} \quad 0 \le t \le t_0 \\ \kappa(Q^\lambda_{\max}(\bar{\eta}_t)) - \|\kappa_t - \kappa\|_{\sup} & \text{if} \quad t_0 < t \end{cases}$$

and

$$d_t = \begin{cases} \kappa(Q_{\min}) + \|\kappa_t - \kappa\|_{\sup} & \text{if} \quad 0 \le t \le t_0 \\ \kappa(Q^\lambda_{\min}(\bar{\eta}_t)) + \|\kappa_t - \kappa\|_{\sup} & \text{if} \quad t_0 < t \end{cases}$$

such that a compact, convex set for the capacity sequences can be defined as

$$N_2 := \Big\{ \{y_t\}_{t \in \mathbb{N}_0} \in \ell^\infty : \, c_t \le y_t \le d_t, \, \forall \, t \in \mathbb{N}_0 \Big\}. \tag{3.95}$$

Because of the switching entry distributions, the sets N_1 and N_2 defined here are larger than the ones in Section 3.5. The shape of both sets is illustrated in Figure 3.5.

The final step is to show that τ maps the compact, convex set $N = N_1 \times N_2$ into itself. To this end, let the output/capacity sequence $\{x_t, y_t\}_{t \in \mathbb{N}_0}$ be an element of N. As we can infer from Lemma 3.12, the exit sequence $\{x_t\}_{t \in \mathbb{N}_0}$ provokes an evolution of industry structure $\{\mu_t\}$ that is subject to $\|\mu_t - \lambda\|_{TV} \le \bar{\eta}_s$ for all $t \ge s$. The resulting aggregate industry output is thus bounded by $Q^\lambda_{\min}(\bar{\eta}_t) \le Q_t \le Q^\lambda_{\max}(\bar{\eta}_t)$.

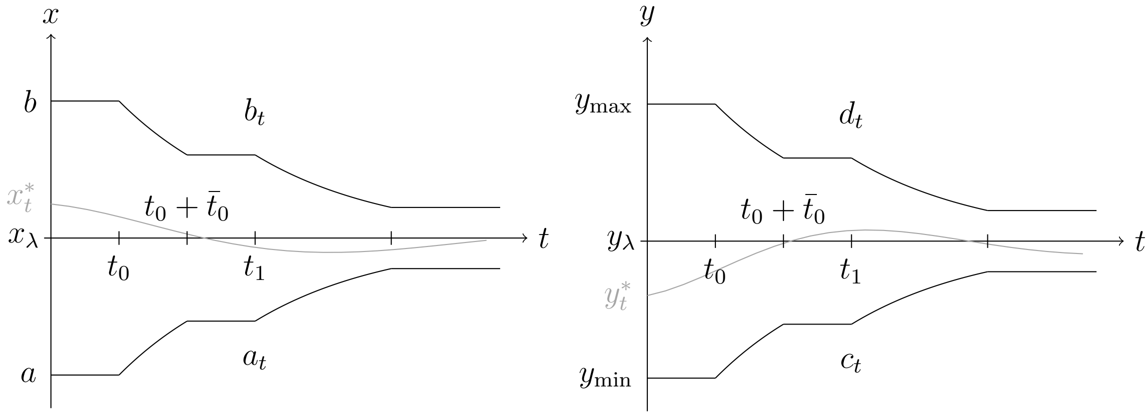

Figure 3.5: Illustration of the sets N_1 (left side), N_2 (right side), and a possible fixed point solution of the mapping τ.

If $z_t = \{Q_j, y_j\}_{j=t}^{\infty}$ marks the output/capacity sequence starting at time t, we have

$$\begin{aligned} c_t &= \kappa(Q^{\lambda}_{\max}(\bar{\eta}_t)) - \|\kappa_t - \kappa\|_{\sup} \\ &\leq \kappa_t(Q^{\lambda}_{\max}(\bar{\eta}_t)) \\ &= \int_a^b v(\varphi, Q^{\lambda}_{\max}(\bar{\eta}_t), \kappa_t(Q^{\lambda}_{\max}(\bar{\eta}_t)))\, dG_t(\varphi) - c_e \\ &\leq \int_a^b v(\varphi, z_{t+1})\, dG_t(\varphi) - c_e \\ &= \widetilde{y}_t \\ &\leq \int_a^b v(\varphi, Q^{\lambda}_{\min}(\bar{\eta}_t), \kappa_t(Q^{\lambda}_{\min}(\bar{\eta}_t)))\, dG_t(\varphi) - c_e \\ &= \kappa_t(Q^{\lambda}_{\min}(\bar{\eta}_t)) \leq d_t. \end{aligned}$$

Consequently, the sequence $\{\widetilde{y}_t\}$ must be an element of N_2. The boundedness of the sequence z_t implies moreover that

$$s^t_L(\bar{\eta}_t, \varphi) \geq v^c(\varphi, z_{t+1}) - \widetilde{y}_t \geq s^t_U(\bar{\eta}_t, \varphi) \tag{3.96}$$

holds for all $\varphi \in S$. But, this leads to

$$\begin{aligned} a_t &= x^t_L(\bar{\eta}_t) \\ &= \inf\left\{\varphi \in S : s^t_L(\bar{\eta}_t, \varphi) \geq 0\right\} \\ &\leq \inf\left\{\varphi \in S : v^c(\varphi, z_{t+1}) - \widetilde{y}_t \geq 0\right\} \\ &= \widetilde{x}_t \\ &\leq \inf\left\{\varphi \in S : s^t_U(\bar{\eta}_t, \varphi) \geq 0\right\} \end{aligned}$$

$$= x_U^t(\bar{\eta}_t) = b_t,$$

and we conclude that $\{\widetilde{x}_t\} \in N_1$. Hence, $\tau(N) \subseteq N$, and the existence of a fixed point follows by Schauder's Fixed Point Theorem. $\square$

In the analysis above, we focused particularly on an increasing sequence $\{\nu_t\}$. But, the results can analogously be deduced for a stochastically decreasing sequence of entry distributions as well.

3.7 Conclusive remarks

In this chapter we have investigated how limited sectoral production capacity may affect the industry dynamics. We have set up a dynamic stochastic framework, in which the available production capacity is distributed and traded among all active firms. If a firm decides to cease production, it releases a marginal capacity unit, which then becomes available to potential entrants. By construction, the emerging capacity price in period t clears the market in the sense that it equates the mass of exiting firms with the mass of entering ones. Hence, the limited capacity supply creates an interdependency between firm entry and exit, which is even more pronounced than in markets where capacity constraints are not an issue. We have illustrated that the direct interaction between entering and exiting firms makes the industry converge to a steady state in the long-run. Moreover, the convergence to a stationary equilibrium can even be sustained for altering entry distributions. Although the adjustment to a steady state seems to be a trivial outcome, it has not formally been proven yet, and does not necessarily apply to every competitive industry as Example 3.1 has shown.

We have identified the fixed entry costs c_e and the discount factor β as two significant parameters affecting the speed of convergence. The speed, at which the industry tends to a stationary equilibrium, increases when either one of those parameters decreases. If both parameters are sufficiently small, the stationary equilibrium is uniquely determined, and the industry will attain this steady state no matter what the current firm size distribution is. The adjustment path, however, depends notably on the starting distribution. This means that different starting distributions will evoke different parameter sequences in a dynamic equilibrium.

A crucial point in the proof has been the separability of period profits, i.e. the possibility to separate a firm-specific productivity term from the aggregate industry output and, thus, from the overall firm size distribution. In our framework, we took a multiplicative structure of the period profits $\pi(\varphi, Q) = g(\varphi)\,h(Q)$ for granted. But, the proof works also for other types of separability, as for instance the additive type $\pi(\varphi, Q) = g(\varphi) + h(Q)$, in a quite similar manner. If no kind of separability is given though, the convergence proof becomes much more complicated.

A common feature of models dealing with firm entry and exit is that incumbents leave the industry whenever their outside option is superior to keeping up production. In our setup, an incumbent's outside option is basically dictated by the willingness of entrants to pay for production capacity. Since each incumbent is holding a marginal capacity unit, this creates a positive outside option for exiting firms. If the scarcity of production factors does not play a role, and firms have theoretically free access to the market, the outside option is often set equal to zero. In such a situation firm turnover is mainly driven by fixed production costs potentially leading to a negative production value for inefficient firms.

We have disregarded fixed production costs in this study. But, fixed production costs exist in many industries and can, to some extent, also be introduced in this framework. As long as the fixed costs are rather low, the constructed mechanism to trade capacity units between entering and exiting firms will still work, and the core findings regarding the convergence to a stationary equilibrium will stay untouched. Nevertheless, the resulting capacity costs $\{y_t^*\}_{t\in\mathbb{N}_0}$ will presumably be smaller than in the equilibrium without fixed costs. A problem might occur when the fixed costs are so high that no firm is willing to enter the industry. In this case the limited capacity does not constitute a scarce production factor anymore, and the capacity price will fall to zero. Moreover, less productive incumbents possibly have a negative expected value because of the high fixed costs. They are better of ceasing production then and will leave the industry anyway. As new firms have no incentive to enter on the other hand, this leads to a shrinkage of the industry in terms of the total mass $\mu_t(S)$, or rather the total amount of capacity units used. Hence, the trade of production capacity fails to work here. To exclude such a scenario, we decided to neglect fixed production costs in our analysis.

4. Impact of the milk quota

The dairy sector in the European Union (EU) is currently passing through a profound adjustment process. Milk prices declined about 25 percent between 2013 and 2015, and this drop in prices puts considerable economic pressure on dairy farms. In response to this development the EU has recently developed bailout plans for milk producers who are in jeopardy (European Commission, 2015). One factor that is cited to be responsible for the price decline is the abolition of the milk quota system, which was effective between 1984 and 2015. Hence, this chapter addresses the question of how production quotas affect the dynamics of an industry, as well as reactions to signals of abolishment in the nearer future.

A strength of the previous modelling has been that the capacity price is determined endogenously and really offsets the mass of entering with the mass of exiting firms. But, the introduced mechanism only works if the industry permanently operates at the sectoral capacity limit, and the demand for production capacity at a price zero exceeds the supply. We have already argued that this assumption is not necessarily fulfilled when firms have to pay fixed costs and the period profits become negative. In this case, net exit of firms is also possible, and the total size of the industry may fall below the capacity limit. Since operational losses and net exit of firms may very well occur in practice though, it is necessary to include such a scenario in an application to the dairy sector. Hence, we must modify the previous dynamic stochastic model using a different mechanism to reflect the limited sectoral capacity under a production quota.

To this end, we pick up the dynamic, stochastic framework of Hopenhayn (1992a) once again, and model a perfectly competitive but heterogeneous industry. Firms are differentiated by their productivity level, which defines the variable production costs and is subject to an AR(1)-process. As every firm can be considered an atomistic production unit in this framework, we must constrain both the total number of production units possibly operating under a quota regime and the maximum output of each unit. To constrain the number of production units the mass of entrants must not exceed the number of available production units. Therefore, the entry costs, as well as the liquidation value of exiting firms, must correlate with the industry structure. A measure that reflects the entry/exit differential and serves as a good proxy for the used number of production units is the total industry mass. While a high value indicates a high appraisal of production capacity, a low value stands for unused capacity units. The overall limitation of production capacity is thus modelled through a sharply increasing, continuous function that links both the entry costs and the liquidation value to the total industry mass. This way, we gain some flexibility and can account for fluctuations in the size of the sector. Nevertheless, it comes at the cost that we have to presume an explicit function and the capacity value is no longer determined by market clearance.

As a part of European Union the West German dairy industry was subject to a production quota for the last 30 years. This offers a perfect opportunity to analyse impacts of different quota schemes (tradable, non-tradable) and the fading out on industry dynamics. We thus apply our theoretic model to this industry and answer two major questions: First, how does the milk quota affect long-run industry dynamics? Second, how do farms react to the recent abolishment of the tradable milk quota? We analyse the long-run effects by comparing the stationary equilibria with and without quota. Surprisingly, we find that the firms' average productivity level under the tradable quota is higher than in the scenario without quota. We further compare the short-run adjustment paths of the industry to the announcement of quota termination, and find increasing output levels directly after the termination of the quota scheme with price effects. These effects, however, vanish in the course of time, and the industry develops towards the long-run equilibrium.

In addition to the analysis of steady-state properties, we consider a finite dynamic equilibrium. This allows us to keep track of short-run effects in greater detail, and to explore how the sector adjusts to the abolishment of the milk quota regime. Moreover, we illustrate how to calibrate a dynamic stochastic entry/exit-model using farm-level panel data, as well as how to calculate a finite equilibrium numerically. From an applied perspective we contribute to the current debate in the EU's agricultural policy. Our model results quantify the impact that abolishing the milk quota has on firm turnover and the farm size distribution in this sector. This assessment is important for deciding whether market interventions are necessary, and how they should be designed.

The remainder of this chapter is structured as follows.[1] In Section 4.1 we present the dynamic stochastic framework and introduce both the stationary and the finite dynamic equilibrium concept. Focus will be placed on incorporating limited sectoral capacity into the farms' value function. The existence of a stationary as well as a finite dynamic equilibrium is proven in Sections 4.3 and 4.4. The model is applied to the Western German dairy sector, and the respective calibration is presented in Section 4.5. After describing the development of the dairy industry over recent decades, we calibrate model parameters and calculate stationary as well as dynamic market equilibria. Based on the equilibrium outcome we will discuss the effect of a milk quota on farms' entry and exit decisions in Section 4.5.4.

4.1 The formal model

The basic setup of our model draws closely upon the seminal papers of Jovanovic (1982) and Hopenhayn (1992a), whose respective approaches explicitly allow us to model endogenous entry and exit of the firms, which is crucial for analysing industry dynamics under production quota constraints.

We consider a perfectly competitive industry with a continuum of firms producing a homogeneous good (milk). Each firm takes the output price as given and chooses its optimal production quantity. The output price will be determined by market clearance. The inverse demand function $D(Q) > 0$ should be continuously differentiable and strictly decreasing. We assume that $\lim_{Q \to +\infty} D(Q) = 0$, and competition takes place in discrete time $t \in \mathbb{N}_0$.

All firms have the same production technology but they differ with respect to their productivity level. That is, we account for firm-specific productivity differences, which may be a result of farm size, capital stock, feed and livestock management or natural conditions. The firm-specific productivity is supposed to be the only source of uncertainty faced by the firms.

We model the firm's individual productivity as a stochastic parameter $\varphi_t \in \mathbb{R}$, which follows the stationary AR(1)-process

$$\varphi_{t+1} = \rho \varphi_t + \varepsilon_{t+1}, \quad \rho \in (0,1) \text{ and } \varepsilon_{t+1} \overset{\text{iid}}{\sim} N(\xi_\varepsilon, \sigma_\varepsilon^2), \tag{4.1}$$

and is assumed to be the same for all incumbents. The realisation of the error term ε_{t+1} is independent across firms and over time. The process as given in (4.1) inherits

[1] Parts of Chapter 4 are taken from Kersting et al. (2016).

the Markov property and is time-homogeneous. Under the hypothesis $\varphi_t = \varphi$, it follows that $\varphi_{t+1} \sim N(\rho\varphi + \xi_\varepsilon, \sigma_\varepsilon^2)$. The density of this normal distribution is denoted by

$$p(\varphi, z) := \frac{1}{\sqrt{2\pi\sigma_\varepsilon^2}} \exp\left(-\frac{(z-(\rho\varphi+\xi_\varepsilon))^2}{2\sigma_\varepsilon^2}\right), \tag{4.2}$$

and the conditional cdf $F(\varphi'|\varphi) = \text{Prob}(\varphi_{t+1} \le \varphi' | \varphi_t = \varphi)$ is given by

$$F(\varphi'|\varphi) = \int_{-\infty}^{\varphi'} p(\varphi, z)\, dz. \tag{4.3}$$

The function $F(\varphi'|\varphi)$ constitutes a probability kernel and is continuous with respect to both arguments. Moreover, it is strictly decreasing in φ if we keep φ' fixed.[2] That is, all active firms can be explicitly distinguished by their current productivity level φ_t. The distribution of these values across all firms thus expresses the state of the industry in period t, which is denoted by the measure $\mu_t : \mathcal{B}(\mathbb{R}) \to \mathbb{R}_+$ defined on the Borel sets of the real numbers.[3] Hence, any changes of the industry structure caused by the stochastic productivity process, as well as entry/exit of firms, translate into changes of μ_t.

We further proceed upon the assumption that firms with higher productivity levels are able to produce any amount of output q at lower costs. This property is represented by a twice continuously differentiable cost function $c(\varphi, q)$, which is monotonically decreasing in φ with the limits

$$\lim_{\varphi \to +\infty} c(\varphi, q) = 0 \quad \text{and} \quad \lim_{\varphi \to -\infty} c(\varphi, q) = \infty, \quad \forall q \ge 0. \tag{4.4}$$

Further, function $c : \mathbb{R}_+^0 \times \mathbb{R} \to \mathbb{R}_+^0$ should satisfy $c(\varphi, 0) = 0$, and

$$\frac{\partial c}{\partial q} > 0 \quad \text{with} \quad \frac{\partial c}{\partial q}(0, \varphi) = 0, \quad \frac{\partial^2 c}{\partial q^2} > 0, \quad \frac{\partial^2 c}{\partial \varphi \partial q} \le 0, \quad \lim_{\bar{q} \to +\infty} \frac{\partial c}{\partial q}(\bar{q}, \varphi) = \infty. \tag{4.5}$$

In each period t of the planning horizon, all active firms have to choose their own optimal production output. If no production quota exists, firms would take the output price $p_t \ge 0$ as well as their current productivity level φ_t as given and maximise

$$\max_{q_t \ge 0} \quad p_t\, q_t - c(\varphi_t, q_t). \tag{4.6}$$

[2]If $\varphi_1 < \varphi_2$, the distribution $F(\cdot|\varphi_2)$ stochastically dominates $F(\cdot|\varphi_1)$.

[3]μ_t does not need to be a probability measure. The total mass $\mu_t(\mathbb{R})$ may be smaller or larger than one, indicating the size of industry.

The first-order condition for a maximum in (4.6) implies that the optimal firm-specific output q_t^* satisfies

$$p_t \leq \frac{\partial c}{\partial q}(\varphi_t, q_t), \quad \text{with equality if } q_t > 0. \tag{4.7}$$

The imposed restrictions on the cost function guarantee that for all valid combinations of p_t and φ_t, a unique solution $q_t^* = q^*(\varphi_t, p_t)$ to (4.7) exists. The firm-specific optimal output possesses the following properties:

Lemma 4.1. *(i) The function $q^*(\varphi, p)$ is continuous and (strictly) monotonic increasing in p and φ. (ii) For all $\varphi \in \mathbb{R}$, we have $q^*(\varphi, p) > 0$ if $p > 0$ and $q^*(\varphi, 0) = 0$. (iii) $q(\varphi, p) \to \infty$ if either $p \to +\infty$ or $\varphi \to +\infty$.*

Proof. All statements follow immediately from the first order condition (4.7) and the assumed structure of the cost function in (4.4), (4.5). □

It is important to mention here that the firm-specific output is hypothetically unbounded in this setup. This means that a firm's output will tend to infinity whenever its productivity level tends to infinity. If the industry is subject to a production quota, however, such a scenario should not be possible. We account for this by an upper production limit $q^{\max} > 0$ that is set to the single firm's output in this case. This boundary is supposed to be an exogenously given parameter that holds for all firms. As a consequence, the optimal firm-specific output level under a production quota must be redefined as $q_t^* = \min\{q^*(\varphi_t, p_t), q^{\max}\}$.

The aggregate industry output $Q_t = Q^s(p_t, \mu_t)$ depends on the structure of the industry μ_t and is given by

$$Q^s(p_t, \mu_t) = \int_{\mathbb{R}} q_t^*(\varphi, p_t)\, d\mu_t(\varphi). \tag{4.8}$$

In case the integral on the right-hand side exists for any output price, we infer from q^*'s properties that $Q^s(p, \mu)$ is continuous and increasing with respect to p.

Production incurs a fixed cost $c_f > 0$, which is the same for all firms and has to be paid at the beginning of each period before a new productivity level is revealed to incumbents according to the Markov process (4.1). Hence, the fixed costs are sunk by the time firms get to know their new realisation of φ_t and choose the production output. A firm's profit per period is then

$$\pi(p_t, \varphi_t) := p_t\, q_t^* - c(\varphi_t, q_t^*) - c_f, \tag{4.9}$$

with q_t^* being the optimal firm-specific output level. The properties of $c(\varphi, q)$ and q^* imply that the function $\pi(\varphi, p)$ satisfies:

Lemma 4.2. *(i) π is continuous in φ and p. (ii) π is strictly increasing in p, and if $p > 0$, it is strictly increasing in φ. (iii) $\pi(\varphi, p) \to \infty$ if either $p \to +\infty$ or $\varphi \to +\infty$. (iv) $\pi(\varphi, p) \to -c_f$ if either $p \to 0$ or $\varphi \to -\infty$.*

Proof. The statements follow from Lemma 4.1 and the presumed cost function. □

At the end of each period firms have the option to leave or (re-)enter the market. As potential entrants must build up production capacity before producing any output, they need to pay entry costs k_t when entering the industry. Although these entry costs are supposed to be the same for all entrants, the respective amount will depend on the non-/existence of a production quota. Having paid these entry costs, each new firm is assigned to a productivity level drawn from the common cdf G, which is supposed to be continuous and have compact support. This implies that all entrants expect the same entry value by the time their exact productivity level is revealed.

In a scenario without production quota the entry costs are simply given by a constant value $k_t := c_e \geq 0$. If either a tradable or non-tradable quota exists, however, the entry costs should be somehow related to the available number of production/quota units. In this continuous framework each firm can be regarded as some kind of marginal production unit that is linked to a marginal quota unit. A variable serving as a measure for the number of active production units is then the total industry mass $\mu_t(\mathbb{R})$. This parameter must be contained under a production quota as the aggregate industry output (4.8) could become too large otherwise, and the production quota designed to constrain exactly this aggregate output level would be ineffective. For this reason, the mass of entering firms must be geared to the availability of production units. If the industry is subject to a quota regime, production units become available to potential entrants merely when incumbent firms decide to leave the industry and release their associated quota units. This means that exiting firms create the supply of production units that are demanded by new firms.

The parameter $\mu_t(\mathbb{R})$ is directly determined by the balance between exiting and entering firms; it increases whenever the number of entrants exceeds the mass of exiting firms, and vice versa. Therefore, the value $\mu_t(\mathbb{R})$ can be interpreted as a proxy for the availability of production units at the sector level. We utilise this feature and define the quota-constrained entry costs by $k_t := c_e + k(\mu_t(\mathbb{R}))$, with k

being a positive, continuous, and non-decreasing function. Hence, the entry costs are composed of two parts: entering firms have to pay the constant part c_e matching the unconstrained entry costs, and a variable part $k(\mu_t(\mathbb{R}))$ reflecting the additional quota costs. The monotonicity of k implies that the entry costs increase with respect to the number of firms willing to enter the industry. When we apply the model in Section 4.5, we determine $\mu_t(\mathbb{R}) = 1$ as a rough upper boundary and define k as a nearly perfectly elastic function around this point. Hence, the case $\mu_t(\mathbb{R}) < 1$ refers to a situation with unused quota units, while $\mu_t(\mathbb{R}) > 1$ indicates an overflow of production.

If the production quota cannot be traded among firms, the presumed structure of the entry costs just builds an entry barrier. In this case the quota costs represent a kind of penalty afforded by new firms to enter the industry. The total entry costs k_t thus coincide with the maximum value that entrants are willing to pay. As long as the quota is not tradable, incumbent firms cannot benefit directly from other firms' willingness to enter. Exiting firms get no compensation for ceasing production and releasing their associated quota units. This changes, of course, when the quota becomes tradable. The previously-defined quota costs can then be captured by exiting firms because they can sell their quota to potential entrants. This creates a positive liquidation value for incumbents, which we define by $r_t := k(\mu_t(\mathbb{R}))$. If the quota is either non-tradable or the industry is not subject to a production quota, exiting firms do not receive any compensation payment, and the liquidation value is set equal to zero, $r_t := 0$.

A firm bases its entry/exit decision on the expected discounted future profits. The discount rate for all firms is supposed to be $0 \leq \beta < 1$. If the sequence $y = \{p_t, r_t\}_{t \in \mathbb{N}_0}$ denotes the output prices and potential liquidation values for all periods, the value of an incumbent with productivity φ at time t can be defined recursively by

$$v_t(\varphi, y) = \pi(\varphi, p_t) + \beta \max \left\{ r_{t+1}, \int_{\mathbb{R}} v_{t+1}(\varphi', y) dF(\varphi'|\varphi) \right\}, \quad \forall\, t \in \mathbb{N}_0, \tag{4.10}$$

and is composed of the current profits plus the optional liquidation or continuation value.

A firm stays in the industry as long as its continuation value offsets the liquidation value, r_{t+1}. The continuation value indicates the expected future profits conditioned on the firm's current productivity level. The exit-point x_t describes the critical threshold for being indifferent between staying in or leaving the market,

$$x_t := \inf \left\{ \varphi \in \mathbb{R} : \int_{\mathbb{R}} v_{t+1}(\varphi', y) dF(\varphi'|\varphi) \geq r_{t+1} \right\}. \tag{4.11}$$

The assumptions made on the stochastic process and the period profits imply that all firms with a productivity level above the exit-point $\varphi_t \geq x_t$ stay in the industry, while all firms with a lower productivity level $\varphi_t < x_t$ take the exit compensation and quit. If the infimum in (4.11) does not exist, we are in a situation where no exit occurs in period t and we formally set $x_t = -\infty$.

The expected profits of a firm willing to enter the industry at the end of period t are given by

$$v^e_{t+1}(y) = \int_{\mathbb{R}} v_{t+1}(\varphi, y) dG(\varphi). \tag{4.12}$$

We denote the mass of firms, which decide to enter at time t and start production in the following period, by M_t. As this is the only kind of investment available to any firm, this may also include established firms building up additional production units. An increasing number of active firms will lead to a higher aggregate industry output, and result in a lower market price. New firms will be entering the industry as long as their expected future profits cover the entry costs, that is, in an equilibrium we obtain $v^e_{t+1} \leq k_{t+1}$. This condition must hold with equality if $M_t > 0$.

Due to the large number of firms in the industry (recall that firms are assumed to constitute a continuum), we do not have to deal with aggregate uncertainty. The frequency distribution of productivity levels in upcoming periods is completely specified by the stochastic productivity process and the entry/exit behaviour of firms. For a given exit-point x_t and entry-mass M_t, the industry structure in period $t+1$ is

$$\mu_{t+1}((-\infty, \varphi']) = \int_{\varphi \geq x_t} F(\varphi'|\varphi) d\mu_t(\varphi) + M_t\, G(\varphi'). \tag{4.13}$$

If both μ_t and G have Lebesgue densities $m_t(z)$ and $g(z)$, the state of the sector μ_{t+1} can also be characterised by its density

$$m_{t+1}(z) = \int_{\varphi \geq x_t} p(\varphi, z) m_t(\varphi) d\varphi + M_t\, g(z). \tag{4.14}$$

4.2 Equilibrium definitions

As a direct consequence of (4.13), both industry output and market price follow deterministic sequences. Firms are atomistic and cannot affect price by the choice of their output quantity. However, firms have perfect information about the decisions of others and are thus able to foresee the development of output prices. In a dynamic equilibrium firms adjust their output as well as their entry/exit decisions to the anticipated prices. These output prices, in turn, must be reinforced by the behaviour of firms. Keeping this in mind, we define a dynamic stochastic equilibrium as follows:

Definition 4.1 (Dynamic Equilibrium)**.** Given a starting distribution μ_0, a dynamic equilibrium consists of an infinite sequence of measures $\{\mu_t^*\}$, and sequences $\{p_t^*\}$, $\{Q_t^*\}$, $\{x_t^*\}$, $\{M_t^*\}$ containing the market prices, aggregate industry output, exit-points and entry-masses such that for each period $t \in \mathbb{N}$ the following conditions are satisfied:

(i) The output market is cleared
$p_t^* = D(Q_t^*)$,
$Q_t^* = Q^s(p_t^*, \mu_t^*)$.

(ii) The exit-rule (4.11) holds with x_t^*.

(iii) No more firms have an incentive to enter the industry, i.e. $v_t^e(y^*) \leq k_t$.

(iv) μ_t^* is determined recursively by (4.13).

One objective of this paper is to analyse the difference between a quota-constrained and a quota-free industry based on the long-run industry dynamics. Therefore, we focus on the concept of a stationary equilibrium, which is defined as the steady state of the dynamic model laid out above.

Definition 4.2 (Stationary Equilibrium)**.** A stationary equilibrium consists of a vector $(\mu^*, p^*, Q^*, x^*, M^*)$ such that the constant sequences $p_t = p^*$, $Q_t = Q^*$, $x_t = x^*$ and $M_t = M^*$ form a dynamic equilibrium for the starting distribution $\mu_0 = \mu^*$.

The definition indicates that a stationary equilibrium is a particular dynamic equilibrium starting from the stationary distribution $\mu_0 = \mu^*$. Notably, it is a state that still exhibits firm entry and exit. A common assumption is that the industry will sooner or later adjust to this steady state no matter what the current status is. Since the industry will remain in this steady state once it has been achieved, it allows us to analyse the long-run industry dynamics.

Next to those long-run properties, we also want to find out how firms react to the abolition of a production quota. For this purpose, we take the firm size distribution that emerges in a quota-constrained stationary equilibrium as starting distribution $\mu_0 := \mu^*$ for a dynamic equilibrium analysis. As the previously-defined infinite dynamic equilibrium is not implementable, however, we must restrict the time horizon to $t = 0, ..., T < \infty$ and calculate a dynamic equilibrium in this finite framework. Notice that because of this restriction the value function at the end of competition is just equal to the profits generated in the final period $v_T(\varphi, y) = \pi(\varphi, p_T)$. A finite dynamic equilibrium can thus be defined as follows:

Definition 4.3 (Finite Dynamic Equilibrium)**.** Given a starting distribution μ_0, a dynamic equilibrium consists of a finite sequence of measures $\{\mu_t^*\}$ and vectors $\mathbf{p}^*, \mathbf{Q}^*, \mathbf{x}^*, \mathbf{M}^*$ containing the market prices, aggregate industry output, exit-points and entry-masses for each period such that the conditions (i)-(iv) of an infinite dynamic equilibrium are satisfied for all times $t = 1, ..., T$.

To some extent, the equilibrium outcome will be affected by the assumed length of the planning horizon. Since we consider a finite time horizon, the function $v_t(\varphi, \mathbf{p}^*)$ is essentially a discounted sum of expected future profits. The value of a firm at time t will therefore depend on the number of time periods that are still to come. At the beginning, firms take the industry development over the whole time span into consideration, while they base their entry/exit decision on just a few upcoming periods at the end of competition. An extension of the time horizon by one period may thus have a strong impact on the value of a firm in the final periods. As firms discount future profits by the factor $\beta < 1$, however, the impact on a firm's value in the first periods is less harsh and will possibly diminish in the long-run. For this reason, we expect results to stabilise if the time horizon tends to infinity. However, the numeric effort to calculate an equilibrium in this case will be enormous.

In the next two sections we prove both the existence of a stationary equilibrium and the existence of a finite dynamic equilibrium in this continuous setup. The proofs themselves also demonstrate how to calculate such equilibria as we transform the dynamic equilibrium concept into a simple fixed-point problem. For numerical reasons, however, it is convenient to discretise the state space and calculate both the stationary and the finite dynamic equilibrium in a discrete framework. More information on the calculation can be found in Section 4.5, where the model is finally applied to the Western German dairy sector.

4.3 Existence of an applied stationary equilibrium

A crucial part of the stationary equilibrium is that the industry structure does not vary over time, although firms consistently leave and enter the industry. Therefore, we first show the existence and uniqueness of an invariant distribution

$$\mu((-\infty, \varphi']) = \int_{\varphi \geq x} F(\varphi'|\varphi)\, d\mu(\varphi) + M \cdot G(\varphi') \tag{4.15}$$

with respect to the given exit-point and entry-mass (x, M).

For any Borel set $A \in \mathcal{B}(\mathbb{R})$ let $P(\varphi, A)$ denote the probability kernel that corresponds with the underlying AR(1)-process. It can be derived from the conditional cdf F by

$$P(\varphi, A) = \int_A \mathbf{1}(\varphi')\, dF(\varphi'|\varphi) = \int_A p(\varphi, z)\, dz. \tag{4.16}$$

This stochastic kernel describes the time homogeneous one-step transition probability. If a firm possesses the current productivity value $\varphi_t = \varphi$, the value $P(\varphi, A)$ reflects the probability that the subsequent productivity level φ_{t+1} belongs to the set A. The two-step transition probability $\text{Prob}(\varphi_{t+2} \in A|\varphi_t = \varphi)$ can be derived as

$$\begin{aligned} P^2(\varphi, A) &= \int_{\mathbb{R}} P(\varphi', A)\, dF(\varphi'|\varphi) \\ &= \int_A \int_{\mathbb{R}} p(\varphi, \varphi')\, p(\varphi', z)\, d\varphi'\, dz. \end{aligned}$$

If we compute this integral, we will see that the two-step transition probability can also be characterised by a Normal distribution. In general, we have:

Lemma 4.3. *The n-step transition probability is given by the Normal distribution*

$$P^n(\varphi, \cdot) = N\left(\varphi\rho^n + \xi_\varepsilon \sum_{j=0}^{n-1} \rho^j, \sigma_\varepsilon^2 \sum_{j=0}^{n-1} \rho^{2j}\right). \tag{4.17}$$

Proof. By definition of $P(\varphi, \cdot)$ the equality is trivial for $n = 1$. We show that the statement is true for any n by induction. Hence, we assume that the statement holds for $n-1$. Let $\phi(z|\xi, \sigma^2)$ denote the pdf of a Normal distribution with mean ξ and variance σ^2. Using Lemma A.14 we get

$$
\begin{aligned}
P^n(\varphi, A) &= \int_{\mathbb{R}} P(\varphi', A)\, dF^{n-1}(\varphi'|\varphi) \\
&= \int_A \int_{\mathbb{R}} \phi \left(\varphi' \;\middle|\; \varphi\rho^{n-1} + \xi_\varepsilon \sum_{j=0}^{n-2} \rho^j, \sigma_\varepsilon^2 \sum_{j=0}^{n-2} \rho^{2j} \right) \, p(\varphi', z)\, d\varphi'\, dz \\
&= \int_A \phi \left(z \;\middle|\; \rho \left[\varphi\rho^{n-1} + \xi_\varepsilon \sum_{j=0}^{n-2} \rho^j \right] + \xi_\varepsilon \,,\; \rho^2\sigma_\varepsilon^2 \sum_{j=0}^{n-2} \rho^{2j} + \sigma_\varepsilon^2 \right) dz \\
&= \int_A \phi \left(z \;\middle|\; \varphi\rho^{n} + \xi_\varepsilon \sum_{j=0}^{n-1} \rho^j, \sigma_\varepsilon^2 \sum_{j=0}^{n-1} \rho^{2j} \right) dz.
\end{aligned}
$$

This had to be shown. □

As $\rho < 1$, the distribution $P^n(\varphi, \cdot)$ converges for $n \to \infty$ to the Normal distribution $P^\infty(\varphi, \cdot) = N\left(\frac{\xi_\varepsilon}{1-\rho}, \frac{\sigma_\varepsilon^2}{1-\rho^2}\right)$. This limiting distribution does not depend on the value φ anymore, which means that it is the same for all firms. Based on this result, we can now prove the existence and uniqueness of an invariant distribution.

Lemma 4.4. *Let G be a continuous cdf with compact support. For any combination of exit-point x and entry-mass M there exists a unique stationary distribution μ satisfying*

$$\mu((-\infty, \varphi']) = \int_{\varphi \geq x} F(\varphi'|\varphi)\, d\mu(\varphi) + M \cdot G(\varphi'). \tag{4.18}$$

Proof. Assume that a vector (x, M) of exit-point and entry-mass is given. For any Borel set $A \in \mathcal{B}(\mathbb{R})$ we define the stochastic kernel

$$\hat{P}_x(\varphi, A) = \begin{cases} P(\varphi, A) & \text{if } \varphi \geq x \\ 0 & \text{if } \varphi < x. \end{cases} \tag{4.19}$$

This defines a linear operator on the space of bounded measures as

$$\hat{P}_x\mu(A) = \int_{\mathbb{R}} \hat{P}_x(\varphi, A)\, d\mu(\varphi). \tag{4.20}$$

Moreover, let $\nu(A) = \int_A \mathbf{1}(\varphi)\, dG(\varphi)$ be the probability measure that is induced by the cdf G. With the help of these definitions, the condition for a stationary distribution (4.18) can be rephrased as

$$\mu = \hat{P}_x\, \mu + M \cdot \nu. \tag{4.21}$$

As suggested by Kolmogorov and Fomin (2012), Chapter 23, Theorem 4, a stationary distribution can be specified by the infinite series

$$\mu = M \sum_{t=0}^{\infty} \hat{P}_x^{\,t} \nu. \tag{4.22}$$

Here, $\hat{P}_x^{\,t}$ denotes the $t-$fold application of the operator $\hat{P}_x$, and $\hat{P}_x^{\,0} = Id.$ is defined as the identity operator on the space of bounded measures. It is easy to see that this measure satisfies the stationary condition (4.21) as for any Borel set $A \in \mathcal{B}(\mathbb{R})$

$$\begin{aligned}
\hat{P}_x\, \mu(A) + M \cdot \nu(A) &= M \sum_{t=1}^{\infty} \hat{P}_x^{\,t} \nu(A) + M \cdot \nu(A) \\
&= M \left[\sum_{t=1}^{\infty} \hat{P}_x^{\,t} \nu(A) + \hat{P}_x^{\,0} \nu(A) \right] \\
&= \mu(A).
\end{aligned}$$

To complete the proof, we must verify that the series in (4.22) converges. By assumption, the measure ν has compact support $C \subset \mathbb{R}$. Lemma 4.3, thus, implies that for any $\delta > 0$ there exists a number $n_0 \in \mathbb{N}$ such that $\|P^n(\varphi, \cdot) - P^\infty(\varphi, \cdot)\|_{TV} < \delta$ for all $\varphi \in C$ and $n \geq n_0$. As the limiting distribution P^∞ is a Normal distribution, we conclude $P^\infty(\varphi, B) > 0$ for the set $B = (-\infty, x)$. Hence, if δ is chosen small enough, we can find an $\varepsilon > 0$ such that $P^n(\varphi, B) > \varepsilon$ for all $\varphi \in C$ and $n \geq n_0$. For any Borel set $A \in \mathcal{B}(\mathbb{R})$ we can infer:

$$\begin{aligned}
0 &\leq M \sum_{t=0}^{\infty} \hat{P}_x^{\,t} \nu(A) \\
&\leq M \sum_{t=0}^{\infty} \hat{P}_x^{\,t} \nu(\mathbb{R}) \\
&\leq M \left[\sum_{t=0}^{n_0} \hat{P}_x^{\,t} \nu(\mathbb{R}) + \sum_{n=0}^{\infty} (1-\varepsilon)^n \right] < \infty.
\end{aligned}$$

The last inequality holds since the first sum is finite, and the second one tends to the value $\frac{1}{\varepsilon}$. Hence, the series in (4.22) must converge. This concludes the proof. □

The Lemma allows us to define a mapping $(x, M) \mapsto \mu(x, M) := M \sum_{t=0}^{\infty} \hat{P}_x^{\,t} \nu$ that assigns the stationary distribution μ to any given vector (x, M). This mapping possesses the following properties:

Lemma 4.5. *The mapping $\mu(x, M)$ has the following properties:*

(i) It is continuous in both arguments.

(ii) It is (strictly) decreasing with respect to the exit-point x. If $x_1 < x_2$, and the measures $\mu_1 = \mu(x_1, M)$, and $\mu_2 = \mu(x_2, M)$, we have that $\mu_1(A) > \mu_2(A)$ for all non-null sets $A \in \mathcal{B}(\mathbb{R})$.

(iii) It is (strictly) increasing with respect to the entry-mass M. If $M_1 < M_2$, and the measures $\mu_1 = \mu(x, M_1)$, and $\mu_2 = \mu(x, M_2)$, we have that $\mu_1(A) < \mu_2(A)$ for all non-null sets $A \in \mathcal{B}(\mathbb{R})$.

Proof. **ad(i).** The continuity of $\mu(x, M)$ with respect to M is trivial. Therefore, we fix a value $M > 0$ and prove the continuity with respect to x. By induction with respect to t we first show that any term $\hat{P}_x^t \nu$ is continuous in x. We begin with the case $t = 1$. Let $x_n \to x$ be a sequence converging to x for $n \to \infty$. Moreover, let $B(x, \varepsilon) = \{y \in \mathbb{R} : |x - y| < \varepsilon\}$ denote an $\varepsilon-$neighbourhood around x. If $x_n \in B(x, \varepsilon)$, we have that

$$\begin{aligned}\left|\hat{P}_{x_n}\nu(A) - \hat{P}_x\nu(A)\right| &\leq \int_{B(x,\varepsilon)} P(\varphi, A)\, d\nu(\varphi)\\ &\leq \nu(B(x, \varepsilon)) \;\to 0 \text{ for } \varepsilon \to 0.\end{aligned}$$

As ν is a continuous measure, the latter expression tends to zero whenever $\varepsilon \to 0$. We conclude from this that $\|\hat{P}_{x_n}\nu - \hat{P}_x\nu\|_{TV} = 2 \sup_{A\in\mathcal{B}(\mathbb{R})} |\hat{P}_{x_n}\nu(A) - \hat{P}_x\nu(A)| \to 0$ for $n \to \infty$, and the term $\hat{P}_x\nu$ is continuous in x.

Now, we prove the continuity of $\hat{P}_x^t \nu$ provided that $\hat{P}_x^{t-1} \nu$ is continuous with respect to x. To simplify the notation we introduce the measures $\Theta := \hat{P}_x^{t-1} \nu$ and $\Theta_n := \hat{P}_{x_n}^{t-1}\nu$, and presume that $\|\Theta_n - \Theta\|_{TV} \to 0$ for $n \to \infty$. If $x_n \in B(x, \varepsilon)$, this implies:

$$\begin{aligned}&\left|\hat{P}_{x_n}^t\nu(A) - \hat{P}_x^t\nu(A)\right|\\ &= \left|\int_{\varphi\geq x_n} P(\varphi, A)\, d\Theta_n(\varphi) - \int_{\varphi\geq x} P(\varphi, A)\, d\Theta(\varphi)\right|\\ &= \left|\int_{x_n}^{\infty} P(\varphi, A)\, d\Theta_n(\varphi) - \int_{x_n}^{\infty} P(\varphi, A)\, d\Theta(\varphi) + \int_{x_n}^{\infty} P(\varphi, A)\, d\Theta(\varphi) - \int_{x}^{\infty} P(\varphi, A)\, d\Theta(\varphi)\right|\end{aligned}$$

$$\leq \left| \int_{x_n}^{\infty} P(\varphi, A)\, d\Theta_n(\varphi) - \int_{x_n}^{\infty} P(\varphi, A)\, d\Theta(\varphi) \right| + \left| \int_{x_n}^{\infty} P(\varphi, A)\, d\Theta(\varphi) - \int_{x}^{\infty} P(\varphi, A)\, d\Theta(\varphi) \right|$$

$$\leq \|\Theta_n - \Theta\|_{TV} + \Theta(B(x, \varepsilon)) \;\; \to 0 \text{ for } \varepsilon \to 0 \text{ and } n \to \infty.$$

By assumption, the first term tends to zero for $n \to \infty$. As Θ is a continuous measure, the latter term tends to zero for $\varepsilon \to 0$. But, this means that $\hat{P}_x^{\,t} \nu$ is continuous in x. Accordingly, any finite sum $\sum_{t=0}^{T} \hat{P}_x^{\,t} \nu$ is continuous with respect to x. Since the aforementioned series converges for $T \to \infty$ (cf. proof of Lemma 4.4), the same applies to the mapping $\mu(x, M) = M \sum_{t=0}^{\infty} \hat{P}_x^{\,t} \nu$.

ad(ii). The monotonicity statement follows immediately from the representation of $\mu(x, M)$ as infinite series and the definition of the operator $\hat{P}_x$ in (4.20).

ad(iii). The monotonicity of $\mu(x, M)$ with respect to M is clear. □

In the final part of this section we can now utilise the previous results and show the existence of a stationary equilibrium.

Theorem 4.6 (Existence of a stationary equilibrium)**.** *Let G be a continuous cdf with compact support, and both q^* and π be integrable with respect to Normal distributions. A stationary equilibrium exists in the infinite framework if $D(0) =: p^{max} < \infty$, and the AR(1)-process is stationary ($\rho < 1$).*

The proof to this Theorem is based upon Brouwer's Fixed Point Theorem (see Appendix A.1), which states that every continuous function, mapping a compact, convex subset of $\mathbb{R}^n$ into itself, possesses a fixed point. We utilise this property and proceed as follows: First, we construct a continuous mapping $(x, M) \mapsto \tau(x, M)$ such that a fixed point of this mapping coincides with the stationary equilibrium. Second, we define a compact, convex subset N that is mapped into itself by τ. According to Brouwer's Theorem this implies the existence of a stationary equilibrium. We distinguish between a scenario with and without capacity constraints.

Proof. (*Without capacity constraints*)
Step I. Let a tuple (x, M) of exit-point and entry-mass be given. Based on the corresponding stationary distribution $\mu(x, M)$, we can derive the market clearing output price p^{mc} as solution to the equation

$$p = D\left(\int_{\mathbb{R}} q^*(\varphi, p)\, d\mu(\varphi) \right). \tag{4.23}$$

The imposed assumptions on $D(Q)$ and $q^*(\varphi, p)$ guarantee that a unique solution p^{mc} exists. Due to the continuity of $\mu(x, M)$ (cf. Lemma 4.5), the market clearing output price must be a continuous function of the exit-point x and entry-mass M as well. The monotonicity of $\mu(x, M)$ induces, moreover, that p^{mc} is increasing in x and decreasing with respect to M.

Assuming that this output price emerges in all time periods, the firm's value function is given by the equation

$$v(\varphi, p^{mc}) = \pi(\varphi, p^{mc}) + \beta \max\left\{0, \int_{\mathbb{R}} v(\varphi', p^{mc})\, dF(\varphi'|\varphi)\right\}. \tag{4.24}$$

Since the function π is supposed to be integrable with respect to any Normal distribution and $\beta, \rho < 1$, a unique solution to this equation (4.24) exists. The solution is a continuous function v that possesses the same properties as the function π. In particular, it is upward sloping with respect to the price p^{mc} and strictly increasing in the productivity parameter φ.[4] As long as $p^{mc} > 0$ we have the limiting behaviour

$$\lim_{\varphi \to +\infty} v(\varphi, p^{mc}) = \infty \quad \text{and} \quad \lim_{\varphi \to -\infty} v(\varphi, p^{mc}) = -c_f. \tag{4.25}$$

According to Lemma A.15 the same must apply to the continuation value

$$v^c(\varphi, p^{mc}) = \int_{\mathbb{R}} v(\varphi', p^{mc})\, dF(\varphi'|\varphi). \tag{4.26}$$

This allows us to derive a new exit-point $\widetilde{x}$ by the exit-rule

$$\widetilde{x} := \inf\left\{\varphi \in \mathbb{R} : v^c(\varphi, p^{mc}) \geq 0\right\}. \tag{4.27}$$

Due to the continuity and strict monotonicity of the continuation value with respect to φ, the exit-point $\widetilde{x}$ must be a continuous function of the market clearing output price p^{mc}. A new entry-mass $\widetilde{M}$ can be specified by

$$\widetilde{M} := \min\left\{M^{max}, \max\left\{0, v^e(p^{mc}) - c_e + M\right\}\right\}, \tag{4.28}$$

where

$$v^e(p^{mc}) = \int_{\mathbb{R}} v(\varphi, p^{mc})\, dG(\varphi) \tag{4.29}$$

[4]The existence of the value function v, as well as its properties, can essentially be derived by the same argument as in the proof to Theorem 3.1. As the presumed state space $S = \mathbb{R}$ is not compact though, the stationarity of the stochastic process ($\rho < 1$) is necessary to guarantee the existence of a solution.

denotes the expected value of entering firms, and M^{max} stands for a maximum possible amount of entrants.

Referring to the derived exit-point (4.27) and entry-mass (4.28), we can now define the mapping $\tau(x, M) := (\widetilde{x}, \widetilde{M})$. By construction τ maps a combination (x, M) of exit-point and entry-mass to another combination $(\widetilde{x}, \widetilde{M})$ of exit-point and entry-mass. Moreover, a fixed point of this mapping constitutes a stationary equilibrium in terms of Definition 4.2. The mapping $\tau : \mathbb{R}^2 \to \mathbb{R}^2$ is continuous as the output price p^{mc} is continuous with respect to x, M, and both variables $\widetilde{x}, \widetilde{M}$ are continuous functions of p^{mc}.

Step II. According to Brouwer's Theorem the existence of a stationary equilibrium is proven if we can find a compact, convex subset $N \subset \mathbb{R}^2$ that is mapped into itself by $\tau(N) \subseteq N$. We begin to determine a maximum possible entry-mass $M^{max} \geq 0$. To this end we presume an industry structure $\mu = M^{max}\nu$, and compute the value M^{max} such that the market clearing output price under this industry structure satisfies $v^e(p^{mc}) = c_e$. A higher entry-mass cannot occur in a stationary equilibrium because $v^e(p^{mc}) < c_e$ for any stationary distribution

$$\mu(x, M^{max}) = M^{max} \sum_{t=0}^{\infty} \hat{P}_x^t \, \nu.$$

By assumption, the equilibrium output price is bounded by $p^* \in (0, p^{max}]$. Hence, solving the exit-rule (4.27) for $p^{mc} = p^{max}$ yields the minimum possible exit threshold x^{min}. The monotonicity of the mapping $\mu(x, M)$ implies that a minimum justifiable output price p^{min} arises for the stationary distribution $\mu(x^{min}, M^{max})$. Applying the exit-rule to this minimum price allows us to derive a maximum possible exit-point x^{max}.

A compact, convex subset $N \subset \mathbb{R}^2$ can now be defined as the cartesian product

$$N := [x^{min}, x^{max}] \times [0, M^{max}]. \tag{4.30}$$

This set is mapped into itself by τ since every element $(x, M) \in N$ leads to a market clearing output price $p^{mc} \in [p^{min}, p^{max}]$. By the way N has been designed this implies $(\widetilde{x}, \widetilde{M}) \in N$, and we can conclude the existence of a stationary equilibrium. □

Proof. (*With capacity constraints*)
The proof is carried out completely analogously to the one without capacity constraints.

Step I. We look for a continuous mapping τ, whose fixed point reflects a stationary equilibrium with capacity constraints. As before, we assume that a combination (x, M) of exit-point and entry-mass is given, and derive the corresponding stationary distribution $\mu(x, M)$ as well as the market clearing output price p^{mc}. Although the individual output boundary q^{max} comes into play here, the output price is still uniquely determined and continuous in (x, M). Based on the stationary distribution μ we also compute the capacity constrained entry costs $\bar{k} = c_e + k(\mu(\mathbb{R}))$ and the firms' liquidation value r. Depending on the tradability of production units between firms the liquidation value is either $r = 0$ (non-tradable capacity) or $r = k(\mu(\mathbb{R}))$ (tradable capacity). Both entry costs $\bar{k}$ and liquidation value r form a continuous function of the vector (x, M) because the function k is supposed to be continuous.

Compared to the unconstrained case, the value function turns into

$$v(\varphi, p^{mc}, r) = \pi(\varphi, p^{mc}) + \beta \max \left\{ r, \int_{\mathbb{R}} v(\varphi', p^{mc}, r) \, dF(\varphi'|\varphi) \right\}. \tag{4.31}$$

A continuous, monotone solution v to this equation exists by the same reasoning as in the unconstrained case. But, the limiting behaviour alters slightly and is now given by

$$\lim_{\varphi \to +\infty} v(\varphi, p^{mc}, r) = \infty \quad \text{and} \quad \lim_{\varphi \to -\infty} v(\varphi, p^{mc}, r) = -c_f + \beta r. \tag{4.32}$$

for all $p^{mc} > 0$. Based on the modified continuation value

$$v^c(\varphi, p^{mc}, r) = \int_{\mathbb{R}} v(\varphi', p^{mc}, r) \, dF(\varphi'|\varphi), \tag{4.33}$$

we derive a new exit-point $\widetilde{x}$ by

$$\widetilde{x} := \inf \{ \varphi \in \mathbb{R} : v^c(\varphi, p^{mc}, r) \geq r \}. \tag{4.34}$$

The equilibrium entry condition $v^e(p^{mc}, r) \leq \bar{k}$ is used to deduce a new entry-mass

$$\widetilde{M} := \min \left\{ M^{max}, \max \left\{ 0, v^e(p^{mc}, r) - \bar{k} + M \right\} \right\}. \tag{4.35}$$

As before, $v^e(p^{mc}, r)$ signifies the expected value of entering firms, while M^{max} marks an upper mass of entrants that can possibly arise in a stationary equilibrium.

Being a composition of continuous functions, the mapping $\tau(x, M) := (\widetilde{x}, \widetilde{M})$ must be continuous itself. To prove the existence of a fixed point, and thus the existence of a stationary equilibrium in the constrained case, we just need to find a compact, convex subset $N \subset \mathbb{R}^2$ such that $\tau(N) \subseteq N$.

Step II. The maximum entry-mass $M^{max} \geq 0$ is determined in the same way as for the unconstrained scenario. This means, we compute the value M^{max} such that $v^e(p^{mc}, r) = \bar{k}$ for the industry structure $\mu = M^{max}\ \nu$. By definition both the entry-costs and the liquidation value are (strictly) increasing with respect to the entry-mass. As the slope of $\bar{k}$ is steeper than the one of $v^e(p^{mc}, r)$ though, any entry-mass $M > M^{max}$ leads to $v^e(p^{mc}, r) < \bar{k}$ and cannot be justified in a stationary equilibrium.

A minimum exit-value x^{min} can be specified by solving the exit-rule (4.34) for $p^{mc} = p^{max}$ and $r = 0$. The stationary distribution $\mu(x^{min}, M^{max})$ generates the minimum market clearing output price p^{min} as well as the maximum liquidation value r^{max}. If we insert these values into the exit-rule (4.34), we can derive the maximum exit-point x^{max}. Hence, the compact, convex subset

$$N := [x^{min}, x^{max}] \times [0, M^{max}] \subset \mathbb{R}^2 \tag{4.36}$$

is mapped into itself as $\tau(N) \subseteq N$. According to Brouwer's Fixed Point Theorem this implies the existence of a stationary equilibrium. □

The previous Theorem guarantees that the presented model possesses at least one stationary equilibrium. While we cannot fully exclude multiple steady states in the capacity constrained scenario, the stationary equilibrium is uniquely determined in the unconstrained case. Why is that? As the value function is strictly increasing with respect to the output price, there exists at most one price p^{mc} such that the unconstrained entry condition $v^e(p^{mc}) = c_e$ holds with equality. Given this output price there is exactly one value x^* obeying the exit-rule. According to Lemma 4.5 the mapping $\mu(x^*, M)$ must be strictly decreasing with respect to M. Therefore, only one entry-mass M^* can generate the stationary distribution $\mu(x^*, M^*)$ that leads to the market clearing output price p^{mc}. This implies the uniqueness of the unconstrained stationary equilibrium.

Although we focus on stationary equilibria with positive firm entry and exit in our further analysis, it is worth mentioning that theoretically a steady state without entry/exit can also occur in this framework. If the fixed entry costs c_e are so high, for instance, that they cannot be covered by the highest possible expected value $v^e(p^{max})$, no firm has an incentive to enter the industry, and the only possible stationary equilibrium consists of an empty industry.

4.4 Existence of a finite dynamic equilibrium

A necessary condition for the existence of a finite dynamic equilibrium is the integrability of $q^*(\varphi, p)$ and $\pi(\varphi, p)$ with respect to the relevant measures. If both functions are integrable with respect to any Normal distribution, the properties in Lemma 4.2 translate one-to-one to the value function $v_t(\varphi, \mathbf{p}, \mathbf{r})$, and also to the continuation value

$$v_t^c(\varphi, \mathbf{p}, \mathbf{r}) = \int_{\mathbb{R}} v_t(\varphi', \mathbf{p}, \mathbf{r})\, dF(\varphi'|\varphi). \tag{4.37}$$

Hence, the continuation value is continuous and strictly increasing in φ. Furthermore, the limits

$$\lim_{\varphi \to +\infty} v_t^c(\varphi, \mathbf{p}, \mathbf{r}) = \infty \quad \text{and} \quad \lim_{\varphi \to -\infty} v_t^c(\varphi, \mathbf{p}, \mathbf{r}) = -c_f \tag{4.38}$$

hold for every price vector $\mathbf{p} > 0$ and liquidation vector $\mathbf{r} = 0$. Having these properties at hand, we can now prove the subsequent proposition.

Theorem 4.7 (Existence of a dynamic equilibrium)**.** *Let μ_0, G be continuous distributions with compact support, and both q^* and π be integrable with respect to Normal distributions. A dynamic equilibrium exists in the finite framework if the demand function satisfies $D(0) =: p^{max} < \infty$.*

The additional assumptions made in Theorem 4.7 imply that all integrals, which will be considered in the following, exist. Just like for the stationary equilibrium, the proof relies on Brouwer's Fixed Point Theorem. The theorem states that every continuous function, mapping a compact, convex space into itself, has a fixed point. Hence, we will define a continuous mapping and illustrate, in the first step, that a fixed point of this mapping constitutes a dynamic equilibrium. In a second step, we will specify a compact, convex subset and show that the previously defined function maps this set into itself.

Proof. (*Without capacity constraints*)
We show the existence of a dynamic equilibrium in the case without capacity constraints first. This implies $r_t = 0$, $k_t = c_e$ for all t, and

$$v_t(\varphi, \mathbf{p}) = \pi(\varphi, p_t) + \beta \max\left\{0, \int_{\mathbb{R}} v_{t+1}(\varphi', \mathbf{p})\, dF(\varphi'|\varphi)\right\}. \tag{4.39}$$

Step I. By condition (i) of Definition 4.3, the equilibrium prices p_t^* must clear the output market in every single period, i.e. the equality

$$p_t = D(Q^s(p_t, \mu_t)) = D\left(\int_{\mathbb{R}} q^*(\varphi, p_t) d\mu_t(\varphi)\right) \tag{4.40}$$

must be satisfied. The properties of $D(Q)$ and $q^*(\varphi, p)$ imply that for any given industry structure μ_t a unique solution p_t^{mc} to (4.40) exists (Intermediate Value Theorem). Considering the industry dynamics (4.13), p_t^{mc} will be a continuous function of all previous exit-points $x_0, ..., x_{t-1}$ and entry-masses $M_0, ..., M_{t-1}$.[5] Furthermore, the market clearing output price p_t^{mc} is upward sloping in x_j and downward sloping with respect to M_j (this holds for any $j \in \{0, ..., t-1\}$).

Now, we construct an operator $\tau : \mathbb{R}^{2T} \to \mathbb{R}^{2T}$ that maps a given vector $(\mathbf{x}, \mathbf{M})$ of exit-points and entry-masses to another vector $(\widetilde{\mathbf{x}}, \widetilde{\mathbf{M}})$. For any vector $(\mathbf{x}, \mathbf{M}) \in \mathbb{R}^{2T}$, containing a number of exit-points x_t and entry-masses M_t, we can derive the resulting industry structures μ_t according to the transition rule (4.13). The market clearing output prices p_t^{mc} are then determined by the equality (4.40). Given this output price vector $\mathbf{p^{mc}}$, we can define the mapping $\tau(\mathbf{x}, \mathbf{M}) = (\widetilde{\mathbf{x}}, \widetilde{\mathbf{M}})$ by:

$$\widetilde{x}_t := \inf\left\{\varphi \in \mathbb{R} : \int_{\mathbb{R}} v_{t+1}(\varphi', \mathbf{p^{mc}}) dF(\varphi'|\varphi) \geq 0\right\}, \tag{4.41}$$

$$\widetilde{M}_t := \min\left\{M_t^{max}, \max\left\{0\ ,\ v_{t+1}^e(\mathbf{p^{mc}}) - c_e + M_t\right\}\right\}. \tag{4.42}$$

The value M_t^{max} determines an upper boundary to the entry-mass, which can possibly arise in an equilibrium and will be specified later on. It is trivial that (4.41) coincides with the exit-rule of our equilibrium definition, and equation (4.42) rephrases the entry condition (iii). Hence, the values $\widetilde{x}_t$ represent the critical productivity thresholds under the price vector $\mathbf{p^{mc}}$. According to (4.38), the function $v_t^c(\varphi, \mathbf{p})$ tends to infinity for $\varphi \to +\infty$. Therefore, the infimum in (4.41) does indeed exist for any output price vector. The solution $\widetilde{x}_t$ will be a continuous function of $\mathbf{p^{mc}}$ as the continuation value is continuous and strictly increasing with respect to φ. It is evident that the same applies to the values $\widetilde{M}_t$. Hence, the mapping τ, which is a composition of continuous functions, must be continuous itself. The way we have constructed τ implies, moreover, that a fixed point of this mapping describes an equilibrium.

Step II. In the remainder we will specify a compact, convex subset $N \subset \mathbb{R}^{2T}$ such that $\tau(N) \subseteq N$, i.e. τ maps N into itself. First, recall that the output prices,

[5] Recall that $\mu_0, F(.|\varphi)$ and G are all supposed to be continuous distributions

which may occur in an equilibrium, are bounded by $p_t^* \in (0, p^{max}]$. This allows us to determine lower boundaries for the exit-points. We define the price vector $\mathbf{p^{max}} := (p^{max}, ..., p^{max})$, which would for instance arise in an empty industry. The corresponding solutions of the exit-rule with respect to $\mathbf{p^{max}}$ determine the minimum attainable values x_t^{min}.

Next, we define the upper boundary M_t^{max} for the entry-masses that can materialise in an equilibrium. Note first, that the expected value of entrants can be written as

$$v_t^e(\mathbf{p}) = \sum_{j=t}^{T} \beta^{j-1} \int_{\mathbb{R}} \pi(\varphi, p_j)\, d\bar{\mu}_j(\varphi), \tag{4.43}$$

with $\bar{\mu}_t \equiv G$ and $\bar{\mu}_j$ being the distribution of a firm's productivity in period j. These measures depend implicitly on the firm's optimal exit decisions (with respect to any given price vector $\mathbf{p}$), and the total mass $\bar{\mu}_j(\mathbb{R})$ displays the probability of still being active in period j. In an equilibrium, the firm's exit policy must coincide exactly with the exit-points $x_t, ..., x_{T-1}$.

Now, we turn this around and define a firm's productivity distribution in period j as an explicit function of given exit-points $x_t, ..., x_{j-1}$. The distribution is denoted by the measure

$$\lambda_j((-\infty, \varphi']) = \int_{x_t}^{\infty} \cdots \int_{x_{j-1}}^{\infty} F(\varphi'|\varphi_{j-1})\, dF(\varphi_{j-1}|\varphi_{j-2}) \cdots dF(\varphi_{t+1}|\varphi_t) dG_(\varphi_t). \tag{4.44}$$

This implies $\lambda_j = \bar{\mu}_j$ whenever the exit-points $x_t, ..., x_{j-1}$ represent an optimal exit policy (as they do, for instance, in an equilibrium).

For any given vector of exit-points $\mathbf{x} = (x_0, ..., x_{T-1})$ there exists an entry-mass M_0 such that

$$v_1^e(\mathbf{p^{mc}}, \mathbf{x}) = \sum_{j=1}^{T} \beta^{j-1} \int_{\mathbb{R}} \pi(\varphi, p_j^{mc})\, d\lambda_j(\varphi) \leq c_e. \tag{4.45}$$

The reason for this is simply that all output prices $p_1^{mc}, ..., p_T^{mc}$ tend to zero if the exit-points are fixed and $M_0 \to \infty$. Hence, we define $\bar{M}_x := \inf\{M_0 \geq 0 : v_1^e(\mathbf{p^{mc}}, \mathbf{x}) \leq c_e\}$ as the smallest entry-mass satisfying (4.45). Furthermore, $\lambda_j((-\infty, \varphi']) \to 0$ if any exit-point $x_k \to \infty$ and $k \in \{t, ..., j-1\}$. Therefore, we can find exit values $\bar{x}_{j-1}$ such that

$$\int_{\mathbb{R}} \pi(\varphi, p^{max})\, d\lambda_j(\varphi) \leq \frac{c_e}{T} \tag{4.46}$$

for any $x_{j-1} \geq \bar{x}_{j-1}$.

Recall that the market clearing output price in period t is a continuous function of all previous exit-points and entry-masses. Hence, the function $v_1^e(\mathbf{p^{mc}}, \mathbf{x})$ is also continuous with respect to M_0 and $\mathbf{x}$. On the compact subset

$$X := [x_0^{min}, \bar{x}_0] \times \cdots \times [x_{T-1}^{min}, \bar{x}_{T-1}],$$

there must be a maximum value $M_0^{max} = \sup_{\mathbf{x} \in X} \bar{M}_x < \infty$ that satisfies the inequality in (4.45) for every exit vector $\mathbf{x} \in [x_0^{min}, \infty) \times \cdots \times [x_{T-1}^{min}, \infty)$. The values $M_1^{max}, ..., M_{T-1}^{max}$ are determined by exactly the same procedure.

To determine maximum attainable exit values x_t^{max}, we need to calculate the minimum output prices first. It is clear that the output prices are minimised if the aggregate output is maximised. This is the case if no exit takes place and the maximum amount of firms M_t^{max} enters the industry in each period t. Hence, the entry/exit-vector $(\mathbf{x}, \mathbf{M}) = (-\infty, ..., -\infty, M_0^{max}, ..., M_{T-1}^{max})$ yields the minimum justifiable output prices $(p_1^{min}, ..., p_T^{min})$, which are given by market clearance (4.40). By taking the minimum over all $t = 1, ..., T$ we can also determine an absolute minimum price $p^{min} > 0$ that serves as a lower boundary to all market clearing output prices. Solving the exit-rule for the constant price vector $\mathbf{p^{min}} = (p^{min}, ..., p^{min})$ thus yields the maximum possible exit values x_t^{max}.

With all those exit and entry values at hand, we define the subset $N \subset \mathbb{R}^{2T}$ by the cartesian product

$$N := [x_0^{min}, x_0^{max}] \times \cdots \times [x_{T-1}^{min}, x_{T-1}^{max}] \times [0, M_0^{max}] \times \cdots \times [0, M_{T-1}^{max}]. \qquad (4.47)$$

Obviously, this compact, convex set is mapped into itself by the operator τ. If we take any $(\mathbf{x}, \mathbf{M}) \in N$, the resulting output prices p_t^{mc} will be in $[p^{min}, p^{max}]$. Due to the construction of N and the monotonicity of the continuation value $v_t^c(\varphi, \mathbf{p})$, which was mentioned right at the beginning of the proof, the resulting exit-points $\widetilde{x}_t$ must lie inside the interval $[x_t^{min}, x_t^{max}]$. Furthermore, the calculated entry values $\widetilde{M}_t$ are surely between 0 and M_t^{max}. This implies $(\widetilde{\mathbf{x}}, \widetilde{\mathbf{M}}) \in N$, and we have indeed $\tau(N) \subseteq N$.

Summing up, we have argued that $\tau : N \to N$ is a continuous mapping on a compact space. In compliance with Brouwer's Theorem this mapping possesses a fixed point. The fixed point essentially represents a finite dynamic equilibrium and, thus, the theorem is proven. □

Proof. (*With capacity constraints*)
Now, we turn to the scenario where the total production capacity is limited to the

sector level, and show the existence of a dynamic equilibrium in this case. The proof is carried out in the same fashion as the previous one. But, we request that the entry costs are $k_t = c_e + k(\mu_t(\mathbb{R}))$, with $k(x)$ being an upward sloping continuous function satisfying $k(0) = 0$ and $\lim_{x\to\infty} k(x) = \infty$. If production capacity is tradable among firms, we define the liquidation value as $r_t = k(\mu_t(\mathbb{R}))$. Otherwise, we set $r_t = 0$. Let the vector $\mathbf{r} = (r_1, ..., r_T)$ comprise these exit premiums. The value function thus alters into

$$v_t(\varphi, \mathbf{p}, \mathbf{r}) = \pi(\varphi, p_t) + \beta \max \left\{ r_{t+1}, \int_{\mathbb{R}} v_{t+1}(\varphi', \mathbf{p}, \mathbf{r})\, dF(\varphi'|\varphi) \right\}. \tag{4.48}$$

Note, that the continuity and monotonicity properties with respect to φ and $\mathbf{p}$ remain unchanged. Moreover, the firm's value at t increases with respect to all subsequent exit premiums $r_{t+1}, ..., r_T$.

Step I. We construct a continuous mapping $\tau : \mathbb{R}^{2T} \to \mathbb{R}^{2T}$. For any entry/exit-vector $(\mathbf{x}, \mathbf{M})$ the industry structures μ_t and market clearing output prices p_t^{mc} are derived as before. This time, however, we also calculate the capacity values k_t and r_t based on $\mu_t(\mathbb{R})$. The operator $\tau(\mathbf{x}, \mathbf{M}) = (\widetilde{\mathbf{x}}, \widetilde{\mathbf{M}})$ is then determined by the exit/entry-rules

$$\widetilde{x}_t := \inf \left\{ \varphi \in \mathbb{R} : \int_{\mathbb{R}} v_{t+1}(\varphi', \mathbf{p^{mc}}, \mathbf{r})\, dF(\varphi'|\varphi) \geq r_{t+1} \right\} \tag{4.49}$$

$$\widetilde{M}_t := \min \left\{ M_t^{max} \, , \, \max \left\{ 0 \, , \, v_{t+1}^e(\mathbf{p^{mc}}, \mathbf{r}) - k_{t+1} + M_t \right\} \right\} \tag{4.50}$$

The constant M_t^{max} is the maximum amount of firms that will possibly enter the industry by the end of period t. The exact value M_t^{max} will be specified later on. Recall that τ is again a continuous mapping, and the dynamic equilibrium is characterised as a fixed point of τ.

Step II. The challenge is once more to specify a compact, convex subset $N \subset \mathbb{R}^{2T}$ such that $\tau(N) \subseteq N$. Utilising the constant price vector $\mathbf{p^{max}}$ allows us to compute a M_t^{max} such that

$$v_{t+1}^e(\mathbf{p^{max}}, \mathbf{r}) = \int_{\mathbb{R}} v_{t+1}(\varphi, \mathbf{p^{max}}, \mathbf{r})\, dG(\varphi) \leq c_e + k(M_t^{max}). \tag{4.51}$$

Here, we presume that new firms enter an empty industry, and $r_j = k(M_t^{max})$ for all $j = 1, ..., T$. Firms having entered the industry in period t and paid capacity costs $k(M_t^{max})$ can, thus, recapture the same (discounted) value as exit premium in prospective periods. The discount factor $\beta < 1$ guarantees that a solution to equation (4.51) exists. Because $v_t^e(\mathbf{p}, \mathbf{r}) \geq v_{t+1}^e(\mathbf{p}, \mathbf{r})$ for any constant vectors $\mathbf{p}, \mathbf{r} \geq 0$, we will have $M_{t-1}^{max} \geq M_t^{max}$.

By the same approach as in the unconstrained case, we compute minimum justifiable output prices with the vector $(\mathbf{x}, \mathbf{M}) = (-\infty, ..., -\infty, M_0^{max}, ..., M_{T-1}^{max})$. The minimum of the market clearing output prices constitutes the lower boundary p^{min}. We can also determine an upper boundary for the exit premium by

$$r^{max} = k\left(\sum_{j=1}^{T} M_j^{max}\right). \tag{4.52}$$

This implies minimum and maximum values for the exit-points by

$$x_t^{min} := \inf\left\{\varphi \in \mathbb{R} : \int_{\mathbb{R}} v_{t+1}(\varphi', \mathbf{p^{max}}, \mathbf{r^{max}})\, dF(\varphi'|\varphi) \geq 0\right\}, \tag{4.53}$$

$$x_t^{max} := \inf\left\{\varphi \in \mathbb{R} : \int_{\mathbb{R}} v_{t+1}(\varphi', \mathbf{p^{min}}, \mathbf{0})\, dF(\varphi'|\varphi) \geq r^{max}\right\}. \tag{4.54}$$

The compact, convex subset $N \subseteq \mathbb{R}^{2T}$ is set up in the same way as in the unconstrained proof:

$$N := [x_0^{min}, x_0^{max}] \times \cdots \times [x_{T-1}^{min}, x_{T-1}^{max}] \times [0, M_0^{max}] \times \cdots \times [0, M_{T-1}^{max}]. \tag{4.55}$$

For any entry/exit-vector $(\mathbf{x}, \mathbf{M}) \in N$, the resulting output prices p_t^{mc} must be in $[p^{min}, p^{max}]$, and the exit premium satisfies $r_t \in [0, r^{max}]$. Consequently, the exit-points $\widetilde{x}_t$, which are related to $\mathbf{p}^*$ and $\mathbf{r}$, will lie inside the interval $[x_t^{min}, x_t^{max}]$. Due to the definition of $\widetilde{M_t}$, the vector $(\widetilde{\mathbf{x}}, \widetilde{\mathbf{M}}) = \tau(\mathbf{x}, \mathbf{M})$ will indeed be an element of the set N. Hence, we have $\tau : N \to N$, and can apply Brouwer's Fixed Point Theorem to conclude that a finite dynamic equilibrium exists in the constrained case. $\square$

4.5 Dynamics in the Western German dairy industry

4.5.1 Stylised facts: agricultural policy and development of the dairy farm size distribution

In 1984, the EU introduced the milk quota system as a reaction to the tremendous excess supply caused by massive price support from the 1970s onwards, e.g., intervention prices for butter and powder. In the initial years of the quota scheme, production rights were not transferable between farms. This restriction has been relaxed over time, from family transfers and regional but rental transfers–some attached to grassland-transfer–to official sales within auctions for Eastern and Western Germany from 2000 until 2015, though first restricted to be traded at the state-level only. With the 2003 reform of the Common Agricultural Policy (CAP), direct payments were decoupled from production levels. Considerable reductions in price regulations and other market interventions have led to increased milk price volatility. Within the health check of the reform of 2003, the fading out of the milk quotas was stipulated. Since then, price supports and other market interventions have been reduced with a stepwise enhancement of the overall production quantity limitation, to have ended in April 2015. These policy changes have likely contributed to falling milk prices at that time (cf. Figure A.1 in Appendix A.3). Hence, farms were exposed to further pressure to adjust, for instance, their production quantity, their size, or even decide about continuing production.

A frequently used indicator for industry dynamics is the development of farm-size distribution over time. The latter can be measured either by the number of cows or the milk output per farm. Here, we start with the number of cows per farm: in Western Germany, the number of dairy farms declined from 1,216,700 in 1960 to 90,200 in 2010 (Statistisches Bundesamt), while the average farm size increased, viz. from an average of 5 cows per farm in 1960 to 43 cows per farm in 2010 (Statistisches Bundesamt), with considerable increases in farm productivity at the same time. The average milk yield per cow increased from 3.6 tons per cow and year in 1964 to 6.9 in 2009. The rather strong consolidation process accompanied with an altered farm size distribution as visualised in Figure 4.1. While the share of the small farms (less than 10 cows per farm) sharply declined over time, the medium (10-49 cows) and large (more than 50 cows) farms increased in number and share of total number of

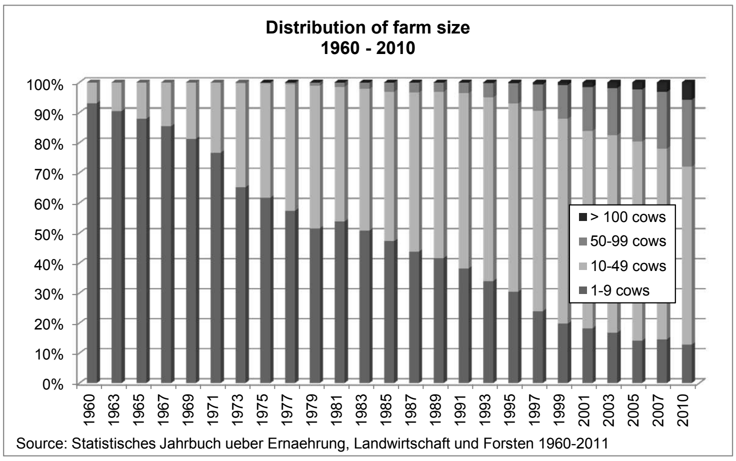

Figure 4.1: Dairy farm size distribution in Western Germany 1960-2010
Source: Statistisches Jahrbuch über Ernährung, Landwirtschaft und Forsten 1960-2011

dairy farms. The share of the large and very large farms particularly increased in more recent years (starting in the mid 1990s).

Figure 4.2 illustrates changes in the farm size distribution between the years 2000 and 2008 based on farm-specific data from the Farm Accountancy Data Network (FADN) for both the number of cows (left-hand side) and milk output (right-hand side). We estimated a kernel density based on a Gaussian kernel and used a distribution in natural logs for specialised dairy farms only. The right-shift of both distributions confirms the growth in average farm size and the industry's consolidation process, also for the shorter time period. Since the market has been influenced by the milk quota for more than 30 years, this begs the following questions: To what extent has the development been provoked by the quota, and how would the distribution look without quota limitations?

The empirical literature does not yet provide a clear answer on how the quota affected industry dynamics, or what will happen in the nearer future. Some argue that structural change in the dairy sector might be accelerated after the quota removal, where this effect is expected to be stronger the tighter the transfer rules of the milk quota in the quota period are (Bailey, 2002). Nevertheless, even in EU Member States where the quota trade scheme is rather well organised, e.g. the United Kingdom (UK), the milk quota scheme could have been proven to foster inefficient production structures (Colman, 2000; Colman et al., 2002). Moreover, as Oskam and Speijers (1992) show, considerable increases in the capital costs of farms that bought or leased quota might hinder investments in efficient production structure. Richards (1995) and Richards and Jeffrey (1997) even find evidence that the milk

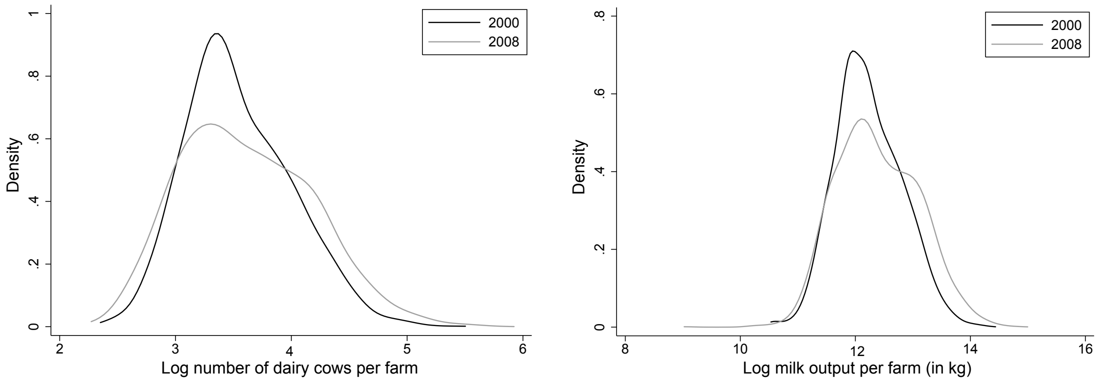

Figure 4.2: Dairy farm size distribution in 2000 and 2008 for Western Germany
Source: EU-FADN-DG AGRI 1997-2011

quota scheme reduces the investment rate of dairy farms in Canada, thus hindering farm growth and the necessary adaptations of technical progress. Thus, it is undisputed that even tradable quotas have an impact on the dairy production industry dynamics and their removal will influence farmers' decision-making.

In what follows, we will examine this effect in greater detail and calculate and compare the stationary equilibrium for two scenarios: the equilibrium of a capacity-constrained sector, that is, with production quotas, and an unconstrained sector, which reflects a potential situation after the production quota scheme. For this, we first calibrate the model to the German dairy sector as a benchmark.

4.5.2 Model calibration

Figure 4.2 indicates that the Western German dairy sector has obviously not been in a stationary equilibrium during recent years. We use 2003 as a reference point to calibrate our model to the Western German dairy sector. At this time the milk quota was already tradable among farms. Further, in 2003 the decoupling of direct payments from production was decided. The data is provided by the EU-FADN-DG AGRI 1997-2011, and the sample contains information on 1,500 specialised dairy farms between 1997 and 2011.

Cost function

We assume a single output and multiple input production technology. Output q is raw milk and assumed to be Cobb-Douglas in inputs $n_1, ..., n_k$ with a stochastic

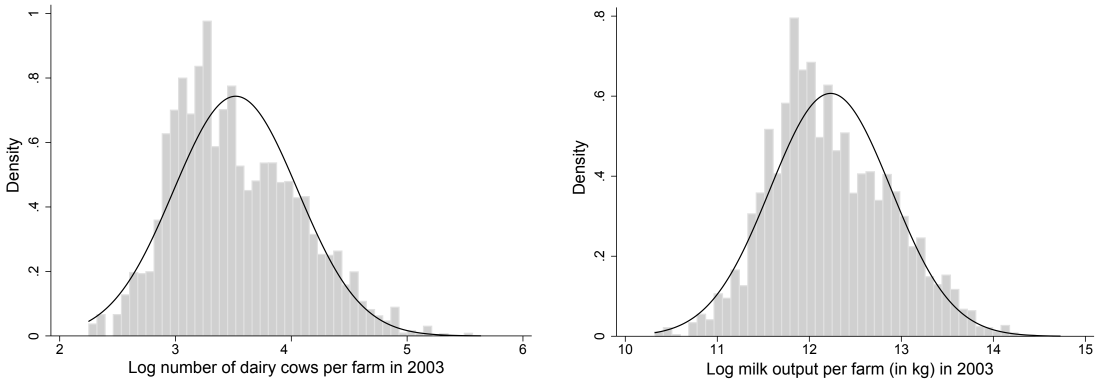

Figure 4.3: Distribution of dairy cows (left) and milk output (right) across farms in 2003 with adjusted normal distributions
Source: EU-FADN-DG AGRI 1997-2011

productivity component denoted by φ:

$$q = c\, n_1^{\alpha_1} \dots n_k^{\alpha_k}\, e^{\varphi}. \tag{4.56}$$

We use the estimates of Petrick and Kloss (2012) to calibrate the production function for Western German dairy farmers with the inputs labour, land, working capital, fixed capital and number of cows since these authors refer to the same data base. The cost function is given by:

$$c(\varphi, q) = h \left(\frac{q}{\exp(\varphi)} \right)^{\frac{1}{\alpha}}, \tag{4.57}$$

with constant term h and $\alpha = \sum \alpha_j$. For a given productivity φ the optimal firm-specific output level in period t is thus given by

$$q^*(\varphi, p_t) = \left(\frac{\alpha\, p_t}{h} \right)^{\frac{\alpha}{1-\alpha}} e^{\frac{\varphi}{1-\alpha}}. \tag{4.58}$$

This functional form implies that the optimal output level q^* follows a log-normal distribution if φ is normally distributed, that is,

$$\mu_t^{\varphi} = N(\xi_{\varphi}, \sigma_{\varphi}^2) \quad \Leftrightarrow \quad \mu_t^{q} = LN(\xi_q, \sigma_q^2). \tag{4.59}$$

Starting distribution of productivity levels across farms

We calibrate a starting distribution μ_0^{φ} using the farm-level output distribution in 2003. Figures 4.3 and 4.4 suggest that milk output is close to being log-normally distributed across farms. To fit a log-normal distribution μ_0^q to the observed firm-specific output values q_{it} (for farms $i = 1, ..., n$), we refer to (4.58), where the cost function parameter h is adjusted such that the corresponding distribution of productivity values φ_{it} is centered-normal: $\mu_0^\varphi = N(0, \sigma_\varphi^2)$.

Distribution of new farms

The nature of the farm accountancy data does not allow us to clearly define new firms. Therefore, we assume that the group of new firms is part of the group of investing firms (cf. Section 4.1), and select those farms from the sample that have increased their number of cows by at least 20% from 2002 until 2003. These farms' milk output distribution is similarly approximated, that is, by a log-normal one (see right side of Figure 4.4). In a second step, the output distribution is used for the normal distribution of productivity levels: $G = N(\xi_g, \sigma_g^2)$.

Demand function

The demand function for milk is supposed to have constant price elasticity η. Generally speaking, an isoelastic demand function is given by $Q(p) = b\, p^{-\eta}$, or the corresponding inverse demand function $D(Q) = \left(\frac{Q}{b}\right)^{-\frac{1}{\eta}}$. Here, we refer to Thiele (2008) to calibrate the demand for dairy products in Germany; this author reports a price elasticity of about $\eta = 1.00$. Parameter b will be adjusted such that the market clearing output price matches the observed one in 2003. Given that the average milk price is $p_{2003} = 0.32$ €/kg and the observed distribution of milk output denoted by μ_{2003}^q, the 2003 calibrated demand function is given by

$$p_{2003} = \left(\frac{1}{b}\int_0^\infty y\, d\mu_{2003}^q(y)\right)^{-\frac{1}{\eta}}. \tag{4.60}$$

Productivity process

According to the formal model, a farm's productivity is assumed to follow an AR(1)-process:

$$\varphi_{it} = \rho\, \varphi_{i,t-1} + \varepsilon_{it}, \quad \text{with} \quad \rho \in (0,1) \quad \text{and} \quad \varepsilon_{it} \sim N(\xi_\varepsilon, \sigma_\varepsilon^2). \tag{4.61}$$

As the output levels are optimally chosen with respect to a farm's productivity level φ_{it}, the process parameter ρ may be estimated by means of the milk output per farm. Let q_{it} denote the milk yield per farm i and year t; we normalise the values

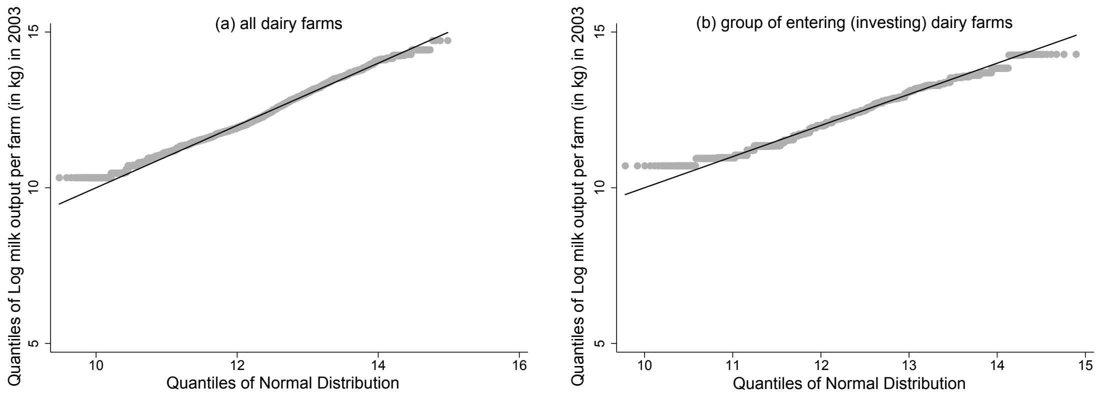

Figure 4.4: Quantiles of a normal distribution plotted against the quantiles of farms' log milk output (in kg) in 2003. The left side contains all farms while the right side displays new farms only, i.e., those which increased their number of cows by at least 20%.
Source: EU-FADN-DG AGRI 1997-2011

by the average farm level production over time $\bar{q}_t = \frac{1}{n}\sum_{j=1}^{n} q_{jt}$ as follows:

$$\widetilde{q_{it}} = \frac{q_{it}}{\bar{q}_t} = \frac{\exp\left(\frac{\varphi_{it}}{1-\alpha}\right)}{\frac{1}{n}\sum_{j=1}^{n}\exp\left(\frac{\varphi_{jt}}{1-\alpha}\right)}. \tag{4.62}$$

For simplicity we let $\bar{e}_t = \frac{1}{n}\sum_{j=1}^{n}\exp\left(\frac{\varphi_{jt}}{1-\alpha}\right)$. Taking logs on both sides yields

$$\begin{aligned}
\log(\widetilde{q_{it}}) &= \log\left[\frac{\exp\left(\frac{\varphi_{it}}{1-\alpha}\right)}{\bar{e}_t}\right] \\
&= \frac{\varphi_{it}}{1-\alpha} - \log(\bar{e}_t) \\
&= \frac{\rho\varphi_{i,t-1}+\varepsilon_{it}}{1-\alpha} - \log(\bar{e}_t) + \rho\log(\bar{e}_{t-1}) - \rho\log(\bar{e}_{t-1}) \\
&= \rho\log(\bar{e}_{t-1}) - \log(\bar{e}_t) + \rho\left[\frac{\varphi_{i,t-1}}{1-\alpha} - \log(\bar{e}_{t-1})\right] + \frac{\varepsilon_{it}}{1-\alpha} \\
&= \widetilde{a} + \rho\log(\widetilde{q_{i,t-1}}) + \widetilde{\varepsilon_{it}},
\end{aligned}$$

with $\widetilde{\varepsilon_{it}} \sim N\left(0, \frac{\sigma_\varepsilon^2}{(1-\alpha)^2}\right)$. We apply this log-log specification to the FADN data and estimate the parameters $\widetilde{a}, \rho$ and σ_ε^2 using a dynamic panel estimator in line with Arellano and Bond (1991).[6]

[6]The dynamic model applied to panel data causes the lagged dependent to be endogenous because of unobserved farm-specific effects. Accordingly, we use the second- and higher-order lags of the output variables as instruments and estimate the model by the generalised method of moments (cf. Bond (2002) for a similar model).

Fixed costs

We infer the fixed costs c_f from the development of the output distribution over time. For simplicity we presume the costs to be constant over incumbents. To estimate the average fixed costs we use the definition of the critical productivity level in (4.11). Given that the transition from μ_t^q to μ_{t+1}^q is induced by passing the exit-point x_t, our idea is to specify the fixed costs by estimating the critical productivity level as a threshold. By (4.11), each firm at the critical productivity level $\varphi = x_t$ must be indifferent between staying in or leaving the industry. Fixed costs are derived such that the continuation value of the indifferent firms exactly offsets the liquidation value, which is zero in the case without quota.

First we must calibrate the critical threshold level. Given that our base year is 2002, we log-normally approximate the distribution of milk output in 2002, denoted by μ_{2002}^q. Accordingly, the 2002 productivity levels follow a normal distribution μ_{2002}^φ. Second, according to (4.13) and (4.14) the distribution of productivity levels across farms is uniquely determined by the current distribution, the stochastic productivity process, and the entry/exit of farms. If $\mu_{2002}^\varphi = N(\xi, \sigma^2)$, then μ_{2003}^φ is given by the density function

$$m_{2003}^\varphi(z) = \frac{1}{\sqrt{2\pi(\sigma_\varepsilon^2 + \sigma^2\rho^2)}} \, e^{-\frac{(z-(\rho\xi+\xi_\varepsilon))^2}{2(\sigma_\varepsilon^2+\sigma^2\rho^2)}} (1 - \hat{F}(x_{2002})) + \frac{M_{2002}}{\sqrt{2\pi\sigma_g^2}} \, e^{-\frac{(z-\xi_g)^2}{2\sigma_g^2}}, \quad (4.63)$$

with $\hat{F}$ being the density function of a $N\left(\frac{\sigma^2\rho(z-\xi_\varepsilon)+\sigma_\varepsilon^2\xi}{\sigma_\varepsilon^2+\sigma^2\rho^2}, \frac{\sigma_\varepsilon^2\sigma^2}{\sigma_\varepsilon^2+\sigma^2\rho^2}\right)$ random variable. The function constitutes a probability density as long as the mass of entry equals the mass of exit. Presuming that $M_{2002} = \mu_{2002}^\varphi((-\infty, x_{2002}))$, the density function depends solely on the critical exit threshold x_{2002} for the given parameters.

Third, we will estimate the critical threshold level x_{2002} using the method of maximum likelihood (ML). To set up the likelihood function we use the observed and independent farm-specific output values in 2003 $(q_{1,2003}, ..., q_{n,2003})$ and the corresponding productivity levels $\varphi_{1,2003}, ..., \varphi_{n,2003}$ based on equation (4.58):

$$L(x_{2002}) = \prod_{k=1}^{n} m_{2003}^\varphi(\varphi_{k,2003}; x_{2002}). \quad (4.64)$$

Maximising the likelihood function in (4.64) yields the estimate of the threshold level.

The continuation value of a firm with productivity level $\varphi = x_{2002}$ must coincide with its liquidation value. If no tradable quota exists, firms giving up production will not be able to generate any positive liquidation value. That is why $r_{t+1} = 0$ must hold for all periods. Further, according to (4.11) the following equality must hold:

$$\int_{\mathbb{R}} v_{2003}(\varphi', y) dF(\varphi' | x_{2002}) = 0. \tag{4.65}$$

Since the continuation value contains the expected future profits, which in turn include fixed costs and depend on future output prices, we choose the level of fixed costs c_f such that (4.65) is satisfied. For simplicity, we assume a constant price sequence $y = \{p_{2003,0}\}$ consisting of the observed milk price in 2003.

Entry and quota costs

Given that the data contains a wide range of investment levels, determining a single and reliable value for farm-specific entry costs remains a challenge. For this reason we utilise the following three values, $c_e = 0/10,000/20,000$ € in our calculations. The low entry costs may describe a situation in which farms possess stable capacities that are unused due to quota constraints but can easily be converted to new production units. In contrast to this, the medium and high entry costs may refer to a case in which such unused stable capacities do not exist but must be built from scratch by new firms.

If the industry is constrained by a quota, new firms must buy production rights in addition to the constant entry costs. To determine quota costs in the model, we use the relation that under a tradable quota the quota costs coincide with the liquidation value. Referring to the theoretical model, the total industry mass $\mu_t(\mathbb{R})$ is determined by the balance of farms leaving and entering industry. For this reason, we define the quota costs as an upward-sloping function of the industry mass

$$k(\mu_t(\mathbb{R})) = \frac{c_f}{2} \exp\big[100\,(\mu_t(\mathbb{R}) - 1)\big]. \tag{4.66}$$

The function k rises extremely if the mass of new farms outweighs the exiting ones, and the total industry mass exceeds one. At this point, entry is no longer profitable, and hence $\mu_t(\mathbb{R}) = 1$ serves as a rough upper boundary to the industry's size. Under production quotas, entry costs are composed of the quota costs and the constant part c_e, that is, $k_t = c_e + k(\mu_t(\mathbb{R}))$. The exit premium r_t, which is the value that firms receive while selling their production rights in the case of exiting the sector, equals (4.66). By this definition we guarantee that the quota confines the total size of the industry.

Upper production boundary (for the quota case)
The maximum firm-specific output is bounded under a production quota. As the upper boundary must hold for all farms, it does not seem appropriate to simply take the maximum observable output value from the data as a proxy for this. In this case, the upper boundary would be so high that too many farms are left with the chance to grow and increase their output, thus making the production quota ineffective. Therefore, as a compromise we opted for a value $q^{\max} = 600\, t$ that is above the farms' average milk output, but below the maximum observable output level.

Table 4.1 summarises the calibrated parameter values. The estimated stochastic productivity process is $\rho = 0.99$, that is, close to one and almost a random walk. Together with the rather low volatility measure $\sigma_\varepsilon^2 = 0.0009$, this indicates that farms operate at a rather stable productivity level over time. That is, a farm with productivity $\varphi_t = \bar{\varphi}$ in period t will likely achieve a similar productivity level in the consecutive period $t + 1$. Note that one time period represents one year.

The fixed costs (c_f) value captures all costs in the production process that do not vary in production output such as fixed insurance rates, expenditures for pensions, as well as depreciation of machinery and buildings stemming from previous investments. Related to prices and output quantities observed, the specified fixed costs seem to be at a relatively low level. Still, every farm not able to cover this absolute value with its production profits will be forced to leave the industry. One should keep in mind that those fixed costs are also paid by new farms once they have entered the industry. Thus, the additional but constant entry costs c_e, comprise all additional investments that are obligatory to set up production but are not covered by the fixed costs and do not involve quota cost.

4.5.3 Equilibrium calculation

Because of the numerical effort it is convenient to calculate both the stationary and the dynamic equilibrium in a discrete framework. This requires us to determine a finite state space for the firm-specific productivity values and transform the originally continuous setup into a discrete one. Here, we apply the method described by Tauchen (1986).

In relation to the estimated distribution of new firms $G = N(\xi_g, \sigma_g^2)$ and the

	Functional form	Specified parameters
Demand function	$D(Q) = \left(\frac{Q}{b}\right)^{-\frac{1}{\eta}}$	$b = 81,470;\ \eta = 1.00$
Cost function	$c(\varphi, q) = h\left(\frac{q}{\exp(\varphi)}\right)^{\frac{1}{\alpha}}$	$h = 0.0376;\ \alpha = 0.86$
Productivity process	$\varphi_{t+1} = \rho\,\varphi_t + \varepsilon_{t+1}$, and $\varepsilon_{t+1} \sim N(\xi_\varepsilon, \sigma_\varepsilon^2)$	$\rho = 0.99;\ \xi_\varepsilon = -0.0027;\ \sigma_\varepsilon^2 = 0.0009$
Starting distribution	$\mu_0 = N(\xi_0, \sigma_0^2)$	$\xi_0 = 0.0000;\ \sigma_0^2 = 0.0085$
Distribution of new firms	$G = N(\xi_g, \sigma_g^2)$	$\xi_g = 0.0150;\ \sigma_g^2 = 0.0105$
Discount factor	–	$\beta = 0.9$
Fixed costs	–	$c_f = 4,602$ €
Entry costs (free access)	–	$c_e = 0/10,000/20,000$ €
Max. possible output	–	$q^{\max} = 600\,t$

Table 4.1: Utilised functional forms and estimated parameters
Source: Estimations based on Farm Accountancy Data Network

volatility of the AR(1)-process, we define the discrete state space as a (row) vector $s = (s_1, ..., s_n)$ consisting of $n = 1,000$ equidistant points inside the interval $[-0.4, 0.6]$. If $w = |s_{j+1} - s_j|$ denotes the distance between two consecutive states, the distribution of new firms is reflected by a stochastic (row) vector $g = (g_1, ..., g_n)$, with

$$g_j = \begin{cases} G\left(s_j + \frac{w}{2}\right), & \text{if} \quad j = 1 \\ G\left(s_j + \frac{w}{2}\right) - G\left(s_j - \frac{w}{2}\right), & \text{if} \quad 1 < j < n \\ 1 - G\left(s_j - \frac{w}{2}\right), & \text{if} \quad j = n. \end{cases} \tag{4.67}$$

The stochastic productivity process turns into a $n \times n$ transition matrix $P = (p_{ij})$ with conditional probability $p_{ij} = \text{Prob}(\varphi_{t+1} = s_j | \varphi_t = s_i)$. Based on the estimated AR(1)-process, we define the transition probability for any fixed productivity level $\varphi = s_i$ by

$$p_{ij} = \begin{cases} F\left(s_j + \frac{w}{2} | s_i\right), & \text{if} \quad j = 1 \\ F\left(s_j + \frac{w}{2} | s_i\right) - F\left(s_j - \frac{w}{2} | s_i\right), & \text{if} \quad 1 < j < n \\ 1 - F\left(s_j - \frac{w}{2} | s_i\right), & \text{if} \quad j = n. \end{cases} \tag{4.68}$$

Here, F denotes the conditional cdf that refers to the estimated AR(1)-process (see (4.3)). The state of the industry in period t is represented by a (row) vector $\mu^t = (\mu_1^t, ..., \mu_n^t)$ in which the value μ_j^t displays the mass of firms with productivity level $\varphi_t = s_j$. The total mass of the industry, which was denoted by $\mu_t(\mathbb{R})$ in the continuous framework, is now equal to $\mu^t(s) = \sum_{j=1}^n |\mu_j^t|$. The industry dynamics can be determined as follows. For any entry-mass M and exit-point $s_{i-1} < x \leq s_i$, we define the matrix

$$P_x = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \\ p_{i1} & \cdots & p_{in} \\ \vdots & & \vdots \\ p_{n1} & \cdots & p_{nn} \end{bmatrix}. \tag{4.69}$$

The industry structure in period $t+1$ is then given recursively by

$$\mu^{t+1} = \mu^t \cdot P_x + M \cdot g. \tag{4.70}$$

The aggregate industry output that arises from an industry structure μ^t and output price p_t is equal to

$$Q^s(\mu^t, p_t) = \mu^t \cdot \begin{bmatrix} q_t^*(s_1, p_t) \\ \vdots \\ q_t^*(s_n, p_t) \end{bmatrix}. \tag{4.71}$$

Hence, the market clearing output price $p_t^{mc}(\mu^t)$ is the unique solution to the equation $p_t = D(Q^s(\mu^t, p_t))$.

To calculate a dynamic equilibrium in the finite time horizon $T < \infty$, we apply the same fixed point concept as in the existence proof (see Section 4.4). This means, we construct a mapping $\tau : \mathbb{R}^{2T} \rightarrow \mathbb{R}^{2T}$ such that a fixed point of this mapping corresponds with the dynamic equilibrium. The numerical approach presented here can be used to compute a quota-constrained as well as a quota-free equilibrium.

For any given starting distribution μ^0 and entry/exit-vector $(\mathbf{x}, \mathbf{M}) \in \mathbb{R}^{2T}$, we can determine the industry structures $\mu^1, ..., \mu^T$ according to the (discrete) transition rule (4.70). Based on these structures, we can immediately derive the market clearing output prices $\mathbf{p^{mc}} = (p_1^{mc}, ..., p_T^{mc})$ and the entry costs $k_1, ..., k_T$. By definition $k_t = c_e$ if no quota exists, and $k_t = c_e + k(\mu^t(s))$ if the industry is subject to a quota. The firms' liquidation value is thus given by $r_t = k(\mu^t(s))$ in case of a tradable quota and by $r_t = 0$ otherwise. This allows us to specify the firms' value for every period t and productivity level s_i by means of backward induction. The value function in the last period is reflected by the (column) vector

$$v_T = \begin{bmatrix} \pi(s_1, p_T^{mc}) \\ \vdots \\ \pi(s_n, p_T^{mc}) \end{bmatrix}. \tag{4.72}$$

With this vector at hand, we can work backwards and specify the value function for all other times $t = 1, ..., T-1$ by

$$v_t = \begin{bmatrix} \pi(s_1, p_t^{mc}) \\ \vdots \\ \pi(s_n, p_t^{mc}) \end{bmatrix} + \beta \max \left\{ \begin{bmatrix} r_{t+1} \\ \vdots \\ r_{t+1} \end{bmatrix}, P \cdot v_{t+1} \right\}. \tag{4.73}$$

With respect to these vectors, we can now derive new exit-points $\widetilde{x}_t$ and entry-masses $\widetilde{M}_t$.

The firms' continuation value is given by the vector $v_{t+1}^c = P \cdot v_{t+1}$. This means, the i-th entry of the vector v_{t+1}^c represents the continuation value of a firm with current productivity level $\varphi_t = s_i$. By definition, the exit-point $\widetilde{x}_t$ refers exactly to the smallest value $s_i \in \{s_1, ..., s_n\}$ such that $v_{t+1}^c(s_i, \mathbf{p^{mc}}) \geq r_{t+1} > v_{t+1}^c(s_{i-1}, \mathbf{p^{mc}})$. The new entry-masses are computed as $\widetilde{M}_t = \max\{0\,,\, g \cdot v_{t+1} - k_{t+1} + M_t\}$. This defines the mapping $\tau(\mathbf{x}, \mathbf{M}) = (\widetilde{\mathbf{x}}, \widetilde{\mathbf{M}})$. The finite dynamic equilibrium can be derived with any algorithm that is suitable to find a fixed point of τ.

The approach to calculate a stationary equilibrium pretty much resembles the procedure for the finite dynamic case. Again, we construct a mapping τ, whose fixed point is equal to a stationary equilibrium. In contrast to the finite dynamic case, however, the dimension of this mapping reduces significantly as $\tau : \mathbb{R}^2 \to \mathbb{R}^2$.

Any combination of exit-point and entry-mass $(x, M) \in \mathbb{R}^2$ is associated to a stationary distribution by the condition $\mu = P_x \cdot \mu + M \cdot g$. If I denotes the $n \times n$ identity matrix, the stationary distribution can simply be computed with the formula

$$\mu = M \cdot g \cdot (I - P_x)^{-1}. \tag{4.74}$$

For the stationary industry structure μ we can derive the market clearing output price p^{mc}, the entry costs $k = c_e$ (quota-free) or $k = c_e + k(\mu(s))$ (quota-constrained), and the liquidation value r.

The value function is represented by a n-dimensional (column) vector v that contains the firm value for every productivity level $s_1, ..., s_n$ and satisfies the equation

$$v = \begin{bmatrix} \pi(s_1, p^{mc}) \\ \vdots \\ \pi(s_n, p^{mc}) \end{bmatrix} + \beta \max \left\{ \begin{bmatrix} r \\ \vdots \\ r \end{bmatrix}, P \cdot v \right\} \tag{4.75}$$

Due to the contraction property ($\beta < 1$), we can choose an arbitrary initial vector and iterate on the value function until a sufficient degree of convergence is reached. Analogous to the finite dynamic equilibrium, the firms' continuation value is $v^c = P \cdot v$, and the new exit-point $\widetilde{x}$ is the smallest value $s_i \in \{s_1, ..., s_n\}$ such that $v^c(s_i, p^{mc}) \geq r > v^c(s_{i-1}, p^{mc})$. The entry-mass $\widetilde{M} = \max\{0\ ,\ g \cdot v - k + M\}$ completes the mapping $\tau(x, M) = (\widetilde{x}, \widetilde{M})$. It is obvious that a stationary equilibrium coincides with the fixed point of the mapping τ. Hence, the stationary equilibrium can be computed with any suitable fixed point algorithm.

4.5.4 Findings

First we present the steady-state results. We have computed stationary equilibria for three scenarios, which reflect three different eras of the milk market policy in the EU. First is a scenario with a non-tradable milk quota that was effective in the early years. Second is a scenario with a tradable milk quota that came later, and third is a deregulated market without any production quota, which has been prevailing since April 2015. The calculations were carried out with the calibrated parameter values as presented in Table 4.1, and we applied the numerical method described above.

The stationary equilibrium results are presented in Table 4.2, and include the following: the product price p^*, the aggregate milk output Q^* (projected for Western Germany), the size of the dairy industry represented by the measure $\mu(\mathbb{R})$, the mass of firms entering the dairy industry M^*, the critical exit-point x^* defining the productivity threshold whether firms stay in the dairy industry, the average productivity $\bar{\varphi}$, as well as the average output of the firms $\bar{q}$. Further, we present the turnover rate that relates the mass of entering firms to the total size of the industry, and thus captures the long-run industry dynamics.

The level of milk prices as shown in Table 4.2 is rather close to observed milk prices in the EU. This shows that the calibration of the model produces reasonable results. As expected the price (aggregate output) is higher (lower) under quota regimes compared to an unregulated sector. The price (quantity) effects are higher for a non-tradable quota. We note that the differences in prices between the three scenarios depend on the size of the entry costs. The lower the entry costs are, the more pronounced are the differences in prices and quantities. This finding can be explained by the fact that entry costs have a similar effect to quota costs.

Comparing the turnover rates reveals that the industry dynamics are slowed down by a non-tradable quota. Over all levels of entry costs, the steady state distribution under a non-tradable production quota reveals the lowest average productivity level. This holds particularly under low or medium entry cost levels, where output prices

Scenario	Output Price p^* (ct/kg)	Agg. Output Q^* (Mio. t)	Total Ind. Mass $\mu(\mathbb{R})$	Entry Mass M^*	Exit Point x^*	Avg. Prod. $\bar{\varphi}$	Avg. Output $\bar{q}$ (in t)	Turnover Rate (in %)
$\mathbf{c_e = 0}$ €								
Non-trad. quota	29.007	24.548	1.015	0.041	-0.047	0.108	276.7	4.04
Trad. quota	27.222	26.158	1.014	0.246	0.101	0.165	295.1	24.26
No quota	25.198	28.258	1.427	0.351	0.102	0.165	226.6	24.60
$\mathbf{c_e = 10,000}$ €								
Non-trad. quota	29.025	24.533	1.020	0.040	-0.049	0.107	275.1	3.92
Trad. quota	28.016	25.417	1.006	0.088	0.024	0.146	289.1	8.75
No quota	27.318	26.066	1.008	0.078	0.015	0.141	296.0	7.74
$\mathbf{c_e = 20,000}$ €								
Non-trad. quota	29.113	24.459	1.012	0.038	-0.052	0.105	276.5	3.75
Trad. quota	28.757	24.762	0.997	0.051	-0.024	0.121	284.2	5.12
No quota	28.456	25.024	0.867	0.042	-0.028	0.119	330.4	4.84

Table 4.2: Stationary equilibrium outcome for the calibration shown in Table 4.1
Source: Author's own calculations

are considerably higher compared to the scenario without quota. This in turn leads to a relatively small exit-value, that is, even rather unproductive firms keep producing, whereas under free market conditions only more productive firms would stay in the industry. Accordingly, a lower turnover rate is observed under a non-tradable quota. Surprisingly, we find the opposite result for the tradable quota regime at least for medium and high entry costs. That is, the tradable quota does not reduce the speed of adjustment within the industry as compared to the non-tradable quota. This can be explained by the higher critical productivity (exit-point) that prevails under tradable quota. The larger exit-points reflect the opportunity to sell quota, which in turn increases the incentive for less productive firms to leave the industry.

Also noteworthy is the higher average firm-specific productivity level under a tradable quota regime compared to a scenario without a quota. This can again be explained by the increased incentive of unproductive firms to exit induced by the quota value serving as an exit premium. Moreover, the hurdle rate is higher for entering firms due to quota acquisition costs. An interesting outcome can be observed for the average firm-specific output level. When the entry costs are zero, the average output level under a production quota is higher than in the scenario without a quota. This is somewhat astonishing since the firm-specific output is bounded under a quota regime. As the optimal firm-specific output is an upward-sloping function of the output price, however, the significantly lower price causes firms to produce less output in the equilibrium. This is the reason why the unconstrained output level is lower on average.

The entire distribution of the productivity levels is displayed in Figure 4.5. The area below these curves coincides exactly with the total industry mass. Apparently,

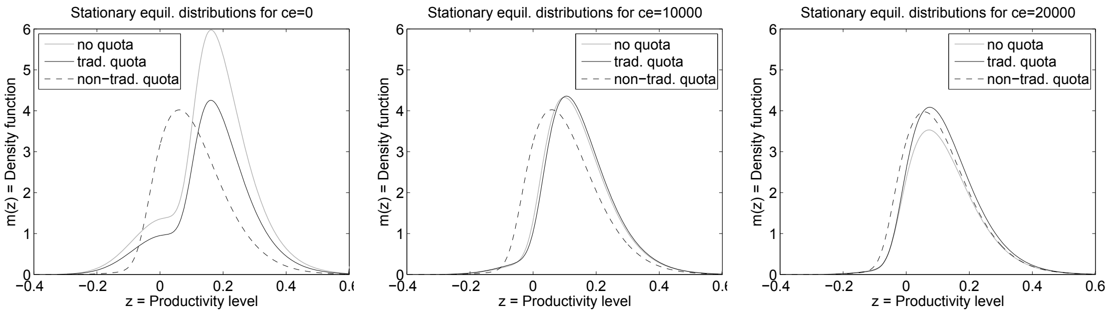

Figure 4.5: Stationary equilibrium distributions of productivity levels φ
Source: Author's own calculations

the distribution of firm-specific productivity for a non-tradable quota is not significantly affected by the level of the entry costs. In contrast to this, the productivity distribution in the no-quota case is rather sensitive to changes in the entry costs. If the entry costs are zero, the stationary distribution without a quota differs notably from the quota-constrained distributions. The graphs illustrate, however, that the differences between a quota-constrained and a quota-free stationary equilibrium decline as the entry costs increase. This finding is also supported by the results in Table 4.2. We therefore conclude that the impact of a production quota on the long-run equilibrium diminishes with increasing entry costs and can be be interpreted in the long-run as additional entry costs.

Relating these findings to the empirical distribution of the Western German milk production industry, consolidation processes have been noted over time. Zimmermann and Heckelei (2012) note that high prices, as was the case under the quota regime, might lead to industry structure conservation. In contrast to their empirical estimation approach, we find differences in the industry structure under different quota schemes, which can be traced back to the production limitation.

Next, we turn to the short-run adjustment of the industry. Tables 4.3–4.5 depict the evolution of the industry under a fading quota regime for the different levels of entry costs. Starting with the steady state equilibrium under a tradable quota, farmers anticipate that production quotas will be ceased after five years. This means that the time lag between the announcement and the actual end of the milk quota regime is five years, which mimics the situation that dairy farmers faced in the EU. We decided to restrict the total time horizon to 100 periods when calculating the dynamic equilibrium. Although it would be possible to consider even longer time horizons, we find that the additional effect, which a time extension would have on the crucial first 20 periods, is extremely small. We thus refrain from presenting more than 20 periods.

Period t	Output Price p_t^* (ct/kg)	Agg. Output Q_t^* (Mio. t)	Total Ind. Mass $\mu_t(\mathbb{R})$	Entry Mass M_t^*	Exit Point x_t^*	Avg. Prod. $\bar{\varphi}_t$	Avg. Output $\bar{q}_t$ (in t)	Quota Costs (ct/kg)
Stat. equil. (Trad. quota)	27.222	26.158	1.014	0.246	0.101	0.165	295.1	7.4
1	27.209	26.170	1.016	0.236	0.097	0.166	294.6	9.2
2	27.209	26.170	1.014	0.227	0.096	0.167	295.3	7.4
3	27.213	26.166	1.011	0.220	0.095	0.168	296.1	5.5
4	27.223	26.157	1.007	0.212	0.094	0.168	297.3	3.5
5	27.242	26.139	0.999	0.205	0.094	0.169	299.3	1.6
6	25.155	28.307	2.446	1.676	0.102	0.077	132.4	0
7	25.154	28.307	2.086	1.033	0.102	0.107	155.2	0
8	25.154	28.307	1.903	0.742	0.102	0.125	170.1	0
9	25.155	28.307	1.799	0.594	0.102	0.136	180.0	0
10	25.155	28.307	1.734	0.514	0.102	0.143	186.8	0
11	25.155	28.307	1.689	0.466	0.102	0.148	191.8	0
12	25.155	28.307	1.656	0.437	0.102	0.151	195.6	0
13	25.155	28.307	1.631	0.417	0.102	0.154	198.6	0
14	25.155	28.307	1.610	0.403	0.102	0.155	201.1	0
15	25.155	28.307	1.593	0.393	0.102	0.157	203.3	0
16	25.155	28.307	1.579	0.386	0.102	0.158	205.1	0
17	25.155	28.307	1.566	0.380	0.102	0.159	206.8	0
18	25.155	28.307	1.556	0.375	0.102	0.160	208.2	0
19	25.155	28.307	1.546	0.371	0.102	0.161	209.5	0
20	25.155	28.307	1.537	0.368	0.102	0.161	210.7	0
Stat. equil. (No quota)	25.198	28.258	1.427	0.351	0.102	0.165	226.6	0

Table 4.3: Dynamic equilibrium outcome for entry costs $c_e = 0$ €
Source: Author's own calculations

Table 4.3 describes the connecting path from the stationary equilibrium with tradable quota to the one without quota. We find that in the first five periods, a quota has a positive value to farms which is, however, declining. This development could actually be observed in the years prior to the quota being abolished in the EU. Due to the depreciating quota value, the exit option becomes less attractive to incumbents, and the critical productivity threshold for staying in the industry decreases in the last quota constrained periods.

As long as the industry is subject to the tradable production quota, the output price stays at a relatively high level. Moreover, the output price may even increase if the industry approaches the final periods of the quota regime. In period six, the first year without a quota, however, the output price drops significantly. If the entry costs are zero, this price decline is the result of the large mass of new firms and the increased total industry mass. Tables 4.4 and 4.5 show that the price effect is less pronounced for higher entry costs. In this case, the total size of the industry shrinks right after the quota removal since the number of exiting firms exceeds the mass of entrants. Hence, the declining output price is not provoked by the entry of additional production units but by the repeal of the firm-specific output boundary. Very productive farms whose output has been constrained by the quota can now extend their production. This leads to an increased aggregate industry output and

Period t	Output Price p_t^* (ct/kg)	Agg. Output Q_t^* (Mio. t)	Total Ind. Mass $\mu_t(\mathbb{R})$	Entry Mass M_t^*	Exit Point x_t^*	Avg. Prod. $\bar{\varphi}_t$	Avg. Output $\bar{q}_t$ (in t)	Quota Costs (ct/kg)
Stat. equil. (Trad. quota)	28.016	25.417	1.006	0.088	0.024	0.146	289.1	2.7
1	28.011	25.421	1.009	0.087	0.023	0.146	288.4	3.6
2	28.017	25.415	1.007	0.082	0.022	0.146	288.9	2.9
3	28.028	25.405	1.004	0.078	0.021	0.146	289.6	2.2
4	28.046	25.390	1.000	0.072	0.019	0.146	290.6	1.5
5	28.074	25.364	0.992	0.064	0.017	0.147	292.4	0.7
6	27.318	26.066	0.943	0.016	0.015	0.153	316.4	0
7	27.317	26.066	0.962	0.063	0.015	0.148	310.1	0
8	27.317	26.066	0.969	0.072	0.015	0.146	307.8	0
9	27.318	26.066	0.973	0.075	0.015	0.145	306.5	0
10	27.318	26.066	0.976	0.076	0.015	0.145	305.5	0
11	27.318	26.066	0.979	0.077	0.015	0.144	304.6	0
12	27.318	26.066	0.981	0.077	0.015	0.144	303.9	0
13	27.318	26.066	0.984	0.077	0.015	0.144	303.2	0
14	27.318	26.066	0.985	0.077	0.015	0.143	302.7	0
15	27.318	26.066	0.987	0.078	0.015	0.143	302.2	0
16	27.318	26.066	0.989	0.078	0.015	0.143	301.7	0
17	27.318	26.066	0.990	0.078	0.015	0.143	301.3	0
18	27.318	26.066	0.991	0.078	0.015	0.143	300.9	0
19	27.318	26.066	0.992	0.078	0.015	0.142	300.5	0
20	27.318	26.066	0.993	0.078	0.015	0.142	300.2	0
Stat. equil. (No quota)	27.318	26.066	1.008	0.078	0.015	0.141	296.0	0

Table 4.4: Dynamic equilibrium outcome for entry costs $c_e = 10000$ €
Source: Author's own calculations

goes along with a moderate price drop.

The average productivity levels were found to increase in the last five quota periods, irrespective of entry cost levels. What actually happens after quota removal depends on the size of entry costs, though. If the entry costs are zero, the average productivity level sharply declines after the market has been opened because of the new entrants and the considerably lower output price. But after reaching its peak in the first post-quota period, the industry starts to shrink in terms of production units (net exit of firms). This means more and more unproductive firms cease production, and the average productivity level soars again. In contrast to this, the average productivity level spikes immediately after the quota removal if the entry costs are at either a medium or high level. As less (or even no firms) enter the industry right after the quota has been abolished, the rise in average productivity is simply generated by the exit of the least productive firms. Nevertheless, the average productivity decreases again by the time more firms enter the industry.

The development of average output values resembles the sequence of average productivity levels to some extent. If the entry costs are zero, the average firm-specific output drops extremely after the quota removal and recovers thereafter. When the entry costs are higher, however, the average output level increases up to a peak value before it adjusts to the steady state solution. The latter case has exactly been

Period t	Output Price p_t^* (ct/kg)	Agg. Output Q_t^* (Mio. t)	Total Ind. Mass $\mu_t(\mathbb{R})$	Entry Mass M_t^*	Exit Point x_t^*	Avg. Prod. $\bar{\varphi}_t$	Avg. Output $\bar{q}_t$ (in t)	Quota Costs (ct/kg)
Stat. equil. (Trad. quota)	28.757	24.762	0.997	0.051	-0.024	0.121	284.2	1.0
1	28.753	24.765	0.998	0.052	-0.023	0.121	283.9	1.1
2	28.763	24.756	0.995	0.046	-0.024	0.121	284.6	0.8
3	28.784	24.738	0.990	0.040	-0.026	0.121	285.9	0.5
4	28.825	24.703	0.980	0.032	-0.028	0.122	288.5	0.2
5	28.956	24.592	0.942	0.000	-0.030	0.125	298.8	0.0
6	28.054	25.382	0.900	0.000	-0.021	0.129	322.5	0
7	28.129	25.314	0.871	0.000	-0.023	0.131	332.5	0
8	28.203	25.248	0.845	0.000	-0.024	0.132	342.0	0
9	28.276	25.183	0.821	0.000	-0.026	0.133	350.9	0
10	28.348	25.119	0.799	0.000	-0.027	0.134	359.9	0
11	28.420	25.055	0.776	0.000	-0.027	0.135	369.2	0
12	28.427	25.049	0.788	0.033	-0.028	0.131	363.5	0
13	28.427	25.049	0.797	0.040	-0.028	0.129	359.7	0
14	28.427	25.049	0.803	0.041	-0.028	0.128	356.9	0
15	28.427	25.049	0.808	0.041	-0.028	0.127	354.6	0
16	28.427	25.049	0.813	0.042	-0.028	0.126	352.5	0
17	28.427	25.049	0.817	0.042	-0.028	0.125	350.6	0
18	28.427	25.049	0.821	0.042	-0.028	0.124	348.9	0
19	28.427	25.049	0.825	0.042	-0.028	0.124	347.4	0
20	28.427	25.049	0.828	0.042	-0.028	0.123	346.0	0
Stat. equil. (No quota)	28.456	25.024	0.867	0.042	-0.028	0.119	330.4	0

Table 4.5: Dynamic equilibrium outcome for entry costs $c_e = 20000$ €
Source: Author's own calculations

observed in the European Union after the milk quota removal in April 2015. One reason might be that once the herd size has been determined, farmers tend to produce more to cover fixed costs. As noted by the EU commission's short term outlook (13/2015), this effect seems to be more pronounced in grass-fed production systems, that is, rather affordable forage systems.

Summing up, we can conclude that the short-run adjustments after a quota removal depend on the size of the fixed entry costs. This refers particularly to the mass of entering firms, the average productivity level, and the average firm output, as these parameters follow substantially different adjustment paths for low and high entry costs. The adjustment paths in Tables 4.3–4.5 indicate, however, that irrespective of the entry costs any parameter sequence converges to its unconstrained stationary solution. Hence, the industry tends to the quota-free steady state in the course of time.

4.6 Final remarks

This chapter has examined how a constrained sectoral production capacity in the agricultural industry affects farms' entry and exit decisions. We have presented a method to incorporate tradable/non-tradable production quotas into a dynamic stochastic framework with endogenous entry and exit of firms. The long-run effects of the quota on industry structure are evaluated by comparing steady state equilibria with and without production constraints. However, the concept of a stationary equilibrium does not fully capture the dynamics of an industry. To trace changes of the sector composition and the productivity distribution over time, we additionally calculate finite dynamic equilibria for a fading out scenario of the milk quota, as recently took place in the EU.

Our results have important implications for the economic appraisal of production quotas. Quotas have not only been criticised for negative welfare effects due to price distortions. It has also been argued that the introduction of production quotas hinders adjustment processes in an industry and thus retains inefficient production structures (Colman, 2000). Our results show that this simple view on the effect of production quotas needs to be qualified, particularly when discussing the fading out of production quotas. We find that a non-tradable quota reduces firm turnover and leads to lower average productivity levels, but this is not always true for a tradable quota. A tradable milk quota regime may instead increase the turnover as well as average productivity of firms. This happens if entry costs accrue that reduce firms' willingness to enter the market, irrespective of whether a quota exists or not. In that case the fluctuation of firms is rather low and a tradable quota provides an incentive for firms with low productivity to cease production that does not exist in an unregulated market.

The results of the finite dynamic equilibrium model are particularly useful for understanding the current development in the EU dairy sector. Our model correctly predicts the decline of the quota price during the fading out phase of the quota regime. The model also offers an explanation for the drop in milk prices that could be observed shortly before and after the removal of the quota. The further development of milk prices, however, hinges on the level of entry costs. In the case of low entry costs, the industry adjusts rather quickly to the new steady state equilibrium since new firms readily enter the market. In contrast, under high entry costs new farms are reluctant to enter the market, even though quota costs no longer accrue. The increase of aggregated output rather results from the productivity growth of incumbent farms, and prices eventually recover from the plunge after quota removal. In reality, entry costs vary considerably among firms. While some farms can activate idle production capacities at low marginal costs, others will incur considerable investment costs. Moreover, environmental regulations, such as the new phosphate directives in the Netherlands, constitute a limiting factor to dairy farms that is comparable to a production quota for milk.

In view of the current milk crisis, bailout plans for milk producers in jeopardy may help in the short run but this may also provide incentives for rather unproductive firms to continue. In view of our findings, such short-run market interventions should be discussed more critically since the turnover of firms cannot be prevented, and propping up unproductive firms for too long may only lead to a longer transition phase before reaching the steady state. Our results further contribute to the debate over using voluntary quantity limitations to stabilise milk prices: quantity limitations do not prevent industry dynamics–depending on other costs, dynamics might even be accelerated. Thus, from an economic perspective, the intervention in the short run is not well justified, and further de-regulation should be pushed forward rather than re-activating price support or production limitation schemes.

5. Conclusions

The agricultural industry is a sector that has undergone a significant consolidation process for quite some decades now. Exemplary for this is the decreasing number of farms on the one hand, and the increasing average farm size on the other hand. This thesis has investigated to what extent this development can be explained by limited sectoral production capacity that exists in terms of agricultural land or is generated by a production quota. As many other studies dealing with industry dynamics do not account for the resulting direct interdependency between firm entry and exit, the focus has been on incorporating this feature into a theoretic model. I opted for a dynamic stochastic framework here because it is suitable to consider the industry as a whole, and it simplifies the integration of a sectoral capacity limit. Moreover, the introduction of firm-specific uncertainty allows us to display the heterogeneous industry structure that exists in agriculture. The assumption of uncountable small firms forming the industry (a continuum) secures that changes in the industry structure are really induced by entry/exit decisions on a microeconomic level. A further benefit of this approach is that the uncertainty washes out at the aggregate level, and the evolution of industry structure follows deterministic paths. Changes in the sector's composition can thus be traced in much detail.

The analysis in Chapter 3 has shown that the industry adjusts to a steady state in the long-run no matter what the current industry structure is. Although many studies implicitly assume such a behaviour, it is by no means a trivial result. As I pointed out in Example 3.1, even small violations of the model assumptions may prevent the convergence to a stationary equilibrium and lead to an alternating development instead. Such a development can also occur, of course, if external parameters do not remain constant over time but are subject to variations. Increasing fixed or entry costs, a fluctuating consumer demand, or a changing entry distribution lead to a perpetual adjustment and prevent the persistence of a steady state. The proof illustrates, however, that during a period in which the external environment remains stable there is always the tendency to achieve a steady state. According to Section

3.6, this tendency is still valid when the entry distribution stochastically increases over time and converges to a limiting distribution. The speed, at which the industry approaches the ultimate steady state in this case, is clearly reduced though.

Although some interesting findings could be derived from this dynamic, stochastic modelling, the utilised framework has also some drawbacks. First, the model abstracts from output price stochasticity, which seems unrealistic considering the volatility of agricultural markets in general and milk markets in particular. Aggregate uncertainty could be introduced through demand shocks. However, this would generate real options effects and complicate the firms' decision problem considerably as the firms' value function would depend on two random processes. Therefore, most research studies consider either aggregate or idiosyncratic uncertainty, but do not include both at the same time. Second, we virtually treat production capacity as a homogeneous good, and assume that a unique capacity value emerges in an equilibrium. While this assumption certainly applies to a quota regime, it is less accurate for agricultural land. In fact, a wide range of price levels exists in agricultural land markets as both soil quality and regional circumstances, such as high livestock density or biogas production in a region, play an important role (cf. Hüttel et al., 2013). It is difficult, however, to account for such spatial differences in a dynamic stochastic equilibrium model. An econometric or game theoretic approach might be more suitable to explain such price differences and analyse local industry dynamics.

Although arable land features some anomalies as production factor, its limited availability seems to have a comparable impact on farm entry and exit as a tradable production quota. According to the stationary equilibria in Section 4.5.4, this means particularly that adjustment processes within the agricultural sector can be accelerated, and the firm turnover rate may be higher. These findings are based on the comparison of steady states. The question arises, however, whether the industry is actually in such a steady state, or whether it will at least attain such a state in the future. The observable changes in the firm size distribution indicate that the agricultural industry is rather not in a steady state (see Figure 1.1 or Figure 4.2). One possible explanation for this may be that investments in the agricultural industry are usually made by established farms and not so much by completely new farms. In fact, incumbents often possess different investment options than new firms, particularly when they can benefit from economies of size. The presented dynamic, stochastic equilibrium model, however, does not include economies of size. Any kind of firm, either new or established, must formally enter the industry to set up an additional marginal production unit. If the entry distribution remains constant, the average productivity of such a production unit is the same at all times. This might not be realistic, considering that predominantly established farms invest in additional production units, and also technological progress increases the firms' productivity. A possible way of adapting the model to this feature could be to proceed as in Section 3.6 and allow for shifts in the entry distribution. This could reflect the common observation that the structure in the agricultural industry is still changing.

The agricultural sector is not the only industry operating at a capacity limit. In fact, there are also some other prominent industries that have to deal with a capacity limit. Generally, such a shortage of production capacity arises either from natural reasons or is induced by institutional policies. The arable land area, for instance, constitutes a natural boundary to the total production capacity and falls into the first category. But, also the limited takeoff/landing spots at highly frequented airports can be cited here as they confine the maximum number of airlines operating at that airport (cf. Borenstein, 1988). In contrast to this, the second case occurs whenever governments decide to issue broadcasting licenses or introduce a quota system (cf. Hoppe et al., 2006). Examples for this are the milk quota, of course, or the carbon dioxide (CO_2) emission rights, which are intended to reduce the greenhouse gas emissions and have been imposed essentially on manufacturing and energy generating firms (cf. Ishikawa and Kiyono, 2006; Elliott et al., 2010; Scotchmer, 2011).

These are just few examples of industries that have already been affected by limited sectoral production capacity. Although the milk quota in the European Union was abolished recently, the impact of limited capacity on the industry dynamics might gain even more attention in the future. Particularly industries relying on non-reproducible production factors or natural resources might be concerned by limited production capacity. The introduction of CO_2 certificates by various governments shows that also ecological damages caused by firms belong to this category. Although I particularly focused on the milk quota in Western Germany, the utilised framework, as well as some major results, can surely be transferred to other quota regimes, too. Therefore, this work offers some interesting insights for those policy-makers, who decide on introducing quota systems. Nevertheless, the impact of a production quota on the industry dynamics should not be the only decision criterion when it comes to introducing or abolishing a quota regime. A further important question related to this topic but going beyond the scope of this study is how a quota affects producer and consumer surplus. While a production quota often leads to a higher output price, and thus entails a lower consumer surplus, its impact on the producer surplus is not clearly defined. A deeper investigation of this issue is necessary though for a final assessment of production quota.

A. Appendix

A.1 General mathematical concepts and theorems

Theorem A.1 (Banach Fixed Point Theorem). *Suppose that N is a non-empty, closed set in a Banach space, and the operator $\tau : N \to N$ is contractive, i.e.*

$$\|\tau(x) - \tau(y)\| \leq \rho \|x - y\|$$

for all $x, y \in N$ and some $0 \leq \rho < 1$. Then τ has a fixed point, which means, there exists an element $x \in N$ such that $\tau(x) = x$.

Theorem A.2 (Brouwer Fixed Point Theorem). *Suppose that N is a non-empty, compact, convex subset of $\mathbb{R}^n$, where $n \geq 1$, and that $\tau : N \to N$ is a continuous mapping. Then τ has a fixed point, which means, there exists an element $x \in N$ such that $\tau(x) = x$.*

Theorem A.3 (Schauder Fixed Point Theorem). *Let N be a nonempty, compact, convex subset of a Banach space and $\tau : N \to N$ a continuous mapping. Then τ has a fixed point, which means, there exists an element $x \in N$ such that $\tau(x) = x$.*

Theorem A.4 (Dini's Theorem on Uniform Convergence). *Let K be a compact metric space and $f : K \to \mathbb{R}$ a continuous function. Let $f_n : K \to \mathbb{R}$ $(n \in \mathbb{N})$ be a sequence of continuous function such that $\{f_n\}_{n\in\mathbb{N}}$ converges pointwise to f and*

$$f_{n+1}(x) \geq f_n(x), \quad \forall\, x \in K \text{ and } n \in \mathbb{N}.$$

Then the convergence is uniform, meaning $\|f_n - f\|_{\sup} = \sup_{x\in K} |f_n(x) - f(x)| \to 0$ for $n \to \infty$. The same statement applies to a decreasing sequence $f_{n+1} \leq f_n$.

The total variation norm (TV-norm) is defined for signed measures $\mu \in \mathcal{M}(S,\mathcal{B})$. Signed measures, in general, do not need to be positive. They may assign either positive or negative values to sets $A \in \mathcal{B}(S)$. The definition of the TV-norm rests upon the next theorem.

Theorem A.5 (Jordan Decomposition Theorem)**.** *Let $\mu \in \mathcal{M}(S,\mathcal{B})$ be a signed measure and $S_+, S_- \in \mathcal{B}(S)$ a disjoint decomposition of S such that $\mu(A) \leq 0$ for $A \subseteq S_-$ and $\mu(A) \geq 0$ for $A \subseteq S_+$. In this case uniquely determined non-negative measures μ^+ and μ^- exist that satisfy:*

(i) $\mu = \mu^+ - \mu^-$

(ii) $\mu^+(A) = \mu(A \cap S_+)$ *and* $\mu^-(A) = \mu(A \cap S_-)$

Definition A.1 (Total variation norm)**.** Let $\mu \in \mathcal{M}(S,\mathcal{B})$ be a signed measure with Jordan decomposition (μ^+, μ^-). The measure

$$|\mu| := \mu^+ + \mu^- \tag{A.1}$$

is called variation of μ. The total variation is defined as the value

$$\|\mu\|_{TV} := |\mu|(S) = \mu^+(S) + \mu^-(S). \tag{A.2}$$

An equivalent formulation of the total variation norm is

$$\|\mu\|_{TV} = \sup\left\{\sum_{k=1}^{n} |\mu(A_k)| : A_1, ..., A_n \text{ form finite disjoint decomposition of S}\right\}.$$

Definition A.2 (Mutually singular measures)**.** Let $\mu, \lambda \in \mathcal{M}^+(S,\mathcal{B})$ be finite and non-negative measures. If there are disjoint sets $A, B \in \mathcal{B}$ such that $\mu(C) = \mu(C \cap A)$ and $\lambda(C) = \lambda(C \cap B)$ for all $C \in \mathcal{B}$, then μ and λ are mutually singular. **Notation:** $\mu \perp \lambda$.

Lemma A.6. *Let $\mu_1, \mu_2 \in \mathcal{M}^1(S, \mathcal{B})$ be probability measures. Then there are non-negative measures λ and α_1, α_2 such that:*

(i) $\mu_k = \lambda + \alpha_k$, *for* $k = 1, 2$ *and* $\alpha_1 \perp \alpha_2$

(ii) $\|\mu_1 - \mu_2\|_{TV} = 2 \sup_{A \in \mathcal{B}} |\mu_1(A) - \mu_2(A)| \leq 2$

Proof. A proof of these statements can, for instance, be found in Stokey et al. (1989). □

Definition A.3 (Strong convergence of probability measures)**.** Let $\{\mu_t\}_{t \in \mathbb{N}}$ be a sequence of probability measures. We say that μ_t converges strongly to the probability measure μ if $\|\mu_t - \mu\|_{TV} \to 0$ for $t \to \infty$.

Definition A.4 (Weak convergence of probability measures)**.** Let $\{\mu_t\}_{t \in \mathbb{N}}$ be a sequence of probability measures. We say that μ_t converges weakly to the probability measure μ if

$$\lim_{t \to \infty} \int_S f \, d\mu_t = \int_S f \, d\mu \tag{A.3}$$

for all bounded, continuous functions f.

Lemma A.7. *(i) When F_t, F mark the cumulative distribution functions associated with μ_t, μ, an equivalent formulation of weak convergence is given by: $F_t(x) \to F(x)$ for all continuity points x of F. (ii) Strong convergence implies weak convergence.*

Definition A.5 (First order stochastic dominance)**.** Let $\mu_1, \mu_2 \in \mathcal{M}^+(S, \mathcal{B})$ be bounded, non-negative measures that have the same total variation $\mu_1(S) = \mu_2(S)$. We say that μ_1 stochastically dominates μ_2 if

$$\int_S f \, d\mu_1 \geq \int_S f \, d\mu_2 \tag{A.4}$$

for any bounded, measurable, and increasing function f. (**Notation:** $\mu_1 \overset{FOSD}{\geq} \mu_2$) When μ_1 and μ_2 are particularly probability measures with cdfs F_1 and F_2, first order stochastic dominance is equivalent to $F_1(x) \leq F_2(x)$ for all $x \in S$.

Lemma A.8. *Let F, G be cumulative distribution functions such that $G(x) < F(x)$ for all $x \in (a, b)$. That means, G strictly dominates F. If $f(x)$ is a strictly increasing (continuous) function, we have:*

$$\int_a^b f(x)\, dF(x) < \int_a^b f(x)\, dG(x). \tag{A.5}$$

Proof. Integrating both expressions by parts yields

$$\int_a^b f(x)\, dF(x) = f(x)F(x)\Big|_a^b - \int_a^b f'(x)F(x)\, dx$$

and

$$\int_a^b f(x)\, dG(x) = f(x)G(x)\Big|_a^b - \int_a^b f'(x)G(x)\, dx.$$

This implies:

$$\begin{aligned}\int_a^b f(x)\, dG(x) - \int_a^b f(x)\, dF(x) &= f(x)\left(G(x) - F(x)\right)\Big|_a^b - \int_a^b f'(x)\left(G(x) - F(x)\right) dx \\ &= \int_a^b \underbrace{f'(x)}_{>0}\underbrace{\left(F(x) - G(x)\right)}_{>0}\, dx \\ &> 0\end{aligned}$$

But, this had to be shown. □

A.2 Lemmas and proofs used in Chapter 3

Lemma A.9. *Let $\lambda \in \mathcal{M}^1(S, \mathcal{B})$ be a fixed probability measure. For all probability measures μ that satisfy $\|\mu - \lambda\|_{TV} \leq \eta$ the following inequalities hold:*

(i) $\mu + \frac{\eta}{2}\,\delta_b \overset{FOSD}{\geq} \lambda + \frac{\eta}{2}\,\delta_a$

(ii) $\mu + \frac{\eta}{2}\,\delta_a \overset{FOSD}{\leq} \lambda + \frac{\eta}{2}\,\delta_b$

Notice here that δ_a and δ_b denote the Dirac measures in a and b.

Proof. Recall that the stochastic ordering $\overset{FOSD}{\geq}$ is actually defined for probability measures. This concept is extended here to bounded measures that have the same total mass. Without loss of generality let $\|\mu-\lambda\|_{TV} = \eta$. According to Lemma (A.6) there exist measures $\gamma, \alpha_1, \alpha_2$ such that $\mu = \gamma + \alpha_1$, $\lambda = \gamma + \alpha_2$ and $\alpha_1 \perp \alpha_2$. We have

$$
\begin{aligned}
\|\mu - \lambda\|_{TV} &= \|\alpha_1 - \alpha_2\|_{TV} \\
&= \alpha_1(S) + \alpha_2(S) \\
&= 2\,\alpha_1(S)
\end{aligned}
$$

and obtain $\alpha_1(S) = \frac{\eta}{2}$. Because $\alpha_1(S) \cdot \delta_b \overset{FOSD}{\geq} \alpha_1, \alpha_2 \overset{FOSD}{\geq} \alpha_1(S) \cdot \delta_a$, we can conclude that

$$
\begin{aligned}
& \gamma + \alpha_1(S) \cdot \delta_b + \alpha_1 \overset{FOSD}{\geq} \gamma + \alpha_2 + \alpha_1(S) \cdot \delta_a \\
\Leftrightarrow \qquad & \mu + \frac{\eta}{2}\,\delta_b \overset{FOSD}{\geq} \lambda + \frac{\eta}{2}\,\delta_a.
\end{aligned}
$$

This proves statement (i). The inequality (ii) follows by the same argument. □

Lemma A.10. *The derivative of the capacity function κ is bounded by:*

$$
\frac{g_F\, h'_{\min}}{(1-\beta)^2} \leq \kappa'(Q) \leq 0, \quad \textit{for all} \quad Q \in [Q_{\min}, Q_{\max}]. \tag{A.6}
$$

Proof. We fix an aggregate output level $\bar{Q} \in [Q_{\min}, Q_{\max}]$. Recall that the function κ has been implicitly defined as solution to $v^e(Q, \kappa(Q)) - \kappa(Q) = c_e$. This implies $\frac{d}{dQ}\Big(v^e(\bar{Q}, \kappa(\bar{Q})) - \kappa(\bar{Q})\Big) = 0$. By means of the chain rule that is equivalent to

$$
\begin{aligned}
0 &= \begin{bmatrix} \frac{\partial v^e}{\partial Q}(\bar{Q}, \kappa(\bar{Q})) & \frac{\partial v^e}{\partial y}(\bar{Q}, \kappa(\bar{Q})) \end{bmatrix} \cdot \begin{bmatrix} 1 \\ \kappa'(\bar{Q}) \end{bmatrix} - \kappa'(\bar{Q}) \\
&= \frac{\partial v^e}{\partial Q}(\bar{Q}, \kappa(\bar{Q})) + \frac{\partial v^e}{\partial y}(\bar{Q}, \kappa(\bar{Q})) \cdot \kappa'(\bar{Q}) - \kappa'(\bar{Q})
\end{aligned}
$$

and we can conclude

$$
\kappa'(\bar{Q}) = \frac{\frac{\partial v^e}{\partial Q}(\bar{Q}, \kappa(\bar{Q}))}{1 - \frac{\partial v^e}{\partial y}(\bar{Q}, \kappa(\bar{Q}))}. \tag{A.7}
$$

The boundaries for $\frac{\partial v^e}{\partial Q}$ and $\frac{\partial v^e}{\partial y}$, which have been derived in Lemma 3.2, thus yield the statement. □

Lemma A.11. *The following inequalities hold for all $\bar{Q} \in [Q_{\min}, Q_{\max}]$:*

(i) $0 \leq \int_a^b u(\varphi, \bar{Q}, \kappa(\bar{Q}))\, dG(\varphi) \leq \frac{g_F\, h(Q_{\min})}{1-\beta}$

(ii) $\frac{g_F\, h'_{\min}}{(1-\beta)^2} \leq \frac{d}{dQ} \left(\int_a^b u(\varphi, \bar{Q}, \kappa(\bar{Q}))\, dG(\varphi) \right) \leq 0$

Proof. **ad(i).** The first inequality is clear. According to the definition, we have the function $u(\varphi, \bar{Q}, \kappa(\bar{Q})) = \max\left\{\kappa(\bar{Q}), v^c(\varphi, \bar{Q}, \kappa(\bar{Q}))\right\}$. Since both terms inside the max option are bounded above by the value $\frac{g_F\, h(Q_{\min})}{1-\beta}$ (which follows directly from Lemma 3.3), this implies statement (i).

ad(ii). Recall that for any constant sequence of output and capacity values $\bar{z} = \{Q, y\}_{t\in\mathbb{N}}$, the solution to $v^c(x, \bar{z}) = y$ is represented by a continuous function $x = \gamma(Q, y)$ (see also (3.18) for this). Therefore, we can rephrase the function

$$\int_a^b u(\varphi, Q, y)\, dG(\varphi) = \int_a^{\gamma(Q,y)} y\, dG(\varphi) + \int_{\gamma(Q,y)}^b v^c(\varphi, Q, y)\, dG(\varphi). \tag{A.8}$$

Taking the partial derivatives of this expression with respect to Q and y yields:

$$\frac{\partial}{\partial Q} \left(\int_a^b u(\varphi, \bar{Q}, \bar{y})\, dG(\varphi) \right) = \int\limits_{\gamma(\bar{Q},\bar{y})}^{b} \frac{\partial v^c}{\partial Q}(\varphi, \bar{Q}, \bar{y})\, dG(\varphi) \tag{A.9}$$

$$\frac{\partial}{\partial y} \left(\int_a^b u(\varphi, \bar{Q}, \bar{y})\, dG(\varphi) \right) = G\big(\gamma(\bar{Q}, \bar{y})\big) - G(a) + \int\limits_{\gamma(\bar{Q},\bar{y})}^{b} \frac{\partial v^c}{\partial y}(\varphi, \bar{Q}, \bar{y})\, dG(\varphi). \tag{A.10}$$

Combining both parts and utilising the chain rule allows us to calculate the gradient

$$\begin{aligned}
0 &\geq \frac{d}{dQ} \left(\int_a^b u(\varphi, \bar{Q}, \kappa(\bar{Q}))\, dG(\varphi) \right) \\
&= \int\limits_{\gamma(\bar{Q},\bar{y})}^{b} \frac{\partial v^c}{\partial Q}(\varphi, \bar{Q}, \bar{y})\, dG(\varphi) + \kappa'(\bar{Q}) \left[G\big(\gamma(\bar{Q}, \bar{y})\big) - G(a) + \int\limits_{\gamma(\bar{Q},\bar{y})}^{b} \frac{\partial v^c}{\partial y}(\varphi, \bar{Q}, \bar{y})\, dG(\varphi) \right] \\
&\geq \frac{g_F\, h'_{\min}}{1-\beta} \Big(G(b) - G(\gamma(\bar{Q})) \Big) + \frac{g_F\, h'_{\min}}{(1-\beta)^2} \Big[G(\gamma(\bar{Q})) - G(a) + \beta \Big(G(b) - G(\gamma(\bar{Q})) \Big) \Big] \\
&= \frac{g_F\, h'_{\min}}{(1-\beta)^2} \underbrace{\Big(G(b) - G(a) \Big)}_{=1}.
\end{aligned}$$

Here, the first inequality is trivial. The second inequality follows from Lemma 3.2 and Lemma A.10. This concludes the proof. □

Lemma A.12. *The partial derivatives of the functions s_L and s_U with respect to η are bounded by*

$$\text{(i)}\ \ 0 \leq \frac{\partial s_L}{\partial \eta}(\bar{\eta}, \bar{\varphi}) \leq -\frac{h'_{\min}\,(Q_{\max} - Q_{\min})}{h^2(Q_{\max})}\left[\frac{\beta\,(2-\beta)\,g_F\,h(Q_{\min})}{(1-\beta)^2} + \frac{c_e}{2}\right]$$

$$\text{(ii)}\ \ \frac{h'_{\min}\,(Q_{\max} - Q_{\min})}{h^2(Q_{\max})}\left[\frac{\beta\,(2-\beta)\,g_F\,h(Q_{\min})}{(1-\beta)^2} + \frac{c_e}{2}\right] \leq \frac{\partial s_U}{\partial \eta}(\bar{\eta}, \bar{\varphi}) \leq 0$$

Proof. The statement is straightforward if we recall the definitions of $Q^{\lambda}_{\min}(\eta), Q^{\lambda}_{\max}(\eta)$ and apply the boundaries in Lemma A.11 to the first derivative. □

Lemma A.13. *Let $\{a_t\}_{t\in\mathbb{N}_0}$ and $\{b_t\}_{t\in\mathbb{N}_0}$ be bounded sequences such that at $\leq b_t$. If the distance $|a_t - b_t| \to 0$ for $t \to \infty$, then the cartesian product $N = \prod_{t=0}^{\infty}[a_t, b_t]$ is a compact, convex subset in the space of bounded sequences ℓ^∞.*

Proof. Note, first of all, that the space of bounded sequences is a Banach space. A subset N of a Banach space is compact if it is closed and has a finite covering of ε-balls (see Zeidler (1995) for instance). That is for every $\varepsilon > 0$ there is a finite number of vectors $V_1, ..., V_J \in N$ such that

$$\min_{1\leq j\leq J} \|\zeta - V_j\|_\infty \leq \varepsilon \quad \text{for all} \quad \zeta \in N.$$

In the space of bounded sequences, the norm of an element $\zeta = \{x_t\}_{t\in\mathbb{N}_0} \in \ell^\infty$ is defined as $\|\zeta\|_\infty = \sup_t |x_t|$.

To prove the compactness of the aforementioned cartesian product, we fix an $\varepsilon > 0$ and determine a $T \in \mathbb{N}$ such that $|a_t - b_t| < \varepsilon$ for all $t \geq T$. Obviously, the finite cartesian product $C = \prod_{t=0}^{T}[a_t, b_t] \subset \mathbb{R}^{T+1}$ is compact and, hence, possesses a finite covering of ε-balls. Consequently, a finite number of vectors $W_1, ..., W_J \in C$ can be found such that each element of C is in some ε-ball around W_j. By setting $V_j = (W_j, a_{T+1}, a_{T+2}, ...) \in N$ for each $j = 1, ..., J$, we can accordingly create a finite covering of ε-balls for the subset $N = \prod_{t=0}^{\infty}[a_t, b_t] \subset \ell^\infty$. This means for every sequence $\zeta \in N$ there is at least one element V_j with $\|\zeta - V_j\|_\infty \leq \varepsilon$. But, this implies the compactness of N.

To show the convexity of N, we assume that $\zeta_1, \zeta_2 \in N$ with $\zeta_1 = \{x_t\}_{t\in\mathbb{N}_0}$ and $\zeta_2 = \{y_t\}_{t\in\mathbb{N}_0}$. For any value $\alpha \in (0,1)$, we have $a_t \leq \alpha x_t + (1-\alpha) y_t \leq b_t$. Hence, the element $\alpha\zeta_1 + (1-\alpha)\zeta_2$ must be in the set N as well and the convexity is proven. □

A.3 Lemmas and proofs used in Chapter 4

Lemma A.14. *Let $\phi(z|\xi, \sigma^2)$ denote the pdf of a Normal distribution with mean ξ and variance σ^2. If $p(\varphi, z) = \frac{1}{\sqrt{2\pi\sigma_\varepsilon^2}} \exp\left(-\frac{(z-(\rho\varphi+\xi_\varepsilon))^2}{2\sigma_\varepsilon^2}\right)$ defines the probability density derived from an AR(1)-process, the following equality holds:*

$$\int_{\mathbb{R}} p(\varphi, z)\, \phi(\varphi|\xi, \sigma^2)\, d\varphi = \phi(z|\rho\xi + \xi_\varepsilon, \rho^2\sigma^2 + \sigma_\varepsilon^2). \tag{A.11}$$

Proof. It is fairly easy to verify that

$$\begin{aligned} p(\varphi, z)\, \phi(\varphi|\xi, \sigma^2) &= \phi(z|\rho\varphi + \xi_\varepsilon, \sigma_\varepsilon^2)\, \phi(\varphi|\xi, \sigma^2) \\ &= \phi(z|\rho\xi + \xi_\varepsilon, \rho^2\sigma^2 + \sigma_\varepsilon^2)\, \phi\left(\varphi \,\middle|\, \frac{\rho\sigma^2(z-\xi_\varepsilon) + \xi\sigma_\varepsilon^2}{\rho^2\sigma^2 + \sigma_\varepsilon^2}, \frac{\sigma^2\sigma_\varepsilon^2}{\rho^2\sigma^2 + \sigma_\varepsilon^2}\right). \end{aligned}$$

As the first term does not depend on φ anymore, and the second one is a Normal density integrating up to one, this implies the equality in (A.11). □

Lemma A.15. *Let $p(\varphi, z) = \frac{1}{\sqrt{2\pi\sigma_\varepsilon^2}} \exp\left(-\frac{(z-(\rho\varphi+\xi_\varepsilon))^2}{2\sigma_\varepsilon^2}\right)$ denote the probability density derived from an AR(1)-process, and let the function $f : \mathbb{R} \to \mathbb{R}_+$ be continuous, increasing, and integrable with respect to any Normal distribution. The function*

$$g(\varphi) := \int_{\mathbb{R}} f(z)\, p(\varphi, z)\, dz \tag{A.12}$$

is then continuous, increasing, and integrable, too.

Proof. The monotonicity of g follows from Lemma A.8, and the integrability with respect to Normal distributions follows from Lemma A.14. To show the continuity in φ, we assume that $\varphi_n \to \varphi$ for $n \to \infty$. Based on this sequence, we define the functions $h_n(z) = f(z)\, p(\varphi_n, z)$. As f is supposed to be continuous, this implies $h_n(z) \to f(z)\, p(\varphi, z)$ for all z. For $\bar{\varphi} = \sup_{n\in\mathbb{N}} \varphi_n$ we further have

$$0 \leq h_n(z) \leq f(z)\, p(\bar{\varphi}, z).$$

Since the function $f(z)\, p(\bar{\varphi}, z)$ is integrable with respect to z, we can apply the Theorem on Dominated Convergence and get

$$\lim_{n\to\infty} g(\varphi_n) = \lim_{n\to\infty} \int_{\mathbb{R}} h_n(z) dz = \int_{\mathbb{R}} f(z)\, p(\varphi, z)\, dz = g(\varphi). \tag{A.13}$$

This proves the continuity of g. □

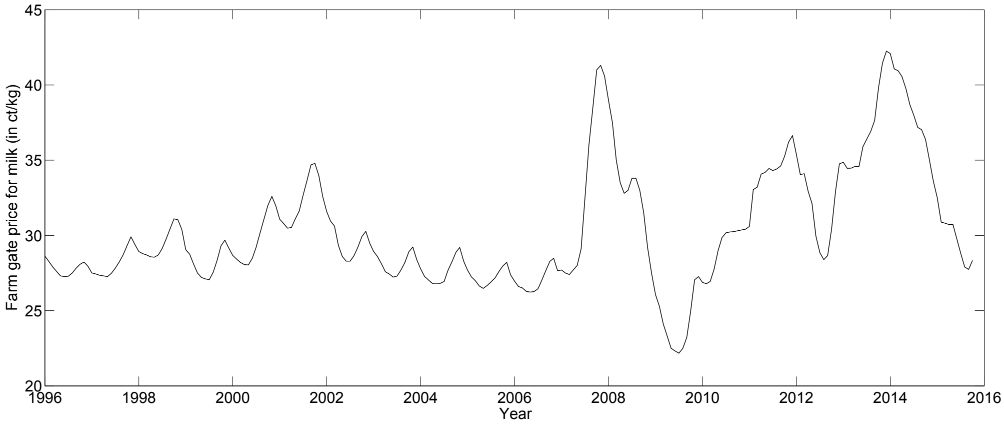

Figure A.1: Farm gate price for milk in Germany
Source: Zentrale Markt- und Preisberichtsstelle (1996-2010), Bundesanstalt für Landwirtschaft und Ernährung (2011-2015)

References

Adamopoulos, T. and D. Restuccia (2014). The size distribution of farms and international productivity differences. *The American Economic Review 104*(6), 1667–1697.

Arellano, M. and S. Bond (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The Review of Economic Studies 58*(2), 277–297.

Asplund, M. and V. Nocke (2006). Firm turnover in imperfectly competitive markets. *The Review of Economic Studies 73*(2), 295–327.

Bailey, A. (2002). Dynamic effects of quota removal on dairy sector productivity and dairy farm employment. In D. Coleman (Ed.), *Phasing out milk quotas in the EU. Main report.* University of Manchester, Centre for Agricultural, Food and Resource Economics.

Balmann, A., K. Dautzenberg, K. Happe, and K. Kellermann (2006). On the dynamics of structural change in agriculture: Internal frictions, policy threats and vertical integration. *Outlook on Agriculture 35*(2), 115–121.

Balmann, A., K. Kataria, and O. Musshoff (2013). Investment reluctance in supply chains: An agent-based real options approach. *Journal of Mathematical Finance 3*, 1–10.

Barichello, R. R. (1995). Overview of Canadian agricultural policy systems. In R. Loyns, R. Knutsen, and K. Meilke (Eds.), *Understanding Canada/United States Grain Disputes: Proceedings of First Canada/U.S. Agricultural and Food Policy Systems Information Workshop*, pp. 37–59. Winnipeg: Friesen Printers.

Bento, P. (2014). Niche firms, mass markets, and income across countries: Accounting for the impact of entry costs. *Journal of Economic Dynamics and Control 48*(0), 147–158.

Besanko, D. and U. Doraszelski (2004). Capacity dynamics and endogenous asymmetries in firm size. *The RAND Journal of Economics 35*(1), 23–49.

Blackwell, D. (1965). Discounted dynamic programming. *The Annals of Mathematical Statistics 36*(1), 226–235.

Bond, S. R. (2002). Dynamic panel data models: A guide to micro data methods and practice. *Portuguese Economic Journal 1*(2), 141–162.

Borenstein, S. (1988). On the efficiency of competitive markets for operating licenses. *The Quarterly Journal of Economics 103*(2), 357–385.

Caballero, R. J. and R. S. Pindyck (1996). Uncertainty, investment, and industry evolution. *International Economic Review 37*(3), 641–662.

Cabral, L. M. B. (1997). Competitive industry dynamics: A selective survey of facts and theories. Technical report, Mimeo, London Business School.

Caves, R. E. (1998). Industrial organization and new findings on the turnover and mobility of firms. *Journal of Economic Literature 36*(4), 1947–1982.

Chavas, J.-P. (2001). Structural change in agricultural production: Economics, technology and policy. *Handbook of agricultural economics 1*, 263–285.

Colman, D. (2000). Inefficiencies in the UK milk quota system. *Food Policy 25*(1), 1–16.

Colman, D., M. Burton, D. Rigby, and J. Franks (2002). Structural change and policy reform in the UK dairy sector. *Journal of Agricultural Economics 53*(3), 645–663.

Das, S. and S. P. Das (1997). Dynamics of entry and exit of firms in the presence of entry adjustment costs. *International Journal of Industrial Organization 15*(2), 217–241.

Dixit, A. (1989). Entry and exit decisions under uncertainty. *Journal of Political Economy 97*(3), 620–638.

Doraszelski, U. and A. Pakes (2007). Chapter 30: A framework for applied dynamic analysis in IO. In M. Armstrong and R. Porter (Eds.), *Handbook of Industrial Organization, Volume 3*, Volume 3 of *Handbook of Industrial Organization*, pp. 1887–1966. Elsevier.

Dunne, T., S. D. Klimek, M. J. Roberts, and D. Y. Xu (2013). Entry, exit, and the determinants of market structure. *The RAND Journal of Economics 44*(3), 462–487.

Elliott, J., I. Foster, S. Kortum, T. Munson, F. P. Cervantes, and D. Weisbach (2010). Trade and carbon taxes. *The American Economic Review 100*(2), 465–469.

Ericson, R. and A. Pakes (1995). Markov-perfect industry dynamics: A framework for empirical work. *The Review of Economic Studies 62*(1), 53–82.

Esö, P., V. Nocke, and L. White (2010). Competition for scarce resources. *The RAND Journal of Economics 41*(3), 524–548.

European Commission (2015). Fact sheet: Comprehensive package of measures.

Fariñas, J. C. and S. Ruano (2005). Firm productivity, heterogeneity, sunk costs and market selection. *International Journal of Industrial Organization 23*(7-8), 505–534.

Feil, J.-H. and O. Musshoff (2013). Modelling investment and disinvestment decisions under competition, uncertainty and different market interventions. *Economic Modelling 35*(0), 443–452.

Feldman, M. and C. Gilles (1985). An expository note on individual risk without aggregate uncertainty. *Journal of Economic Theory 35*(1), 26–32.

Foltz, J. D. (2004). Entry, exit, and farm size: Assessing an experiment in dairy price policy. *American Journal of Agricultural Economics 86*(3), 594–604.

Féménia, F. and A. Gohin (2011). Dynamic modelling of agricultural policies: The role of expectation schemes. *Economic Modelling 28*(4), 1950–1958.

Gomes, J. F. (2001). Financing investment. *The American Economic Review 91*(5), 1263–1285.

Hanazono, M. and H. Yang (2009). Dynamic entry and exit with uncertain cost positions. *International Journal of Industrial Organization 27*(3), 474–487.

Hopenhayn, H. A. (1992a). Entry, exit, and firm dynamics in long run equilibrium. *Econometrica 60*(5), 1127–1150.

Hopenhayn, H. A. (1992b). Exit, selection, and the value of firms. *Journal of Economic Dynamics and Control 16*(3-4), 621–653.

Hopenhayn, H. A. and E. C. Prescott (1992). Stochastic monotonicity and stationary distributions for dynamic economies. *Econometrica 60*(6), 1387–1406.

Hoppe, H. C., P. Jehiel, and B. Moldovanu (2006). License auctions and market structure. *Journal of Economics & Management Strategy 15*(2), 371–396.

Hüttel, S., M. Odening, K. Kataria, and A. Balmann (2013). Price formation on land market auctions in east germany – an empirical analysis. *German Journal of Agricultural Economics 62*(2), 99–115.

Ishikawa, J. and K. Kiyono (2006). Greenhouse-gas emission controls in an open economy. *International Economic Review 47*(2), 431–450.

Jorgenson, D. W. and M. P. Timmer (2011). Structural change in advanced nations: A new set of stylised facts. *Scandinavian Journal of Economics 113*(1), 1–29.

Jovanovic, B. (1982). Selection and the evolution of industry. *Econometrica 50*(3), 649–670.

Judd, K. L. (1985). The law of large numbers with a continuum of IID random variables. *Journal of Economic Theory 35*(1), 19–25.

Kersting, S., S. Hüttel, and M. Odening (2016). Industry dynamics under production constraints – the case of the EU dairy sector. *Economic Modelling 55*, 135–151.

Kolmogorov, A. N. and S. V. Fomin (2012). *Introductory real analysis.* Courier Corporation.

Leahy, J. V. (1993). Investment in competitive equilibrium: The optimality of myopic behavior. *The Quarterly Journal of Economics 108*(4), 1105–1133.

Leombruni, R. and M. Richiardi (2005). Why are economists sceptical about agent-based simulations? *Physica A: Statistical Mechanics and its Applications 355*(1), 103–109.

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica 71*(6), 1695–1725.

Miao, J. (2005). Optimal capital structure and industry dynamics. *The Journal of Finance 60*(6), 2621–2659.

Novy-Marx, R. (2007). An equilibrium model of investment under uncertainty. *Review of Financial Studies 20*(5), 1461–1502.

Oskam, A. and D. Speijers (1992). Quota mobility and quota values: Influence on the structural development of dairy farming. *Food Policy 17*(1), 41–52.

Petrick, M. and M. Kloss (2012). Drivers of agricultural capital productivity in selected EU member states. *Factor Markets Working Paper 30*, 1–41.

Piet, L., L. Latruffe, C. Le Mouël, and Y. Desjeux (2012). How do agricultural policies influence farm size inequality? The example of France. *European Review of Agricultural Economics 39*(1), 5–28.

Richards, T. J. (1995). Supply management and productivity growth in Alberta dairy. *Canadian Journal of Agricultural Economics 43*(3), 421–434.

Richards, T. J. and S. R. Jeffrey (1997). The effect of supply management on herd size in Alberta dairy. *American Journal of Agricultural Economics 79*(2), 555–565.

Scotchmer, S. (2011). Cap-and-trade, emissions taxes, and innovation. *Innovation Policy and the Economy 11*(1), 29–54.

Stokey, N., R. Lucas, and E. Prescott (1989). *Recursive Methods in Economic Dynamics.* Harvard University Press.

Sutton, J. (1991). *Sunk costs and market structure: Price competition, advertising, and the evolution of concentration.* The MIT press.

Syverson, C. (2004). Market structure and productivity: A concrete example. *Journal of Political Economy 112*(6), 1181–1222.

Tauchen, G. (1986). Finite state markov-chain approximations to univariate and vector autoregressions. *Economics letters 20*(2), 177–181.

Thiele, S. (2008). Elastizitäten der Nachfrage privater Haushalte nach Nahrungsmitteln – Schätzung eines AIDS auf Basis der Einkommens- und Verbrauchsstichprobe 2003. *Agrarwirtschaft (German Journal of Agricultural Economics) 57*(5), 258–268.

Verikios, G. and X. Zhang (2013). Structural change in the australian electricity industry during the 1990s and the effect on household income distribution: A macro–micro approach. *Economic Modelling 32*, 564–575.

Weiss, C. R. (1999). Farm growth and survival: Econometric evidence for individual farms in Upper Austria. *American Journal of Agricultural Economics 81*(1), 103–116.

Wolf, C. A. and D. A. Sumner (2001). Are farm size distributions bimodal? Evidence from kernel density estimates of dairy farm size distributions. *American Journal of Agricultural Economics 83*(1), 77–88.

Zeidler, E. (1995). *Applied Functional Analysis - Applications to Mathematical Physics*. Applied Mathematical Sciences. Springer New York.

Zepeda, L. (1995). Asymmetry and nonstationarity in the farm size distribution of Wisconsin milk producers: An aggregate analysis. *American Journal of Agricultural Economics 77*(4), 837–852.

Zimmermann, A. and T. Heckelei (2012). Structural change of european dairy farms – a cross-regional analysis. *Journal of Agricultural Economics 63*(3), 576–603.

Hiermit versichere ich, dass ich die vorliegende Dissertation selbständig und nur unter Verwendung der angegebenen Quellen und Hilfsmittel verfasst habe.

Berlin, den 29. Juni 2016