European University Institute

Department of Economics

Essays on Innovations, Technology and the Labor Market

Stephan Fahr

Thesis submitted for assessment with a view to obtaining the degree of
Doctor of Economics of the European University Institute

Florence, June 2007

EUROPEAN UNIVERSITY INSTITUTE
Department of Economics

Essays on Innovations, Technology and the Labor Market

Stephan Fahr

Thesis submitted for assessment with a view to obtaining the degree of Doctor of Economics of the European University Institute

Examining Board:

Prof. Omar Licandro, EUI Supervisor
Prof. Salvador Ortigueira, EUI
Prof. Diego Comin, New York University
Prof. Franck Portier, University of Toulouse

Fahr, Stephan (2007), Essays on Innovations, Technology and the Labor Market
European University Institute DOI: 10.2870/44384

Contents

Acknowledgements

Conducting – and eventually “producing” research is a complex task. For it to be successful, numerous different ingredients are required. Next to institutional settings, infrastructure and financial support by the European University institute and the "Deutscher Akademische Austauschdienst" (DAAD), it is the presence of individual people that allows to make the seemingly impossible possible.

Foremost I am grateful for the support and patience I have received from my supervisor Omar Licandro. Without his conviction about the possible success of my research itinerary this thesis would probably have stalled at its very beginning. I also benefited from his experience and his efforts to integrate me into the wider research community. This allowed me to acquire research skills in other institutions such as the New York University with excellent lectures by Tom Sargent (a special thank to him) and the Royal Economic Society as well as to forge contacts with talented researchers.

I am thankful to Salvador Ortigueira who contributed in important ways to the research path taken towards labor markets. His very thorough analysis helped me to improve many specific issues of the model. Morten Raven’s fast comprehension, strong sense for economic intuition and knowledge of the literature helped me substantially for my research work - thanks to him too. I am also grateful to Rainer Klump for his backing of my EUI application and for the collaboration during my Diploma thesis, through him I learnt to enjoy theoretic modeling.

Obviously research needs networks, and the European University Institute makes it possible to interact with a large number of young co–researchers on similar topics. I am grateful for the discussions I had with Giovanni Callegari. Due to him many rough initial ideas developed into more viable projects. But I am also strongly indebted to Markus Poschke for critical and thorough examinations of my research. Thank you to the many co–researchers Mario Mariniello, Renato Faccini, Clara Barrabés Solanes, Aitor Erce Dominguez, Tomasz Kozluk, Christopher Milde, Andrea Gallice, Matthias Hertweck and also to Rocío Sánchez Mangas as well as Philip

Sauré all of whom made both economics and life in Florence a more pleasant experience.

Researchers from other departments within the EUI contributed to a wider view on economic and non–economic issues, recalling me that economic thinking is one specific way of addressing issues in society. The continuous discussions with Volker Balli enriched my view on methodological terms, but more importantly his presence in crucial moments was invaluable. Special thanks go also to the extended "family" consisting of Ana Avalos with her fresh ideas and Chiara de Franco for the cultural enrichment. I thank Wim van Aken for having shared with me many aspects of life in New York in an innovative spirit and with a very accute sense of humor.

Thanks to the assistance of Giovanni Stanghellini for helping at various stages in my development and to learn in coping with deadlines.

Most importantly I am deeply thankful to Ilaria Montemagno for offering me her love and patience in a relationship of a never withering rose with its blossom and spines. Her continuity and presence are essential to me.

Thank you to my entire family, Sabine and Wilhelm for their confidence and unconditional support, Philipp, Milena and Boris for the continuous encouragement, exchange and presence during the entire development of the thesis. I wish Philipp all the best towards completing his personal dissertation.

Introduction

Innovative ideas, development of skills and job search – these are not only the constitutive elements in the life of a doctoral researcher but also of this dissertation. The three chapters presented in the following apply the concept of general equilibrium to topics related to technological innovations and labor markets. Thereby, the thesis combines strands of the industrial organization literature with the growth literature - and includes a chapter with business cycle analysis.

The first chapter "Explaining vacancy-unemployment volatility over the business cycle: the role of on-the-job-search" focuses on the influence of job search by employed and unemployed workers on the volatility of unemployment and vacancy postings over the business cycle. It is well-known that the standard Mortensen–Pissarides matching model generates too little volatility of the vacancy–unemployment ratio over the business cycle. The chapter investigates the role of optimal job search by unemployed and productivity–dependent search by employed workers on the fluctuation of the labor market tightness. The main finding is that variable search allows for an amplification of the vacancy–unemployment volatility by a factor of two compared to the original matching model. The model is innovative by shifting the focus towards the worker's side instead of the firm in generating fluctuations. Preceding work has mainly focussed on wage rigidities to increment firm profit's volatility which then translates into higher vacancy—unemployment volatility. Here the worker's side augments the volatility of total search effort. From a modelling perspective, the paper introduces explicitly the valuation of time as a determinant in the job searching process.

The second chapter "Competition and Growth in a Cournot setup with imitation" analyzes the influence of imitation costs as a proxy for competition and growth in a neo-Schumpeterian growth model. In the traditional Schumpeterian framework, less competition implies larger profits and more innovation, which contradicts the general view of competition being performance enhancing. In order to reproduce the empirical hump–shaped relationship between competition and growth the model exploits the fact that incumbents need to replace themselves when innovating. In fact, if the current profits of incumbent firms are too high, they have no incentives to innovate. Outside firms that may enter through imitation reduce these profits, reestablishing in this way the incentives for innovation. Imitation costs are chosen as measure of competition which distinguishes the model from other setups and bears the advantage that these costs can be influenced by policy makers either through taxes or subsidies, or through patent legislation. The model includes heterogeneous firms, free entry and technological obsolescence as well as non–drastic innovations combined with Cournot competition to determine prices and quantities. High imitation costs reduce entry which increases the value of incumbent firms, but reduces also their incentives for innovation. By lowering imitation costs increases the number of incumbents and reduces profits which in itself reduces innovation incentives (Schumpeterian effect), but at the same time the industry is shifted to a more innovative market structure as more firms engage into a race for the next innovation (composition effect). The two opposing effect leads to a hump–shaped relationship between competition and growth. Too low imitation costs imply too little profits for innovation, while too high imitation costs imply too much profits for innovation.

The third chapter "Diffusion of technologies with skill heterogeneity and productivity increments" models technology diffusion focussing on the requirements of skills for the adoption of a technology and the evolution of its productivity applied to a General Purpose Technology like information technology. A technology is identified by two characteristics which evolve along the diffusion path: minimum skill requirement and productivity level. An R&D firm owning monopoly rights on a technology maximizes profits by improving both of the characteristics. The model unveils the complementarity of skill requirement and productivity during the maturation of a new technology, a reduction in skill requirements increases the market size which is the more profitable the more productive the technology is and productivity enhancements have a larger return the larger the lower the required skills for the technology. The framework yields an S-shaped diffusion pattern which is the result of the complementarity between productivity and skill requirement. No specific distributional assumption is needed for the skill distribution of the population. In addition, applying this setup to General Purpose Technologies, provides a rationale for the observed growing wage differentials – between users of the old and new technologies – as well as for the productivity slowdown –an initial phase of reduced output growth due to increased R&D activity.

Chapter 1

On–the–job search and v–u volatility

1.1 Introduction

The Mortensen–Pissarides matching model has become an important workhorse in the business cycle literature to incorporate the mechanisms of the labor market. At its core the model formulates a macroeconomic matching function which reflects the frictional and timely process for workers and employers in forming employment relationships. In addition it includes a microeconomic foundation for wage formation based on a bilateral bargain of a surplus generated by the match, see Pissarides (1985) and Pissarides (2000) for a detailed description. But by analyzing new data on vacancies Hall (2005) and Shimer (2005a) find that the model fails to explain the variability of its two main variables, unemployment and posted vacancies by employers. The difference between the data and the calibrated version of Shimer is by an order of magnitude. This means that contrary to its analytical appeal, the matching framework has difficulties in explaining the variability of its two relevant labor market variables over the business cycle. As main source for the lack of vacancy–unemployment variability Hall and Shimer identify insufficient profit variability of the firm. This affects directly the job–creation mechanism by the firm as employment contracts require vacancy opening, which itself depends on future expected profits by the firm.

This new finding triggered numerous articles generating remedies to the basic model by increasing firm profit volatility. The attempts to reconcile the model with the data can be divided into three interrelated categories. The first one introduces wage rigidity within the bargaining set of workers and employers in order to attribute the surplus volatility towards the firm instead of dividing it in a more equal manner between workers an employers. The second category of models formulates alternative micro–founded representations of the model's bargaining or informational structure leading either to optimal longer–term wage contracts or an information flow from workers to employers that lead to wage smoothing. Finally, the third category, which we pursue here as well, introduces procyclical search intensity by workers leading to higher job–filling probabilities for open vacancies, which increases the variability of expected gains for the employer of an open vacancy.

In order to augment the firm's profit variability the first two solution methods alter the continuous wage renegotiation of the original model. Continuous wage negotiation leads to excessive procyclical variability in wages and reduces the amount of profits that are attributed to firms in an upswing. By smoothing wages over the business cycle variability of profits is increased at the expense of wage variability. Therefore most of the recent literature has concentrated on wage rigidity, either directly imposed as in Hall (2005) or through a different bargaining scheme such as the one proposed by Hall and Milgrom (2005). As a result these features do earn higher variability of the vacancy–unemployment ratio, which has further been shown by Gertler and Trigari (2005). But this source of variability is not sufficient to account for the large discrepancy between the empirical findings vacancy–unemployment variability and the model's predictions.

The features introduced in this paper is variable job search by unemployed and productivity–dependent on–the–job search by employed workers. By including variable search we uncouple the empirical labor market tightness, the vacancy–unemployment (v/u) ratio, from the model's relevant choice variable, vacancy–search ratio (v/s). In the original model the observable and the model's variable are identical, tying total search amount strictly to the number of unemployment. By separating the two variables it is possible to have falling unemployment without necessarily reducing the total search amount. This leaves the firms with larger incentives to create vacancies as the job–filling probability remains larger. It has been shown that variable search by unemployed or employed and productivity–independent on–the–job–search taken separately can neither generate the magnitude of volatility nor the observed impulse responses of the business cycle variables to productivity shocks[1]. The intuition to this is as follows. We can subdivide changes of total search amount into changes of per–capita search intensity, representing the intensive margin, and changes in the composition of searchers (employed or unemployed), being the extensive margin. Including variable search only for unemployed increases only the intensive margin. But the total number of unemployed workers is too small to generate big fluctuations unless the search variation leads to indeterminacy of the model as shown by Hashimzade and Ortigueira (2005)[2]. By extending variable search also to employed workers we add a source of volatility to search activity which creating variability on the intensive margin as well as on the extensive margin (unemployment-employment movement) and earning higher vacancy-unemployment volatility.

In existing models of on–the–job search, such as Mortensen and Nagypal (2005) or Shimer (2005a), the total number of job–seekers is increased by introducing variable, but identical, search intensity for all employed workers. In such a framework workers of all productivities, low or high productivity, search with the same intensity[3]. The problem is the higher rejection–rate of new

[1] See for this Nagypal (2004) and Krause and Lubik (2004b).

[2] Krause and Lubik (2004b) have determined areas of indeterminacy and non–existence for parameters of the standard matching model.

[3] or in a more refined version, workers search with a constant level up to a threshold productivity level, above which they do not search at all (see Pissarides (2000))

matches by high–quality workers when matched to lower productivity jobs. The rejection of successful matches by high productivity workers decreases the effective job–filling rate for firms and hence the incentives for employers to open vacancies as the costs for opening vacancies are wasted. Mortensen and Nagypal (2005) show that a positive productivity shock may even decrease the effective job–filling rate due to the rejection problem by employed workers, which implies that the labor market variables become insulated to productivity shocks.

As a solution to the rejection–dilemma and to address the tightness variability problem we therefore extend the basic model by including productivity–dependent on–the–job search in addition to variable search intensity by unemployed workers. In equilibrium we obtain that all unemployed workers search with identical time–varying intensity, and on–the–job searchers search dependent on the productivity with which they are currently employed. Unemployed workers search most and accept all offered jobs, while search intensity by employed workers decreases with their match–specific productivity level, as the probability of finding better jobs decreases. In the pool of job–searchers the probability of obtaining a better match for an employed worker is higher the lower the current productivity. The bulk of search amount is carried out by unemployed and low–productivity workers, which on the aggregate reduces the probability of a rejected match from the point of view of the firm. As a consequence employers are more inclined to open vacancies compared to the standard on–the–job model.

The model increases the variability of the empirical labor market tightness compared to the standard model. The main reason lies in the distinction of unemployment and search amount. In an upswing the employer has larger incentives to open new vacancies due to higher expected profits which subsequently increases job formation and reduces unemployment. In a standard matching model the reduction in unemployment strongly counteracts the increased profits for the firm as it becomes harder to find workers, the job–filling rate declines. With endogenous search intensity for unemployed and on–the–job search by employed workers total search does not decrease drastically with the formation of new employment relationships although unemployment reduces implying a larger job–filling rate for the firm. Hence, vacancies are larger than before while unemployment is reduced leading to a stronger volatility in the observable vacancy–unemployment ratio. In this way the increased search amount overcomes partially the initial difficulty of the Mortensen–Pissarides model.

On–the–job–search models face numerous conceptual difficulties, either at the level of bargaining or at the level of job separations, which I address in the model. The first one already mentioned is the rejection by employed workers of lower productive jobs reducing the incentives for firms to open vacancies. By making on–the–job search productivity dependent we are able to reduce the importance of the rejection rate in reducing incentives to vacancy–opening. The other three problems mentioned in the literature are outside options that depend on the employment history, bidding races between former and future employers and higher quit rates of on–the–job searchers.

Firstly, it might be that on–the–job searchers as opposed to unemployed searchers could negotiate higher wages with their new employer due to the fact that they have a higher outside option due to their current employment. For such a setup it is necessary to abandon the hypothesis of continuous bargaining between worker and employer and adopt a specific contract at the moment of the match formation. But this would lead to a wage formulation which depends on the entire employment history of the individual making not only the endogenous employment distribution but also the wage of each individual a state variable. Next to the fact that we assume continuous re–bargaining possibilities between employer and employee we also assume throughout the model that wage bargaining takes place only after the worker has been matched with an employer and after the old job has been definitively quit. The worker once matched to a new employer and after the match–specific productivity has been observed, can not return to his former occupation. In this way both types of workers, formerly employed and unemployed, have the same outside option, the value of being unemployed.

Secondly, higher wages for on–the–job searchers are also generated due to outbidding possibilities by potential new employers. A recurrent topic in the on–the–job search literature is the possibility of the current employer to offer higher wages, similar to efficiency wages, to prevent job–to–job transitions by the worker. In such a framework an employer presents a new job–match to his current employer and asks for higher wages to in exchange for remaining on the job. With a continuous bargaining setup there is no space these contracts would be renegotiated once the other job has been filled by a different job–seeker.

Finally, the search activity of employed workers leads to higher quit rates compared to a situation without on–the–job search. As a job value for the firm consists of the per–period profits to the firm and the expected duration of the employment relationship, on–the–job search reduces the expected profits and therefore reduces the amount of opened vacancies. In general such a situation would lead to lower wages for workers as the capital gain to the employer is lower. In this paper we assume a that a continuum of firms exist and that each of them is characterized by a continuum of jobs. On–the-job search leads to productivity gains within the firm which increases profits for the firms. The law of large numbers leads to a situation in which each worker lost through on–the–job search is compensated by an equally higher productive worker, benefitting the profits of the firm in this way. Wages are weighted mean between the net productivity of a job and the outside option of the worker. As the searching technology requires time and the outside option consists of increased leisure, a higher on–the–job search intensity also increases the outside option. In this way on–the–job search increase wages of the workers.

One issue presented by Nagypal (2005) nevertheless remains intact: on–the–job search models predict that firms prefer to hire unemployed workers over employed workers because the expected profits upon contact of contacting unemployed workers is higher. The difference in expected profits stems from the rejection of low productivity matches by on–the–job searchers; unemployed workers instead accept every job they become matched to. This fact clearly demonstrates that informational flows are important in the process of job matching.

Introducing productivity–dependent on–the–job search along with variable search into a fully fledged matching model significantly augments the volatility of the vacancy–unemployment ratio. The relative volatility to output is more than doubled compared to the basic version with constant search only by unemployed. Nevertheless comparing the result with empirical data reveals that endogenous search even when correctly accounting for rejections can only contribute a small part to the variability puzzle. But combining on–the–job search with other mechanisms that directly affect profits of firms may already do the job. In this way the model contradicts the finding by Mortensen and Nagypal (2005) (section 6), and allows on–the–job search to have a significant role in the Mortensen-Pissarides framework.

The paper is structured as follows, the next section describes the model, its time structure, the labor market setup , the problems of the household and the firm and wage bargaining. We then describe the calibration strategy and computational issues. Section 1.4 presents the steady state and the dynamic results and we conclude with section 1.5.

1.2 The model

The model presented in this paper incorporates variable search for unemployed and productivity–dependent on–the–job search for employed workers within a Dynamic General Equilibrium model with a Mortensen–Pissarides matching framework for the labor market. The aim is to increase the volatility of vacancies and unemployment in order to better match the model's predictions with the data. To solve some of the conceptual difficulties regarding on–the–job search I present the timing and the information flow within a given period.

Time is discrete. At the beginning of the period workers and employers know wether they are matched or not and know their match–specific productivity. An aggregate productivity shock then determines the precise productivity level of the jobs. This enables workers and employers to engage in wage negotiation or both sides decide to endogenously destroy the match in the case the productivity does not generate a positive surplus. In this case the worker joins the pool of unemployed and the firm makes zero profit. Thereafter all individuals engage in market activity, either working and/or searching for jobs. Only after an exogenous destruction rate hits the existing job relationships and dissolves a given fraction of them, new matches between workers and employers form within the period. The matching probability for each individual depends on his specific search intensity and at the moment of the match the match–specific productivity of the new job is revealed to the worker and to the employer within the period. On–the–job searchers decide wether to accept the new job or rejecting it by comparing their actual productivity with the new one and by anticipating next period's wage bargain outcome. All unemployed workers, instead, do accept the newly formed matches in this period even though they may be lower than current reservation productivity, because the decision to transform the match into an active employment occurs only in the next period. With the new employment status including the relative productivities all parties enter the new period in which an aggregate productivity shock occurs and the new match becomes operable.

These timing assumptions, especially the fact that bargaining takes place only in the period following the match, implies equal outside options for all workers (employed and unemployed) and simplifies the wage distribution to depend entirely on the actual productivity[4].

1.2.1 Labor market

The economy is characterized by a frictional labor market. For employers and workers to become productive, firms need to open vacancies, and employed and unemployed workers need to search for vacant jobs. The number of employer–worker matches occurring in a given time period is characterized by an aggregate matching function representing the matching technology. Only with the formation of a match between a worker and a firm the match–specific productivity

[4] In a continuous time model with continuously renegotiated wages (no fixed wage contracts), the outside option for previously unemployed or employed workers is identical following the moment of the match. The formerly employed can no longer return to his former job and remains therefore with outside option for unemployment.

becomes known to both sides.

Search. The search amount is chosen optimally by unemployed and employed workers in every period. All unemployed workers are ex–ante identical and therefore search with the same effort, while employed workers search depending on their current match–specific productivity. The total search amount is the aggregation of all individuals' search effort, employed or unemployed, taking into consideration the productivity distribution of employed workers:

$$s_t = s_{ut}u_t + n_t \int_{G_t(a_{rt})}^{1} s_{wt}(a)dG_t\,(a)\,. \tag{1.1}$$

The total amount of search units in period t consists of search by each unemployed worker s_{ut} multiplied by their number u_t and of a productivity–dependent search effort $s_{wt}\,(a)$ by employed workers of total number $n_t \int_{a_{rt}}^{\bar{a}} dG_t\,(a)$, with $G_t\,(a)$ characterizing the productivity–distribution of workers matched to an employer in period t, n_t determines the total number of matched workers, and a_{rt} reflects a reservation productivity below which matched workers prefer to be unemployed. In fact, workers follow a reservation strategy: with a match–specific productivity lower than a_{rt} they prefer to be unemployed and consume a higher amount of leisure while searching for new jobs with intensity s_{ut}.

Search technology. Searching for jobs is exclusively a time consuming activity, to capture the timely process of job search and the non–transferability of these costs. Compared to a pecuniary cost, this version leads to differentiated valuation of time across workers as time is not transferrable. Employed workers who work a certain number of hours have a smaller time window than unemployed workers and their marginal value for leisure is higher due to a concave utility function.

The search technology is characterized by decreasing returns in search time. Equivalently, in order to search for jobs with effort s_{it}, the individual of type i faces a convex cost structure. Search time σ_{it} is characterized by the function

$$\sigma_{it} \equiv \sigma(s_{it}),\ \sigma'(s_{it}) > 0,\ \sigma''(s_{it}) > 0.$$

In this paper we will use the parameterization

$$\sigma\,(s_{it}) = g_i s_{it}^{\gamma_i}, \gamma_i > 1,\ \forall i = u, w. \tag{1.2}$$

An individual of type i searching with intensity s_{it} during period t needs to consume σ_{it} of her time, g_i represents a scaling parameter and γ_i determines the convexity of the cost function. Furthermore, we assume that the parameters g_i and γ_i are identical for employed and unemployed workers.

Total search time in the economy consists of search time spent by each unemployed and employed worker

$$\sigma_t = u_t \sigma\left(s_{ut}\right) + n_t \int_{G_t(a_{rt})}^{1} \sigma\left[s_{wt}\left(a\right)\right] dG_t\left(a\right).$$

The total time dedicated to search is the amount of time spent by each unemployed multiplied by the number of unemployed worker in addition to the time spent by employed workers weighted by their productivity distribution.

Matching function. The matching technology follows Pissarides (1985)[5] and takes the form of a time–invariant function with constant returns to scale and decreasing returns to each of the two factors, search effort s_t and posted vacancies v_t. We stick to the most common parameterization and will use a Cobb-Douglas function throughout the paper in order to compare the results with the existing literature[6].

$$M(s_t, v_t) = \min\left[\gamma s_t^{\eta} v_t^{1-\eta}, v_t, 1\right].$$

$M\left(s_t, v_t\right)$ is the number of workers that are matched in period t to a new employer, depending on the aggregate amount of search and the aggregate number of vacancies, with γ being a scaling parameter and η the elasticity of matches to the search amount. The total number of matches can neither exceed the total number of vacancies (all open vacancies would be filled) nor the total population, which we normalize directly to 1, in which case the entire labor force would either find or switch job. At the moment of the match workers draw a match–specific productivity from a time–invariant distribution which is revealed to the employee and the employer.

Job finding and filling rates. The matching function assesses the amount of matches occurring in a given period depending on search effort and vacancies in the respective period. The matching probability per search unit is given by

$$m_t = M(s_t, v_t)/s_t = \gamma\left(v_t/s_t\right)^{1-\eta}.$$

[5]For a detailed illustration of the matching model see Pissarides (2000).

[6]Petrongolo and Pissarides (2001) survey and estimate the different functional forms for matching functions. They can not reject a functional form with constant returns to scale such as the Cobb–Douglas form used here.

The probability for an unemployed worker to be matched to a job is the matching probability per search unit multiplied by the search effort s_{ut} of unemployed workers

$$f_{ut} = s_{ut} m_t, \tag{1.3}$$

where the search effort matching probability m_t is taken as given in a market solution, while the search intensity s_{ut} is chosen optimally by the unemployed workers. Every unemployed worker accepts the match, as the aggregate productivity is not known in the current period but is revealed only in the next period. For a match to last longer than one period, it is necessary that the productivity level of the job is larger than the reservation productivity in the next period, otherwise the worker withdraws from the match preferring to search for a new job while unemployed.

Regarding employed workers, the rate of job–to–job transitions depends on the match–specific productivity level a of the worker for two reasons. The first one is productivity–dependent search effort for on–the–job searchers and the second one is the probability that the new job exhibits a lower productivity than the current one leading to a rejection of the job–offer. This leads to a rate of job–to–job transition of

$$f_{wt}(a) = [1 - Z(a)]\, s_{wt}(a)\, m_t, \tag{1.4}$$

where on-the-job search effort $s_{wt}(a)$ and the rejection probability $Z(a)$, which is the exogenous productivity distribution, depends on the productivity level a. The probability of finding matches that have higher productivity and are hence not rejected are $\int_a^{\bar{a}} z(a)\, da = 1 - Z(a)$, which is independent of the endogenous productivity distribution $G_t(a)$. In order to compare the job finding probabilities between unemployed and employed workers mentioned in the literature such as in Shimer (2005a), I compute the mean job–to–job transition rate

$$\bar{f}_{wt} = \int_{G_t(a_{rt})}^{1} f_{wt}(a)\, dG_t(a),$$

which depends on the exogenous as well as the endogenous productivity distribution through the reservation productivity.

The ex–ante job filling probability for the firm depends on the amount of matches and the number of posted vacancies. Compared to the standard literature the effective job–filling rate includes a correction factor which captures the rejection rate by on–the–job searchers

$$q_t = \frac{M(s_t, v_t)}{v_t} \left\{ 1 - \frac{n_t \int_{G_t(a_{rt})}^{1} s_{wt}(a)\, Z(a)\, dG_t(a)}{s_t} \right\}. \tag{1.5}$$

The effective job–filling probability for a vacancy q_t is the total number of matches M per vacancy v_t taking into account rejected job–to–job matches. The share of those search units stemming from employed workers with productivity higher than the newly extracted productivity needs to be subtracted from the job-filling rate. This depends on the search intensity distribution, the exogenous productivity distributon as well as the endogenous productivity distribution. The effective job–filling probability is identical for all vacancies as the productivity of the resulting job is not known at the moment of vacancy posting. Note that jobs with a productivity lower than a_{rt} are counted as matched although not necessarily as employment. This depends on the aggregate productivity next period.

Job separations and job destructions. A large literature has focussed on job destruction over the business cycle. Most prominently Davis et al. (1996), Caballero and Hammour (1994) and den Haan et al. (2000) regard it as the main driving force in job turnover. In order to compare the results in this paper with theirs, I compute a job destruction and a job separation rate. Job separations occur when the employment relationship is quit either through an exogenous or an endogenous destruction, or a job–to–job transition

$$sep_t = \rho + (1-\rho) G_t(a_{rt}) + (1-\rho) m_t \int_{G_t(a_{rt})}^{1} [1 - Z(a)] s_{wt}(a) \, dG_t(a).$$

Total separation rate is the sum of the exogenous job–destruction rate ρ which affects all jobs independently, the endogenous separations occur to all those remaining jobs with productivity below the reservation productivity a_{rt}, and separations due to effective job–to–job transitions. The second term is due to the reservation strategy of workers and firms: with a productivity lower than the reservation threshold a_{rt} in the current period, the worker prefers to severe the employment relationship in order to consume leisure and search while unemployed.

Job destruction, instead, is only a subset of job separations and includes the exogenous destruction rate as well as the jobs with too low–productivity, but does not consider the movements due to on–the–job search

$$\rho_t = \rho + (1-\rho) G_t(a_{rt}). \tag{1.6}$$

The job destruction rate ρ_t is composed by the exogenous rate ρ in addition to the endogenous rate $G(a_{rt})$ of not exogenously separated matches.

Productivity distribution. Workers and jobs are not per se characterized by a productivity level, only when matched they generate an idiosyncratic productivity. At the moment of matching workers draw a match–specific productivity from a time invariant productivity distribution $Z(a)$. Through job-destruction, job–finding and job–to–job transitions results an endogenous time–variant productivity distribution $n_t G_t(a)$ of workers matched to an employer, where n_t characterizes the total number of employed workers and $G_t(a)$ their distribution across productivities in period t. The cumulative distribution evolves according to

$$\begin{aligned} n_{t+1}G_{t+1}(a) &= u_t s_{ut} m_t Z(a) + (1-\rho) n_t \int_{G_t(a_{rt})}^{G_t(a)} dG_t(\tilde{a}) \\ &\quad - (1-\rho) n_t [1 - Z(a)] m_t \int_{G_t(a_{rt})}^{G_t(a)} s_{wt}(\tilde{a}) \, dG_t(\tilde{a}) . \end{aligned} \tag{1.7}$$

The number of workers matched to an employer with a productivity lower than a in period $t+1$, consists of unemployed workers matched with probability $s_{ut}m_t$ to an employer in the previous period and having drawn a productivity below a from the exogenous distribution $Z(a)$. In addition, all those workers that were employed with a productivity larger than a_{rt} and that have not been exogenously separated with rate ρ remain in the pool of matched workers. The third term identifies on–the–job search. All those workers that have found a new job with higher productivity than a quit the pool of lower–productivity workers.

To retrieve the evolution of the total number of matched workers from equation (1.7) we set the individual productivity to the maximum level, $a = \bar{a}$, and obtain a more familiar law of motion

$$n_{t+1} = u_t s_{ut} m_t + (1-\rho)[1 - G_t(a_{rt})] n_t, \tag{1.8}$$

where the value n_t characterizes the workers in a relationship with an employer and we have ρ as rate of exogenous job destruction and $G(a_{rt})$ as a time varying endogenous job destruction rate. It becomes apparent that on–the–job–search has no direct influence on total employment evolution, but only indirectly through its general equilibrium effect on the reservation productivity a_{rt}. In order to become operative in period $t-1$ a worker needs be matched and in addition needs to have drawn a match–specific productivity larger than the current reservation productivity: $a > a_{rt}$.

The steady state unemployment rate may be computed from the aggregate employment equation (1.8) relating job–finding and job–destruction rates (1.3) and (1.6)

$$u_{ss} = \frac{\rho_{ss}}{f_u + \rho_{ss}}$$

It is the usual relationship between job destruction and job finding when normalizing the entire labor force to 1.

1.2.2 Households

Individuals obtain instantaneous utility from consumption and leisure. Both are time separable to account for the secular constancy of working time and we assume a logarithmic utility function for consumption and a CIES form for leisure for each individual i of the household

$$U(c_t(i), l_t(i)) = \ln c_t(i) + b\frac{[l_t(i) - 1]^{1-\phi}}{1-\phi}, \qquad \phi > 0$$

where b characterizes the attributed weight to leisure and with $\phi = 1$ the utility in leisure takes also a logarithmic form.

Aggregating over all individuals and normalizing their total number to 1, the household maximizes the discounted value of an infinite stream of utility with respect to consumption and leisure

$$\max_{\{c_t\}_{t=0}^{\infty}, \{l_t(i)\}_{t=0}^{\infty}} \sum_{0}^{\infty} \left[\int_0^1 \left(\ln c_t(i) + b\frac{[l_t(i) - 1]^{1-\phi}}{1-\phi} \right) di \right] dt \tag{1.9}$$

subject to three constraints: a budget, an employment and a time constraint. We assume that the household pools income of its members, such that consumption is independent of the employment status of the individual and hence the intertemporal decisions can be dealt with at the aggregate level. In a world of complete markets Andolfatto (1996) and Merz (1995) have shown which assets are needed to obtain identical result between a large unique household and the decentralized version. Such an insurance or pooling scheme is not possible for leisure due to the nature of time, which cannot be transferred between individuals. The intratemporal decision for the individuals may lead to differing levels of leisure depending on their employment status as well as their optimal search intensity.

The household's utility function (1.9) may be rewritten by distinguishing between employed and unemployed workers

$$U_t = \max_{\{c_t\}_{t=0}^{\infty}, \{l_t(i)\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \left[\ln c_t + \frac{b}{1-\phi} \left\{ n_t \int_{G_t(a_{rt})}^1 [l_t(a) - 1]^{1-\phi} dG_t(a) + u_t [l_{ut} - 1]^{1-\phi} \right\} \right],$$

where we have that $u_t = 1 - n_t \int_{a_{rt}}^{\bar{a}} dG_t(a)$ from the discussion in section 1.2.1. The household maximizes this function subject to three constraints: budget, employment and time.

Budget Constraint. The budget constraint applies to the income pooling household, but we could also imagine a perfect risk sharing mechanism between individuals to decentralize such a setup, as seen by Andolfatto and Merz.

$$C_t + I_t \leq \Pi_t + r_t K_t + n_t \int_{G_t(a_{rt})}^{1} w_t(a)\, h dG_t(a) \tag{1.10}$$

Variables in capital letters refer to aggregate quantities applying to the entire household. Resources of the household stem from aggregate profits Π_t from household–owned firms, rents from aggregate capital K_t as well as productivity–dependent wages from employed workers. On the spending side, households use their resources for consumption and investment.[7] Capital evolves according to the law of motion

$$K_{t+1} = I_t + (1-\delta)K_t, \tag{1.11}$$

with next period's aggregate capital increasing thanks to investment I_t but reducing through physical depreciation with rate δ.

Time Constraint. Every individual faces a single time constraint with total time endowment normalized to 1. This may be used for working h hours when employed, enjoy leisure $l_t(i)$ or search for jobs $\sigma_t(i)$.

$$1 = h(a) + l_t(i) + \sigma_t(i) \tag{1.12}$$

We can subdivide individuals into employed and unemployed workers and write for individuals of the two types:

$$\begin{aligned} \text{employed} &: \quad 1 = h(a) + \sigma(s_{wt}(a)) + l_t(a) \\ \text{unemployed} &: \quad 1 = \sigma(s_{ut}) + l_{ut} \end{aligned}$$

The first group works $h(a)$ of their time and has a productivity-dependent on–the–job search time $\sigma[s_{wt}(a)]$, which depends on the search intensity at the different productivities. Within the group of unemployed, all individuals are identical and their search time $\sigma(s_{ut})$ depends on the search intensity s_{ut}. The conditions for the functional form for search time were discussed in section 1.2.1.

[7] There is no governmental sector offering any unemployment benefits or collecting taxes

Labor market constraint. The labor market is characterized by frictions in the matching process of job–seekers and vacancies. The individual takes the matching probability m_t and the probability of job destruction ρ_t as given. This assumption is a major difference to the social planner solutions by Andolfatto (1996) and Merz (1995). While the search externality in those models is taken into consideration during optimization, we analyze a market equilibrium where the individual does not account for the searching externality when optimizing for the search amount.

When expressing the cumulative distribution function for $G_{t+1}(a)$ from (1.7) as a probability distribution function $g_t(a)$, a discontinuity is present at the reservation productivity a_{rt}

$$\begin{aligned} n_{t+1}g_{t+1}(a) &= u_t s_{ut} m_t z(a) + \mathbf{1}_{\{a \geq a_{rt}\}} (1-\rho) n_t \times \\ &\times \left\{ g_t(a) + m_t \left[z(a) \int_{G_t(a_{rt})}^{G_t(a)} s_{wt}(\tilde{a})\, dG_t(\tilde{a}) - [1 - Z(a)]\, s_{wt}(a)\, g_t(a) \right] \right\} \end{aligned} \tag{1.13}$$

The probability density function of matched workers next period consists of three terms. The first term characterizes the number of unemployed workers becoming matched in period t with probability $s_{ut}m_t$ and drawing productivity a from the exogenous distribution function $Z(a)$. The second term refers to the workers that have not been affected by exogenous job destruction ρ, and hence remain employed with the same productivity if their productivity is larger than the reservation productivity $\bar{a}_{rt}$. The third term identifies the effects of on-the-job search activity to the probability distribution: employment with productivity a increases due to workers being matched into productivity a and decreases with the amount of workers currently employed with a finding a higher productivity job with the effective probability $[1 - Z(a)]\, m_t s_{wt}(a)$.

The distribution $n_t g_t(a)$ characterizes the number of workers who are currently matched and start working in the current period. The distribution of workers effectively working differs with respect to this distribution for productivities lower than a_{rt}, who are matched but do not work. By normalizing total labor force to 1, the measure of employed is

$$EMP_t = n_t \int_{G_t(a_{rt})}^{1} dG_t(a) = n_t [1 - G_t(a_{rt})]$$

and total measure of unemployed is

$$u_t = 1 - n_t [1 - G_t(a_{rt})],$$

which is identical to the unemployment rate.

First-order conditions of the household

The household maximizes utility function (1.9) under the three constraints (1.10), (1.12), (1.13) exposed before.[8]

The Euler equation relating consumption in periods t and $t+1$ is

$$\frac{1}{c_t} = \beta E_t \left[\frac{1}{c_{t+1}} \left(1 + r_{t+1} - \delta\right) \right] \tag{1.14}$$

where r_{t+1} is the gross real interest rate, δ the depreciation rate, β the subjective discount factor and E_t the expectation operator.

The intratemporal valuation of leisure for the individuals is

$$\begin{array}{ll} \frac{u_{lt}(a)}{u_{ct}} = bc_t \left[l_t\left(a\right)\right]^{-\phi}, & \text{if employed with productivity } a \\ \frac{u_{ult}}{u_{ct}} = bc_t l_{ut}^{-\phi}, & \text{if unemployed} \end{array} \tag{1.15}$$

The marginal value of leisure in consumption units depends on the level of consumption as well the level of leisure for $\phi \neq 0$. And for $\phi > 1$, the leisure utility function is concave implying that the marginal valuation of leisure decreases with the amount of consumed leisure.

Regarding the value of match between a worker and an employer, note that the matching framework allows for bilateral monopolistic rents between the two sides once matched. The net value of being employed with productivity a for the individual compared to the value of being unemployed is

$$W_t^1\left(a\right) = \beta E_t \left\{ \frac{c_t}{c_{t+1}} \left[\begin{array}{c} w_{t+1}\left(a\right) h\left(a\right) + \left(1-\rho\right) W_{t+1}\left(a\right) \\ +\left(1-\rho\right) m_{t+1} s_{wt+1}\left(a\right) \left[\int_{Z(a)}^{1} \left[W_{t+1}\left(\tilde{a}\right) - W_{t+1}\left(a\right)\right] dZ\left(\tilde{a}\right)\right] \\ -\left(\frac{bc_{t+1}}{1-\phi}\left[l_{ut+1}^{1-\phi} - l_{t+1}\left(a\right)^{1-\phi}\right] + s_{ut+1} m_{t+1} \int_0^1 W_{t+1}\left(a\right) dZ\left(a\right)\right) \end{array} \right] \right\}. \tag{1.16}$$

This value represents the capital gain for a worker holding an employment contract with productivity a for next period net of the unemployment value. It consists of the expected discounted flow value of next period's wage as well as the continuation value in case the working relationship is not severed exogenously. In addition, if the job relationship is not severed the worker finds a higher productivity job with probability $m_{t+1}s_{wt+1}\left(a\right)\left[1 - Z\left(a\right)\right]$ and higher expected value $W_{t+1}\left(\tilde{a}\right)$. The terms in the last line represent the value of an unemployed worker consisting of the leisure gain and the expected value of employment. $W_t\left(a\right)$ is therefore the net gain for a worker of being employed instead of unemployed.

[8] The exact formulation of the household's problem may be found in the appendix.

The worker follows a reservation wage strategy to accept a job (see Pissarides (2000) for a discussion). The net value of employment is therefore correctly characterized by the maximum value between the net matching value and unemployment.

$$W_t(a) = \max\left[W_t^1(a), 0\right] \tag{1.17}$$

The reservation productivity a_{rt}, below which the worker rejects employment equalizes the value of being employed with productivity a_{rt} and the value of being unemployed:

$$s_{ut} m_t \bar{W}_t + \frac{bc_t}{1-\phi}\left(l_{ut}^{1-\phi} - \left[l_t(a_{rt})\right]^{1-\phi}\right) = w_t(a_{rt}) + (1-\rho)\, s_{wt}(a_{rt})\, m_t \bar{W}_t$$

where $\bar{W}_t = \int_{a_{rt}}^{\bar{a}} z(a)\, W_t(a)\, da$. The left hand side represents the sum of leisure value while unemployed and the value of expected future employment, while the right hand side represents the sum of wages with productivity a_{rt} in addition to expected gains from on–the–job search. The last term is new compared to other models and permits the worker to accept lower productivity–jobs because she expects better jobs from on–the–job–search.

Search effort is determined optimally for employed and unemployed workers. Workers in either employment status uses the same search technology (1.2). The optimality condition for unemployed workers

$$\sigma_u'(s_{ut})\frac{u_{lut}}{u_{ct}} = m_t \bar{W}_t$$

equates the marginal costs of searching time $\sigma_u'(s_{ut})$ evaluated in consumption units to the expected income value of employment for a marginal search unit. For each search unit employment occurs with probability m_t. By using the intratemporal valuation of leisure (1.15) and the fact that productivities are drawn from the time invariant distribution $z(a)$ we can rewrite the optimality condition to be

$$bc_t l_{ut}^{-\phi} \sigma_u'(s_{ut}) = m_t \int_0^1 W_t(a)\, dZ(a)\,. \tag{1.18}$$

Similarly to unemployed workers, employed workers choose their productivity–dependent on-the-job search intensity with the marginal condition:

$$\sigma_w'\left[s_{wt}(a)\right]\frac{u_{lat}}{u_{ct}} = (1-\rho)\, m_t \left\{\int_{Z(a)}^1 \left[W_t(\tilde{a}) - W_t(a)\right] dZ(\tilde{a})\right\},$$

where $\frac{u_{lat}}{u_{ct}} = bc_t\left[l_t(a)\right]^{-\phi}$ is the value of leisure when employed with productivity a. The search costs for on–the–job–search evaluated in consumption units are equalized to the expected value gain from a higher productivity job. The valuation of leisure with a concave utility function depends on the employment status and the productivity level while employed.

1.2.3 Firms

Firms produce a homogenous output good using capital and labor. We assume a continuum of firms of mass 1, each offering a continuum of jobs. Firms open vacancies in order to meet workers to form jobs with match–specific productivities. This productivity is not embodied in capital but is fixed for the entire duration of the match. The overall productivity of a job j in firm i depends on the aggregate productivity level A_t, the idiosyncratic productivity level a_{ij} and the average amount of capital allocated to a jobs in firm i. While the aggregate productivity is stochastic over time, the idiosyncratic productivity is fixed over the entire duration of the match. A shock to the aggregate productivity shifts the productivity distribution of jobs, but does not affect the relative productivities of different jobs. The capital stock within a firm is allocated uniformly across jobs, independently of the job's productivity and decreasing marginal products of capital occur at the job level. Capital is rented from households and can be traded frictionless at every point of time. These assumptions imply that the output of a job j in firm i is determined by aggregate and idiosyncratic productivity, the capital intensity per job and hours worked.

$$a_{ij} A_t h^{1-\alpha} k_t^{\alpha}$$

with h being the amount of hours worked in the firm and $k_t \equiv \frac{K_t}{n_t(1-G(a_{rt}))}$ being capital per employed worker. Total output of a firm is the aggregate amount of all jobs

$$Y_{it} = A_t h^{1-\alpha} k_{it}^{\alpha} n_{it} \int_{G_t(a_{rit})}^{1} a dG_t(a), \tag{1.19}$$

where we exploited the facts, that workers' productivity is exclusively match–specific and does not depend on worker's or firm's characteristics. The aggregate capital stock at the firm level is $K_{it} = k_{it} n_t [1 - G(a_{rt})]$. The production function exhibits constant returns to scale to hours worked and capital, and is linear in the number of workers. Average productivity per worker therefore does not depend on the number of workers, but only on the endogenous distribution function of productivities. The decreasing returns to capital or worked hours occur exclusively within a single job.

The economy is characterized by a continuum of identical firms uniformly distributed on the interval $[0, 1]$, each of which can open jobs. In the following we will therefore use the concept of a representative firm.

Firms maximize the discounted stream of profits by renting capital from the household sector and opening vacancies in order to employ workers. Capital renting is done after all information regarding the job match is known (number of matches and match–specific productivities). The firm does not face any hold–up problem which would instead occur if the firm acquired capital and the second hand market would be frictional. As we are focussing on a market outcome firms take the law of motion of productivity-specific employment and the effective job-filling rate q_t as given.

The discounted stream of profits of the firm

$$\Pi_0 = E_0 \sum_{t=0}^{\infty} \beta^t \frac{\lambda_t}{\lambda_0} \left\{ \int_{G_t(a_{rt})}^{1} \left[a A_t h^{1-\alpha} k_t^{\alpha} - r_t k_t - w_t(a) h \right] n_t dG_t(a) - \kappa v_t \right\} \tag{1.20}$$

consists of the workers' productivity of a worker net of capital and labor costs discounted with the consumer's Lagrange multiplier on consumption λ_t/λ_0. Profits of the firm are reduced by the costs of vacancy opening κv_t. The total amount of capital hired by the firm is the average capital per worker multiplied by the number of employed workers: $K_t = k_t n_t [1 - G_t(a_{rt})]$.

The evolution of labor from the point of view of the firm differs to the one by the households as firms face a job-filling probability q_t which is taken as given and defined by (1.5).

$$\begin{aligned} n_{t+1} g_{t+1}(a) &= q_t v_t z(a) + \mathbf{1}_{\{a \geq a_{rt}\}} (1-\rho) n_t \times \\ &\times \left\{ g_t(a) + m_t \left[z(a) \int_{G_t(a_{rt})}^{G_t(a)} s_{wt}(\tilde{a}) dG_t(\tilde{a}) - [1 - Z(a)] s_{wt}(a) g_t(a) \right] \right\} \end{aligned} \tag{1.21}$$

Employment density at productivity a increases with the number of vacancies filled with productivity a and the number of existing employment relationships of productivity a from last period increased by net job–to–job transitions. It can be seen that the distribution of employment at the firm level shifts to higher productivities due to on–the–job search. This is different to existing on–the–job search models with one firm– one job approaches, for which firms had no advantage from on–the–job search.

First-order conditions of the firm

Firms maximize the discounted stream of profits (1.20) subject to the law of motion of labor (1.21) and rent capital from households.[9]

The first-order condition of the firm for the demand of capital per worker is

$$r_t \int_{G_t(a_{rt})}^{1} dG_t(a) = \alpha A_t h^{1-\alpha} k_t^{\alpha-1} \int_{G_t(a_{rt})}^{1} a \, dG_t(a), \tag{1.22}$$

[9] A detailed presentation of the problem may be found in the appendix.

equating total hiring costs of capital per worker to the average productivity of capital. The demand for capital is independent of total employment at the firm level, but does depend on the productivity distribution within the firm and the reservation productivity. In this model employment does not affect the average capital per job.

A job is valuable to the firm as long as it generates profits. The firm's value of a filled job depends mainly on the match–specific productivity a and evolves according to

$$J_t^1(a) = \beta E_t \left\{ \frac{\lambda_{t+1}}{\lambda_t} \left[\begin{array}{l} aA_{t+1}h^{1-\alpha}k_{t+1}^{\alpha} - r_{t+1}k_{t+1} - w_{t+1}(a)\,h + (1-\rho)\,J_{t+1}(a) \\ +(1-\rho)\,m_t s_{wt}(a) \int_{Z(a)}^{1} \left[J_{t+1}(\tilde{a}) - J_{t+1}(a)\right] dZ(\tilde{a}) \end{array} \right] \right\}. \quad (1.23)$$

The job value consists of the discounted productivity next period net of capital and labor costs plus the continuation value taking into consideration possible exogenous separations with rate ρ and job transitions to higher productivities stated in the second line. The fact of using a continuum of firms in the economy with symmetry across firms and a continuum of jobs within each firm, leads to a situation where each on–the–job searcher finds a better job either in the own firm or in another one. The law of large numbers makes that on–the–job–search profits the own firm due to equally incoming jobs with the specific productivity.

Similar to the worker in (1.17), the firm follows a reservation strategy

$$J_t(a) = \max\left[J_t^1, 0\right].$$

The optimal behavior for opening vacancies leads to cost equalization of opening a vacancy κ (marginal and average) and the expected benefits from a vacancy

$$\kappa = q_t \int_{Z(a_{rt})}^{1} J_t(a)\, dZ(a). \quad (1.24)$$

An open vacancy costs κ and is effectively filled with a probability q_t, leading to an average job surplus of $\bar{J}_t \equiv \int_{Z(a_{rt})}^{1} z(a)\, J_t(a)\, da$.

The reservation strategy of the firm implies a reservation productivity level below which the employment relationship is not profitable and the firm optimally withdraws from the match. The marginal condition for this productivity a_{rt} equates the labor capital costs of the job to the benefits in form of output and capital value, including the one of a possible job–to–job transition.

$$w_t(a_{rt})\,h + r_t k_t = a_{rt} A_t h^{1-\alpha} k_t^{\alpha} + (1-\rho)\, m_t s_{wt}(a_{rt})\, \bar{J}_t$$

The costs of a filled job stem from wages for a specific productivity $w_t(a_{rt})\,h$ and capital costs $r_t k_t$, while the gains consist of output value $a_{rt} A_t h^{1-\alpha} k_t^{\alpha}$ and future higher productivity through job–to–job transition $(1-\rho) m_t s_{wt}(a_{rt}) \bar{J}_t$. The value of the job J_t in period t is zero by definition, as we are describing the reservation productivity, i.e. the highest productivity level for which the employer is on the limit between profits and losses to the firm. The term with the value of job–to–job transitions is new in comparison to other models. The worker generates value to the firm through her ambition of finding a better job in the current period with higher discounted pay–off to the firm. In this setup the use of large numbers confers to the firm a positive externality from the worker's search activity, while generally the firm would incur a loss due to the severance of the match. Therefore, the firm has stronger incentives to keep low productive jobs open and lowers the reservation productivity compared to a model without on–the–job search.

1.2.4 Wage determination

The frictional labor market creates rents for the worker and the firm once the two parties have matched and drawn a math–specific productivity larger than the reservation productivity. This surplus needs to be split between the worker and the firm by a bargaining setup. The total surplus for each job consists of the worker's net value $W_t(a)$ and the firm's net value $J_t(a)$, both reflecting the gain compared to the outside option of unemployment or job closure, respectively. We stick to the most common splitting rule, the Nash–bargaining (Nash (1950)) between the worker and the employer. This axiomatic bargain results in maximizing the weighted surplus of the two parties,

$$\max_{w_t} W_t(a)^{\lambda} J_t(a)^{(1-\lambda)}$$

where β reflects the worker's relative bargaining power in the process. The first–order conditions to this problem may be described by

$$W_t(a) = \lambda \left[J_t(a) + W_t(a)\right] \tag{1.25}$$

This rule states that the worker obtains a constant share λ of total surplus $J_t(a) + W_t(a)$ or alternatively that $(1-\lambda) W_t(a) = \lambda J_t(a)$. From this we can recover the wage bill for the employed worker by using the equation for the worker's value (1.16) and the firm's value (1.23) with the sharing rule (1.25) we obtain the expression

$$w_t(a) h = \lambda \left(a A_t h^{1-\alpha} k_t^{\alpha} - r_t k_t\right) + (1-\lambda)\left[\frac{b c_t}{1-\phi}\left(l_{ut}^{1-\phi} - l_t(a)^{1-\phi}\right) + s_{ut} m_t \bar{W}_t\right]. \tag{1.26}$$

Wages are a weighted average between the productivity of the worker net of capital costs and the outside option of the worker consisting of the utility gain due to increased leisure and the expected value for a new job. Rewrite the wage by using the sharing rule (1.25) and the search intensity by unemployed (1.18) we obtain a clear dependence of the wage from labor market conditions m_t/q_t

$$w_t(a)\,h = \lambda\left(aA_th^{1-\alpha}k_t^{\alpha} - r_tk_t + \kappa s_{ut}\frac{m_t}{q_t}\right) + (1-\lambda)\frac{bc_t}{1-\phi}\left(l_{ut}^{1-\phi} - l_t(a)^{1-\phi}\right)$$

The average wage or the expected wage for all employment relationships with productivity above the reservation level a_{rt} is

$$E(w_t|a > a_{rt}) = \int_{G_t(a_{rt})}^{1} w_t(a)\,dG_t(a)$$

1.2.5 Resource constraint and closing the model

Total production can be used for consumption by households, investment and for vacancy payments by firms

$$Y_t = C_t + I_t + \kappa v_t$$

using equation (1.11) for capital accumulation and (1.19) for the production function we have a dynamic budget constraint relating today's capital to tomorrow's

$$A_tn_thk_t^{\alpha}\int_{G_t(a_{rt})}^{1} adG_t(a) = C_t + K_{t+1} - (1-\delta)K_t + \kappa v_t, \tag{1.27}$$

where total capital is given by

$$K_t = k_tn_t\int_{G_t(a_{rt})}^{1} dG(a). \tag{1.28}$$

Aggregate productivity A_t evolves according a first–order auto–regressive process with persistence parameter $0 < \tilde{\rho} < 1$:

$$A_{t+1} = A_t^{\tilde{\rho}}\exp(\varepsilon_{t+1}),$$

and we assume that $\varepsilon_t \sim \mathbf{N}(0, \sigma_a)$. Written in log–form $a_t = \log A_t$:

$$a_{t+1} = \tilde{\rho}a_t + \varepsilon_{t+1}$$

1.2.6 Equilibrium

The market equilibrium is defined as an allocation of $\{C_t, K_t, s_{wt}, s_{ut}, a_{rt}, h, k_t, v_t, n_t, G_t(a), u_t\}$ and prices $\{r_t, w_t(a)\}$ such that:

- $\{C_t, K_t, s_{wt}(a), s_{ut}, h, a_{rt}\}$ solves the household problem (1.9) subject to the budget constraint (1.10), the time constraint (1.12), the law of motion for capital and labor (1.11) and (1.7), respectively.

- Firms choose $\{k_t, v_t\}$ to maximize profits (1.20) subject to the employment flow equation (1.7).

- The laws of motion for the worker's productivity distribution n_tG_t and the number of unemployed workers u_t, are given by (1.7) and $u_t = 1 - n_t$.

- Markets clear. The interest rate r_t equalizes capital demand by the firms and the supply of capital by households following equation (1.22) and (1.28). The aggregate resource constraint (1.27) holds.

- Total wage payments per worker $w_t(a)h$ are determined by Nash bargaining after matches have formed given by the sharing rule (1.25).

1.3 Calibration

The parameter calibration is subdivided into two sets, those related to the standard business cycle values such as depreciation and discount factors, its values being relatively uncontroversial, and a second set relating to the labor market which requires more careful analysis. The calibration assumes a monthly timing interval, and has its empirical counterpart from 1964:1-2005:1.

The first value in the calibration for capital accumulation is the individual monthly discount rate β which we set to 0.9967 in order to reflect an annual real interest rate of $r = 4\%$ and a monthly depreciation rate of 0.87% to match a yearly capital-output ratio of 10. The output elasticity of capital in the production function α is set to match a long–term capital share of 36% as found in American income data within the considered period. Note that the usual derivation of the labor share whn/y is not valid in this model with labor market frictions because part of the original's labor share is used for vacancy posting.

Regarding labor market values, we distinguish between those related to leisure in the utility function, those to the search and matching process, and those to the exogenous distributions. The parameters b and ϕ in the utility function are set in order to obtain a consumption–output share of 67% as is generally assigned to consumption in the United States, and a replacement ratio of 50%. The replacement ratio in this model without taxes or government sector represents exclusively the increased leisure by unemployed workers compared to their average employed counterparts and does not include any pecuniary benefits from unemployment benefits or insurance schemes. These values lead to a working time of roughly one–third of total time endowment, which is a value generally used in the RBC literature and also used by Andolfatto (1996).

Regarding the calibration of parameters for the search and matching technologies, we first normalize the search intensity of unemployed workers without loss of generality to 1[10]. This also facilitates the comparison with the original model of exogenous search intensity by unemployed set to unity. In order to reduce the amount of arbitrarity we set the parameters in the search function for employed and unemployed workers to the same value, i.e. $g_u = g_w$. By normalizing search intensity by unemployed to 1 we obtain a value of 0.16 for g_u and g_w, which in turn implies that unemployed workers search roughly one sixth of their time. Similarly we employ identical convexity of costs $\gamma_u = \gamma_w = 2$ in the search function. The value of $\gamma_u = \gamma_w$ affects the curvature of the search function, which we fix in order to obtain worker's search intensity to be 20% of the one by unemployed, which corresponds to 10% of worker's search time. This value is around the value given by Shimer (2005b) who states an effective matching probability by on–the–job searchers of 15% of the one by unemployed workers.) Quadratic costs seem plausible, and we make sure that we do not run into indeterminacy with these values[11]. The convexity parameter together with the Frisch–elasticity of labor in the utility determine the costs of the search amount and therefore directly affect the variability of search intensity over the business cycle.

For the parameters of the matching function and the wage bargaining we set a match efficiency γ to lead to a monthly job–finding rate of 45% which implies an average unemployment duration of a quarter and represents the value found by Nagypal (2004). The values for the elasticity of vacancies η in the matching function and the bargaining power of workers λ have been strongly discussed in the literature. The generally accepted range of values for η lies between 0.5 and 0.7. These values for the elasticity have been empirically supported by Petrongolo and Pissarides (2001), while Shimer (2005a) chose 0.72, slightly above that window. The problem is that these values in the standard matching model do not reproduce the volatility observed in the data for vacancies and unemployment. A value for the wage bargaining of the worker λ is similarly difficult to pin down and recently Hagedorn and Manovskii (2005) have chosen a very low one,

[10]See Krause and Lubik (2004a) for an extensive description of free parameters and parameter restrictions within the Mortensen–Pissarides matching model.

[11]With smaller values for γ, indeterminacy is much more likely to arise. In order to rule out those regions we stick to a value of 2.

which means that the largest part of the match surplus is attributed to the firm. In order to maintain comparability with other results I use the value 0.5 for the elasticity η and also for the bargaining power λ which allows in addition for social optimality as shown by Hosios (1990). On the firm's side exists a vacancy opening parameter κ which is indeterminate as it appears always in a linear combination with the number of vacancies v, but it strongly affects the job–filling rate for vacancies. Shimer (2005a) states that vacancies are on average filled within 90 days, implying a monthly job–filling rate of 33%.

Finally, the total destruction rate of jobs is calibrated to obtain a job duration of 2.5 years. This implies a monthly job destruction rate of about 3.5%, which needs to be split up in an exogenous and an endogenous part. For this I assume an exogenous monthly destruction rate ρ of 3% and calibrate the standard deviation of the log–normal distribution of productivities to match the remaining destruction probability. The standard deviation the standard deviation of $Z(a)$ is chosen in such a way that in steady state $s_u m Z(a_r)$ equates the necessary 0.5% required for endogenous job–destruction. We normalize the mean of the productivity distribution roughly to 1 and use $\mu_Z = 0$ (it is a log–normal distribution). All parameter values and the steady state values for other variables can be found in Table 1 of the appendix.

1.3.1 Computation

The computational difficulty in the model lies in the endogenous distribution of productivity–specific matches. The entire distribution represents a state variable within the model and is characterized by a discontinuity at the reservation productivity as may be seen in figure !!fig . An identification of the distribution function by its moments would capture the discontinuity only insufficiently, and would not account adequately for its dynamic behavior. In order to compute the evolution of the employment distribution I proceed by discretizing the distribution for a large number of points on the relevant support covering the necessary range for the endogenous and exogenous distribution.

The results regarding distributional variables are computed through numerical integration over the interpolating points. To obtain the dynamic evolution of the system I then linearly approximate every single interpolation point using the algorithm by Klein (2000). But this method requires differentiability of the variables at the steady state. In our model this is not the case at the reservation productivity a_r due to the discontinuity within the endogenous distribution $G_t(a)$ of matched workers.

To circumvent the approximation difficulty within the law of motion for employment (1.13) or (1.21) around the steady state, the reservation productivity a_{rt} is "smoothed out". This implies substituting the employment decision to around the threshold level a_{rt} by a logistic probabibility function. In this way employment acceptance is no longer a binary decision above or below the reservation productivity but is stochastic depending on the variance of the logistic probability function.

An implication of the smoothing method is that an employer–worker match with positive surplus may be endogenously severed though it would have become productive in the original setup, and in the opposite case, a match with negative surplus may remain operative generating negative profits for firms or being sub–optimal for unemployment workers. Although the smooth constrained is not an exact application of economic theory, it nevertheless bears some realistic features. Employment decisions, especially at the margin to unemployment, may be influenced by a number of other factors including noisy signals regarding the characteristics of the job or the employee. This is captured by the approximation proposed

The logistic distribution $\Lambda_{a_{rt}}(a)$ is characterized by two parameters, its mean and its standard deviation. We choose the mean $\mu_{\Lambda,t}$ to be identical to a_{rt} and let it move with the reservation productivity over the business cycle. To quantify the error of the smoothing procedure we will compare the steady state features of an exact specification to versions with different standard deviations of the logistic distribution. The aim is to find a value for σ_Λ which allows for reasonable approximation results without altering the quantitative results compared to a hypothetical thorough solution. In order to assess the quality of the approximation we present robustness checks for three different values in table 1.2 and a graphical presentation on the effects for the endogenous employment distribution in figure 1.2.

1.4 Results

The results for this paper shall be presented in two parts, one regarding the properties of the non–stochastic steady state, and a second part which addresses the dynamic part in form of impulse–response functions and relative variances compared with the data. We compare three scenarios: the first scenario with exogenous search by unemployed workers only, the second with optimal search intensity by unemployed and the third one with optimal search by both unemployed and employed workers.

A main element of the steady state characterization is the endogenous productivity–distribution of matched workers (figure 1.1). It captures those workers that are currently matched to an employer and are able work within the match in the same period. On–the–job search influences the characteristics of the distribution in two ways. Firstly, the endogenous distribution is shifted to the right when compared to the exogenous productivity distribution $Z(a)$ from which match–specific productivities are drawn. This effect increases aggregate productivity as workers climb the productivity ladder, before being exogenously separated. Secondly, the career possibilities for employed workers lets unemployed workers accept matches with lower productivities compared to a situation lacking on–the–job search. The reservation productivity is therefore lower and the support of the endogenous distribution larger.

The steady state values for search time and intensity lead to an average on–the–job search intensity of roughly 20% compared to the search intensity by unemployed. But in terms of search time this represents only 6% of the search time by unemployed. Unemployed workers search 14.9% of their time, which is comparable to the exogenous value assumed by Andolfatto (1996). The average working time of workers is with 38.8% somewhat larger than the generally value of 1/3 used in the real business cycle literature.

Unemployed workers find with a probability of 45% a new employment, while on–the–job searchers change their occupation with an effective probability of 2.74%, these take already into consideration that roughly 7/10 of all matches are rejected by job–to–job transitioners. Figure 1.3 represents graphically job–finding probabilities for unemployed and employed workers including the number of job–to–job transitioning workers. The difference to other models is the productivity–dependent search intensity for employed workers. It leads to a negative relationship between on–the–job search and idiosyncratic productivity. A worker employed with the highest possible productivity $\bar{a}$ has no incentives to search for better jobs as these do not exist. On the contrary a worker employed with the reservation productivity a_r has large probabilities of finding higher productivity jobs. It is the rejection possibility of rejecting a matched job that leads to a downward sloped relationship for on–the–job search. The pool of workers transiting to a new job from their current employment is concentrated at low productivity levels.

The results of the dynamic analysis are made for technology shocks characterized by a monthly persistence of $\tilde{\rho} = 0.983$ which is 95% at the quarterly time horizon. The graphs in figure 1.4 state the impulse responses of different variables to a one-percent increase in aggregate technology A_t.

The results regarding output, capital and consumption are nearly unaltered to other real–business–cycle models, though on–the–job search amplifies the effects on total output slightly.

The labor market variables, vacancies and unemployment differ in this model to the standard model for various reasons. To understand the mechanism it is necessary to grasp the effects on search intensity and the composition effects over time. Search by unemployed workers and employed workers increase with a positive productivity shock due to higher returns from working, so individuals try to move into employment activity by searching more intensively. In other terms, the intratemporal decision for the individual induces her to move to the more profitable activity between consuming leisure and searching. With higher productivity, search has become more profitable than leisure and the individual substitutes away from leisure. The shape of total search does not only mimic the shapes of search intensities, but its curvature over time is reinforced due to a composition effect, where low productivity workers move to higher productivities and hence the drop in total search activity is stronger than the individual search intensities of employed and unemployed workers.

The main difference of this paper compared to other models lies in the separation of search intensity and unemployment. Search intensity increases after a positive technology shock for unemployed and employed workers and unemployment decreases by more than in the models with less variable search. While the reduction in unemployed after a positive shock leads to a reduction in total search intensity in standard models, this is no longer the case with on–the–job search. Even more, in the model with on–the–job–search, unemployment decreases by much more than in the other two models, and nevertheless the search effort is persistently higher. When unemployed workers enter the pool of employment they continue to search for better jobs, keeping total search effort high, which in turn exerts an externality onto vacancy openings. Firms are confronted with higher total search effort from the worker's side, which makes job openings more profitable, if the rejection rate by on–the–job searchers remains low. Thanks to the productivity–dependent search effort it effectively does not harm the matching process to such a degree to reduce the movements in vacancy openings.

By combining the interplay between search variations at the extensive and intensive margin as well as the contributions of changes in vacancies we obtain the result on the two central variables, the theoretical and empirical labor market tightness. The theoretical tightness is the choice variable of firms in the model and exhibits much lower persistence in the model with otjs. But the empirically relevant ratio of vacancies to unemployment increases by more than 3% on impact, much more than in the standard models. This is to attribute to the strong decline in unemployment after a positive technology shock.

Regarding the volatilities reported in table 1.3, no model is able to reproduce the empirical values either for unemployed or vacancy volatilities. But endogenous search does augment the volatility of labor market tightness significantly, especially when applied to on–the–job search. The otjs model doubles the volatility compared to a model of exogenous search. Table 1.3 presents also a version of the model in which search by employed workers is 50% more efficient than that of unemployed. This further improves the volatilities of labor market variables, but at the expense of unrealistic wage volatility. In fact, on–the–job search volatility is strongly linked to the volatility of wages at the different productivity levels and further increments it.

Finally, on–the–job search generates persistence in the correlation of vacancies to unemployment improving the fit of the Beveridge curve to the data as may be seen in table 1.4 or figure 1.5.

1.5 Conclusion

The difficulty of the Mortensen–Pissarides model to reproduce the empirical volatility of the two main labor market variables, unemployment and vacancies, has been revealed by Shimer (2005a) and has triggered a vast research agenda to increase the model's volatility generally by increasing the volatility of firm's profits. This paper contributes to the literature from a slightly different angle. It includes endogenous search effort by unemployed and productivity–dependent search by employed workers and uncouples in this way search intensity from the number of unemployed, which otherwise leads to counterfactual behavior of the vacancy–unemployment ratio. Endogenous search indeed increases the volatility of labor market variables, but not to the extent necessary to account for the empirical volatility. The increase in volatility is due to the procyclical search behavior of workers. Comparing the contribution to volatility by endogenous search with the contributions by rigid wages as in Hall (2005) and Gertler and Trigari (2005) it is apparent that directly affecting firm profits is more effective for business cycle volatilities.

This model reopens the debate on the replacement ratio and gives it a clear economic meaning by including explicitly time as economic variable within the search process and the valuation of the outside option. The outside option of an unemployed worker consists of the gain in leisure time compared to the situation when working. Clearly, the next steps are to introduce a pecuniary insurance scheme for unemployed and introduce variable working hours for employed workers in order to assess how the adjustment to a productivity shocks takes place: at the search margin in order to find better jobs or at the working margin in order to profit from temporarily higher productivity at the current employment.

Furthermore an analysis of the role of preference shocks would be a natural continuation combined with a theory of the allocation of time to work, search and leisure as already emphasized by Hall (1997).

Bibliography

Andolfatto, D. (1996), Business cycles and labor-market search, *American Economic Review*, 86(1), 112–32.

Caballero, R. J. and M. L. Hammour (1994), The cleansing effect of recessions, *American Economic Review*, 84(5), 1350–68.

Davis, S. J., J. Haltiwanger, and J. Schuh (1996), *Job Creation and Job Destruction*, The MIT Press, Boston, MA.

den Haan, W. J., G. Ramey, and J. Watson (2000), Job destruction and propagation of shocks, *American Economic Review*, 90(3), 482–98.

Gertler, M. and A. Trigari (2005), Unemployment fluctuations with staggered nash wage bargaining, *Working Paper*, New York University.

Hagedorn, M. and I. Manovskii (2005), The cyclical behavior of equilibrium unemployment and vacancies revisited, *Working Paper*.

Hall, R. E. (1997), Macroeconomic fluctuations and the allocation of time, *Journal of Labor Economics*, 15(1), S223–50.

Hall, R. E. (2005), Separating the business cycle from other economic fluctuations, *NBER Inc NBER Working Paper*, 11651.
URL `http://www.nber.org/papers/w11651.pdf`

Hall, R. E. and P. R. Milgrom (2005), The limited influence of unemployment on the wage bargain, *NBER Inc NBER Working Paper*, 11245.
URL `http://www.nber.org/papers/w11245.pdf`

Hashimzade, N. and S. Ortigueira (2005), Endogenous business cycles with frictional labour markets, *Economic Journal*, 115(502), C161–75.

Hosios, A. J. (1990), On the efficiency of matching and related models of search and unemployment, *Review of Economic Studies*, 57(2), 279–98.

Klein, P. (2000), Using the generalized schur form to solve a multivariate linear rational expectations model, *Journal of Economic Dynamics and Control*, 24(10), 1405–23.

Krause, M. and T. Lubik (2004a), A note on instability and indeterminacy in search and matching models, *Department of Economics The Johns Hopkins University Economics Working Working Paper*.
URL `http://www.econ.jhu.edu/pdf/papers/WP518lubik.pdf`

Krause, M. U. and T. A. Lubik (2004b), On-the-job search and the cyclical dynamics of the labor market, *Tilburg University Center for Economic Research Discussion Paper*, 92.
URL `http://greywww.kub.nl:2080/greyfiles/center/2004/doc/92.pdf`

Merz, M. (1995), Search in the labor market and the real business cycle, *Journal of Monetary Economics*, 36(2), 269–300.

Mortensen, D. T. and E. Nagypal (2005), More on unemployment and vacancy fluctuations, *IZA Discussion Paper*, DP(1765).

Nagypal, E. (2004), Worker reallocation over the business cycle: The importance of job-to-job transitions, *Working Paper, Northwestern University*.

Nagypal, E. (2005), Labor-market fluctuations, on-the-job search, and the acceptance curse, *Working Paper, Northwestern University*.

Nash, J. (1950), The bargaining problem, *Econometrica*, 18, 155–162.

Petrongolo, B. and C. A. Pissarides (2001), Looking into the black box: A survey of the matching function, *Journal of Economic Literature*, 39(2), 390–431.

Pissarides, C. A. (1985), Short-run equilibrium dynamics of unemployment vacancies, and real wages, *American Economic Review*, 75(4), 676–90.

Pissarides, C. A. (2000), *Equilibrium Unemployment Theory*, Second edition, MIT Press, Cambridge and London.

Shimer, R. (2005a), The cyclical behavior of equilibrium unemployment and vacancies, *American Economic Review*, 95(1), 25–49.

Shimer, R. (2005b), The cyclicality of hires, separations, and job-to-job transitions, *Federal Reserve Bank of St. Louis Review*, 87(4), 493–507.
URL `http://research.stlouisfed.org/publications/review`

1.A Appendix

1.A.1 Calibration values

Description	Parameter	Value	Calibration
Discount factor	β	99.67%	Annual interest rate of 4%
Depreciation rate	δ	0.87%	Annual capital depreciation of 10%
Capital elasticity in output	α	36%	Capital share K/Y of 36%
Relative weight of leisure	b	1.14	Replacement rate is 50%
Elasticity of leisure	ϕ	1.5	
Workers bargaining power	λ	0.5	Exogenously set
Search elast. matching fct.	η	0.5	Hosios condition
Matching function efficiency	γ	0.539	Job–finding rate for unempl.: 45%
Exog. job destruction rate	ρ	3%	Exog. job destr. 8.7% per quarter
Search efficiency	$g_u = g_w$	0.149	Search intensity by unemp. $s_u = 1$
search costs elasticity	$\gamma_u = \gamma_w$	2	Quadratic search costs
vacancy posting costs	κ	0.7652	Effective job–filling rate 33%
Mean of exog.prod. distrib.	μ_Z	0	Normalization: $\exp\left(\mu + \sigma^2/2\right) \approx 1$
Std. dev. of exog. distrib.	σ_Z	0.2	
Mean of logistic distribution	$\mu_\Lambda = a_r$	$a_r = 0.887$	Steady state cut–off productivity
Std. dev. of logistic distrib.	σ_Λ	$0.001, 0.02, 0.05$	

Table 1.1: Calibration of the model's parameters.

1.A.2 Steady state values

Values in brackets [] are used for calibration

	S.S. Value			
Var.	$\sigma_\Lambda = 0.001$	$\sigma_\Lambda = 0.02$	$\sigma_\Lambda = 0.05$	Description (monthly calibration)
$labsh$	60.8%	60.6%	60.0%	Labor share
$capsh$	[36%]	[36%]	[36%]	Capital share
c/y	[67%]	69.6%	69.2%	Consumption–output share
i/y	26.2%	26.2%	26.2%	Investment–output share
$\kappa\ v/y$	4.10%	4.15%	4.59%	Profit share
rep	[40%]	[40%]	[40%]	Replacement rate
ρ_n	1.05%	1.47%	2.85%	Endogenous job destruction
ρ_{ss}	4.05%	4.47%	5.85%	Total steady state job destruction rate
s_u	[1]	[1]	[1]	Search intensity by unemployed (normalized)
$\bar{s}_w$	20.2%	19.7%	19.7%	Average search intensity by employed
σ_u	14.9%	14.8%	14.7%	Search time by unemployed
$\bar{\sigma}_w$	0.89%	0.84%	0.74%	Average search time of employed
$\bar{\sigma}_w/\sigma_u$	5.97%	5.66%	5.04%	Relat. search time employed to unemployed
h_{med}	38.8%	39.0%	39.4%	Average working time
u	8.00%	8.72%	11.2%	Unemployment rate
f_u	[45%]	[45%]	[45%]	Job–finding rate for unemployed
f_e	2.74%	2.53%	2.17%	avg. effective job–to–job transion
q	[33%]	[33%]	[33%]	Effective Job–filling rate
rej	48.8%	48.1%	43.9%	Rate of rejected vacancies (employer's view)
$reje$	69.8%	71.5%	73.9%	Rate of rejection by otj searchers

Table 1.2: Steady State values for different variables. Calculated with 200 gridpoints and a standard deviation of the logistic smoothing function of 0.001, 0.02, 0.05

Relative standard deviations

σ_x/σ_y	y	c	i	emp	u	v	$\boldsymbol{\theta}$	θ_{theo}	w
US data	1	0.50	3.99	0.50	**7.33**	**8.72**	**15.83**	−−	0.47
with exogenous search	1	0.37	2.89	0.047	**0.71**	**0.93**	**1.56**	1.56	0.94
with endogenous search	1	0.37	2.89	0.093	**1.39**	**1.07**	**2.27**	1.47	0.89
with otjs	1	0.36	2.83	0.19	**2.00**	**1.04**	**2.96**	0.88	1.10
with otjs (otjs more efficient)	1	0.36	2.76	0.13	**1.81**	**1.47**	**3.21**	1.13	1.15

Table 1.3: Standard deviations are from log-deviations of quarterly data to HP(1600)–filtered series. Increasing degrees of search intensity from exogenous search to on-the-job search earns higher volatility for unemployment and vacancies.

Beveridge curve

$corr(v_{t+i}, u_t)$	−3	−2	−1	**0**	1	2	3
US data	−0.64	−0.82	−0.94	**−0.94**	−0.81	−0.61	−0.37
with exogenous search	−0.69	−0.84	−0.95	**−0.79**	−0.35	−0.15	−0.02
with endogenous search	−0.66	−0.78	−0.87	**−0.64**	−0.39	−0.20	−0.44
with otjs	−0.56	−0.73	−0.85	**−0.78**	−0.56	−0.38	−0.19

Table 1.4: Beveridge curve. On-the-job search generates a more persistent negative correlation between vacancies and unemployment.

Contemporaneous correlations with output

ρ_{xy}	y	c	i	emp	u	v	θ	w
US data	1	0.79	0.93	0.84	−0.83	0.90	0.88	0.58
with exogenous search	1	0.89	0.99	0.95	−0.95	0.95	0.99	1.00
with endogenous search	1	0.89	0.99	0.96	−0.96	0.85	0.99	1.00
with OTJS	1	0.89	0.99	0.98	−0.98	0.93	0.98	1.00

Table 1.5: Contemporaneous correlations with output. On–the–job search increases employment and unemployment correlations only marginally.

1.A.3 The household's problem

Households maximize the discounted value of aggregate utility subject to the individual's budget, labor and time constraint. The Lagrangian and the first order conditions are stated below

$$\begin{aligned}
\mathcal{L} &= \ln c_t + b\left\{ n_t \int_{G_t(a_{rt})}^{1} \frac{l_t(a)^{1-\phi}}{1-\phi} dG_t(a) + \left[1 - n_t\left(1 - G_t(a_{rt})\right)\right] \frac{l_{ut}^{1-\phi}}{1-\phi} \right\} \\
&+ \lambda_t \left[\pi_t + r_t k_t + n_t \int_{G_t(a_{rt})}^{1} w_t(a)\, h dG_t(a) - c_t - k_{t+1} + (1-\delta) k_t \right] \\
&+ \int_0^1 \psi_t(a) \left[\begin{array}{c} u_t s_{ut} m_t z(a) - n_{t+1} g_{t+1}(a) + \mathbf{1}_{\{a \geq a_{rt}\}} (1-\rho) n_t \times \\ \times \left\{ g_t(a) + m_t \left[z(a) \int_{a_{rt}}^{a} s_{wt}(\tilde{a})\, dG_{t+1}(\tilde{a}) - [1 - Z(a)]\, s_{wt}(a)\, g_t(a) \right] \right\} \end{array} \right] da \\
&+ \int_{G_t(a_{rt})}^{1} \epsilon_t(a) \left[1 - h - l_t(a) - \sigma_w\left(s_{wt}(a)\right)\right] n_t dG_t(a) + \epsilon_{ut} \left(1 - n_t \int_{a_{rt}}^{\bar{a}} dG_t(a)\right) \left[1 - l_{ut} - \sigma_u(s_{ut})\right]
\end{aligned}$$

First–order conditions of the household

Consumption c_t:

$$u_{c_t} = \frac{1}{c_t} = \lambda_t$$

Leisure l_t:

$$\begin{aligned} \epsilon_t(a) &= b l_t(a)^{-\phi} \\ \epsilon_{ut} &= b l_{ut}^{-\phi} \end{aligned} \quad \Longleftrightarrow \quad \begin{aligned} \epsilon_t(a)/\lambda_t &= b c_t l_t(a)^{-\phi} \\ \epsilon_{ut}/\lambda_t &= b c_t l_{ut}^{-\phi} \end{aligned}$$

Capital k_{t+1}:

$$\frac{1}{c_t} = \beta E_t \left[\frac{1}{c_{t+1}} (1 + r_{t+1} - \delta) \right]$$

Productivity–specific employment $n_{t+1} g_{t+1}(a)$ (using $\psi_t(a)/\lambda_t = W_t(a)$, and $u_t = 1 - n_t \int_{G_t(a_{rt})}^{1} dG_t(a)$)

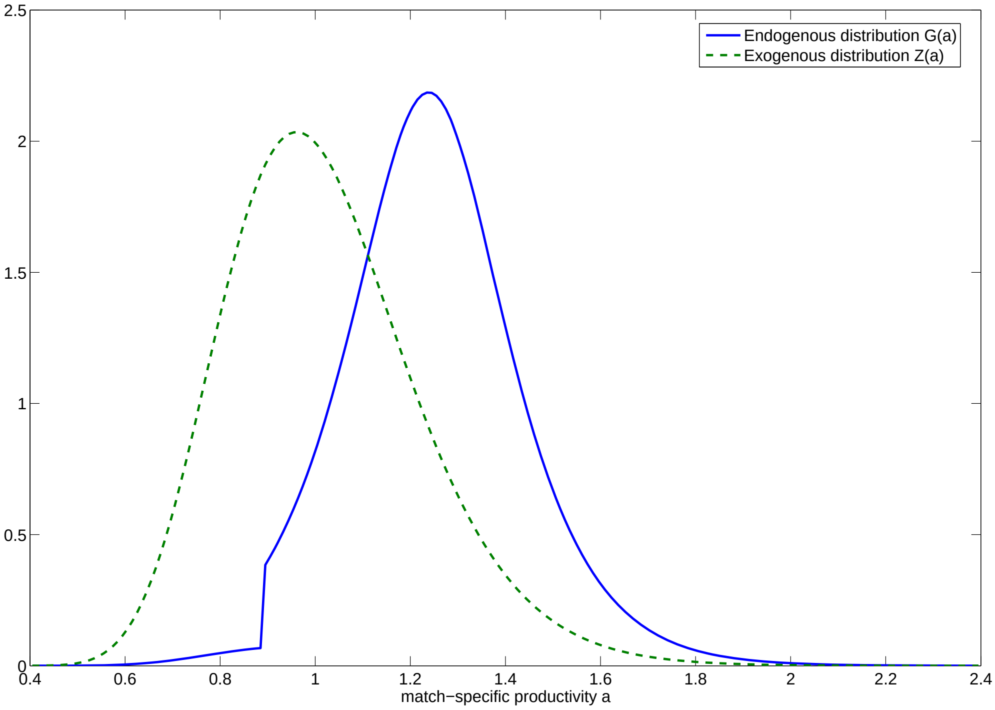

Figure 1.1: Exogenous $z(a)$ and endogenous $g_t(a)$ steady state productivity distribution of matched workers. Workers matched to an employer can start working in the given period. The endogenous distribution is shifted to the right due to on-the-job search and exhibits a kink at the reservation productivity a_r due to endogenous job-destruction decisions.

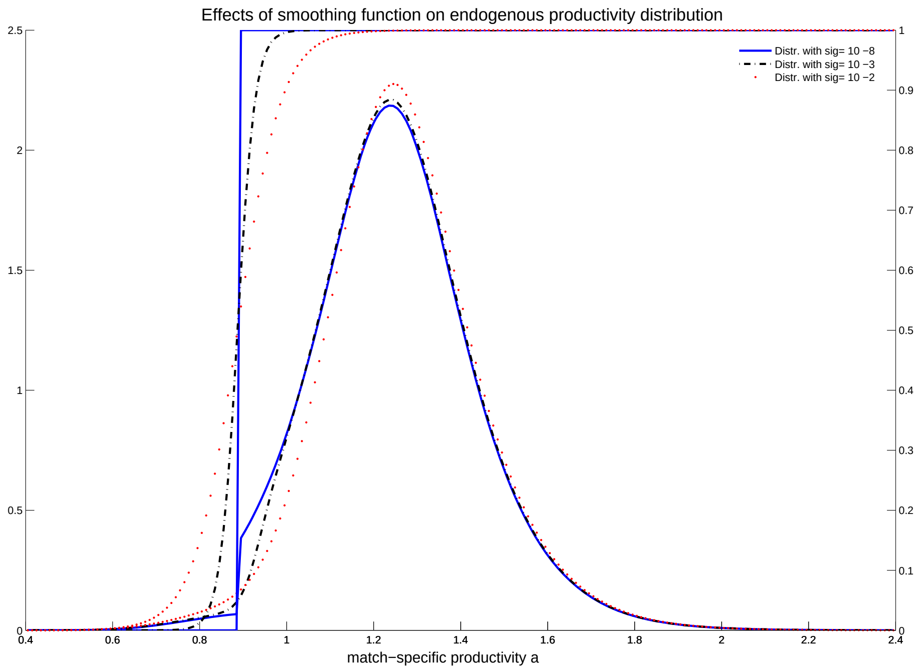

Figure 1.2: Effect of "smoothing out" the employment constrained at the reservation productivity a_r by a logistic c.d.f. of different standard deviations. With larger variance, the endogenous productivity distribution of matched workers becomes smoother and shifts to the right. Numerical computation is done with $\sigma_\Lambda = 10^{-3}$. (left scale: matched worker distribution, right scale: logistic distribution)

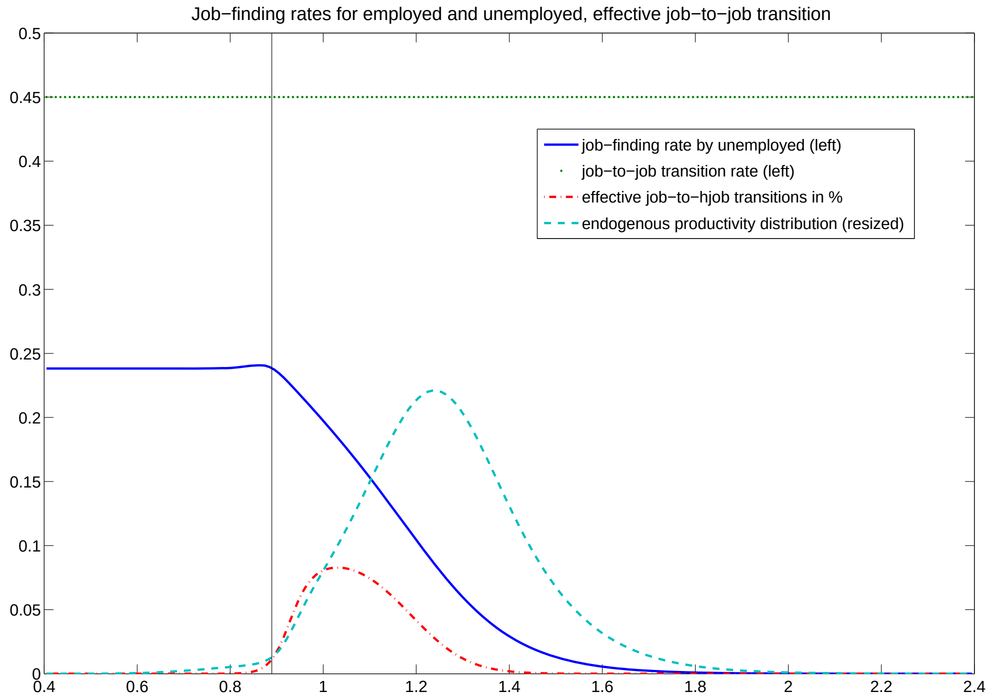

Figure 1.3: The job finding rates of unemployed and employed workers including the effective number (rescaled to 10^-2) of job-to-job transitioners. The bulk of transitioners is concentrated in low-productivity jobs.

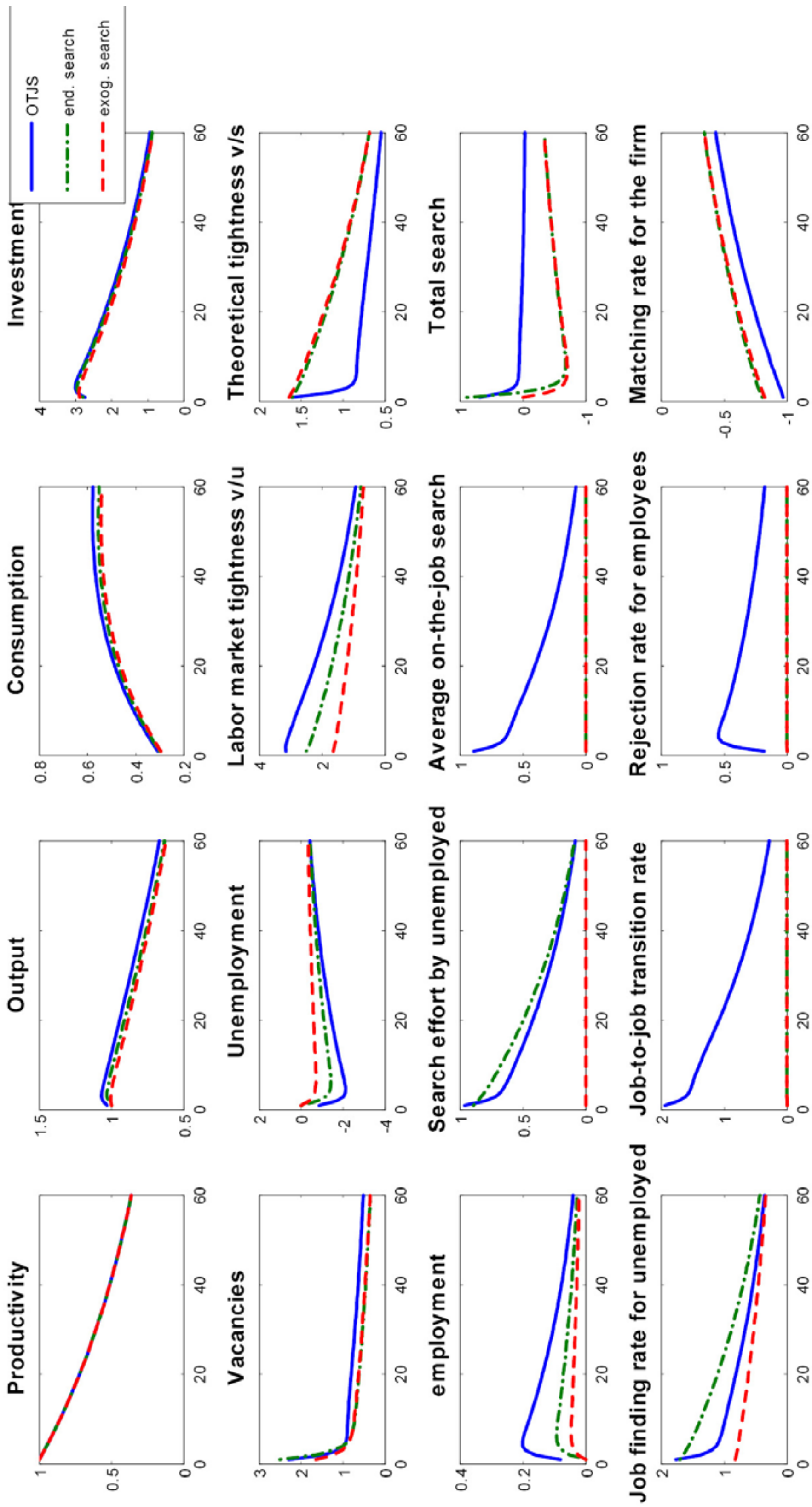

Figure 1.4: Impulse response of selected variables to a 1% positive technology shock within all three distinct models.

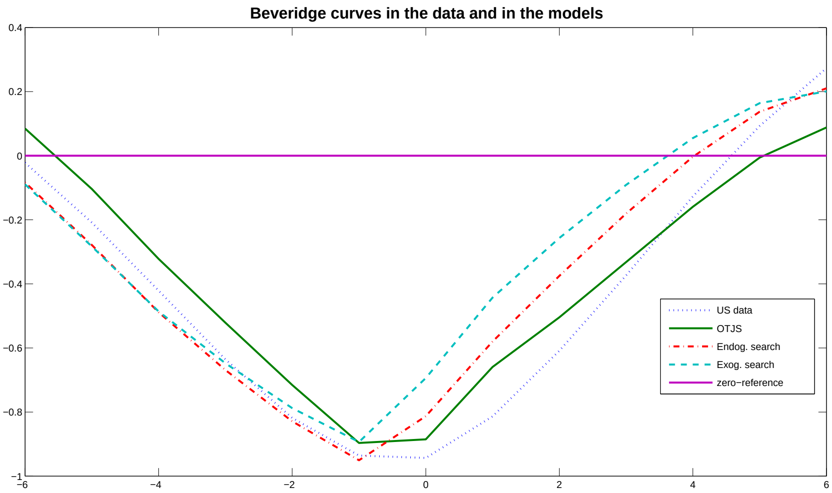

Figure 1.5: Beveridge curve with 6 leads and lags comparing US empirical data and the three distinct models: with exogeneous search, with endogenous search by unemployed and with on–the–job search. The version with on–the–job search generates an evolution most similar to the data.

$$W_t^1(a) = \beta E_t \left\{ \frac{c_t}{c_{t+1}} \left[\begin{array}{c} w_{t+1}(a)\, h + (1-\rho)\, W_{t+1}^1(a) \\ + (1-\rho)\, m_{t+1} s_{wt+1}(a) \int_{Z(a)}^1 \left[W_{t+1}^1(\tilde{a}) - W_{t+1}^1(a) \right] dZ(\tilde{a}) \\ - \left[\frac{bc_{t+1}}{1-\phi} \left(l_{u,t+1}^{1-\phi} - l_{t+1}(a)^{1-\phi} \right) + s_{ut+1} m_{t+1} \int_{Z(a_{rt+1})}^1 W_{t+1}^1(a)\, dZ(a) \right] \end{array} \right] \right\}$$

Reservation strategy of workers

$$W_t(a) = \max\left[0, W_t^1(a)\right]$$

Average Job value $\bar{W}_t$

$$\bar{W}_t = \int_0^1 W_t(a)\, dZ(a) = \int_0^1 W_t(a)\, dZ(a)$$

Search intensity by unemployed s_{ut}

$$\frac{u_{l_{ut}}}{u_{l_{at}}} \sigma_u'(s_{ut}) = bc_t l_{ut}^{-\phi} \sigma'(s_{ut}) = m_t \int_0^1 W_t(a)\, dZ(a) = m_t \bar{W}_t$$

Search intensity by employed s_{wt}

$$bc_t l_t(a)^{-\phi} \sigma'[s_{wt}(a)] = (1-\rho)\, m_t \int_{Z(a)}^1 [W_t(\tilde{a}) - W_t(a)]\, dZ(\tilde{a})$$

Reservation productivity a_{rt}

$$w_t(a_{rt})\, h + (1-\rho)\, s_{wt}(a_{rt})\, m_t \bar{W}_t = \frac{bc_t}{1-\phi} \left[l_{ut}^{1-\phi} - l_t(a_{rt})^{1-\phi} \right] + s_{ut} m_t \bar{W}_t$$

The equation for job-values and search intensities change slightly when "smoothing" the constraint on the reservation productivity with a logistic distribution function:

$$\Lambda_{a_{rt}}(a) = \frac{1}{1 + \exp\left(\frac{\mu_\Lambda - a_{rt}}{\sigma_\Lambda} \right)}$$

The change occurs in the law of motion for matches:

$$\begin{aligned} n_{t+1} g_{t+1}(a) &= u_t s_{ut} m_t z(a) + (1-\rho)\, n_t \Lambda_{a_{rt}}(a) \times \\ &\times \left\{ g_t(a) + m_t \left[z(a) \int_{G_{t+1}(a_{rt})}^{G_{t+1}(a)} s_{wt}(\tilde{a})\, dG_{t+1}(\tilde{a}) - [1 - Z(a)]\, s_{wt}(a)\, g_t(a) \right] \right\} \end{aligned}$$

$$\bar{W}_t^\Lambda = \int_0^1 \Lambda_{a_{rt}}(a)\, W_t^\Lambda(a)\, dZ(a)$$

$$W_t^\Lambda(a) = \beta E_t \left\{ \frac{c_t}{c_{t+1}} \left[\begin{array}{c} w_{t+1}(a)\, h + (1-\rho)\, \Lambda_{a_{rt+1}}(a)\, W_{t+1}(a) \\ + (1-\rho)\, m_{t+1} s_{wt+1}(a) \int_{Z(a)}^1 \Lambda_{a_{rt+1}}(\tilde{a}) \left[W_{t+1}^\Lambda(\tilde{a}) - W_{t+1}^\Lambda(a) \right] dZ(\tilde{a}) \\ - \left(\frac{bc_{t+1}}{1-\phi} \left[l_{u,t+1}^{1-\phi} - l_{t+1}(a)^{1-\phi} \right] + s_{ut+1} m_{t+1} \bar{W}_t^\Lambda \right) \end{array} \right] \right\}$$

$$\frac{u_{l_{ut}}}{u_{l_{at}}}\sigma'_u(s_{ut}) = bc_t l_{ut}^{-\phi}\sigma'(s_{ut}) = m_t \bar{W}_t^{\Lambda}$$

$$bc_t [l_t(a)]^{-\phi}\sigma'[s_{wt}(a)] = (1-\rho)\, m_t \int_{Z(a)}^{1} \Lambda_{a_{rt}}(\tilde{a}) \left[W_t^{\Lambda}(\tilde{a}) - W_t^{\Lambda}(a)\right] dZ(\tilde{a})$$

1.A.4 The firm's problem

The individual firm i maximizes the discounted value of profits by renting capital from the households and opening vacancies subject to the productivity–specific labor market constraints

$$L_i = E_0 \sum_{t=0}^{\infty} \beta^t \frac{\lambda_t}{\lambda_0} \left\{ \begin{array}{l} \int_{G_t(a_{rt})}^{1} \left[a A_t h^{1-\alpha} k_t^{\alpha} - r_t k_t - w_t(a)\, h\right] n_t dG_t(a) - \kappa v_t + \\ \int_0^1 \mu_t(a) \left[\begin{array}{l} q_t v_t z(a) - n_{t+1} g_{t+1}(a) + \mathbf{1}_{\{a \geq a_{rt}\}} (1-\rho)\, n_t \times \\ \times \left\{ g_t(a) + m_t s_{wt}(a) \left[z(a) \int_{a_{rt}}^{a} dG_t(\tilde{a}) - [1 - Z(a)]\, g_t(a) \right] \right\} \end{array} \right] dG_t(a) \end{array} \right\}$$

First–order conditions for the firm

Capital per capita k_t:

$$\alpha A_t h^{1-\alpha} k_t^{\alpha-1} \int_{G_t(a_{rt})}^{1} a\, dG_t(a) = r_t \int_{G_t(a_{rt})}^{1} dG_t(a)$$

Job value to the firm $n_{t+1} g_{t+1}(a)$:

$$J_t^1(a) = \beta E_t \left\{ \frac{\lambda_{t+1}}{\lambda_t} \left[\begin{array}{l} a A_{t+1} h^{1-\alpha} k_{t+1}^{\alpha} - r_{t+1} k_{t+1} - w_{t+1}(a)\, h + (1-\rho)\, J_{t+1}(a) \\ + (1-\rho)\, s_{wt}(a)\, m_t \int_{Z(a)}^{1} [J_{t+1}(\tilde{a}) - J_{t+1}(a)]\, dZ(\tilde{a}) \end{array} \right] \right\}$$

Reservation–productivity strategy of the firm:

$$J_t(a) = \max\left[0, J_t^1(a)\right]$$

Vacancy openings v_t

$$\frac{\kappa}{q_t} = \int_0^1 J_t(a)\, dZ(a) = \bar{J}_t$$

Reservation productivity a_{rt}

$$w_t(a_{rt}) + r_t k_t = a_{rt} A_t h^{1-\alpha} k_t^{\alpha} + (1-\rho)\, m_t s_{wt}(a_{rt}) \int_{Z(a_{rt})}^{1} J_t(\tilde{a})\, dZ(\tilde{a})$$

Rewriting the F.O.C of the firm with a "smoothed" reservation productivity using the "smoothed" law of motion for labor:

$$
\begin{aligned}
n_{t+1}g_{t+1}(a) &= q_t v_t z(a) + (1-\rho)\, n_t \mathbf{\Lambda}_{a_{rt}}(a) \times \\
&\times \left\{ g_t(a) + m_t \left[z(a) \int_{G_t(a_{rt})}^{G_t(a)} s_{wt}(\tilde{a})\, dG_t(\tilde{a}) - [1 - Z(a)]\, s_{wt}(a)\, g_t(a) \right] \right\}
\end{aligned}
$$

$$
J_t^{\Lambda}(a) = \beta E_t \left\{ \frac{\lambda_{t+1}}{\lambda_t} \left[\begin{array}{l} aA_{t+1}h^{1-\alpha}k_{t+1}^{\alpha} - r_{t+1}k_{t+1} - w_{t+1}(a)\, h + (1-\rho)\, \Lambda_{a_{rt+1}}(a)\, J_{t+1}(a) \\ +(1-\rho)\, s_{wt}(a)\, m_t \int_{Z(a)}^{1} \Lambda_{a_{rt+1}}(\tilde{a}) \left[J_{t+1}(\tilde{a}) - J_{t+1}(a) \right] dZ(\tilde{a}) \end{array} \right] \right\}
$$

$$
\frac{\kappa}{q_t} = \int_0^1 \Lambda_{a_{rt}}(a)\, J_t^{\Lambda}(a)\, dZ(a) = \bar{J}_t^{\Lambda}
$$

1.A.5 Searching, matching bargaining equations

Sharing rule between worker and employer

$$
(1-\lambda)\, W_t(a) = \lambda J_t(a)
$$

Bargain between workers and employers over the entire wage bill:

$$
w_t(a)\, h = \lambda \left(aA_t h^{1-\alpha} k^{\alpha} - r_t k_t \right) + (1-\lambda) \left[\frac{bc_t}{1-\phi} \left(l_{ut}^{1-\phi} - l_t(a)^{1-\phi} \right) + s_{ut} m_t \bar{W}_t \right]
$$

Matching technology

$$
M(s_t, v_t) = \gamma s_t^{\eta} v_t^{1-\eta}
$$

Matching rate per search unit

$$
m_t = \frac{M(s_t, v_t)}{s_t} = \gamma \left(\frac{v_t}{s_t} \right)^{1-\eta}
$$

Effective job filling rate (after correcting for rejections))

$$
q_t = \frac{M(s_t, v_t)}{v_t} \left\{ 1 - \frac{n_t \int_{G_t(a_{rt})}^{1} s_{wt}(a)\, Z(a)\, dG_t(a)}{s_t} \right\}
$$

Total search intensity

$$
s_t = s_{u_t} u_t + n_t \int_{G_t(a_{rt})}^{1} s_{wt}(a) dG_t(a)
$$

Search technology

$$
\sigma\left[s_{it}(a) \right] = g_i s_{it}(a)^{\gamma_i}, i = u, w,\ \gamma_i > 1
$$

Employment

$$EMP_t = n_t \left[1 - G_t \left(a_{rt}\right)\right]$$

Unemployment

$$u_t = 1 - n_t \left[1 - G_t \left(a_{rt}\right)\right]$$

Evolution of capital

$$K_{t+1} = \left(1 - \delta\right) K_t + Y_t - C_t - \kappa v_t$$

Aggregate Capital

$$K_t = k_t n_t \int_{G_t(a_{rt})}^{1} dG_t \left(a\right)$$

Evolution of technology

$$A_{t+1} = \tilde{\rho} A_t \exp\left[\varepsilon_t\right], \quad \varepsilon_t \tilde{\ } N\left(0, \sigma_\varepsilon\right)$$

1.B Steady state equilibrium conditions

Interest rate (from Euler equation)

$$r = \frac{1}{\beta} - 1 + \delta$$

Resource constraint

$$Ah^{1-\alpha}k^{\alpha}n \int_{G(a_r)}^{1} adG(a) = c + \delta kn \int_{G(a_r)}^{1} dG(a) + \kappa v$$

Demand for capital:

$$\alpha Ah^{1-\alpha}k^{\alpha-1} \int_{G(a_r)}^{1} adG(a) = r \int_{G(a_r)}^{1} dG(a)$$

Matched worker distribution:

$$\begin{aligned} & \left[\rho + 1_{\{a \geq a_r\}} (1-\rho) m s_w(a) [1 - Z(a)]\right] ng(a) \\ = \quad & mz(a) \left[us_u + 1_{\{a \geq a_r\}} (1-\rho) n \int_{G(a_r)}^{G(a)} s_w(\tilde{a}) dG(\tilde{a}) \right] \end{aligned}$$

Worker's value $W(a)$

$$\begin{aligned} & \{r - \delta + \rho + (1-\rho) s_w(a) m [1 - Z(a)]\} W(a) \\ = \quad & w(a) h + (1-\rho) s_w(a) m \int_{Z(a)}^{1} W(\tilde{a}) dZ(\tilde{a}) - \frac{bc}{1-\phi} \left[l_u^{1-\phi} - l(a)^{1-\phi} \right] - s_u m \bar{W} \end{aligned}$$

Job creation condition (Explicit specification for $J(a)$ and sharing rule $(1-\lambda) W(a) = \lambda J(a)$

are employed)

$$\begin{aligned} \frac{\kappa}{q} &= \int_0^1 J(a) dZ(a) \\ (r - \delta + \rho) \frac{\kappa}{q} &= \int_0^1 \{aAh^{1-\alpha}k^{\alpha} - rk - w(a) h \\ & + (1-\rho) s_w(a) m \frac{(1-\lambda)}{\lambda} \int_{Z(a)}^{1} [W(\tilde{a}) - W(a)] dZ(\tilde{a})\} dZ(a) \end{aligned}$$

Hours worked h

$$\int_{G(a_r)}^{1} w(a) dG(a) = \int_{G(a_r)}^{1} bcl_a^{-\phi} dG(a)$$

Unemployed search intensity s_u

$$\begin{aligned} (1-\lambda) bcl_u^{-\phi} \sigma'(s_u) &= \lambda\kappa \frac{m}{q} \\ (1-\lambda) bcl_u^{-\phi} \sigma'(s_u) &= \lambda\kappa \frac{v}{s} \left\{ 1 - \frac{n \int_{G(a_r)}^{1} s_w(a) Z(a) dG(a)}{s} \right\}^{-1} \end{aligned}$$

On–the–job search intensity $s_w(a)$

$$bcl(a)^{-\phi}\,\sigma'[s_w(a)] = (1-\rho)\,m\int_{Z(a)}^{1}[W(\tilde{a}) - W(a)]\,dZ(\tilde{a})$$

Reservation productivity a_r (employing wage specification and expected job value)

$$\lambda\left[aAh^{1-\alpha}k^{\alpha} - rk - \lambda\frac{bc}{1-\phi}\left(l_u^{1-\phi} - l(a)^{1-\phi}\right)\right] = [\lambda s_u - (1-\rho)\,s_w(a_r)]\,bcl_u^{-\phi}\sigma_u'(s_u)$$

Chapter 2

Imitation and Competition

2.1 Introduction

The relationship between competition and growth is a recurrent topic for growth economists and competition authorities alike. The Schumpeterian growth literature has identified rents from product market competition as the main incentive for innovative activities. Larger expected profits lead to larger returns of R&D investment and more innovations. This stands in stark contrast to the interest of competition authorities to increase product market competition and reduce firm profits for the benefit of consumers. This paper focuses on imitation as an essential part of the relationship between competition and growth. To proxy competition we use imitation costs relative to the costs of innovation and are able to reproduce the empirically found hump–shaped relationship between competition (imitation costs) and growth.

The underlying mechanism of the model exploits the fact that entry of imitators reduces the rents of incumbents and fosters races for new innovations. Incumbents have incentives to innovate for future higher profits due to better technology, but at the same time they need to renounce to their current profit. The fact of self-replacing itself, the so–called Arrow–effect, constitutes an obstacle for innovation incentives. The imitation of a technology reduces incumbent's current profits, per se this reduces the incentives for future innovators, but at the same time the reduction of these profits reduces also the value that is cannibalized. By reducing the incumbent's profits, imitators both reduce incentives for innovation, but reduce also the value that needs to be replaced. Through these two effects, imitation and growth have a hump–shaped relationship.

The empirical findings on the relationship between competition and growth are abundant, but only recently obtained clear results as they are confronted with numerous measurement problems. Competition or innovations are difficult to measure. To measure product market competition the literature has used profits either at firm or industry level, price markups (Lerner index) or the number of firms in the industry. Obviously, these measures are *not* exogenous in a dynamic framework, especially when accounting for entry and exit which make the market structures endogenous.

The first contribution by Blundell et al. (1995) measures innovations by "technologically significant and commercially important innovations" following the definitions of the Science Policy Research Unit (SPRU). The results of the paper appear contradicting themselves. On the one hand dominant firms in an industry (those with larger market shares) tend to innovate more, but on the other hand more concentrated industries are less innovative. This is the first evidence that neither of the two corner solutions, full competition or monopoly, are optimal to foster growth.

More recently Aghion et al. (2005) analyzed the relationship between the price markup over production costs, the Lerner index, as measure of competition and weighted patent citations as measure of innovations to proxy for growth. They conclude that the relationship between competition and growth is of inverted U shape. Their study indicates that too large or too low profits are both detrimental to growth within an industry, indicating that too high profits make firms sluggish in their R&D effort and too fierce competition erodes profits that are a necessary component for innovation incentives.

A third empirical study by Nickell (1996) focuses on total factor productivity instead of using innovations or patents as proxy for growth, and use either industry–wide rents or the number of firms as competition indicators. The results are that firm's profitability decreases with competition but productivity of the industry increases, and in addition higher market shares reduce the level of productivity, as opposed to the result by Blundell et al. (1995). Nickell's results would fit best to the view that competition improves static and dynamic efficiency, but the other papers tend to indicate an in–between solution, neither perfect competition nor too concentrated markets are optimal to foster innovations and growth.

The main difficulty in translating empirical measures such as profit levels or market shares to theoretical concepts of competition is the endogeneity of these measures. In this papers we underline the benefits of imitation costs to proxy competition as compared to other concepts in the theoretical literature.

The variable is exogenous to the measures of competition in product markets, but influences these effectively. Price markups and profits are determined in the model by a Cournot game, hence dependent on the number of incumbent firms. Imitation costs act as entry barrier to the industry, but entry itself and the number of incumbent firms are endogenous. Imitation costs affect profits and entry, but their realization is determined by general equilibrium.

Furthermore, imitation costs can be influenced easily by industrial policies such as taxation, subsidies, entry laws, patent policy or even the enforcement of patent policy as used in an extreme form by Boldrin and Levine (2005). This stands in strong contrast to preference parameters (see Segerstrom (1991) and Mukoyama (2003)) or parameters of product substitution such as in Aghion et al. (1997) or Koeniger and Licandro (2006). These values alter the price markup that producing firms can charge to their costumers (either final good firms or consumers), which in turn alters the profit level of firms. The concept of substitutability has two main drawbacks, it is difficult to measure and it cannot be influenced by policy action. Especially when interpreting it as a preference parameter, the influence policy makers might have seems limited. In contrast, imitation costs are at the reach of policy.

Imitation costs can be varied in a continuous way allowing to understand gradual changes in competition. When using the switch from Cournot to Bertrand competition as a measure to increase competition, only two single competition environments may be analyzed. In addition, the switch between the two types of interaction (prices and quantities) alters profoundly the environment of the economy and is only feasible for symmetric firms as in the studies by Boone (2001) and DenicolÃš and Zanchettin (2003). With heterogeneous firms, as we have here, Bertrand competition implies that the technologically superior firm applies limit pricing which drives the follower out of the market.

In addition to the theoretical advantages, surveys state that imitation costs do play an important role in shaping the market structure in the process of innovation, as they affect the innovation intensities of incumbents as well as the speed of imitation by outsiders. The empirical survey by Mansfield et al. (1981) finds that the majority of products in the sample have been imitated within four years of its innovation. Regarding the costs and timing of imitation, they find that imitation costs are 65% of innovation costs and it takes 70% of the innovation time to bring an imitation to the market. This is evidence for the lower costs and shorter development times for imitations compared to innovations. Another finding in the analysis is the fact, that the existence of patents do not preclude imitators from joining an industry, but they do increase imitation costs as the imitator circumvents the patent legislation. Finally, Mansfield et al. (1981) find the evidence that higher imitation costs (due to legislation, patents, etc.) lead to stronger market concentration by creating higher hurdles for entry. We will use these facts to assess the theoretical implications in the model. These are the facts that our model reproduces entirely.

The model in this paper is based on Mukoyama (2003) and aims to replicate the empirical results by Aghion et al. (2005) and Blundell et al. (1995), in which firms with larger market share invest more, but more concentrated industries generate lower innovation incentives. Intermediate good firms compete in quantities à la Cournot and may operate with different technologies. The step size between technologies is fixed and innovations are non–drastic, such that various technological levels are accommodated within the same industry. Technology is sequential in the sense that outsiders need to enter the market first through imitation and innovate only in a second step. Imitation reduces profits of incumbent firms, alters the market structure of he industry and leads to a continuous turnaround in the industry through entry, which makes other firms obsolete.

We reproduce the hump–shaped relationship between costs of imitation and growth found by Aghion et al. (2005). Two opposing effects generate the hump–shaped relationship between competition and growth: the first effect is the well–known Schumpeterian effect. Lower imitation costs lead to higher entry into the industry reducing in this way the level of achievable profits in the industry and at the same time shortening the length of time of more concentrated industries. The direct implication is a reduction in the returns to innovation and hence of innovation investment. The second effect relates to the Arrow–effect of self–replacement. Firms with high profits (monopolies) do not replace themselves in order not to destroy their own present profits. Imitation reduces these profits and reestablishes the incentives for innovation for the incumbent firms. Hence, imitation alters the market structure by increasing the number of firms in the industry and shifts the economy to a more innovative environment. We call this the composition effect. The interplay of the composition and the Schumpeterian effect leads to the hump–shaped relationship between competition and growth. From the perspective of static efficiency, imitation is undoubtedly good for welfare as it lowers prices, increases productive efficiency (industry leaders are imitated) and augments the amount of goods being consumed. But from a dynamic perspective the results depend on the current level of imitation.

Beyond the innovation–growth relationship the model accounts for several other regularities in the innovation–growth literature. By allowing heterogeneous firms we find that technologically more advanced firms innovate more. These are also the larger firms in the industry, either measured by output or by profits. Their incentives for innovations stems from the escape–competition effect, as successful innovations allows them to apply monopoly pricing, while laggards would still find themselves constrained by the other incumbents. As a direct consequence, leaders in an industry remain leaders with higher probability. Regarding entry and exit, the model replicates continuous entry through imitation and exit through technological obsolescence. In this process higher imitation costs prolong imitation time and therefore lead to more concentrated industries, as also found by Mansfield et al. (1981).

The model inserts itself into the Neo–Schumpeterian literature with its earliest version by Aghion and Howitt (1992). The literature focussed generally on drastic innovations, in which an innovation replaces directly former technologies, while we address non–drastic innovations to allow for technologically heterogeneous incumbents but with homogeneous output goods. In the main Schumpeterian literature patents are required as incentives to pursue innovative R&D

to protect post–innovation profits. Abolishing the idea of patents and allowing for imitation, Segerstrom (1991) and Mukoyama (2003), on which we base the presented model here, show that incentives to pursue innovations still exist due to the first–mover advantage by maintaining technologies secret, which may be quite substantial as shown by Gort and Klepper (1982). The main drawback of the initial models is that only outsiders replace incumbents, as these do not self–replace themselves. We therefore allow for innovation races of incumbents and the coexistence of firms with heterogeneous technology. To allow for heterogeneous firms in markets of homogeneous goods it is necessary to include Cournot competition combined with non–drastic innovations, which received only little attention in the literature. We build here on the work by Barro and Sala-i Martin (1995) and DenicolÃš and Zanchettin (2003), but our proxy of competition is different from theirs.

Finally, comparing with most prominent research strand on competition and growth by Aghion et al. (2005) we allow for entry and exit and use the threat of entry as part of measure of competition. Faster entry decreases the rate of profits (and also the Lerner index) and therefore is an important determinant of competition. In fact, the authors themselves state that "an important extension would be to introduce entry and entry threat as alternative measures of competition." The entry mechanism for imitators exploited in this model may be seen as the counterpart to step–by–step innovations proposed by Aghion et al. (1997).

The paper is structured as follows: the next section describes the model with the setups for production in the final and intermediate goods sector, the innovation and imitation and structure and the household problem in a general form. Section 2.3 reduces the dimensional space of the setup by imposing parameter restrictions and solves the model numerically. Section 2.4 relates competition to growth and unveils numerically the mechanism behind the hump–shaped relationship. Furthermore we relate the proxy for imitation to empirical measures of competition more commonly used in the literature.

2.2 The model

The economy consists of a an intermediate and a final good sector following the basic Schumpeterian literature as in Aghion and Howitt (1992). The final good sector produces a unique output good (numéraire) under perfect competition with a constant returns to scale technology using labor and a composite intermediate good as input. Its output is used for consumption and as input in the intermediate and the R&D sector.

The intermediate sector is subdivided into a continuum of industries $\omega\epsilon\,[0,1]$. Within each industry firms produce an identical good of a given quality with different technologies and compete in quantities à la Cournot[1]. To allow for various technological levels within industries we assume non–drastic innovations, implying that the quality gap between technological leaders and followers is sufficiently small not to make the first follower obsolete. In addition, Cournot competition prevents the technological leader from applying limit pricing as would be the case with Bertrand competition[2].

Innovations are made exclusively by incumbents, outsiders enter the market through imitation. The number of firms in each industry is endogenously determined through free entry by imitation and exit through obsolescence if production costs are higher than the Cournot market price. The market structure and the relative technological position of firms evolves endogenously through innovation races. In fact, for given parameter values which will always be fulfilled in the steady state considered, monopolistically operating firms do not replace themselves, the so called Arrow effect. Imitation reduces the current profits by incumbents, which removes the obstacle to their innovation, leading to an innovation race between incumbents. Entrants imitate the technological leader of an industry and succeed with a Poisson distributed hazard rate. Once successful, the new entrant engages in innovation activity. Technology is therefore cumulative, building first on imitation to enter and then become innovative when in the market.

2.2.1 Product markets

Final good sector

The representative firm in the final good sector uses labor L and a continuum of intermediate goods $x_{kt}\,(\omega)$, indexed by $\omega\epsilon\,[0,1]$ and by their reference quality level k, and combines these in a Cobb–Douglas production function with production elasticity α. Production takes the form

$$Y_t = L_t^{1-\alpha} \int_0^1 q^{\alpha k(\omega)} x_{kt}\,(\omega)^{\alpha}\, d\omega. \tag{2.1}$$

The representative firm of the final good sector uses $x_{kt}(\omega)$ units of intermediate goods of quality index $k\,(\omega)$ from industry ω. The final good sector is characterized by perfect competition and

[1] In this way the final sector demands a good of unique quality from the intermediate sector and the price for the intermediate good depends exclusively on the market structure in the intermediate industry.

[2] In fact, with price competition à la Bertrand and non-drastic innovations the optimal strategy for the technologically advanced producer is to undercut the followers' price by applying limit ricing and drive him out of the market. This leaves a single firm in the market, which earns positive profits.

the representative firm maximizes profits by choosing labor and the intermediate good optimally. The inverse demand for intermediate goods $x_k(\omega)$ by profit maximization under full competition is[3]

$$p(\omega) = \alpha q^{\alpha k(\omega)} x_k(\omega)^{\alpha-1}, \quad (2.2)$$

where prices equate marginal productivity of the intermediate good $p(\omega) = \frac{\partial Y}{\partial x_k(\omega)}$ (we normalized labor $L = 1$). The quantity demanded may be split into a stationary term and another one affected by the quality level k

$$x_k(\omega) = x(\omega) q^{\frac{\alpha}{1-\alpha} k(\omega)}, \quad (2.3)$$

where $x(\omega) \equiv \alpha^{\frac{1}{1-\alpha}} p(\omega)^{-\frac{1}{1-\alpha}}$ is the part of demand independent of the technological level, but varying with the market structure within industry ω through the Cournot price $p(\omega)$. The demand for intermediate goods fluctuates with the market price and increases with successive technological innovations by $q^{\frac{\alpha}{1-\alpha}}$, which represents the output increment between two innovations. This framework allows to disentangle technological factors, which grow along the balanced growth path, from market structure effects, which are stationary.

Intermediate goods sector

The intermediate goods sector is subdivided into a continuum of industries defined by ω and distributed on the support $[0, 1]$. The intermediate good in each industry is characterized by its quality k and produced by a number of active firms in the industry of differing productivity. We call *leader* the firm (or firms) with the current highest productivity and all other firms *followers* as these produce with inferior technology. The technological leader is characterized by lowest production costs, the firm of subsequent technology, the first follower, has q times the leader's costs, the second follower faces q^2 times the cost to produce the same good of quality k. Final product firms demand the composite good of quality index k, and production across intermediate firms within each industry is allocated via Cournot competition. In this way we obtain a quantity that increases with every quality step and a stationary price along the balanced growth path that varies exclusively with market structure of the industry.

Production and Cournot competition. Technological innovations are non–drastic and firms in each industry compete à la Cournot. These two assumptions allow for the simultaneous presence of technologically heterogeneous firms in the industry. Cournot competition is used in this

[3] In order to ease notation we drop the time subscript as long as it does not create confusion.

paper as it allows for gradual technological obsolescence of firms producing homogeneous goods. Bertrand competition might seem more realistic for its price competition in a static setup, but Kreps and Scheinkman (1983) demonstrated that a sequential setup with capacity decisions in a first step and Bertrand competition in a second step is equivalent to Cournot competition. In addition, in a Cournot setup profits within an industry depend on the technological composition of the industry, which allows to order industry setups by profit. Instead, a Bertrand setup for homogeneous goods generates generally either a monopolistic price or limit price with monopolistic or zero profits, respectively.

For Cournot competition with heterogeneous firms to be sensible, two conditions need to hold. First, we require $\alpha < 1$ for the production elasticity, otherwise intermediate firms with linear production costs cannot generate rents. Second, the non-drastic innovation condition, $0 < \alpha q < 1$, needs to hold. This ensures that minimum two firms of heterogeneous productivity can coexist in the market and the resulting Cournot price is larger than the production costs of the first follower. Non–drastic innovations have been less treated in the innovation–growth literature building on Grossman and Helpman (1991) or Aghion and Howitt (1992), but they are suited to account for firm dynamics and gradual obsolescence.

The production costs of intermediate firms is one unit of final goods, the numéraire, but each firm produces a different quality. By assuming that lower quality firms can compensate the lower quality by higher quantity, we obtain that the leader within an industry produces a good with quality $k(\omega)$ with unit costs, while the follower with lower quality requires q units in order to produce a good of equivalent quality.

The profit maximization problem of an intermediate firm of quality s to produce a good of quality $k(\omega)$ is

$$\max_{x_k^s(\omega)} \left[P\left[x_k(\omega)\right] - q^{k(\omega)-s} \right] x_k^s(\omega), \tag{2.4}$$

where $P[\cdot]$ characterizes the inverse demand by the final goods sector, $x_k(\omega)$ is the industry's total output and $x_k^s(\omega)$ is output of a firm of technological level s in an industry characterized by a state–of–the–art technology level of $k(\omega)$. The firm's constant marginal costs $q^{k(\omega)-s}$ are incurred by firm of quality s to produce a quality $k(\omega)$ good and increase with the distance to the industry frontier quality $k(\omega)$. The first–order condition for profit–maximization is

$$P\left[x_k(\omega)\right] + P'\left[x_k(\omega)\right] x_k^s(\omega) - q^{k(\omega)-s} = 0, \tag{2.5}$$

earning the Nash-equilibrium price of the game in quantities.

Summing the first order conditions (2.5) for all firms in an industry $n(\omega) = \sum_{s=0}^{k(\omega)} n_s(\omega)$ (where $n_s(\omega)$ is the number of firms operating with quality s) shows that the equilibrium price $P[x_k(\omega)]$ is independent of the distribution of marginal costs across firms, but depends only on the mean of productivities as found by Bergstrom and Varian (1985)

$$P[x_k(\omega)] = \frac{1}{n}\sum_{s=0}^{k(\omega)} n_s(\omega)\, q^{k(\omega)-s} - P'[x_k(\omega)]\, x_k(\omega)\,. \tag{2.6}$$

This equation simplifies in our model by using the demand function for intermediate goods (2.3) and exploiting the constant quality steps between subsequent technological levels to[4]

$$P(\omega) \equiv P[x_k(\omega)] = \frac{n(\omega)}{n(\omega)+\alpha-1}\sum_{s=0}^{k(\omega)} \frac{n_s(\omega)\, q^{k(\omega)-s}}{n(\omega)}. \tag{2.7}$$

The specification generalizes the price in a Cournot game with symmetric firms. Prices $P(\omega)$ decrease either due to a larger number of active firms $n(\omega)$ in the market, characterized in the first part of the expression, or through a more productive composition of firms within the industry represented in the second part. These two components, the influence of the number of firms and the efficiency effect are relevant later to determine the relative process of different market structures. Due to the fact that prices depend only on the relative productivities within the industry, we observe stationary prices along the balanced growth path, while intermediate output grows.

Innovation and imitation. Innovations are carried out exclusively by incumbent firms within each industry and their intensity is chosen optimally by the firm in order to maximize its present discounted value. We follow Mukoyama (2003) in modelling the innovation structure. Incumbent firms compete in a race for the next innovation using the final good as input. Their innovative R&D activity produces spillovers from which the other firms within the industry benefit, but not potential entrants or firms in other industries. These spillovers have been empirically tested, Spence (1984) and Cohen and Levin (1989) note in their analysis of R&D conduct that externalities exist within industries due to labor mobility and informational exchange. An R&D department cannot withhold workers or information perfectly. These leakages are helpful for competitors in the same field to improve their own innovative activity. The spillovers are necessary to increase the benefits of an innovation race between incumbents. The fact of entry reduces profits

[4] This is a generalization of the industry price found by DenicolÃš and Zanchettin (2003) for an industry structure with Cournot competition, non–drastic innovations and a single producer at every technological level.

by incumbents and would reduce innovation incentives, but thanks to the spillovers, incumbents have higher incentives to innovate and spur growth in this way. In addition, due to spillovers it is possible to employ linear innovation costs. In general, convex costs are used for innovation races (see Tirole (1988) or Barro and Sala-i Martin (1995)) to generate determinate interior solutions. The linear form in our presentation eases calculations to a large extent, but spillovers are necessary to generate sufficient incentives for incumbents.

The hazard rate for innovations follows a memoryless Poisson process identically and independently distributed for each firm. Incumbents face linear innovation costs, increasing with technological level of the industry in order to generate constant innovation hazard rates in steady state. An R&D intensity for innovation I_j by firm j requires costs of

$$c_I \left[I_j, k\left(\omega\right)\right] = q^{\frac{\alpha}{1-\alpha}k(\omega)} a_I I_j, \tag{2.8}$$

where a_I is the innovation cost per unit and total costs grow with $q^{\frac{\alpha}{1-\alpha}}$ for each increment of the quality index $k\left(\omega\right)$. The factor is identical to the increment in output of the intermediate sector for each successive innovation as seen in equation (2.3). The innovating firm chooses its innovation intensity I_j in a Nash game, taking the innovation intensity of other incumbents as given. The innovation intensity therefore maximizes the present discounted value of the firm, $\arg\max_{I_j} V\left(\cdot\right)$.

Thanks to the spillovers the hazard rate h_j for an individual firm j within industry ω is

$$h_j(I_j, I_{-j}) = I_j + \theta I_{-j}$$

and depends on the intensity I_j of the firm itself and the intensity by all other incumbents I_{-j}, weighted by the degree of spillovers in the industry $\theta \epsilon \left(0, 1\right)$. Due to the independence of innovations between firms the rate with which a successful innovation occurs in an industry is the sum of all firms' hazard rates $\sum_j h_j(I_j, I_{-j})$.

Regarding the situation of outsiders we assume that only incumbents are able to generate innovations. Outsiders need to enter the market first by imitating the incumbent leader and innovates in a second step as incumbent. Innovators do not hold an explicit patent for their technology giving imitators the possibility to copy the current state-of-the-art technology and participate in the market. The cost structure for imitation is similar to innovation. It is linear in the imitation effort and increases with the quality level in the industry, but imitators do not benefit from the spillovers that incumbents generate.

The hazard rate for a successful imitation follows a memoryless Poisson process. In order to imitate with a hazard rate C, an imitator incurs costs of

$$c_C\left[k\left(\omega\right)\right]=q^{\frac{\alpha}{1-\alpha}k(\omega)}a_Ca_IC,$$

with $q^{\frac{\alpha}{1-\alpha}k(\omega)}$ as adjustment factor for the sophistication of the incumbent technology, a_C a technology parameter representing relative costs for imitation (copying) relative to innovation costs a_I. We require that unit costs of imitation are smaller than the unit cost of innovations, implying $0<a_C\leq 1$. In this way we account for the fact that innovations are more costly to generate due to the uncertainty about the direction of research as empirically found by Mansfield et al. (1981). The proxy used for competition in this paper is a_C. Relative imitation costs are an entry barrier for outsiders on the one hand, determining in this way the imitation intensity C, but on the other hand, they fix also the value of a firm in the homogeneous duopoly due to free entry.

2.2.2 Households

The demand for final goods is determined by the preferences of a representative household offering labour inelastically to the firms in the final good sector. The representative consumer maximizes utility intertemporally

$$U_0=\int_0^\infty e^{-\rho t}u\left[c\left(t\right)\right]dt \tag{2.9}$$

where ρ is the subjective discount rate and the instantaneous utility function is concave and otherwise well-behaved. We will use a function with constant intertemporal elasticity of substitution $u\left[c\left(t\right)\right]=\frac{c(t)^{1-\sigma}}{1-\sigma}$. The household holds assets of the firms and has access to a perfectly functioning asset market. We can therefore represent the budget constraint of the consumer as an intertemporal budget constraint

$$\int_0^\infty e^{-r(t)t}Y\left(t\right)dt=A\left(0\right)+\int_0^\infty e^{-r(t)t}w\left(t\right)Ldt, \tag{2.10}$$

where the left hand side denotes expenditure with $r\left(t\right)$ the interest rate and $Y\left(t\right)$ the amount of final goods purchased. The right hand side denotes discounted assets in $t=0$ with $A(0)$ initial wealth, $w(t)$ the wage rate and L the labor force (normalized to 1 in the following). Firms of the intermediate sector are held by consumers. Their value equals the sum of all expected rents made by the firms. No accumulation of physical capital exists.

2.3 Model simplification

The preceding section described the general setup with innovation and imitation through free entry and may serve to analyze entry and exit dynamics as well as cross–section distribution of production within industries. To analyze the relationship between competition and growth and maintain heterogeneous firms it is already sufficient to reduce to two different technological levels within each industry, easing calculations in this way. For this, we further restrict the parameter space of the production parameter α and the technological step size q.

We reduce parameters in such a way that maximum two technological levels are active in each industry of the intermediate sector. The price level p with maximum two distinct levels of production costs in the intermediate sector simplifies for each industry industry ω from equation (2.7) to

$$P = \frac{n_k + n_{k-1}q}{n_k + n_{k-1} + \alpha - 1},\ n_k = 0, 1, 2 \tag{2.11}$$

where n_k and n_{k-1} are the number of firms of high and low quality respectively, and the industry specification ω has been dropped to ease notation. From the price equation (3.6) and the first–order condition in the Cournot game (2.5) we obtain the market shares σ_s for firms of technological level s within an industry

$$\sigma_s = \frac{P - q^{k-s}}{(1-\alpha)P}, \tag{2.12}$$

where P is the prevailing price in a given industry ω and s characterizes the technological level of the firm. Finally, profits for a firm of level s are

$$\pi_s = \left[P - q^{k-s}\right] \sigma_s x_k, \tag{2.13}$$

where x_k is total quantity demanded by firms in the final goods sector, and increasing in quality units with each innovation by the factor $q^{\frac{\alpha}{1-\alpha}}$. Profits for firm s can be conceptually divided into a growth component identified by the quality index k through the increasing quantity x_k and into a stationary component identified by costs to produce quality s and the resulting market share σ_s as well as by the market structure of the industry identified by the resulting price level P. Profits in two industries of different quality level, but identical market structures distinguish themselves only by a multiple of the technological step size, while the relative profits between firms within an industry are identical under the same market structure.

2.3.1 Market structures

We now present the different market structures that are possible with two technological levels and explain that the implied parameter restrictions reduce the number of market structures to only three.

1	2	3	4	5	6
•	•	• •	• •	•	• •
	•		•	• •	• •

Table 2.1: Six possible market structures when reducing the parameter space α - q to non–drastic innovations with maximum two technological levels within an industry.

Market structure 1 is a monopoly, market structure 2 is a heterogeneous duopoly with technological difference of q between firms. Market structure 3 is a homogeneous duopoly and consists of firms with identical technology. The other market structures are characterized in a similar way with either 3 or 4 active firms. We will show in the following that only market structures one to three arise endogenously over time. The parameter restriction and free entry for imitators are the two factors that lead to these results.

Within each of these market structures (except for the monopoly) innovations take place by any of the incumbents, if this is profitable. Imitators enter the industry by free entry equating expected profits to the imitation costs. Therefore imitation does not take place into recently imitated market structures $(3, 4, 6)$, but only into industries which arose from a recent innovation $(1, 2, 5)$, the single-leader markets. Free entry assures that only a single imitator enters the market and maximum two leaders operate.[5]

The parameter conditions on α and q in order to allow for only two technological levels require on the one hand non–drastic innovations, characterized by $0 < \alpha q < 1$, and on the other hand $1 < \alpha q^2$. The first condition has been discussed before, and the second condition implies that a leader with two technological steps to the follower actually escapes competition from the laggard and is able to produce in a monopolistic environment because the follower's costs are higher than

[5] Free entry equalizes expected profits from entry to the costs of imitation. The first imitator therefore equilzes the costs of imitation to the expected benefits of being an incumbent in the homogeneous duopoly leaving no further rents to be obtained by other potential entrants.

the monopoly price. In addition, the condition makes the laggard producers in market structures 4 and 6 obsolete. Their production costs are higher than the Cournot price (see appendix).

Using the price equation (3.6) and these conditions leads to a price order of $P_1 > P_2 > P_3$ (alternatively we will use notation $P_m > P_{het} > P_{\text{hom}}$). The order of prices for the different industries calculated with equation (3.6) is determined by two effects as emphasized in equation (2.7). Firstly, more firms in the industry lower prices like in the symmetric Cournot setup, secondly, the productivity composition of active firms influences prices. Firms with higher technology reduce the resulting price by higher market shares within the industry. Which one of the two effects is more important, number of firms or product efficiency, depends on the parameters. The restriction $1 < \alpha q^2$ implicitly weighs the influence from product efficiency stronger than the one from the number of firms and the complete price order is $P_1 > P_2 > P_5 > P_6 > P_4 > P_3$ (see appendix). Hence, assuming $1 < \alpha q^2$ not only reduces the technological to two, but implicitly reduces the number of feasible industry-types to a monopoly, heterogeneous duopoly and homogeneous duopoly.

The binding conditions for the numerical exercise in the paper are the non–drastic innovations $\alpha q < 1$, the possibility to escape competition $1 < \alpha q^2$ and the condition on no self–replacement: $\pi_m (g-1) < a_I [r + C_m (1 - a_C (g-1))]$.

In order to understand how the interplay of innovation and imitation changes an industry's market structure, table 2.2 resumes the transition probabilities between the three relevant market structures (monopoly, heterogeneous and homogeneous duopoly).

From\To	monopoly	heterogeneous duopoly	homogeneous duopoly
monop.	$1 - C_m$	0	C_m
het. duop.	$I_{het1} + \theta I_{het2}$	$1 - C_{het} - (I_{het1} + \theta I_{het2})$	C_{het}
hom. duop.	0	$(1+\theta)(I_{\text{hom}\,1} + I_{\text{hom}\,2})$	$1 - (1+\theta)(I_{\text{hom}\,1} + I_{\text{hom}\,2})$

Table 2.2: Transition matrix for the three different market structures, monopoly, homogeneous duopoly and heterogeneous duopoly. E.g. the probability to switch from a monopoly to homogeneous duopoly occurs with rate C_m, the imitation intensity into monopolistic market structures. Note that an innovation by the follower in a heterogeneous duopoly keeps the market structure unchanged, but increments the technological level of the industry.

The transition matrix represents the probability of an industry to move from a given market structure (on the left) to an alternative structure (top). I_i and C_i define innovative and imitative effort respectively, while θ is the intra-industry spillover for innovations. Assuming that the market structure is monopolistic, the industry becomes imitated with probability C_m or otherwise the market structure remains unchanged.[6] In the heterogeneous duopoly either the leader or the follower may innovate. If the leader succeeds in an innovation due to his own effort and the spillovers from the competitor, the industry becomes monopolistic, while if the follower succeeds, the market structure remains unchanged with inverted relative positions of the two firms and a higher quality level k. A further possibility in the heterogeneous duopoly is the imitation of the leader with probability C_{het} which leads to two leaders. The laggard firm exits the market because the production costs exceed the Cournot price, which makes the industry become a homogeneous duopoly.

Once in a homogeneous duopoly either one of the two incumbent firms is able to innovate. Such an innovation changes the market structure into a heterogeneous duopoly with higher technological level. Recall, due to free entry only a single imitation takes place and the homogeneous duopoly accommodates no third firm as expected profits for entrants have been driven to zero with the first imitation.

Every innovation leads to a production with higher quality level generating higher profits for the new firms, whereas imitation leads to a lowering of incumbents' profits through a more efficient market structure.

2.3.2 Value functions & optimal firm behavior

We describe the value of firms in the three different market structures (monopoly, homogeneous and heterogeneous duopoly) by a value function, which includes current profits and expected future profits weighted by their probability. In addition, we directly present the values in efficiency terms, i.e. corrected for the quality level. This is possible as profits, imitation and innovation costs all increase by the same factor $q^{\frac{\alpha}{1-\alpha}}$ with each innovation step. The index ω will be used to alternatively describe an industry or a market structure but the distinction is clarified in the text stating market structure ω or industry ω in case ambiguity arises. The value functions for the firms in the different market structures are

[6] We exclude the possibility of self–replacement by the single producer by chosing parameters in such a way that the Arrow effect of self–replacement is sufficiently large and the leader reduces its value by engaging in R&D activity (see appendix for the derivation).

$$\begin{aligned}
V_m &= \frac{\pi_m + C_m V_{\text{hom}}}{r + C_m} \\
V_{het1} &= \frac{\pi_{het1} + C_{het} V_{\text{hom}} + (I_{het1} + \theta I_{het2}) g V_m + (I_{22} + \theta I_{21}) g V_{het2} - a_I I_{het1}}{r + C_{het} + (1+\theta)(I_{het1} + I_{het2})} \\
V_{het2} &= \frac{\pi_{het2} + (I_{het2} + \theta I_{het1}) g V_{het1} - a_I I_{het2}}{r + C_{het} + (1+\theta)(I_{het1} + I_{het2})} \\
V_{\text{hom}} &= \frac{\pi_{\text{hom}} + (I_{\text{hom}\,j} + \theta I_{\text{hom},-j}) g V_{het1} + (I_{\text{hom},-j} + \theta I_{\text{hom},j}) g V_{het2} - a_I I_{\text{hom}\,j}}{r + (1+\theta)(I_{\text{hom},j} + I_{\text{hom},-j})}
\end{aligned} \tag{2.14}$$

V_m characterizes the present discounted value of a monopolistic firm which makes monopolistic profits π_m until imitated with probability C_m and finds itself in the market structure of homogeneous duopolies with value V_{hom}. The parameters are chosen in such a way that monopolists do not self-replace themselves. The discount factor is the interest rate r (constant in steady state) increased by the hazard rate for an imitation C_m. The structure of the other value functions for the individual firms are similar. The value for the leader in a heterogeneous duopoly (V_{het1}) earns profits until he becomes either imitated, whereafter the firm would be in a homogeneous duopoly, or innovates himself and operates as a monopolist by escaping competition, or the incumbent competitor innovates and the firm operates as a follower in a heterogeneous duopoly. Regarding the homogeneous duopoly we distinguish the two individual firms by j in order to account for the Nash game in the innovation race, although their value V_{hom} is identical. The probability for a change in market structure depends on the degree of spillovers between incumbent firms. Each innovation leads to an increase in the value by $g = q^{\frac{\alpha}{1-\alpha}}$ due to the technological advance, while imitations imply no technological benefits.

Incumbent firms maximize profits by competing in quantities in the product market, and by choosing optimal innovation effort I_i in the race for new innovations. The two levels of competition are distinct, one at the intratemporal level representing product market competition, the second at the intertemporal level in form of an innovation race. Firms therefore maximize their value in equation (2.14) with respect to innovative activity and then act optimally in the static Cournot game.

The innovation race is a Nash game in which incumbent firms choose the innovation intensity that maximizes their value, $\arg\max_{I_j} V_\omega$, taking as given the innovation intensity by their competitors I_{-j} and imitation activity C_ω by potential entrants.

The outcome of the Nash game, the reaction functions of the individual firms to their rival incumbent's strategy are

$$I_{het1} = \frac{(C_{het}+r)(gV_{het1}-a_I)-(1+\theta)\pi_{het2}}{(1+\theta)(a_I-(1-\theta)gV_{het1})} \quad (2.15)$$

$$I_{het2} = \frac{(C_{het}+r)(gV_m+\theta gV_{het2}-a_I)-(1+\theta)(C_{het}V_{\text{hom}}+\pi_{het1})}{(1+\theta)(a_I-(1-\theta)g(V_m-V_{het2})} \quad (2.16)$$

$$I_{\text{hom}} = \frac{r(gV_{het1}+\theta gV_{het2}-a_I)-(1+\theta)\pi_{\text{hom}}}{(1+\theta)(a_I-(1-\theta)g(V_{het1}-V_{het2})} \quad (2.17)$$

with $g \equiv q^{\frac{\alpha}{1-\alpha}}$ for simplification. The resulting innovation intensities are for each individual firm and are the outcome of the simultaneous Nash game for innovations between the two incumbent firms.

Free entry

Outsiders have free entry into the industry by imitating the most advanced incumbent. An imitation leads to a homogeneous duopoly independently of which market structure of the two single leader markets is imitated. Free entry implies that the present discounted value of a firm in the homogeneous duopoly is equated to the imitation costs incurred when entering the market. Due to the fact that no further rents are possible, an industry is only imitated once, leading to market structures with maximum two leaders. The success for an imitation is Poisson distributed with parameter C, reflecting the imitation intensity. Due to the linear costs for imitation, marginal costs $a_C a_I$ are constant and equate firm value V_{hom},

$$V_{\text{hom}} = a_C a_I. \quad (2.18)$$

Outsiders imitate the monopolist or the leader in a heterogeneous duopoly with identical intensity. The imitation intensity depends entirely on the market structure a firm faces when it successfully imitates,which is the homogeneous duopoly. Using the free entry condition (2.18) combined with the identity of imitation intensities $C_m = C_{het}$ and the innovation intensities (2.15)-(2.17), the value functions (2.14) can be solved for as functions of the model's parameters.

$$\begin{aligned}
V_m &= \frac{a_I(1+g+\theta)(1+\theta+g\theta)+a_C(1+\theta)[1+\theta\left(2-g^2+\theta\right)]}{g^2\left(1+\theta+g\theta\right)} \\
V_{het1} &= \frac{1+\theta+g\theta+a_C(1+\theta)^2}{g\left(1+\theta+g\theta\right)}a_I \\
V_{het2} &= \frac{1+\theta}{1+\theta+g\theta}a_C a_I \\
V_{\text{hom}} &= a_C a_I
\end{aligned}$$

The value of firms in the different market structures is independent of instantaneous profits π_i. This is due to the linear cost function of the innovative and imitative activity which absorb all profits. Firm values are therefore determined by structural parameters of the innovation process, a_I, a_C, θ and the quality increment $g \equiv q^{\frac{\alpha}{1-\alpha}}$. The higher imitation or innovation costs, the more valuable incumbent firms are, which is the influence of the general equilibrium structure of the model.

The first–order conditions combined with the free entry conditions permitted to identify the innovation intensities of firms in the duopolies and the intensity for copying a monopolistic firm. The setup does not pin down the intensity with which a leader in a heterogeneous duopoly is imitated. We will therefore analyze the results for different values.

2.3.3 Resource constraint

Labor is used for production of final goods, whereas the final good is used for consumption, innovation, imitation and intermediate good production. Output of intermediate industries increases through each innovation with a step size of $q^{\frac{\alpha}{1-\alpha}}$ (conditioning for identical market structures). The final good production (2.1) requires intermediate goods from every industry, hence aggregating over industries leads to a distribution of technological levels across intermediate sectors which we summarize by the index Q. This index increases by the measure $q^{\frac{\alpha}{1-\alpha}}$ if an innovation has taken place in *every* industry. To analyze the steady state, we employ output in efficiency terms Y_e, which aggregates output in efficiency terms from all intermediate sectors (see also Barro and Sala-i Martin (1995), section 6)

$$Y(t) = \left[\int_0^1 q^{k(\omega,t)} d\omega \right]^{\frac{\alpha}{1-\alpha}} Y_e(t) = Q(k,t)^{\frac{\alpha}{1-\alpha}} Y_e(t),$$

where $Q \equiv \left[\int_0^1 q^{k(\omega,t)} d\omega \right]^{-\frac{\alpha}{1-\alpha}}$ is the economy–wide quality index. It consists of the distribution of the state–of–the–art quality levels across all intermediate industries. The resource constraint for the final good in efficiency terms is

$$Y_e = c_e + x_e + c_{Ie} + c_{C_e}.$$

The use of output in efficiency units is subdivided into consumption c_e, the production of intermediate goods x_e, and to cover innovation and imitation, c_{I_e} and c_{C_e}, respectively. In more detail, when subdividing the use into the three different market structures we have

$$\begin{aligned} Y_e &= c_e + \\ &\alpha_m(x_{e1} + a_C a_I C) + \\ &\alpha_{het}\left[(\sigma_{het1} + q\sigma_{het2})x_{e2} + a_I(I_{het1} + I_{het2}) + a_C a_I C\right] + \\ &\alpha_{\text{hom}}(x_{e3} + 2I_{\text{hom}}), \end{aligned} \tag{2.19}$$

where α_m, α_{het}, α_{hom} are the shares of industries operating with the respective market structures and the bracketed terms express the use of intermediates in the three different industries. Final output is used for consumption, for production and imitation in a monopoly, production, imitation and imitation in a heterogeneous duopoly and for production and imitation in a homogeneous duopoly.

2.3.4 Steady State Equilibrium

The equilibrium of the economy is defined by the utility maximization of households (2.9) subject to their budget constraint (2.10). Final output firms maximize their profits by using (2.1) and demanding labor and intermediate input goods while taking prices as given. Intermediate firms maximize their per period profits by choosing the supplied quantity optimally in a Cournot game (2.4) and choose the innovation intensity to maximize their present discounted value (2.14), expressed by (2.15)–(2.16). Free entry through imitation is determined by condition (2.18). Finally the resource constraint (2.19) holds.

In order to obtain constant growth rates for the economy, we require additionally a steady state condition. Innovations and imitations in the intermediate industries occur at random moments in time. A steady state equilibrium requires that the share of industries of a given market share remains constant. Due to the continuum of intermediate industries we can apply the law of large numbers and require that the share of industries switching from a given market structure to another one be exactly replaced by switches from other market structures to the given one. If a given number of industries operating monopolistically become imitated in the time period dt, switching hence to a homogeneous duopoly, the same number of industries is required to switch from the heterogeneous duopoly to a monopoly. Characterizing by α_m, α_{het}, α_{hom} the share of industries operating as monopoly, heterogeneous or homogeneous duopoly, the steady state conditions may be written as

$$\begin{aligned} \text{monopoly:} \quad & \alpha_m C = \alpha_{het}\left(I_{het1} + \theta I_{het2}\right) \\ \text{heterog. duop.:} \quad & \alpha_{het}\left[\left(I_{het1} + \theta I_{het2}\right) + C\right] = \alpha_{\text{hom}} 2(1+\theta) I_{\text{hom}} \\ \text{homog. duop.:} \quad & \alpha_{\text{hom}} 2(1+\theta) I_{\text{hom}} = \left(\alpha_m + \alpha_{het}\right) C \end{aligned} \tag{2.20}$$

The left hand side characterizes the exit from the respective market structure and the right hand side characterizes entry into it for any given small time window. Note that the innovation by a follower in the heterogeneous duopoly changes the technological level of the industry, but not the market structure. In addition to the steady state conditions (2.20) we have that all shares of the different market structures sum up to

$$1 = \alpha_m + \alpha_{het} + \alpha_{\text{hom}}.$$

The steady state fractions of prevailing market structures in a stationary equilibrium depend exclusively on the relative innovation and imitation effort combined with the degree of spillovers for innovations between firms.

Their exact form can be found in the appendix, but the influence of innovation and imitation on steady state market shares is listed in table 2.3. Obviously, the levels of innovation and imitation themselves are not exogenous, but their effects in a partial equilibrium style help to understand the economic mechanism.

	I_{het1}	I_{het2}	I_{hom}	C	θ
α_m	+	+	+	−	+
α_{het}	−	−	+	±	±
α_{hom}	0	0	−	+	−

Table 2.3: The effects of innovation I, imitation C and spillovers θ onto the share of industries operating in the three different market structures (monopoly, homogeneous and heterogeneous duopoly).

An increase in R&D effort by the leader in the heterogeneous duopoly (I_{het1}) leads to a higher share of monopolies and reduces the share of heterogeneous duopolies and leaves the share of homogeneous duopolies unchanged. An increase in the innovation activity by the follower I_{het2}

has a similar effect. The new issue in this paper is the fact that imitation represented by the R&D activity C shifts the market structure towards homogeneous duopolies, which are the innovative market structures, while the effect on the heterogeneous duopoly is ambiguous, and depends on the level of innovation and imitation, as we will see in the numerical results. Imitation influences the composition of the economy's market structure and tilts it towards the innovative structures, away from monopolies (composition effect). On the other hand, the value function indicate that imitation reduces the value of monopolies and firms in the heterogeneous market structure through by increasing the discount factor (Schumpeterian effect), equation (2.14). The numerical exercise shall deliver the quantitative importance of these two effects. Finally, the spillover favors innovation and hence the occurrence of more concentrated industries towards monopolies

2.3.5 Growth of the economy

In the preceding sections we identified that the increase in intermediate output (equation (2.3)) for subsequent quality levels with identical market structure is $q^{\frac{\alpha}{1-\alpha}}$. The increment for profits is identical to this factor as may be derived from (2.4), as well as innovation and imitation costs. To obtain the growth rate of the economy we exploit the characteristics of the Poisson distribution for the irregular occurrences of imitation and innovation. The growth rate of the economy is the share of a given market structure multiplied by the probability of an imitation or innovation and taking into account the output increment due to technological step and the price change

$$\begin{aligned}\dot{Y}/Y = \; & \alpha_m C\left[\left(\frac{1+\alpha}{2\alpha}\right)^{\frac{1}{1-\alpha}}-1\right]+ \\ & \alpha_{het} C\left[\left(\frac{1+q}{2}\right)^{\frac{1}{1-\alpha}}-1\right]+ \\ & \alpha_{het}\left(I_{het1}+\theta I_{het2}\right)\left[\left(\frac{\alpha\left(1+q\right)}{1+\alpha}\right)^{\frac{1}{1-\alpha}}g-1\right]+ \\ & \alpha_{het}\left(I_{het2}+\theta I_{het1}\right)\left[g-1\right]+ \\ & \alpha_{\hom}2\left(1+\theta\right)I_{\hom}\left[\left(\frac{2}{1+q}\right)^{\frac{1}{1-\alpha}}g-1\right]\end{aligned} \tag{2.21}$$

Equation (2.21) states the steady state growth rate of final goods, which is identical to the growth rate of the intermediate good, as may be derived from (2.1). The growth rate composes itself of the sum of output increments for every industry structure, weighted by the shares of industries operating in a given market structure. Output of each industry changes through two different sources, prices and technology effects. The first line states the output increases due to imitation. These are pure price effects representing static improvements as imitators enter the market of monopolies and heterogeneous duopolies. Due to the price decline by the extra competitor in the market, the Cournot price is lower in the new market structure and output is larger. The second source of growth is innovation in the industries currently operating duopolistically. Their contribution is a mix of the technological improvement and the price change. Innovations lead to a higher quality level, expanding intermediate good production in this way, but may lead to less firms in their market or a more unfavorable technological composition which offset the technological advantages due to price increases.

2.4 Competition and Growth

Competition is difficult to define. Variables used in the literature to measure competition have been firm or industry profits, the number of firms, switches from Bertrand to Cournot competition, the Lerner index relating profits to marginal cots, entry barriers and others. It is even more difficult to translate these measures into theoretical concepts which have an influence on competition without determining the endogenous outcome too closely. The main interest of this paper is to reproduce the empirically found hump shaped relationship between competition and growth using as unique proxy for competition in the model entry costs in form of imitation costs, a_C.

In fact, the relative imitation costs affect the imitation intensity, which in turn affects the share of duopolies in the economy as seen from table 2.3, which themselves are the market structure that permit innovation to take place. In addition to this direct effect, a_C has a general equilibrium effect through free entry, it fixes the value of firms in the homogeneous duopoly and the value of firms in the market. The general equilibrium is an important extension compared to other models such as Aghion et al. (2006) dealing with entry threat in a partial equilibrium manner. The advantage of imitation costs lies in the fact that it is measurable (although imperfectly) and can be used within as a model parameter for competition without affecting the outcome in a trivial way. Also, industrial policy can directly influence its value, either financially through subsidies or taxes, entry barriers to industries or even patent policies. In fact, the degree of patent protection would have a direct effect on the imitation costs with stricter patent protection forcing imitators to circumvent existing patents which necessarily increases their costs as described in Mansfield et al. (1981).

In this paper we concentrate on the costs of imitation relative to innovation costs captured by a_C and we show how it translates to economy–wide profits and the average number of firms per industry. The reference values for the numerical exercise is characterized by the values in table 2.4 and uses an annual time horizon.

Param	Value	
r	0.07	Interest rate/ rate of time preference
α	0.75	Production elasticity
q	1.15	Quality step size between innov.
θ	0.5	Innovation spillovers
a_I	0.3	Innovation costs

Table 2.4: Reference parameters for the numerical exercise.

The example values are chosen to reflect some empirical values without being a strict calibration. Interest rates are set to 7%, the production elasticity gives is set to 0.75 and two subsequent technologies have 15% difference in quality, which earns an output increase of $g = q^{\frac{\alpha}{1-\alpha}} = 1.52$ between innovations. Although this might seem big, it is nevertheless a non–drastic innovation step size. Regarding θ and a_I, they are chosen in order to obtain positive innovation and imitation intensities as well as positive market shares for the three market structures. The innovation cost a_I reflects 9 years of monopoly profits.

We choose $a_C = 0.7$ in order to describe a steady state economy with its market shares, innovation and imitation intensities and its growth rate.

With the parameter values of table 2.4 the market structure with the largest share in the economy is the monopoly, followed by the homogeneous duopoly. Note that the general evolution of market structure is from monopoly to homogeneous duopolies through imitation and then through innovations back to monopolies. Analyzing the different intensities of innovation and imitation clarifies that especially the leader in the heterogeneous duopoly escapes from that market structure shifting steady state market structures towards a monopolistic one. The innovation intensity by the leader, expressing the escape–competition effect in the heterogeneous duopoly, is

Variable (in %)	
α_m	75.8
α_{het}	3.32
α_{hom}	20.9
I_{het1}	17.8
I_{het2}	0.45
I_{hom}	1.00
C	0.79
γ	0.85

Table 2.5: Steady state values for different variables using the refernce parameters and an imitation cost ac=0.7

larger than the innovative intensity in the homogeneous duopoly. This slightly alters the result found by Aghion et al. (1997), in which the situation of neck-to-neck generates largest innovative incentives, while a leader–follower structure is less conducive to innovation. But their model does not allow for entry and exit and therefore does not permit a leader to escape competition from the follower altogether.

2.4.1 The costs of imitation as proxy for competition

This model uses imitation costs as proxy for competition. We use the relative price of imitation to innovation to identify the ease of entry. In this way we do not directly alter the general entry costs for firms but only the part affecting imitation. Note that the imitation costs fix the value of the homogeneous duopoly in the general equilibrium and therefore affects the value of firms. This is different to models of entry threat that take a partial equilibrium, without taking into account that free entry affects the value of firms in all market structures.

Figure 2.1 presents the relationship between imitation costs relative to innovation costs and the growth rate of the economy in percentage points. The figure represents a hump–shaped relationship between competition expressed as costs of imitation and growth in the economy (higher imitation costs represent lower competition). The graph plots the parameter space for which innovations, the shares of the three market structures are positive, i.e. $\alpha_m, \alpha_{het}, \alpha_{\text{hom}} > 0$ and innovation and imitation rates are positive $I_{het1}, I_{het2}, I_{\text{hom}}, C > 0$. The dotted part on the

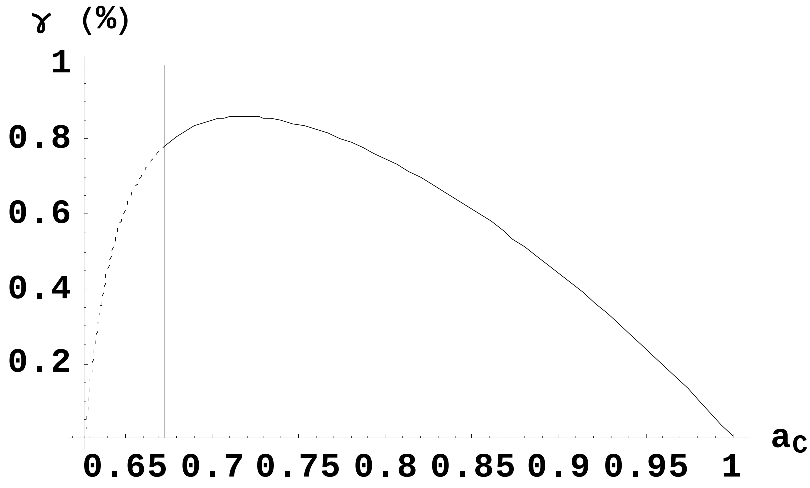

Figure 2.1: Relationship between imitation costs relative to innovation costs and growth. The dotted line, marks an area, where optimal innovation intensities of laggards in the heterogeneous duopoly would be negative.

left is stated for completeness, in that area the optimal innovation intensity by the follower (I_{het}) in a heterogeneous duopoly is negative.

This graph reflects the empiircal findings by Nickell (1996), Blundell et al. (1995) and more recently Aghion et al. (2005). Low competition through higher imitation costs and implicitly lower imitation leads to lower growth rates, while too fierce competition due to low entry barriers also decreases growth. Imitation costs generate two different contemporaneous effects on the innovation process, the Schumpeterian effect and a composition effect.

In order to better understand the mechanisms with which the hump–shaped relationship is generated, we present figures 2.2 and 2.3 which plot the innovation and imitation intensities and the shares of the different market structures, respectively. The levels of innovation by firms in different market structures varies strongly. The leader in a heterogeneous duopoly is by far the strongest innovator, which is due to the possibility to escape competition altogether by becoming a monopolist. For the rest, innovation and imitation exhibit roughly the same magnitude. The effect of imitation costs on innovation intensities is positive, while it is negative on imitation intensity C. With higher costs of imitation, less imitation takes place and the incentives for innovation increase. This is exactly the Schumpeter effect.

The composition effect which counteracts the Schumpeterian effect establishes that with less competition due to higher imitation costs the economy consists mainly of monopolists which is a

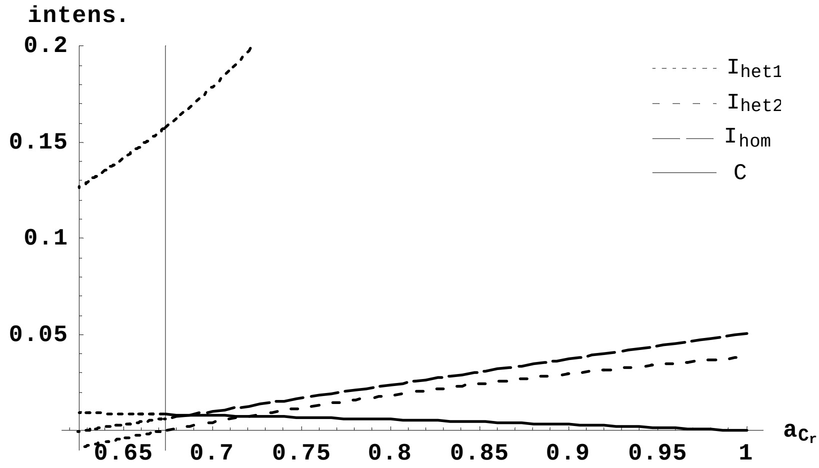

Figure 2.2:

non–innovative market structure.

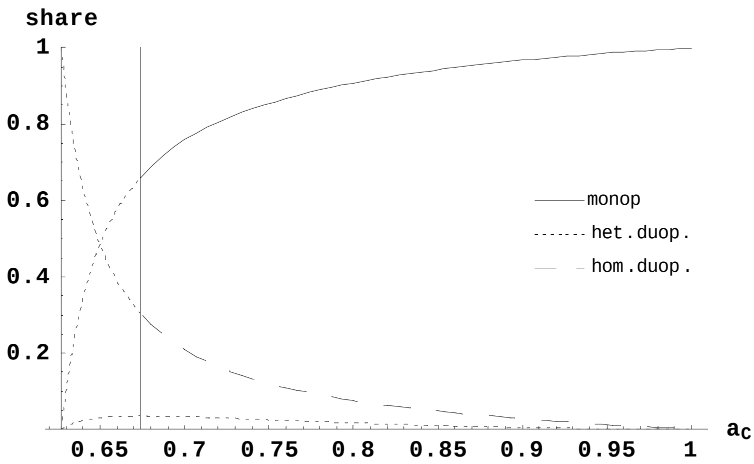

Figure 2.3: Effect of imitation costs $a_{C_{r}}$ on innovation and imitation intensity for the firms in the duopolies and for imitators into the monopolistic structure.

Regarding the three market shares, most changes with respect to variations in imitation costs occur for the monopoly and the homogeneous duopoly. Higher imitation costs increases the share of monopolies at the expense of homogeneous duopolies. This is the composition effect. The larger imitation costs reduce imitation incentives (as seen in figure 2.2) and consequently lower entry which reduces the share of homogeneous duopolies, while at the same time innovation intensities increase and market structures shift towards monopolies in steady state. From figure

2.3 it becomes clear that heterogeneous duopolies serve only as an intermediate step, but their role reflects well the hump–shaped relationship for growth. Higher imitation costs intensify at first innovation of homogeneous duopolies which lead to heterogeneous duopolies. As imitation costs increase further, the innovative R&D activity permits the leader in these duopolies to become a monopolist, while at the same time imitation intensity decreases and the predominant market structure becomes monopolistic.

The costs of imitation have opposing effects on innovation activity and on the share of duopolies. While innovation intensity increases with a_C, the share of duopolies decreases due to lower imitation intensity. These two offsetting effects, the Schumpeter effect and the composition effect, generate the hump–shape relationship between imitation costs and growth.

Relating to other measures of competition

In order to relate our findings to measures of competition used in the empirical literature figures 2.4 and 2.5 relate growth to the average economy–wide Lerner index and the mean number of firms per industry.Figure 2.4 draws the relationship between growth and product market competition

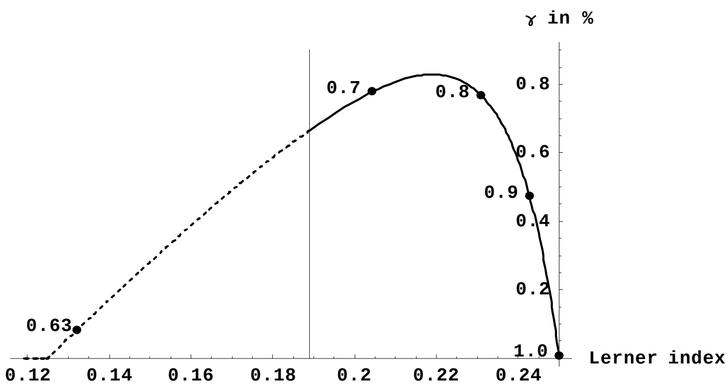

Figure 2.4: Hump–shaped relationship between average Lerner index per firm and the steady state growth rate. Higher competition is represented on the left. The numerical values given in the graph for different points are the relative imitation costs. The numerical values used for the calibration are given in table 2.4.

using the Lerner index[7]. This is the result in the paper by Aghion et al. (2005). Our model with imitation costs can reproduce their findings by offering a different mechanism. Competition is

[7] The Lerner index is the most widely used index to assess competition. It is a measure of price markup relative to production costs.

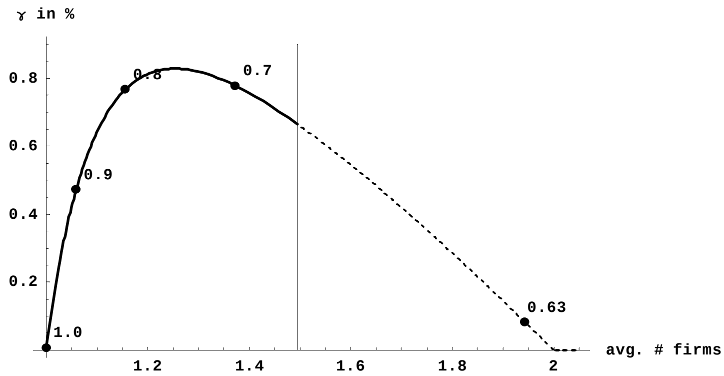

Figure 2.5: Hump–shaped relationship between average number of firms per industry and the steady state growth rate. The numerical values taken are given in table 2.4.

now measured by the Lerner index, which is an endogenous outcome of the model. The Lerner index is proportional to imitation costs (stated in the graph) and therefore exhibits a similar pattern (but not identical) of a hump–shaped relationship. The higher imitation costs, the more monopolies develop which exhibit a larger markup than duopolies. Similarly, if we used the average profits over sales or over production leaves the competition–growth relationship would remain unchanged.

Alternatively we take the average number of firms per industry (obviously limited between 1 and 2) as measure of competition and obtain the hump–shaped relationship.

Figure 2.5 states how an increase in the number of firms relates to growth. An initial increase of the average number from pure monopolists to duopolists first increases growth and then reduces it again. The number of firms is endogenous, such that no direct causal relationship can be established between the number of firms and growth. Only through the imitation mechanism and the variation in imitation costs one can interpret the link. In addition, the number of firms does not reflect in what technological relationship they stand (homogeneous or heterogeneous duopoly).

2.5 Conclusion and further research

This paper has addressed the relationship between competition and growth in a Schumpeterian framework with Cournot competition and free entry by imitation. It uses imitation costs relative to innovation costs as proxy for competition. In this way the main focus of competition lies on the entry effect of outsiders instead of the static interaction in product markets, which has been fixed to Cournot competition. The model is able to reproduce the hump–shaped relationship between competition and growth found by different empirical studies. The model incorporates two effects when varying imitation costs. On the one hand, lower imitation costs ease entry and reduce the expected returns of an innovation. On the other hand entry reduces actual profits by incumbent firms which increases their net return to innovative R&D because the value that needs to be self–replaced reduces. In addition, imitation leads to more firms in the market that generate spillovers for the other firms fostering growth in this way. The interplay of the two effects, the Schumpeterian and the composition effect, lead to a hump–shaped relationship between competition and growth.

The model includes free entry, allows for continuous firm dynamics and endogenous exit due to technological obsolescence. The linear setup with innovation spillovers permits a relatively simple presentation but reduces strongly the feasible parameter space. The numerical calculations obtain furthermore that in heterogeneous duopolies the leader innovates substantially more to obtain monopolistic rents, which is different to other competition–growth models. By comparing the incentives for a laggard and a leading firm, the leading firm benefits strongly from escaping competition altogether and becoming a monopolist through a single innovation, while the laggard would need to innovate twice to escape competition which is relatively costly. This result seems surprising as the replacement value for the leading firm is higher than for the follower. In this model, imitation and innovation are substitutes to each other when changing imitation costs reflecting the trade off between entry barrier and Schumpeterian effect.

An important path for continuation would be the analysis of competition over the product cycle, from possibly large quality improvements in the early phase of a technology to smaller and smaller steps. The framework presented here may be extended to account for a larger number of firms within each industry to generate complex firm dynamics and shed light on the shake–out phenomenon that accompanies most industry evolutions.

Bibliography

Aghion, P., N. Bloom, R. Blundell, R. Griffith, and P. Howitt (2005), Competition and innovation: An inverted-u relationship, *Quarterly Journal of Economics*, 120(2), 701–28.
URL http://mitpress.mit.edu/catalog/item/default.asp?ttype=4&tid=8

Aghion, P., R. Blundell, R. Griffith, P. Howitt, and S. Prantl (2006), The effects of entry on incumbent innovation and productivity, *NBER Inc NBER Working Paper*, 12027.
URL http://www.nber.org/papers/w12027.pdf

Aghion, P., C. Harris, and J. Vickers (1997), Competition and growth with step-by-step innovation: An example, *European Economic Review*, 41(3-5), 771–82.

Aghion, P. and P. Howitt (1992), A model of growth through creative destruction, *Econometrica*, 60(2), 323–51.

Barro, R. J. and X. Sala-i Martin (1995), *Economic growth*, Advanced Series in Economics, McGraw-Hill, New York, London and Montreal.

Bergstrom, T. C. and H. R. Varian (1985), When are nash equilibria independent of the distribution of agents' characteristics?, *Review of Economic Studies*, 52(4), 715–18.

Blundell, R., R. Griffith, and J. Van Reenen (1995), Dynamic count data models of technological innovation, *Economic Journal*, 105(429), 333–44.

Boldrin, M. and D. K. Levine (2005), Perfectly competitive innovation, *Working Paper*.
URL http://www.dklevine.com/papers/pcibasic14.pdf

Boone, J. (2001), Intensity of competition and the incentive to innovate, *International Journal of Industrial Organization*, 19(5), 705–26.
URL http://www1.elsevier.com/homepage/sae/econworld/econbase/indor/frame. htm

Cohen, W. M. and R. C. Levin (1989), Empirical studies of innovation and market structure, in: R. Schmalensee and R. D. Willig (eds.), *Handbook of industrial organization*, VOL 2, 1059–1107, North-Holland, Amsterdam, Oxford and Tokyo.
URL `Elsevier Science, New York`

DenicolÃš, V. and P. Zanchettin (2003), Competition and growth in a neo-schumpeterian model.

Gort, M. and S. Klepper (1982), Time paths in the diffusion of product innovations, *Economic Journal*, 92(367), 630–53.

Grossman, G. M. and E. Helpman (1991), *Innovation and growth in the global economy*, MIT Press, Cambridge, MA and London.

Koeniger, W. and O. Licandro (2006), On the use of substitutability as a measure of competition, *B.E. Macroeconomics*, 6(1), 1–9, b.E.-Journals-in-Macroeconomics:-Topics-in-Macroeconomics.

Kreps, D. M. and J. A. Scheinkman (1983), Quantity precommitment and bertrand competition yield cournot outcomes, *Bell Journal of Economics*, 14(2), 326–37.

Mansfield, E., M. Schwartz, and S. Wagner (1981), Imitation costs and patents: An empirical study, *Economic Journal*, 91(364), 907–18.

Mukoyama, T. (2003), Innovation, imitation, and growth with cumulative technology, *Journal of Monetary Economics*, 50(2), 361–380.

Nickell, S. J. (1996), Competition and corporate performance, *Journal of Political Economy*, 104(4), 724–46.

Segerstrom, P. S. (1991), Innovation, imitation, and economic growth, *Journal of Political Economy*, 99(4), 807–27.

Spence, M. (1984), Cost reduction, competition, and industry performance, *Econometrica*, 52(1), 101–21.

Tirole, J. (1988), *The theory of industrial organization*, MIT Press, Cambridge, MA and London.

2.A Appendix

2.A.1 Parameter restrictions

We resume in this part the parameter restrictions imposed in order to generate the results obtained.

Non–drastic innovations

Condition: $q < 1/\alpha$

The price for two heterogeneous firms in the market is P_2 and the costs of the leader is 1 while q for the follower. In order to have minimum two firms of different quality in the market we need to have $P_2 > q$. The price can be determined from equation (3.6).

$$P_2 = \frac{1+q}{1+\alpha} > q \Leftrightarrow 1 > \alpha q$$

Escape competition

Condition $q > 1/\sqrt{\alpha}$:

A leader in a heterogeneous duopoly becomes a monopolist once he successfully innovates. This is equivalent to the fact that two quality steps are sufficient to make the follower obsolete. In such a case the price within the industry is lower than the costs of the laggard with two quality steps behind q^2.

$$P_{2step} = \frac{1+q^2}{1+\alpha} < q^2 \Leftrightarrow 1 < \alpha q^2$$

Ordering of prices

Condition $q > \frac{2}{1+\alpha}$:

The prices in the six different market structures can be calculated with equation (2.7). This obtains

•	• •	• •	• • •	• • •	• • • •
P_1	P_2	P_3	P_4	P_5	P_6
$\frac{1}{\alpha}$	$\frac{1+q}{1+\alpha}$	$\frac{2}{1+\alpha}$	$\frac{2+q}{2+\alpha}$	$\frac{1+2q}{2+\alpha}$	$\frac{2+2q}{3+\alpha}$

In order to reduce the market structures to only three, we need to show that market structures with prices P_4, P_5 and P_6 are not feasible. The condition for this is $q > \frac{2}{1+\alpha}$. Regarding market structures 4 and 6 the Cournot price is lower than the costs of the follower:

$$P_4 = \frac{2+q}{2+\alpha} > q \Leftrightarrow \frac{2}{1+\alpha} < q$$

$$P_6 = \frac{2+2q}{3+\alpha} > q \Leftrightarrow \frac{2}{1+\alpha} < q$$

Regarding market structure 5, it can only evolve from either market 4 or 6, which themselves are not feasible. The final ordering of prices is then

Comparison	under condition
$P_1 > P_2$	$1 > \alpha q$
$P_2 > P_5$	$1 > \alpha q$
$P_5 > P_6$	$q > 1 > \alpha$
$P_6 > P_4$	$q > \frac{2}{1+\alpha}$
$P_4 > P_3$	$q > \frac{2}{1+\alpha}$

The condition $q > \frac{2}{1+\alpha}$ is included in the condition $q > 1/\sqrt{\alpha}$ as $1/\sqrt{\alpha} \geqslant \frac{2}{1+\alpha}$for $\alpha \epsilon \, (0, 1)$ To obtain the order of prices stated, overall three conditions are necessary

1. $\alpha < 1$: for monopolistic to be possible
2. $\alpha q < 1$: for non–drastic innovations to exist
3. $1 < \alpha q^2$: two consecutive innovations lead to monopolistic market structure and maximum three endogenous market structures

2.A.2 Conditions for self–replacement

A monopolist does not self–replace itself if the value of the firm is lower with innovative R&D activity than without.

The value of a monopolist without R& D activity is:

$$V_m = \frac{\pi_m + C_m V_{\text{hom}}}{r + C_m}$$

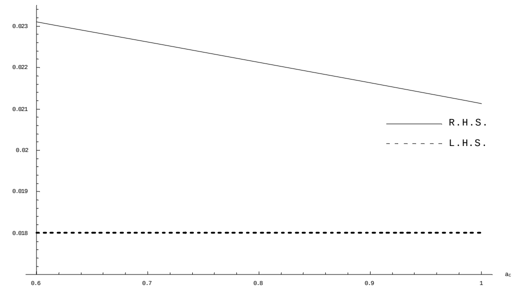

Figure 2.6: Comparison of costs (continuous line) and benefits (dashed line) of self–replacement for a monopolist using the paramters considered in the numerical example in the paper.

Instead, the value of the firm with R&D activity would be:

$$V_m^{RD} = \frac{\pi_m + C_m V_{\text{hom}} + I_m g V_m^{RD} - a_I I_m}{r + C_m + I_m}$$
$$V_m^{RD} = \frac{\pi_m + C_m V_{\text{hom}} - a_I I_m}{r + C_m - I_m(g-1)}$$

By using the free entry condition $V_{\text{hom}} = a_C a_I$, the condition for no self–replacement is

$$V_m^{RD} < V_m$$
$$\frac{\pi_m + C_m a_C a_I - a_I I_m}{r + C_m - I_m(g-1)} < \frac{\pi_m + C_m a_C a_I}{r + C_m}$$
$$\pi_m (g-1) < a_I[r + C_m(1 - a_C(g-1))] \tag{2.22}$$

This condition states that the monopolist does not self–replace itself the lower monopolistic profits and the higher innovation costs, imitation intensity and the lower imitation costs. We plot the left hand and right hand side of the equation in figure 2.6.

The right hand side of equation (2.22) is always larger than the left hand side for the parameters considered in the numerical example in the paper. Hence the monopolist does not self–replace himself.

2.A.3 Optimal innovation and imitation

The value functions for the four different types of firms (monopolist, leader and follower in the heterogeneous duopoly and the homogeneous duopoly) in extensive form are

$V_{em}(k) = \frac{\pi_m q^{\frac{\alpha}{1-\alpha}k} + C_m V_{e\,\hom}(k)}{r+C}$

$V_{ehet1}(k) = \frac{\pi_{het1} q^{\frac{\alpha}{1-\alpha}k} + C_{het} V_{e\,\hom}(k) + (I_{het1}+\theta I_{het2}) V_{em}(k+1) + (I_{het2}+\theta I_{het1}) V_{ehet2}(k+1) - a_I q^{\frac{\alpha}{1-\alpha}k} I_{het1}}{r + C_{het} + (1+\theta)(I_{21}+I_{22})}$

$V_{ehet2}(k) = \frac{\pi_{het2} q^{\frac{\alpha}{1-\alpha}k} + (I_{het2}+\theta I_{het1}) V_{ehet1}(k+1) - a_I q^{\frac{\alpha}{1-\alpha}k} I_{het2}}{r + C_{het} + (1+\theta)(I_{het1}+I_{het2})}$

$V_{e\,\hom}(k) = \frac{\pi_{\hom 1} q^{\frac{\alpha}{1-\alpha}k} + (I_{\hom 1}+\theta I_{\hom 2}) V_{ehet1}(k+1) + (I_{\hom 1}+\theta I_{\hom 2}) V_{ehet2}(k+1) - a_I q^{\frac{\alpha}{1-\alpha}k} I_{\hom 1}}{r + (1+\theta)(I_{\hom 1}+I_{\hom 2})}$

The specification accounts for the quality level k of the intermediate good. With each innovation the output of the intermediate goods increases by the factor $q^{\frac{\alpha}{1-\alpha}}$. Profits, innovation and imitation costs increase by the same factor. To obtain stationary values for the firms in steady state, we transform the problem by dividing for the increment $q^{\frac{\alpha}{1-\alpha}k}$.

$$V_\omega = q^{-\frac{\alpha}{1-\alpha}k} V_{e,\omega}(k), \quad \omega = m, het_1, het_2, \hom .$$

This earns the firm values in intensive form presented in equation (2.14):

$$\begin{aligned}
V_m &= \frac{\pi_m + C_m V_{\hom}}{r + C_m} \\
V_{het1} &= \frac{\pi_{het1} + C_{het} V_{\hom} + (I_{het1} + \theta I_{het2}) q^{\frac{\alpha}{1-\alpha}} V_m + (I_{22} + \theta I_{21}) q^{\frac{\alpha}{1-\alpha}} V_{het2} - a_I I_{het1}}{r + C_{het} + (1+\theta)(I_{het1} + I_{het2})} \\
V_{het2} &= \frac{\pi_{het2} + (I_{het2} + \theta I_{het1}) q^{\frac{\alpha}{1-\alpha}} V_{het1} - a_I I_{het2}}{r + C_{het} + (1+\theta)(I_{het1} + I_{het2})} \\
V_{\hom} &= \frac{\pi_{\hom} + (I_{\hom j} + \theta I_{\hom,-j}) q^{\frac{\alpha}{1-\alpha}} V_{het1} + (I_{\hom,-j} + \theta I_{\hom,j}) q^{\frac{\alpha}{1-\alpha}} V_{het2} - a_I I_{\hom j}}{r + (1+\theta)(I_{\hom,j} + I_{\hom,-j})}
\end{aligned}$$

Firms maximize their value by choosing the innovation intensity, taking other incumbents' innovative activity and imitation intensity as given. This is the typical setup of an innovation (patent) race as in Tirole (1988).

$$\begin{aligned}
&\arg\max_{I_i} V_{\omega,i} \\
&\text{given } \sum_{-j} I_{-j},\ C_\omega.
\end{aligned}$$

The best reaction function of the an incumbent is independent of the others' intensity due to the linear setup of costs. We employ $g = q^{\frac{\alpha}{1-\alpha}}$:

$$\begin{aligned} I_{het1} &= \frac{(1+\theta)\,\pi_{het2}-(C_{het}+r)\,(gV_{het1}-a_I)}{(1+\theta)\,[g\,(1-\theta)\,V_{het1}-a_I]} \\ I_{het2} &= \frac{(1+\theta)\,\pi_{het1}-(C_{het}+r)\,[g\,(V_{het1}+\theta V_{het2})-a_I]+C_{het}\,(1+\theta)\,V_{\text{hom}}}{(1+\theta)\,[g\,(1-\theta)\,(V_{het1}-V_{het2})-a_I]} \\ I_{\text{hom}} &= \frac{(1+\theta)\,\pi_{\text{hom}}-r\,[g\,(V_{het1}+\theta V_{het2})-a_I]}{(1+\theta)\,[g\,(1-\theta)\,(V_{het1}-V_{het2})-a_I]} \end{aligned}$$

By inserting innovation intensities into the value functions and employing the free entry condition (2.18) imitation intensity C_m is identified:

$$C_m = \frac{a_I r\left[a_{C_r}\,(1+\theta)\left[(1+\theta)^2-g^2\theta\right]+(1+g)\,(1+\theta)^2+g^2\theta\right]-\pi_1 g^2\,(1+\theta+g\theta)}{a_I(1+g+\theta)\left[1+g\,(1+\theta)\,(1-a_{C_r})+a_{C_r}\left((1+\theta)^2-g^2\right)\right]}$$

The intensity of imitation into the monopolistic market structure is identical to the one into the heterogeneous setup because any imitation leads to a homogeneous duopoly. With $C_m = C_{het}$ innovation intensity simplifies to:

$$\begin{aligned} I_{het1} &= \frac{(C_{het}+r)(gV_{het1}-a_I)-(1+\theta)\pi_{het2}}{(1+\theta)(a_I-(1-\theta)gV_{het1})} \\ I_{het2} &= \frac{(C_{het}+r)(gV_m+\theta gV_{het2}-a_I)-(1+\theta)(C_{het}V_{\text{hom}}+\pi_{het1})}{(1+\theta)(a_I-(1-\theta)g(V_m-V_{het2})} \\ I_{\text{hom}} &= \frac{r(gV_{het1}+\theta gV_{het2}-a_I)-(1+\theta)\pi_{\text{hom}}}{(1+\theta)(a_I-(1-\theta)g(V_{het1}-V_{het2})} \end{aligned}$$

2.A.4 Steady state market shares

The steady state is obtained when market shares remain constant over time. The conditions for this are equation (2.20) combined with $1=\alpha_m+\alpha_{het}+\alpha_{\text{hom}}$. The solution for the market shares is

$$\begin{aligned} \alpha_m &= \frac{2(1+\theta)I_{\text{hom}}\,(I_{het1}+\theta I_{het2})}{(C+I_{het1}+\theta I_{het2})\,[C+2I_{\text{hom}}\,(1+\theta)]} \\ \alpha_{het} &= \frac{2(1+\theta)CI_{\text{hom}}}{(C+I_{het1}+\theta I_{het2})\,[C+2I_{\text{hom}}\,(1+\theta)]} \\ \alpha_{\text{hom}} &= \frac{C}{C+2(1+\theta)I_{\text{hom}}} \end{aligned}$$

From which one can derive table 2.3.

Chapter 3

Innovations and diffusion

3.1 Introduction

This paper addresses the issue of technological diffusion in an economy of heterogeneous skills when technologies are characterized by minimum skill requirements. The endogenous growth literature identifies innovations as the key feature to promote growth, but their introduction does not occur instantaneously to the entire population but requires time and resources to completely diffuse in the economy. The S-shaped diffusion curves are a stylized fact since Griliches (1957) for nearly all technologies including General Purpose Technologies, such as electricity and the semiconductor see David (1991) and Jovanovic and Rousseau (2005). In addition to time the process of diffusion is not an entirely free process, it requires productive resources in form of investment, adoption costs or learning time (Greenwood and Yorukoglu (1997) and Jovanovic and Lach (1989)).

Skilled labor plays a double role in the process of technological improvements. In a first step skilled labor is needed to invent new technologies, an activity generally computed in research departments. In a second step, skilled labor is needed to implement new technologies for productive usage and to further improve it by a process of continuous adaptation. Nelson and Phelps (1966) relate the speed of adoption to the stock of skilled labor in the economy and postulate that skilled labor has advantages in innovating and adopting new technologies. This is the reason why firms employing a larger share of skilled labor are able to adopt new technologies sooner as found in the empirical studies by Bartel and Lichtenberg (1987) and Doms et al. (1997).

In the model presented here, a technology is only operable for a worker fulfilling the technology's minimum skill requirement. This characteristic is an exclusive restriction. Workers are characterized by heterogeneous skills distributed over a continuous support and instead of workers engaging in a costly learning process to operate the new technology, it is the technology that is further improved by purposeful investment to become easier to use. The higher the skill level of a worker, the earlier she adopts a new technology and, once adopted the new technology, she is able to improve it further, either by increasing the productivity level or by reducing the minimum required skill. Higher skilled labor fullfills a double role in the diffusion porcess as formulated by Nelson and Phelps: they are the ones to adopt a technology earlier and at the same time thy support the diffusion process by developing the technology further.

The diffusion mechanism combines productivity enhancements and ease of use. For a firm owning monopoly rights on the technology lower skill requirements translate into larger market sizes and increases the profits the monopolist may earn. At the same time, the more productive the technology the more profitable are reductions in the skill level. This mechanism is closely related to the concept of Directed Technical Change by Acemoglu (1998) by which the sector with larger market size is favored because it allows to reap higher profits. In our model this market size is not exogenous, but is endogenously determined through purposeful investment by the monopolist. Productivity level and minimum skill requirement together determine total profit and are complementary to one another: the larger the market size, the larger the incentives to improve the technology.

3.1.1 Stylized Facts

The model shall be applied to the diffusion of a General Purpose Technology (GPT), specifically the information technology. The main stylized facts regarding the diffusion of GPTs have been collected by David (1991) and more recently by Jovanovic and Rousseau (2005). The main findings we want to reproduce here concentrate on the diffusion curve and the time length. The usage of electricity measured as share of total horsepower diffused across the productive sector between 1894 and 1930 in the usual S-shape, that is within 36 years. Regarding the IT sector diffusion is taking place since 1971, when one percent of total capital stock was IT capital. The onset of adoption by households occurred some years later in form of electric services or personal computers respectively. We hence focus on time horizons which cover long time spans.

David (1991) unveils a strong co-movement between the diffusion pattern of electric motors and the evolution of labor productivity in the economy. The explanation put forward is twofold. The first obvious reason is that the new technology's higher productivity influences the economy–wide productivity level only insofar the technology has diffused, lower diffusion rates implies smaller effects on output. The second reason regards the productivity in the usage of electricity, the new technology. This does not remain constant after the introduction of the GPT, but increases slowly with time. Rosenberg (1976) evaluates that productivity gains of an innovation occur already with the onset of diffusion, but more important gains are made along its diffusion.

The evolution of wages along the diffusion path of electricity and information technology is very different. The recent IT revolution has been characterized by a growing wage differential in the U.S. between skilled and less skilled workers. Acemoglu (2002) and Goldin and Katz (1999) document the college premium in wages to have decreased in the 1970's and strongly increased since then. The offered explanations for this phenomenon are directed technical change by Acemoglu (2002) or alternatively Violante (2002). We are able to generate the growing wage differential as a phenomenon due to the diffusion process.

3.1.2 Related literature

The leading model in explaining the diffusion of General Purpose Technologies (GPTs) has been developed by Helpman and Trajtenberg (1998) distinguishing between the arrival of a GPT and its practical implementation in an industry. The idea is that a GPT needs a specification for each industry to be implement for the production sector. Two implications earn particular remarks. The model predicts an immediate downturn of GDP due to the shift from productive activity to research activity directly following the discovery of the GPT. This is at odds with the observations by David (1991) that a new GPT requires some years to have macroeconomic effects. The second fact is that research activity itself only makes up a small fraction of total employment, it is therefore difficult to imagine that the sector may generate sizeable slumps in GDP. But nevertheless it is important to notice that the introduction of a GPT leads to the deviation of resources from purely productive activities to adoption activities in some form, which is foregone output.

On the basis of Helpman and Trajtenberg's model Aghion and Howitt (1998b) introduces the notion of social learning, leading to an epidemic diffusion pattern of GPT. Social learning consists of agents observing other agents already using the new technology. The number of agents met is constant over time, but the probability that these agents have adopted the technology under the condition that oneself has not, decreases over time. The model foresees three steps: the arrival of the GPT, cost–free discovery or observation of templates and the implementation in the corresponding sector. The main shortfall of the model is the exogenous and cost-free process of diffusion disregarding of agent's behavior. Our diffusion process instead is driven by a purposeful investment activity of a monopolist into the characteristics of the technology.

Greenwood and Yorukoglu (1997) develop a model which uses learning at the plant level as main driving force for productivity gains and adoption incentives. In addition, the productivity gains are exclusively achieved by skilled workers as we lay it out as well. Our model is more stylized than the one by Greenwood and Yorukoglu, and the main difference is that in our model adoption costs are entirely incurred by the side offering the technology in form of development costs instead of the workers. This is already sufficient to obtain an S-shaped diffusion curve and productivity increments along diffusion.

Most models focus on mechanisms on the adopter's side when analyzing diffusion. To generate either S-shaped diffusion curves or discrete moments of technology switching these models include the learning–by–doing model by Parente (1994), informational models such as Jovanovic and Nyarko (1996) or epidemic models such as Aghion and Howitt (1998b) or training costs. In this model, similar to the one by Mukoyama (2004), we focus on the supply side. It is the innovating R&D sector which subsequently improves the performance of the technology or its usability for workers. In this sense the paper is complementary to the demand driven diffusion.

Similar to Mukoyama (2004) we use the concept of minimum skill requirements for operating a technology, but place the diffusion of new technologies in a general equilibrium model which allows for less stringent assumptions. The diffusion process in the model by Mukoyama (2004) is heavily based on the exogenous skill distribution in the population. With a log–normal distribution curve he obtains the S-shaped diffusion curve which would be linearly shaped when using a uniform distribution of skills.

We present a model which generates an S-shaped diffusion curve without this being primarily dependent on the distribution of skills in the economy, although a single–peaked skill distribution would reinforce the logistic character of the diffusion process. For this effect to take place it is essential that the costs for reducing the skill requirements are convex and that the development is constrained in finding workers who can already use the new technology. In this way slower diffusion takes place at the early stage. In addition, if the technology's productivity improves along diffusion and workers' wages are related to this, development costs increase along diffusion. we are able to reproduce in a stylized way the diffusion of the information technology and show that productivity improvements and diffusion reinforce each other as the technology matures.

3.2 The Model

The economy consists of three sectors, a final goods sector producing a homogeneous good, an intermediate goods sector consisting of a continuum of industries operating monopolistically and employing labor, and an R&D sector innovating new technologies and improving existing ones using sufficiently skilled labor.

The final goods sector aggregates output from intermediate firms to a final homogenous output good by a Dixit–Stiglitz aggregator. Firms in the intermediate sector have the choice of two technologies, either to employ a mature technology usable by any worker in the economy or a new technology operable only by sufficiently skilled workers. The technologies are characterized by two parameters, productivity level and level of skill requirement. The first parameter determines output per worker while the second is a parameter determining the availability of a technology to workers of different skills. We follow Aghion and Howitt (1998a) in the setup of final and intermediate goods production.

The R&D sector accomplishes three tasks, innovation of new technologies, productivity enhancement and reduction of skill requirements for existing technologies. Only workers with the ability to use the new technology are employed in the R&D sector, i.e. workers that have already adopted the new technology. Labor is allocated between the intermediate sector and the R&D sector by an arbitrage condition for wage. We will use the terms skilled/unskilled workers equivalently to the terms adopters/non–adopters.

3.2.1 Goods sector

Final goods

The final goods production is characterized using a continuum of intermediate goods $x(i)$ characterized by productivity $A(i)$

$$Y = \left\{ \int_0^1 A(i)^\alpha x(i)^\alpha \, di \right\}^{1/\alpha}. \tag{3.1}$$

Firms in the sector aggregate the quality–weighted output of the intermediate sector using a Dixit–Stiglitz aggregator with elasticity α. We assume that there exist two different technologies, H and L, with which the production sector may operate with productivities A_H and A_L characterizing high and low productivity level, respectively. Workers in the intermediate sector employ either one of the two technologies, conditioning on the skill requirement of the technology and the skill level of the worker. We assume a continuous skill distribution for workers $F(\theta)$ with skill parameter $\theta \epsilon [0;1]$. The high technology is operated by workers with skill levels larger or equal to the skill requirement of the technology $\theta \geqslant \theta_{it}$, wile those workers with a lower skill level operate the lower technology. Hence, the population is divided into two groups, those workers that have adopted the new technology, $1 - F(\theta_{it})$, and those that have not adopted it yet, $F(\theta_{it})$. The two groups are only divided with respect to the technology they use, their skill level does not affect their productivity. Furthermore, not all skilled workers are available for production in the intermediate sector, a share n_t is employed in the R&D sector and therefore not all industries are active in the intermediate sector and final goods production is aggregated over a smaller number of industries. Final output (3.1) simplifies to

$$\begin{aligned} Y_t &= \left\{ \int_0^{F(\theta_{it})} A_{Lt}^{\alpha} x_{Lt}^{\alpha}(i)\, di + A_{H,t}^{\alpha} \int_{F(\theta_{it})}^{1-n_t} x_{Ht}^{\alpha}(i)\, di \right\}^{1/\alpha}, \\ y_t &= Y_t/A_{Lt} = \left[F(\theta_{it})\, x_{Lt}{}^{\alpha} + \rho_t^{\alpha} \left[1 - F(\theta_{it}) - n_t\right] x_{Ht}{}^{\alpha} \right]^{1/\alpha}. \end{aligned} \tag{3.2}$$

The production employs productivity–weighted output from the intermediate sector originating from the two types of technology. The share $F(\theta_{it})$ of intermediate firms employs the low productive technology with lower skilled workers, and $[1 - F(\theta_{it}) - n_t]$ employs the new technology with higher skilled workers. n_t characterizes the share of adopters employed in the R&D sector, which are not available for intermediate goods production. The second line rewrites the production function as output in efficiency units based on the productivity level of the mature technology: $y_t \equiv Y_t/A_{L,t}$, where ρ_t characterizes the relative productivity between the two technologies $\rho_t = A_{Ht}/A_{Lt}$.

Firms in the final sector are competitive and demand their inputs from the intermediate sector. The amount is determined by profit maximization $\max_x [Y_t - p(i)\, x(i)]$ and we obtain with Shephard's Lemma that the price of inputs equals their marginal product $p(i) = \partial Y/\partial x(i)$. The inverse demand function for the goods of the two types of technologies L, H relative to the lower technological level is[1]

$$\begin{aligned} p_L &= A_L^{\alpha} \left(\frac{Y}{x_L}\right)^{1-\alpha} = A_L \left(\frac{y}{x_L}\right)^{1-\alpha} \\ p_H &= A_H^{\alpha} \left(\frac{Y}{x_H}\right)^{1-\alpha} = A_L \rho^{\alpha} \left(\frac{y}{x_H}\right)^{1-\alpha} \end{aligned} \tag{3.3}$$

where x_L, x_H is the quantity demanded from intermediate firms using the old or the new technology, respectively.

Intermediate goods

The sector of intermediate goods uses a linear production function with labor and operates monopolistically within the sector i

$$x(i) = L(i). \tag{3.4}$$

The production of high technology goods is possible only for workers with worker skills larger than the skill requirement of the technology, i,e, when $\theta \geqslant \theta_{it}$, while workers with low skills are precluded from the possibility to use the new technology. Firms in the intermediate sector are

[1] In order to ease reading we drop the time subscript when it does not represent any source of confusion.

therefore subdivided by their operating technology into two groups, employing either workers having adopted the new technology or not. Along the diffusion path of a new technology the R&D sector reduces skill requirements and the number of users increases as technology matures.

The demand for intermediate goods originates from the final goods sector. The intermediate sector maximizes profits with respect to quantities taking into account the inverse demand (3.3) by the final goods sector.

$$\pi_i^{int} = \arg\max_{x(i)} \left[p\left[x\left(i\right)\right] x\left(i\right) - w\left(i\right) x\left(i\right)\right], \tag{3.5}$$

which earns the familiar expression for prices in the intermediate sector.

$$p\left(i\right) = \frac{w\left(i\right)}{\alpha}. \tag{3.6}$$

Prices are a markup over marginal labor costs. Combining the inverse demand (3.3) with equation (3.6) for prices earns a demand dependent on wages

$$x_L = \left(\frac{\alpha}{\omega_L}\right)^{\frac{1}{1-\alpha}} y, \tag{3.7}$$

$$x_H = \left(\frac{\alpha\rho^{\alpha}}{\omega_H}\right)^{\frac{1}{1-\alpha}} y, \tag{3.8}$$

where $\omega_L \equiv w_L/A_L$ and $\omega_H \equiv w_H/A_L$ are wages relative to the lower technologies productivity level. Inserting prices and quantities into equation (3.5) for profit maximization and using the production function in the intermediate sector (3.4) we obtain profits for each intermediate firm either using the low or the high technology.

$$\pi_L^{int} = (1-\alpha) L^{\alpha} Y^{1-\alpha} \tag{3.9}$$

$$\pi_H^{int} = \rho^{\alpha} (1-\alpha) L^{\alpha} Y^{1-\alpha}. \tag{3.10}$$

Profits of intermediate firms depend on total output as well as the technological level of production. These profits by the intermediate sector are used to rent a production licence from the R&D sector for either technology. An intermediate firm is prepared to pay up to its entire profits to obtain a licence to use the production technology. Every period the R&D sector rents a production licence for the low and high productive technology to all firms in the intermediate sector. Depending on the skill level of its workers, the intermediate firm demands either a licence for the low or the high technology and total profits of intermediate firms dedicated to licences is the sum of profits (3.9) and (3.10) by all firms of a given technology.

$$\begin{aligned}\pi_L^{RD} &= \sum_L \pi_L^{int} = F(\theta_i)\,(1-\alpha)\,y^{1-\alpha} \\ \pi_H^{RD} &= \sum_H \pi_H^{int} = [1 - F(\theta_i) - n]\,(1-\alpha)\,y^{1-\alpha}\rho^{\alpha}\end{aligned}$$

We normalize labor in each industry to 1, and $F(\theta_i)$ of firms acquire licence for the low technology, while a share $[1 - F(\theta_i) - n]$ uses the new technology, with n being the share of workers in the R&D sector. The relative total profits of firms with the two technologies depend on the relative productivity characterized by ρ and the relative size of the two sectors.

$$\frac{\pi_H^{RD}}{\pi_L^{RD}} = \frac{(1 - F(\theta_i) - n)\,\pi_H^{int}}{F(\theta_i)\,\pi_L^{int}} = \rho^{\alpha}\frac{1 - F(\theta_i) - n}{F(\theta_i)} \tag{3.11}$$

Profits of intermediate firms are entirely used for the purchase of licences in equilibrium and represent revenues for R&D firms. The relative revenues for the R&D sector offering either one or the other technology depends positively on the relative productivity of the two technologies and the market size of these. Acemoglu (1998) introduced directed technical change stating that there is a market size effect for research activity. The larger one sector, the more profitable is innovation within this sector. The R&D activity in this model can take two forms for a given technology, R&D enhances productivity or it reduces the minimum skill requirement. Depending on which activity is more profitable to increase total revenues, the R&D sector allocates resources either in one or the other activity.

Resource constraint and wages

Before presenting the R&D sector, we consider the resource constraint and wages in the economy. The economy is endowed with 1 of labor, employed in the R&D sector and either in high or low quality goods production in the intermediate sector, L_H or L_L respectively. Labor in the research sector is employed either for productivity enhancements L_P, reduction of the minimum skill requirement L_θ or for the innovation of new technologies L_I.

$$1 = L_L + L_H + L_P + L_\theta + L_I$$

The labor market is strictly segmented between the two skill groups but operates within each skill group competitively, skilled and unskilled workers are paid their marginal product. We assume that workers who have the choice of adopting (those with a skill level equal or higher than the skill requirement $\theta \geqslant \theta_i$) adopt the new technology if wages paid with the new technology are equal or higher than with the old technology. By dividing the labor force in the productive sector into adopters and non–adopters we obtain $L_L = F(\theta_i)$ and $L_H = [1 - F(\theta_i) - n_P - n_\theta - n_I]$, where n_P, n_θ, n_I represent the share of the labor force employed in the respective research activities.

The equations for labor demand in the high and low technology sectors (3.7), (3.8), combined with the production function in the intermediate sector obtains a reformulation of the final goods production function (3.2).[2]

$$y = \{F(\theta_i) + (1 - F(\theta_i) - n_P - n_\theta - n_I)\,\rho^\alpha\}^{1/\alpha} \tag{3.12}$$

Final good production depends on the relative size of the skilled and unskilled labor force as well as the relative productivity ρ of the new technology. Production of final goods decreases with the amount of workers in the research and development sector, $\partial y/\partial n < 0$. The minimum skill requirement of the new technology θ_i determines the share of adopters and non-adopters, and if the two technologies have differing productivities this affects also total output. A lower skill requirement of the new technology leads to higher output in the final sector $\partial y/\partial \theta_i < 0$.

Relative wages of workers using the two technologies are obtained by combining the equations on the amount of labor employed with the two technologies, equations (3.7), (3.8).

$$\begin{aligned} \omega_L &= \alpha y^{1-\alpha} && (3.13)\\ \omega_H &= \rho^\alpha \alpha y^{1-\alpha} && (3.14)\\ \frac{\omega_H}{\omega_L} &= \rho^\alpha \end{aligned}$$

Wages are affected by total output and the relative productivity ρ. Output in the final goods sector, in turn, is directly influenced the relative technological level with $\partial y/\partial \rho > 0$. The relative wage, instead, is determined exclusively by the relative productivity of the two sectors $\rho = A_H/A_L$. In addition, final output production depends on the share of workers using the new technology taking into account that some of them are employed in the final goods sector. The larger the size of the research sector, the lower are wages due to the fact that total output decreases.

[2] One unit of labor produces one unit of intermediate goods. Due to the normalization of labor and the distribtion of industries on $[0, 1]$, every sector requires 1 unit of labor.

3.2.2 The Research and Development sector

The R&D sector is divided into three activities. It innovates new technologies, improves existing technologies and reduces the required skills for them. It uses as input for all tasks labor that already adopted the newest technology and pays an identical wage to the high technology productive sector (arbitrage condition).

The innovative activity is characterized by free entry similar to Aghion and Howitt (1992) with labor determined in general equilibrium. We consider this research as basic research for a new technology and the patent on the discovered technology is sold to a firm developing the technology further by augmenting productivity and reducing skill requirements. In turn, the developing firm rents out production licences to the intermediate goods sector as seen before.

We define some variables: new technologies are innovated with productivity $A_{it} = b_i$ and skill requirement θ_i. We call b_i the *technological level* of the product which remains fixed throughout time while its *productivity* A_i increases thanks to productivity enhancing R&D activity. The *relative productivity* level, i.e. the productivity level compared to its initial level is denoted by $\rho_i = A_i/b_i$. The *skill requirement* θ_i is an exclusion restriction for usage of the technology. Workers are distributed exogenously along a single–dimensional skill variable with cumulative distribution function $F(\theta)$. All workers with skills higher than θ_i ($\theta \geq \theta_i$) may use the technology to produce the intermediate good, while all those with less skills can not. A higher θ_i of the technology requires higher skills from the worker, which translates into a smaller market size.

Innovation

The main task of the research sector is to innovate new technologies with a higher technological level than the current best technology. We follow Aghion and Howitt (1992) in the incentive structure of innovations. The research sector employs labor that has already adopted the latest production technology and each worker innovates the new technology with a Poisson probability of λ. The costs of innovation are labor costs w_H and the benefits are the expected value of all future profits from the developing firm that buys the patent rights from the innovator. The market for innovations is characterized by free entry, and the research arbitrage condition can be formalized by the equalization of costs and expected benefits:

$$w_H = \lambda V\left(b_{i+1}, \rho_{i+1}, \theta_{i+1}\right), \tag{3.15}$$

where the value of the new technology $i+1$ depends on the technological level b_{i+1}, the relative productivity ρ_{i+1} and its initial skill requirement θ_{i+1}. In addition to its new technological level, the new technology requires higher skills $\theta_{i+1} > \theta_i$, which restricts the number of potential users similar to Mukoyama (2004). Even if the new technological level b_{i+1} is identical to the current productivity level A_i the required skills are larger and the size of population using the technology is smaller. We assume that a new innovation requires skills that are an increasing function of the currently operated technology, $\theta_{i+1} = g(\theta_i)$, with $g'(\theta_i) > 0$. As long as the skill requirement is larger than the maximum skill level in the population $\theta_{i+1} > \bar{\theta}_j$, no innovative activity takes place as no worker would be able to operate the new technology, i.e. $V(\cdot, \cdot, \bar{\theta}_{i+1}) = 0$.

The level of skill requirements for the next technology is independent of its productivity and independent of the productivity of the current best technology. As a consequence, higher skill requirements of a new technology decreases the value of an innovation $\partial V\left(b_i, \rho_i, \theta_i\right)/\partial\theta_i < 0$ and through the skill link between current and next technology $g\left(\cdot\right)$, innovative activity increases as the current technology diffuses in the economy.

A new technology starts with the productivity level A_{i+1}, equal to its technological level $A_{i+1} = b_{i+1}$. The relative productivity level is hence 1 $\left(\rho_{i+1} = A_{i+1}/b_{i+1} = 1\right)$. We can rewrite the research arbitrage condition (3.15) as:

$$\begin{aligned} w_H &= \lambda b_{i+1} V\left(1, \rho_{i+1}, \theta_{i+1}\right) \\ \omega_H &= \lambda \gamma_{i+1} v\left(1, g(\theta_i)\right) \end{aligned} \tag{3.16}$$

with $\omega_H = w_H/b_i$ and $\gamma_{i+1} = b_{i+1}/b_i$ used in the second equation. Moreover, the arguments of the value function in intensive form have been substituted by their values at the moment of innovation related to the characteristics of the current technology, $\rho_{i+1} = 1$ and $\theta_{i+1} = g\left(\theta_i\right)$.

Productivity enhancement

The development sector invests in productivity enhancements of the active technologies by using labor from the pool of skilled workers, those that have already adopted the new technology. Only these workers are able to improve the technology because non-adopters, not being able to use it, cannot improve it. The technology for improving the current technology from relative productivity ρ to $\rho' > \rho$ next period may be expressed in terms of labor requirement L_P

$$\rho_i' = \rho_i + a_P \frac{L_P^{1/\phi_P}}{\rho_i}, \tag{3.17}$$

where ρ_i represents the current productivity level of the technology compared to its initial level, a_P the productivity of the productivity enhancement technology, and $\phi_P > 1$ determines the degree of decreasing returns in the activity. Improving the productivity of a technology is more labor consuming the larger its current productivity relative to its technology, i.e. the larger ρ_i.

The labor costs for productivity enhancements consist of the amount of labor employed multiplied by the wage rate, here expressed in intensive variables, $\omega_H = w_H/b_i$. With the productivity enhancing technology (3.17) costs may be expressed in the step size of improvements $(\rho_i' - \rho_i)$

$$C_P\left(\rho_i, \rho_i'; w_H\right) = w_H \left(\rho_i \frac{\rho_i' - \rho_i}{a_P}\right)^{\phi_P}, \tag{3.18}$$

where ω_H is the wage of skilled labor in intensive units and L_P is labor employed for productivity enhancement. The cost function is proportional to wages, increasing in the productivity advances already made ρ_i, and convex in the step size of the productivity increase $\rho_i' - \rho_i$. The fact that improvements are more costly the higher productivity levels already are, ensures a slower productivity growth the maturer the technology becomes, and convex costs in the productivity step size make fast productivity gains costly. The wage rate for skilled labor is entirely determined by the productive sector of intermediate goods. They change with the productivity advances the R&D sector generates.

The revenue for firms in the R&D sector are the profits of intermediate firms as in (3.10). The intermediate profits are entirely captured by renting production licences to the firms in that sector, where the price for the licence depends proportionally on the productivity level ρ_i of the technology and the market size is determined by the skill requirement.

Reduction of skill requirement

Next to the productivity increments the firm in the development sector expands the number of potential users of the technology by reducing the skill requirement for the technology. Every reduction in skills enlarges the market size but leaves the technology's productivity unaffected. The reduction is induced by purposeful R&D activity carried out by workers who already produce with the technology. The improvements are made in discrete time and are characterized by a technology with decreasing returns to labor intensity based on the current level of skill requirement

θ_i. The specification for the skill–reduction technology takes the form

$$\theta_i' = \theta_i - a_\theta L_\theta^{1/\phi_\theta}, \tag{3.19}$$

where θ_i and $\theta_i^{'}$ represent the current and next period's skill requirement, a_θ is the productivity of the R&D sector in skill-reducing activity and $\phi_\theta > 1$ reflects the degree of decreasing returns in labor L_θ. The corresponding cost function consists of wages paid for workers employed in the activity:

$$C_\theta\left(\theta_i - \theta_i'\right) = w_H L_\theta = w_H \left(\frac{\theta_i - \theta_i'}{a_\theta}\right)^{\phi_\theta}. \tag{3.20}$$

The costs are convex in the size of skill reduction and vary with the wage rate ω_H of skilled labor in the productive sector. The benefits for the R&D sector reducing the skill requirement is to enlarge the number of workers using the technology. The two activities of the development sector generate higher profits for those firms due to higher productivities of the technology or a larger market size. The costs for either activity depend on the level of wages in the productive sector and increase along the diffusion path.

3.3 Value functions

With the R&D sector and the goods sector described above we focus on the recursive optimization problem by the developing firms using value functions. For this, we simplify the setup of the model and specifically address the diffusion of a General Purpose Technology (GPT). Such a technology fits the described setup best because it is a single technology applicable transversally to all sectors. We assume that maximum two technologies coexist at each point in time and that the discovery of the GPT occurs only after the former technology has completely diffused. This means, the old GPT is denoted by $i-1$ and complete diffusion implies $\theta_{i-1} = 0$, such that all workers are able to use the technology. This assures also that older technologies $i-2,\ i-3, ...$ have become obsolete (GPT $i-1$ has highest technology in comparison to all the others). The new GPT i is characterized by a minimum required skill level of θ_i at its introduction.

Regarding productivity, we assume that the improvements have halted for the old GPT because the decreasing returns in this activity make it no longer profitable, hence $\rho_i' - \rho_i = 0$. But the firms in the R&D sector holding the patents for GPT i reaps the profits without improving the technology. At the introduction of the new GPT its productivity is identical to the old one, i.e. $b_i = p_{i-1} = \rho_{i-1} b_{i-1}$, and obviously the relative productivity is 1, $\rho_i = A_i/b_i = 1$. Finally, the introduction of a subsequent GPT $i+1$ can not occur during the diffusion of the actual technology i, implying that labor is not employed in the innovative R&D activity, $n_I = 0$. All these assumptions are made to analyze the diffusion of a GPT without the influence of other factors.

The problem of the monopolistic firm developing the new technology either by improving the productivity or by reducing skills and the one owning the patent of the low technology can be expressed by

$$\begin{aligned} V_H\left(b_i, \rho_i, \theta_i\right) &= \max_{\rho_i', \theta_i'} \left\{ \Pi_H^{RD}\left(b_i, \rho_i, \theta_i, \rho_i', \theta_i'\right) + \beta V_H\left(b_i, \rho_i', \theta_i'\right) \right\} \\ V_L\left(b_{i-1}, \rho_{i-1}; \theta_i\right) &= \Pi_L^{RD}\left(b_{i-1}, \rho_{i-1}, \theta_i\right) + \beta V_L\left(b_{i-1}, \rho_{i-1}; \theta_i\right) \end{aligned}$$

An R&D firm owning the patent for the high technology with technological level b_i, relative productivity ρ_i and skill requirement θ_i maximizes its value by choosing next period's productivity level and the level of skill requirement subject to the cost functions (3.18) and (3.20).The patent holder of the old technology, instead, has no maximization problem to solve, but earns the renting price for the licence Π_L^{RD}. Both firms discount the future with the factor $\beta < 1$. The profits of an R&D firm owning the patent either for the high or the low technology Π_H^{RD} $\left(\Pi_L^{RD}\right)$ are

$$\begin{aligned} \Pi_H^{RD}\left(b_i, \rho_i, \theta_i, \rho_i', \theta_i'\right) &= \left(1 - F\left(\theta_i\right) - n_{\theta_i} - n_{\rho_i}\right) \Pi_H^{int}\left(b_i, \rho_i\right) \qquad (3.21) \\ &\quad -C_P\left(b_i, \rho_i, \rho_i'; w_H\right) - C_\theta\left(b_i, \theta_i, \theta_i'; w_H\right), \\ \Pi_L^{RD}\left(b_{i-1}, \rho_{i-1}; \theta_i\right) &= F\left(\theta_i\right) \Pi_L^{int}\left(b_{i-1}, \rho_{i-1}\right). \qquad (3.22) \end{aligned}$$

Profits consist of the revenues Π_H^{int} $\left(\Pi_L^{int}\right)$ from renting out the production licences to the share of intermediate firms operating with the high or low technology respectively, subtracted by any development costs used for their improvement. Total revenues for the high technology derive from the share of firms employing sufficiently skilled workers $(1 - F(\theta_i) - n_{p,i} - n_{\theta,i})$ reduced by the costs for productivity enhancements and skill reduction, C_P, C_θ respectively. Note that these costs depend on wages which are determined in general equilibrium. Earnings for the low technology licence derives from the share of firms using the old technology $F\left(\theta_i\right)$ multiplied by the profits of a single firm in the intermediate sector.

By rewriting $w_H = b_i \omega_H$ and using equation (3.9) and (3.10) in intensive form, i.e. normalized to the technological level $b_i = b_{i-1}\rho_{i-1}$ we obtain $\Pi_H^{RD}\left(b_i, \rho_i, \theta_i, \rho_i', \theta_i'\right) = b_i \pi_H^{RD}\left(\rho_i, \theta_i, \rho_i', \theta_i'\right)$ and $\Pi_L^{RD}\left(b_{i-1}, \rho_{i-1}; \theta_i\right) = b_{i-1}\rho_{i-1}\pi_L^{RD}\left(\theta_i\right) = b_i \pi_L^{RD}\left(\theta_i\right)$. Similarly the cost functions can be expressed by wages relative to the technological level b_i, $w_H = b_i \omega_H$:

$$\begin{aligned}\pi_L^{RD}(\theta_i) &= F(\theta_i)\ \pi_L^{int} \\ \pi_H^{RD}\left(\rho_i, \theta_i, \rho_i', \theta_i'\right) &= \left(1 - F(\theta_i) - n_\theta - n_\rho\right) \pi_H^{int}(\rho_i) - C_P\left(\rho_i, \rho_i'; \omega_H\right) - C_\theta\left(\theta_i, \theta_i'; \omega_H\right) \\ &= \left(1 - F(\theta_i) - n_P - n_\theta\right)(1-\alpha) L^\alpha y^{1-\alpha} \rho_i^\alpha \\ &\quad -\omega_H \left[\left(\rho_i \frac{\rho_i' - \rho_i}{a_P}\right)^{\phi_1} + \left(\frac{\theta_i - \theta_i'}{a_\theta}\right)^{\phi_2}\right]\end{aligned}$$

where the cost functions for development activity (3.18) and (3.20) have been inserted and also output is expressed relative to the technological level, $y = Y/b_i$.

Employing the profit functions in intensive form allows also to rewrite the value functions in intensive form and reduce the number of state variables.

$$\begin{aligned}v_H(\rho_i, \theta_i) &= \max_{\rho_i', \theta_i'} \left[\pi_H^{RD}\left(\rho_i, \theta_i, \rho_i', \theta_i'\right) + \beta v_H\left(\rho_i', \theta_i'\right)\right] \\ v_L(\theta_i) &= \pi_L^{RD}(\theta_i) + \beta v_L\left(\theta_i'\right)\end{aligned} \tag{3.23}$$

The value functions simplify substantially and are characterized by only two state variables, relative productivity and skill requirement. These functions are identical for any technological level b_i, though the value of each firm holding a patent changes with the technological level b_i. The profits by intermediate firms π_H^{int} and π_L^{int} are given by equations (3.10) and (3.10), and we have assumed $L = 1$, wages in intensive form ω_H of adopters are the marginal product of firms in the intermediate the sector using the new technology as in equation (3.14).

3.4 Analytical results

Having set up the cost functions for R&D activity and the value functions for firms in the development sector, we may generate a first result regarding improvements in productivity and skill reduction. The size of productivity improvements increases with the market size for the technology.

In order to show this, we simplify the setup further and assume that the monopolist improves the productivity level of the technology only once, and the minimum required skill level is taken as given[3]. The firm value of equation (3.23) depends on the infinite discounted flow of profit from

[3] The firm hence optimizes only w.r.t. to the productivity level. In addition we disregard the general equilibrium effects of R&D employment on wages of high skilled workers and the effects it might have on profits, hence $\pi_H^{int}(\rho_i) = \rho_i^\alpha \pi$.

intermediate firms. As the productivity increase occurs only once and the required skill level is exogenous no investment decisions are required.

$$\begin{aligned} & \max_{\rho_i'} \left\{ \pi^{RD}\left(\rho_i, \rho_i'; \theta_i\right) + \left[1 - F\left(\theta_i\right)\right] \sum_{t=1}^{\infty} \beta^t \pi^{int}\left(\rho_i', \rho_i'\right) \right\} \\ = & \max_{\rho_i'} \left\{ \left(1 - F\left(\theta_i\right) - \left(\rho_i \frac{\rho_i' - \rho_i}{a_P}\right)^{\phi_1}\right) \rho_i^{\alpha} \pi - \omega_H \left(\rho_i \frac{\rho_i' - \rho_i}{a_P}\right)^{\phi_1} + \left[1 - F\left(\theta_i\right)\right] \frac{\beta}{1-\beta} \rho_i'^{\alpha} \pi \right\} \end{aligned}$$

The first line expresses the value of the firm divided into current and future profits. Future per–period profits are constant with a fixed share of firms $[1 - F(\theta_i)]$ using the new technology and a constant productivity level of ρ_i'. The second line replaces the cost function for productivity enhancements with equation (3.18). The first–order condition of profit maximization w.r.t ρ_i' is

$$\frac{2\rho_i^2}{a_P^2}\left(\rho_i' - \rho_i\right)\left(\rho_i^{\alpha}\pi + \omega_H\right) = \alpha[1 - F(\theta_i)] \frac{\beta}{1-\beta} \rho_i'^{(\alpha-1)} \pi, \tag{3.24}$$

where the left hand side represents costs of improving the technology from ρ_i to ρ_i', which consist of the foregone profits due to smaller market size and the labor costs. The right hand side represents future profits with higher productivity ρ_i'. The development costs are independent of the minimum required skill level θ_i, while the benefits on the right hand side do depend on the size of the market $[1 - F(\theta_i)]$, with $F'(\theta_i) > 0$. A higher θ_i, i.e. a higher minimum skill level, reduces the benefits of an innovation and consequently incentives for productivity enhancements increase with the market size. There is hence a complementarity between productivity growth and market size. Noting that the market size of the new technology is negatively related to the market size of the old technology we can interpret this as the effect of directed technical change: the larger the number of intermediate firms of a given technology, the larger the investment in productivity increments by the monopolist.

Next to the market size mechanism, the first order condition (3.24) reflects also the decreasing returns to productivity improvements. A higher level of productivity ρ_i increases costs on the left hand side for given θ_i without direct compensation on the side of gains. As a consequence the developing firm reduces the step size $(\rho_i' - \rho_i)$ of productivity enhancements the further the technology has matured.

3.5 Numerical results

The two analytical results obtained before will be encountered again in the numerical example presented here. To recapitulate, productivity improvements increase with market size, but decrease with the productivity level already reached. The numerical exercise intends to mimic the diffusion of IT technology which started in 1971 when using the definition by Jovanovic and Rousseau (2005)[4].

For the numerical example we use the following parameterization

α	0.7	elasticity of substitution between interm. goods
β	0.9	discount rate
a_P	0.12	cost factor productivity increments
a_θ	0.18	cost factor skill requirements
ϕ_P	2	cost convexity of productivity requirements
ϕ_θ	2	cost convexity of skill increments
$F(\theta_i)$	θ_i	uniform skill distribution

Table 3.1: Numerical values for the parameters in the model

α represents the degree of substitutability between the goods of the intermediate industries in the final good sector and $1/\alpha$ is the mark-up over wages for the intermediate sector. β represents the yearly discount factor, set to lead to an internal rate of return to investment of 10%. As a major difference to Mukoyama (2004) we assume a uniform distribution of skills $F(\theta_i)$ in the population. We use such a distribution in order to abstract from possible influences of single peaked distributions such as the frequently used lognormal distribution. The effects in this model do not rely on the specific form of the distribution function but are generated through the endogenous interactions between productivity enhancements and skill reduction in the model. ϕ_P and ϕ_θ have been chosen to generate quadratic costs for productivity increments and skill reduction and the cost parameters of the development sector a_θ and a_P are chosen to obtain a diffusion time which is similar to the one of the two GPTs, electricity (36 years) and IT (since 1971).

We solve the numerical problem by value function iteration with a support for skills $\theta \epsilon [0, 1]$ and a support for productivity $\rho \epsilon [1, 2.2]$.

3.5.1 S-shaped diffusion curve and productivity gains

The numerical results unveil the interconnection of market size and productivity improve-

[4] In order to compare the two General Purpose technologies electricity and IT the authors use as starting date of difussion, the moment at which one percent of total capital in the US economy was IT capital, which was 1971.

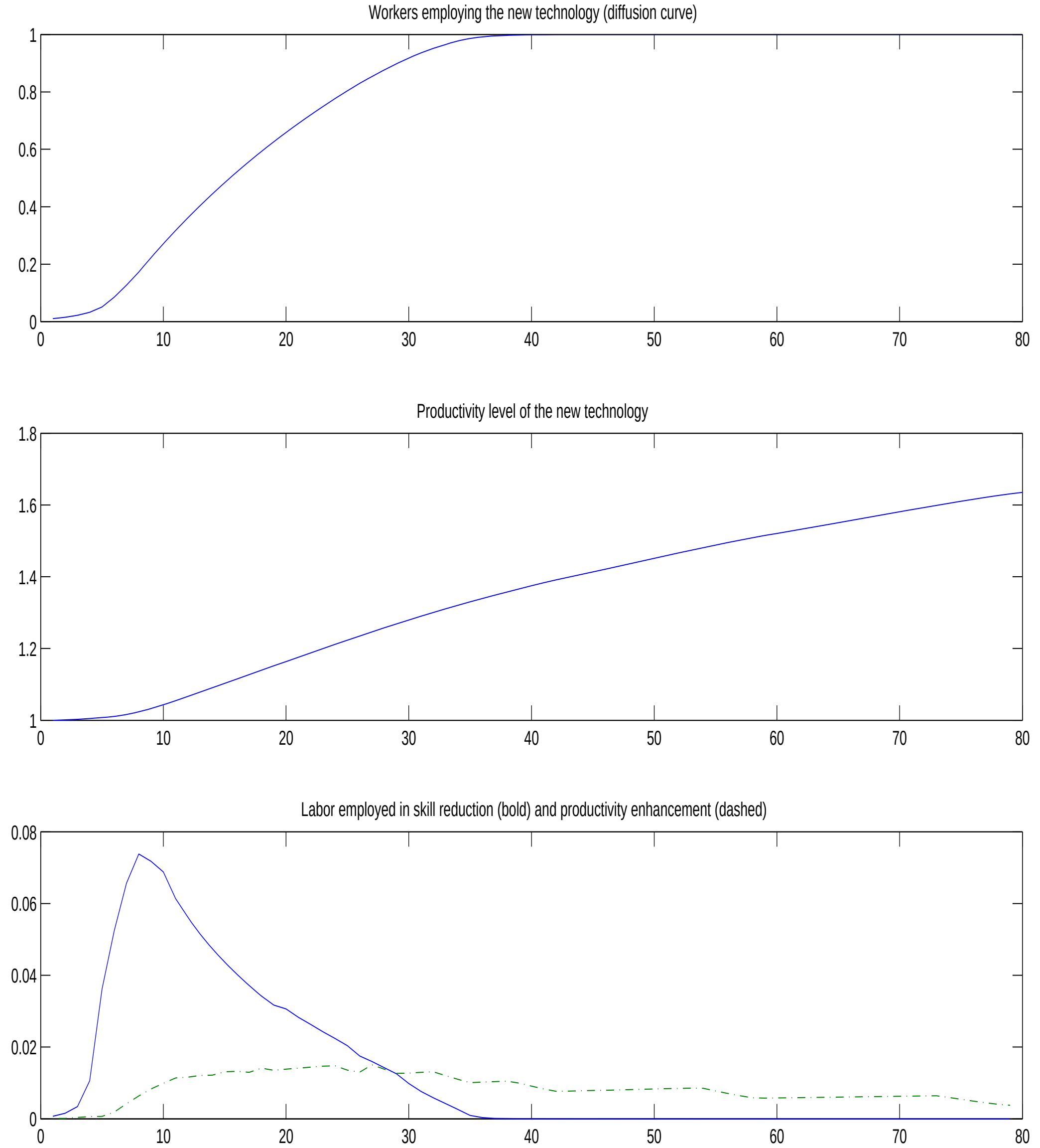

Figure 3.1: The evolution of different variables along the diffusion path over a time horizon of 70 years. Diffusion occurs S-shaped, productivity increases continuously and only slows down when the technology has completely diffused. Skill–reduction is occurs in the first 35 years, productivity enhancements occur on a longer horizon.

ments. The new technology diffuses in an S-shaped manner even though we are assuming a uniform distribution of skills in the population. Single peaked distributions such as the normal or lognormal distribution would lead to variations in the exact form of the diffusion path, but does not change its S-shaped characteristic. The shape of the diffusion curve is determined by two different effects at the beginning and at the end of diffusion. The slow ontake is achieved by the resource constraint of skilled labor with small market size for the technology. A young technology is only known to few workers, as a consequence only few workers can be employed in the development sector to decrease the minimum skill requirement. At the same time only few intermediate firms can operate with the new technology which reflects a small market size for the technology and hence only small per–period profits for the developing firm. As the number of adopting workers increases, profits increase, development is no longer labor constraint and the speed of diffusion increases. Once the technology has diffused to part of the population the productivity is also enhanced.

Comparing the costs for skill reduction at early and later stages reveals that labor costs increase as the productivity of adopters in the intermediate sector increases and the developing sector needs to offer similar wages. This induces firms to reduce the step size as for skill reduction as the technology matures (note that the cost function for the step size is quadratic). The S-shaped diffusion curve is hence obtained, not from a specific distribution of skills in the economy, but through the interplay of little skilled labor and low market size at the beginning and increasingly expensive diffusion costs at the mature stage of the technology.

Regarding the evolution of productivity, it does not increase directly with the introduction of the technology. Only once the technology has sufficiently diffused it becomes profitable for the firm to invest in productivity enhancement. Their benefits increase with the diffusion of the technology due to its larger market size which is especially visible during the early stages of diffusion. The lower the skill requirement the larger become the incentives to increase the productivity and hence the firm invests more into the development of productivity. With the maturing and the complete diffusion also productivity growth slows down due to the fact that it becomes more and more difficult to further improve the technology. Two elements characterize the productivity evolution, at early stages it is strongly conditioned by the market size complementarity, while at later stages the increased costs to improve a mature technology slow down productivity gains.

These explanations show that productivity increments strongly depend on the diffusion level of the technology and at the same time, the exact diffusion pattern hinges on the improvements

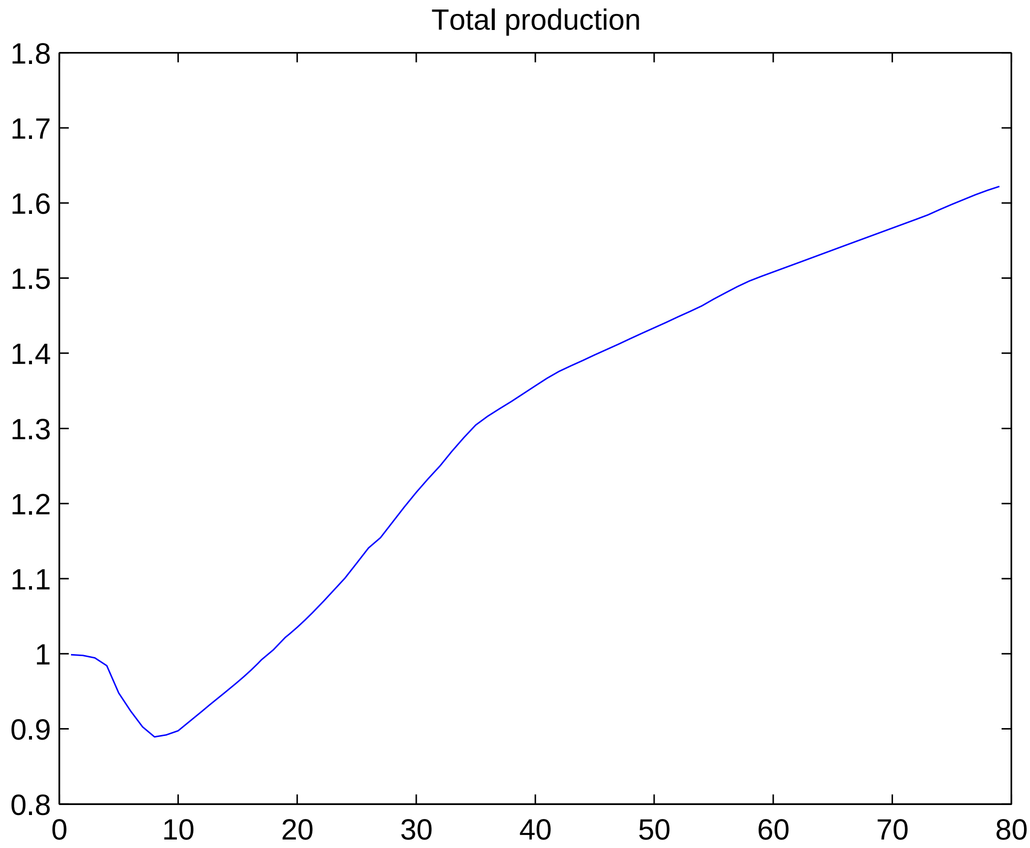

Figure 3.2: Evolution of final good production along the diffusion path. The slump is due to a large number working in the development sector improving the new technology.

in productivity.

3.5.2 Final production output

Output of final goods as defined in (3.1) is influenced by the productivities of the two technologies employed, the degree of diffusion of the new technology and the share of population employed in the development sector. The larger the productivity and the lower the skill requirement of the new technology the larger is the output level of the final good. In addition the share of workers employed in the development sector negatively affects output and reduces the number of operative industries, which in turn has effects on the profit level of R&D firms.‘i‘

The numerical exercise shows that the small R&D activity at the beginning of the diffusion process leaves the level of total output unaltered and only with the increase in R&D activity more labor is diverted from the productive sector to the development sector inducing a slump in final output. With the productivity increase and the reduction in development activity final output recovers and picks up growth. As the technology matures growth slows down and eventually it becomes too costly to improve the GPT further leading to a halt in growth. We observe the lowest levels of production after 8–9 years after the introduction if the GPT, which is roughly the time frame of the information technology covering the late seventies and beginning of the eighties .

3.6 Concluding remarks

This paper provides a foundation for diffusion processes with firms holding rights on a general purpose technology by combining investment decisions in productivity improvements and market size. The diffusion of new technologies and their central role in the growth process has been described in general terms by Rosenberg (1976) and specifically for General Purpose Technologies by David (1991) and Jovanovic and Rousseau (2005). A striking feature is the timely process for technologies to diffuse and the contemporaneous gains in productivity that accompanies this process. In this paper we were able to reproduce some of the stylized facts and generate the a mechanism by which the degree of diffusion and the productivity level are complementary to each other. The technology gains in productivity the more it has diffused and the larger the productivity the faster it diffuses. This is due to the fact that profits for the developing firm increase with market size and the productivity level.

The combined assumption of productivity increments with the restriction that only adopters may further improve the technology leads to an S-shaped diffusion curve. It is therefore not necessary to assume a specific skill distribution within the population, but a single peaked distribution such as a normal or lognormal distribution further enforces the S-shaped pattern.

Applying the setup to the diffusion of IT, the numerical example is able to reflect the long–run diffusion pattern of a General Purpose Technology which requires more than 30 years. Productivity gains to the are slow at the beginning, gain momentum and decrease again as diffusion is completed, but nevertheless the productivity gains continue for a long horizon.

This model is evidently only a stylized version, the diffusion of a GPT implies continuous innovation as well as a much richer a pattern of entry and exit of firms along the path, instead of being controlled by a single monopolist as assumed here. In addition, this model does not consider issues regarding the problems of adoption and their costs on the demand side. But nevertheless it allows to unveil the complementarity between diffusion and productivity improvements.

Bibliography

Acemoglu, D. (1998), Why do new technologies complement skills? directed technical change and wage inequality, *Quarterly Journal of Economics*, 113(4), 1055–89.

Acemoglu, D. (2002), Directed technical change, *Review of Economic Studies*, 69(4), 781–809.

Aghion, P. and P. Howitt (1992), A model of growth through creative destruction, *Econometrica*, 60(2), 323–51.

Aghion, P. and P. Howitt (1998a), *Endogenous growth theory*, Problems and Solutions by Cecilia Garcia-Penalosa in collaboration with Jan Boone, Chol-Won Li, and Lucy White. Coordinated by Maxine Brant-Collett, MIT Press, Cambridge and London.

Aghion, P. and P. Howitt (1998b), On the macroeconomic effects of major technological change, in: E. Helpman (ed.), *General purpose technologies and economic growth*, 121–44, MIT Press, Cambridge and London.

Bartel, A. P. and F. R. Lichtenberg (1987), The comparative advantage of educated workers in implementing new technology, *Review of Economics and Statistics*, 69(1), 1–11.

David, P. (1991), Computer and dynamo. the modern productivity paradox in a not-too-distant mirror, in: *Technology and Productivity. The Challenge for Economic Policy*, 315–347, Organization for Economic Co-operation and Development, Paris.

Doms, M., T. Dunne, and K. R. Troske (1997), Workers, wages, and technology, *Quarterly Journal of Economics*, 112(1), 253–90.

Goldin, C. and L. F. Katz (1999), The returns to skill in the United States across the twentieth century, *NBER Inc NBER Working Paper*, 7126.

Greenwood, J. and M. Yorukoglu (1997), 1974, *Carnegie Rochester Conference Series on Public Policy*, 46, 49–95.

Griliches, Z. (1957), Hybrid corn: An exploration in the economics of technological change, *Econometrica*, 25, 501–522.

Helpman, E. and M. Trajtenberg (1998), A time to sow and a time to reap: Growth based on general purpose technologies, in: E. Helpman (ed.), *General purpose technologies and economic growth*, 55–83, MIT Press, Cambridge and London.

Jovanovic, B. and S. Lach (1989), Entry, exit, and diffusion with learning by doing, *American Economic Review*, 79(4), 690–99.

Jovanovic, B. and Y. Nyarko (1996), Learning by doing and the choice of technology, *Econometrica*, 64(6), 1299–1310.

Jovanovic, B. and P. L. Rousseau (2005), General purpose technologies, *NBER Inc NBER Working Paper*, 11093.

Mukoyama, T. (2004), Diffusion and innovation of new technologies under skill heterogeneity, *Journal of Economic Growth*, 9(4), 451–79.

Nelson, R. and E. Phelps (1966), Investment in humans, technological diffusion, and economic growth, *American Economic Review*, 56, 69–75.

Parente, S. L. (1994), Technology adoption, learning-by-doing, and economic growth, *Journal of Economic Theory*, 63(2), 346–69.

Rosenberg, N. (1976), *Perspectives on technology*, Cambridge University Press Cambridge [Eng.

Violante, G. L. (2002), Technological acceleration, skill transferability, and the rise in residual inequality, *Quarterly Journal of Economics*, 117(1), 297–338.