Three Essays on
Political Institutions, Inequality,
and Economic Growth

Inaugural-Dissertation

zur Erlangung des Grades eines Doktors

der Wirtschafts- und Gesellschaftswissenschaften

durch die

Rechts- und Staatswissenschaftliche Fakultät

der

Rheinischen Friedrich-Wilhelms-Universität Bonn

Vorgelegt von Ling Shen

aus Hangzhou, V. R. China

10.08.2006

Dekan:　　　　　Prof. Dr. Johannes Köndgen

Erstreferent:　　　Prof. Monika Merz, Ph.D.

Zweitreferent:　　Prof. Dr. Jürgen von Hagen

Tag der mündlichen Prüfung 10.08.2006

Diese Dissertation ist auf dem Hochschulschriftenserver der ULB Bonn
http://hss.ulb.uni-bonn.de/diss-online elektronisch publiziert.

To Yun Chai

Acknowledgements

Ling Shen
10.08.2006 Bonn

Contents

Appendix 117

References 126

List of Figures

List of Tables

Chapter 1

Introduction and Summary

Why are some countries much richer than others? Why do some economies grow faster than others? Economists have asked these trite yet crucial questions for more than one century. Traditional neoclassical growth theory, following Solow (1956), Cass (1965) and Koopmans (1965), emphasizes the effect of factor accumulation on economic performance. In these models, cross-country differences in income per capita are due either to differences in the saving rate or other exogenous parameters. More recent research endogenizes steady-state growth and technical progress. For instance, Lucas (1988) emphasizes the externality arising from human capital in production, and Aghion et al. (1992) shed light on the role that destructive innovation plays for economic growth. Although these traditional economic studies provide much insight into the mechanics of economic growth, some economists don't agree that these theories provide a fundamental explanation for economic growth. For instance, North and Thomas (1973, p. 2) argue: "the factors we have listed (innovation, economies of scale, education, capital accumulation etc.) are not causes of growth; they are growth." In their view, a given economic institution provides the fundamental explanation for economic growth, because it shapes incentives for key actors in a society. In particular, institutions influence investments in physical and human capital as well as the development of new technologies.

Acemoglu et al. (2004b) further point out that economic institutions are determined by political institutions and the distribution of resources. Hence, the relevant sources of economic growth can be traced back to political institutions and income or wealth inequality.

This thesis consists of three essays which investigate the role that inequality or political institutions play for economic growth. Before summarizing each essay, we will motivate their respective theme.

Table 1.1 presents the possible linkage between political institutions and economic growth. We collect nominal GDP-data in USD from the United Nations. The civil liberty index is taken from Freedom House, a non-profit organization which publishes surveys detailing state of civil liberties, political rights, and economic freedom every year. According to Freedom House, civil liberties allow for the freedom of expression and belief, associational and organizational rights, the rule of law, and personal autonomy without interference from the state. The Civil liberty index is a widely used measurement of the kind of political institutions surrounding the economic environment. The index ranges from 1.0 to 7.0, where 1.0 reflects the most free and 7.0 the least free rating. There are 157 countries whose data are available. Since there is no commonly accepted criterion for what constitutes a developing country, we simply define developing countries as those countries whose nominal GDP per capita falls below the average level of these 157 countries in any given year. In 1970, there are 113 developing countries and 44 developed countries. In the second part of the table, we classify all 157 countries into 6 categories according to their average civil liberty index between the years 1972 and 1974 (C.L. in Table 1.1).

Table 1.1: Political Institutions and Economic Growth

	Observations (total 157)	Civil Liberty Index 1972-74	GDP per capita [USD]		Growth rate 70-04 [% p.a.]	Variance of Growth rate
			1970	2004		
Developing countries	113	4.56	300	2370	4.78	6.41
Developed countries	44	2.81	2446	28550	6.66	4.71
C.L. 7~6	42	6.38	442	2764	4.37	8.03
C.L. 5.9~5	27	5.22	472	4767	4.46	6.28
C.L. 4.9~4	18	4.28	414	3104	4.44	5.49
C.L. 3.9~3	23	3.26	781	6917	6.00	4.32
C.L. 2.9~2	28	2.08	1175	14278	6.07	5.51
C.L. 1.9~1	19	1.02	2733	34977	7.46	0.89
Average		4.07	881	6418	5.84	6.62

Notes: Nominal GDP per capita is taken from the United Nations Statistics Division, and the Civil Liberty index (C.L.) originates from Freedom house. See the text for definitions and further details.

Our findings are as follows. First, developing countries tend to have a more dictatorial political institution than developed countries. Table 1.1 shows that the average value of the civil liberty index among developing countries is 4.56, whereas that of developed countries is 2.81. According to the index, most west European nations and inhabitants of the United States of America enjoy much more freedom than those living in China, other former communist countries, South Korea, Iran, and many Arab or African countries. The index takes on 1.0 for the former group of countries; it lies between 6.0 and 7.0 for the latter group. Second, the long-run growth rate of developing countries is lower than that of developed countries (4.78% vs. 6.66% p.a. between 1970 and 2004), and the variance of the long-run growth rate among developing economies is higher than that of rich

countries (6.41 vs. 4.71). This implies no absolute convergence of income in the world. Some developing countries even have a negative long-run growth rate. E.g., the nominal GDP of the Democratic Republic of the Congo decreased from 350 USD in 1970 to 115 USD in 2004. Other developing countries are indeed approaching the status of being developed. E.g., the nominal GDP of South Korea increased from 291 USD in 1970 to 14,266 USD in 2004. Finally, we find from the second part of table 1.1 that less freedom tends to accompany lower income, a lower long-run growth rate, and a higher variance of the growth rate. In particular, the most dictatorial countries (C.L. $7 \sim 6$) have a variance of more than 8, whereas the most democratic countries (C.L. $1.9 \sim 1$) tend to have a very stable long-run growth rate (variance 0.89). These figures suggest that a benevolent dictator is much more important for stimulating growth in dictatorial countries than a good president is in a democracy. Hence, it is important to ask which role the dictatorial government plays for the economic performance in developing countries, and why some dictators are benevolent and others are not. Mainstream economic theory has paid little to no attention to this aspect.

Chapter 2 formulates a game-theoretic model between a dictator and the people to find underlying determinants of dictatorial behavior. At first, we assume that the engine of economic growth is private investment. It can increase the productivity of individuals who invest and also the aggregate technology level. Then we define a good dictator as one who gives social transfer to the people and thereby stimulates future investment and output. The bad dictator just taxes her citizens and keeps tax revenue for her own consumption. The degree of goodness is measured by the amount of money transferred from the dictator to citizens. The so-called good dictator decides to give social transfer to stimulate private investment to obtain more income due to taxes in the future. Hence, both good and bad dictators are motivated

by the same goals. However, the citizens' response to the dictator's behavior determines whether a dictator should act as a good or bad dictator. Being good or bad is not an inherent property of the dictator. The dictator and the citizens always act in their own best interest.

This chapter builds on McGuire and Olson (1996). According to these authors, a good dictator implements growth-enhancing policies in order to increase taxes in the future. Our research extends their work by emphasizing the risk involved for a dictator with choosing a growth-enhancing policy: while such policies can raise additional tax revenues in the short-run, they also increase the likelihood of a revolution which can lead to the eventual overthrow of the dictator. Chapter 2 has three main findings. First, the return of private investment has a negative effect on the behavior of dictators. In a country where citizens can earn much through private investment, the dictator has little incentive to pass on social transfer to citizens. Second, contrary to McGuire and Olson (1996) we find that a long life-time of a dictator does not always induce her to act benevolently. With a longer life-time, she will be more concerned with the likelihood of a revolution. Finally, we distinguish two different effects of economic performance on democratization. If a good economic performance is achieved by technological progress, then it will lead to a speedy democratization. This result coincides with the empirical research of Barro (1997). However, if a country becomes richer because of more natural resources, then the good economic performance impedes political transition. This result is consistent with Ross (2001), who finds that oil impedes democracy.

The thesis' second main topic is the relationship between economic growth and income or wealth inequality. Although the wealth inequality is more relevant in theoretical modelling, most empirical studies use income inequality data as a proxy for wealth inequality because of the scarcity of

available data on the distribution of wealth. "It is generally argued that this is unlikely to be a major problem since both measures of inequality generally vary together in cross-sections." (Aghion et al. 1999).

Table 1.2: South Korea vs. Philippines

	S. Korea	Philippines
Nominal GDP p.c. 70 (UN)	291	183
Nominal GDP p.c. 04 (UN)	14266	1059
Nominal growth rate 70-04 p.a.	11.44%	5.16%
Real GDP p.c. 70 (BL)	1680	1369
Real GDP p.c. 89 (BL)	6206	1726
Real Growth rate 70-89 p.a.	6.88%	1.22%
Civil Liberty Index 70-72(BL)	6	5.333
Population 70(UN)	31.9 Mio.	36.6 Mio.
Average schooling years 70(BL)	5.583	4.833
Enrollment ratio for Primary Edu. 70(BL)	100%	100%
Enrollment ratio for Secondary Edu. 70(BL)	42%	46%
Enrollment ratio for High Edu. 70(BL)	8%	19.9%
Population share in Capital 60(Lu)	28%	27%
GDP share of Industry 60(Lu)	20%	28%
Ratio of Gov Consumption to GDP 70-74(BL)	4.2%	11.7%
Ratio of Investment to GDP 70-74(BL)	24%	14%
Gini 65(DS)	0.34	0.51
Gini 71(DS)	0.36	0.49

Notes: The term XX behind a variable denotes a specific year (e.g., Nominal GDP p.c. 70 is nominal GDP per capita in 1970), and the term XX-XX denotes a period. GDP is in USD. The sources of data are in parentheses. UN is United Nations Statistics Division. BL is Barro and Lee Database (1994). Lu is Lucas (1993). DS is Deininger and Squire (1996). See the text for further details.

This theme is motivated by the following puzzle which Lucas (1993) and Bénabou (1996) raised. Table 1.2 shows the main economic aggregates of two East Asian countries: the Philippines and South Korea. In the early 1970s, they were similar in their GDP per capita, the average education level, the size of their population, and the degree of urbanization and

industrialization. According to the civil liberty index, their political institutions are also very similar. Yet, they have had dramatically different growth rates over the last 30 years. While South Korea grew at 11.44% per year, the Philippines have only grown 5.16%. Today, South Korea's GDP per capita is about 14 times higher than that of the Philippines.

Observed differences in the growth rates can be explained by the theory in Chapter 2. The dictator of South Korea is more benevolent than the one in the Philippines in that she spent less on consumption than the Philippines (the ratio of government consumption to GDP is 4.2% in S. Korea and 11.7% in the Philippines), and induced more investment (the ratio of investment to GDP is 24% in S. Korea versus 14% in the Philippines). Moreover, there also are reasons other than different dictators. Looking beyond first moments, we find significant differences in the countries' income distribution. The Gini coefficient was seventeen percentage points higher in the Philippines than in South Korea in 1965 (13 percentage points higher in 1971). This difference equals 1.8 standard deviations in the world distribution of the Gini coefficient, or 2.5 among East-Asian countries. Similar results can be found if we use the Gini coefficient for farmland (Taylor and Hudson 1972), which is a proxy for wealth inequality.

This fact suggests that the answer to the puzzle, why two similar economies could grow so differently, may lie beyond the representative agent framework. In other words, the distribution of income/wealth, instead of the average level, matters. Since the early 1990s, more and more cross country evidence (Berg and Sachs 1987, Persson and Tabellini 1994, Alesina and Rodrik 1994, Clark 1995) supports the prediction that income inequality has a negative impact on long-run growth rates.

Such a result comes as a surprise both with regard to traditional theories in the field, and with regard to the channels through which inequality might affect economic growth. According to the comprehensive survey of Bénabou (1996), there are three main channels: (i) imperfect capital markets, (ii) political economy aspects, and (iii) social conflict. All of these channels address the impact of the wealth/income distribution on investment. We will refer to them as the supply-side effects. Although these theories – the imperfect capital market theory in particular – are considered as plausible by many economists, the evidence in table 1.2 raises doubt concerning the prediction that more equally distributed initial income in South Korea in 1960's leads to a higher level of human capital investment which, in turn, induced a higher long run growth rate. In 1970, the Philippines had a somewhat higher enrollment ratio for secondary and higher education. That implies that inequality may affect output through other channels besides the supply-side effects. In the literature, Murphy et al. (1989) at first introduce, and then Zweimüller et al. (2000, 2005) extend the demand-side effect, i.e., inequality can affect output through the demand for consumption goods.

Chapters 3 and 4 follow the work of Zweimüller et al. (2000, 2005) and illustrate the demand channel through which inequality affects growth. The main idea is based on the vertical differentiated goods market, which was originally introduced by Shaked and Sutton (1982, 1983). Profit of innovation determines its incentive. The profit of a new differentiated good comes from the willingness to pay and the market share. Both of them will be affected by the distribution of income. Inequality may supply enough rich consumers to buy new luxury or higher quality goods. But on the other side, inequality induced by a relative small market size impedes also the spread of new or better quality goods. Hence, the effect of distribution on innovation is not *a priori* clear.

We assume an economy with two kinds of individuals, the poor and the rich. Hence, the Gini coefficient is decomposed into two variables, namely, the relative wealth of the poor and the population share of the poor. The purpose of our research is to show that these two variables might have different effects on economic growth. Thus, the simple regression of the Gini coefficient on the long-run growth rate is able to generate neither an unambiguous empirical result, nor a useful policy recommendation.

In chapter 3 we only focus on the demand-side effect. We introduce the interdependent relationship between the relative wealth of the poor and the population share of the poor in an overlapping generations model. We assume the poor can become rich through education. If the wealth difference between the poor and the rich becomes larger, then more individuals have incentives to undergo education. Hence, the population share of the rich increases. This incentive consideration reflects the traditional argument that inequality is growth-enhancing. However, traditional theory considers the supply-side effect. I.e., the larger human capital investment, the greater output is. We concentrate on the effect of distribution on the demand for high quality goods. Since there are only two kinds of consumers, at most two top quality goods can survive in the quality goods' market. Hence, chapter 3 considers an oligopolistic market for quality goods. After a successful innovation, the newly invented quality goods enter into the market, the current best quality good becomes the second best one and the current second best quality good is driven out of the market. Hence, the return of a successful innovation comes from two parts: profits of the best quality goods until one innovation succeeds and profits of the second best quality goods until two higher qualities are invented.

Our results show that there exists a separating equilibrium, where the best quality good is sold to the rich and the second best good to the poor. In this

equilibrium, a decrease in the relative wealth of the poor implies that more individuals undergo education and become rich. Hence, the best quality good has a larger market share and the return of innovation increases. Consequently, the innovation rate rises. In this sense, inequality is good for the innovation rate. If the population share of the poor increases, the best quality supplier faces a smaller market share. Hence, there is less incentive to invent. The innovation rate decreases. The inequality induced by a large population share of the poor is bad for the innovation rate.

In chapter 4, we combine the above quality improvement model and the neoclassical production function. Thus, the impact of wealth inequality on economic growth is through the supply of human capital as well as the demand for better quality goods. According to the imperfect capital markets theory, larger inequality of wealth leads to a lower aggregate supply of human capital. Hence, the supply-side effect is negative. This result comes directly from the assumption of the neoclassic production function. Similarly as in chapter 3, we investigate the demand-side effect of wealth inequality in quality improvement model. In the separating equilibrium, the relative wealth of the poor has a negative effect on economic growth. And in the pooling equilibrium, the effect is positive. Hence, in general, there is a non-linear relationship between the relative wealth of the poor and economic growth. This result is partly consistent with recent empirical findings (Chen 2003). Contrary to the relative wealth of the poor, the population share of the poor has either negative effect, in the separating equilibrium, or no effect, in the pooling equilibrium, on economic growth.

The results of chapters 3 and 4 have an important implication for economic policy. In a country where the separating equilibrium is overwhelming and the goal of government policy is to achieve both an increase in economic growth and a decrease in inequality, one should consider decreasing the

population share of the poor but not directly redistributing from the rich to the poor.

Finally, chapter 5 addresses some questions for future research.

Chapter 2

When will a Dictator Be Good?

2.1　Introduction

Economists have realized the importance of political institutions in shaping economic performance. Most academic studies of political economy (e.g. Shepsle and Weingast 1995, Cox 1997, Persson and Tabellini 2000, 2003) focus on the democratic political system, where formal political institutions, such as the constitution, the rule of law, and the election system, are already well advanced. However, few studies shed light on dictatorship, although most people on earth live in such regimes.[1] A puzzling phenomenon in dictatorial economies is that they can achieve dramatically different economic growth rates. While East Asian dragons have grown 8-10% per year for almost 30 years, many African countries suffered from recessions in the same period, although both East Asia and African countries are controlled by some dictators. (The study of East Asia, see Collins and Bosworth 1996; the study of Africa, see Easterly and Levine 1997)

A simple comparison between dictators in East Asian dragons and those in African or South American countries implies that the behavior of

---

[1] Recent works in this line include e.g. Wintrobe (1998), Bueno de Mesquita et al. (2003), Acemoglu (2003) and Acemoglu et al. (2004a).

autocracies might be important for the fortune of nations. (For formal research on the relationship between the political institution and economic growth, see Acemoglu et al. 2004b, and Glaeser et al. 2004) The good dictator invests in public education and infrastructure, establishes the rule of law to encourage private investment, subsidizes R&D, and so on. However, the bad one simply transfers a large fraction of social wealth to herself. One classic case of the bad dictator is Mobutu Sese Seko in the Democratic Republic of the Congo from 1965 to 1997. According to Acemoglu et al. (2004a), in the 1970s, 15-20 percent of the operating budget of the state went directly to Mobutu. In 1977 Mobutu's family took 71 million USD from the National Bank for personal use and by the early 1980s his personal wealth was estimated at 5 billion USD. In 1980, GDP of Congo is only 14.7 billion USD according to the databank of UN.

Here, the good dictator invests more in public projects than the bad, although both are willing to tax citizens. It is of interest to ask why some dictators are good and others are bad.[2] This question is important for economists, because the type of dictator determines the kind of economic performance observed. It is also important for politicians, since good economic performance induces early democratization, according to the Lipset/Aristotle hypothesis,[3] which states that prosperity stimulates democracy. Although the impact of democracy on economic performance is far from reaching a consensus among economists,[4] the reverse causality--the Lipset/Aristotle hypothesis--has shown strong empirical regularity in many empirical studies (e.g. Barro 1999, Boix and Stokes 2003).

---

[2] Sah (1991) believes that dictatorship is a risky investment.
[3] We owe this terminology to Barro 1999.
[4] Barro (1997) points out that there is a non-linear relation between democracy and economic growth. Whereas democracy is growth enhancing in the young period, it is bad for further economic growth when democracy exceeds beyond a certain point.

The present article assumes a dichotomic world, where democracy is defined by the one-person-one-vote majority voting system (Huntington 1991, Schumpeter 1947) and dictatorship (or autocracy, or non-democracy, we treat all as equal for simplicity) means that one person holds all political power. We provide a theoretical model to illustrate underlying determinants of a dictator's behavior. Furthermore, we emphasize the trade-off faced by the dictator between economic benefits from a growth-enhancing policy in the short run and the shorter life-time of a dictator due to earlier revolution, which is induced by economic growth in the long-run. This simple model is based on three important components.

First, we argue that economic growth is generated by decentralized investment. Individuals' investment increases their private productivity. This private investment has a positive external effect on the aggregate technology level. The more individuals invest, the higher the aggregate technology level.

Second, consistent with the literature, we assume that the political power affects economic performance through the redistribution policy. The redistribution policy in the current model is summarized by a two-dimensional vector, which consists of the tax rate and the social transfer. A dictator can invest in public education, infrastructure or provide direct subsidies to individuals. All of them can be considered as the social transfer, which encourages individuals to invest. Hence, the social transfer policy measures the goodness of the dictator. Throughout the current paper, the terms bad, good and better are merely shorthand for statements about the amount of the social transfer. Following individuals' production, a dictator sets the tax rate and then collects tax revenue. Hence, the promise to reduce

the tax rate in a dictatorship isn't credible[5]. This assumption simplifies the analysis and enables us to concentrate on the key question of this paper: Why do some dictators transfer more to citizens, thereby inducing higher growth rates, while others concern themselves more with their own consumption and thus less social transfer. We argue that both good and bad dictators behave to benefit themselves. The citizens' response to the dictator's behavior determines whether a dictator should act as a good or bad one. A good dictator implements a social transfer policy because she can tax more. In the short run this is the economic benefit from a growth-enhancing policy.

Third, democracy is growth-enhancing in the current model, because it protects decentralized investment from expropriative taxation. Hence, it is better than any dictatorship under scrutiny.[6] In a dictatorship, the higher the aggregate technology level, the greater the taxed income is. In turn, citizens have greater incentives of political transition. Nevertheless, the ruler impedes this political transition because the loss of political power coincides simultaneously with the loss of economic benefits. A good dictator encourages higher private investments, thereby inducing a higher aggregate technology level in the future. Consequently, democracy is more attractive to citizens. It leads to earlier democratization, which constitutes the cost to a good dictator.

---

[5] According to Acemoglu (2000), democratization is the strategic decision of political elites to prevent revolution. As long as elites hold political power, the citizen can not trust that elites will undergo a pro-citizen redistribution for ever. Hence, citizens would like to revolt if the revolution condition is satisfied. For the elite, it is better to democratize when she faces the risk of revolution. I follow his idea and assume that the taxation is after the private investment. Hence, the promise to reduce the taxation is incredible, as long as the dictatorship does not change.
[6] See Proposition 3, assumption A.2 ensures that democracy is better than a dictatorship in the current model.

We find that the dictator is good if the highest tax rate is sufficiently high and the return rate of private investment is sufficiently low. The goal of the dictator to foster economic growth is to tax more in future. If the highest tax rate is great enough, the dictator expects to tax more. Hence, she is faced with a large incentive to become good. On the other hand, if the return rate of private investment is higher (e.g., because of more oil or other natural resources), then the initial investment level is higher. Hence, the dictator has lower incentives to encourage private investment. In this sense, oil and other natural resources have a negative effect on the behavior of a dictator.

Contrary to McGuire and Olson (1996), we point out that the longer life-time does not always give the dictator the incentive to do better. Their paper considers only the benefit of public investment (similar to social transfer), whereas the current paper emphasizes the trade-off between economic benefits in the short run and the shorter life-time in the long-run. If citizens face a higher revolution cost, i.e., the dictator can live longer, then her positive social transfer policy can generate more benefits for her, in turn, she has a higher incentive to be a good one. This is the argument of McGuire and Olson (1996). Furthermore, by recognizing this effect, we point out, that her positive social transfer will induce a higher economic growth rate in the long-run, which leads to an earlier revolution. If the dictator has a longer life-time, she will be more concerned with the negative effect of her social transfer policy. Hence, her social transfer is not necessarily larger, if she lives longer.

Another novel result is that we illustrate the different effects of good economic performance on democratization. If the return rate of the private investment increases due to a new discovery of natural resources, such as oil, then more individuals will invest. In turn, the country can achieve good economic performance. However, good economic performance does not

imply inducing sooner democratization, vice versa, citizens have lower incentive to revolt and the dictator has also lower incentive to be good. If good economic performance is achieved by the higher technology level, then we can observe the Lipset/Aristotle hypothesis. Hence, this simple model is consistent not only with respect to the empirical results of Barro (1997), but also that of Ross (2001), which finds that oil impedes democracy.

The present paper connects two different strands of the literature. First, the literature of political economy studies taxation and public investment by dictators (e.g., McGuire and Olson 1996) facing the potential contest of other political groups (e.g., Tornell and Lane 1999, Collier 2001, Konrad 2002). However, this literature does not correlate developments in a dictatorial nation with potential democratization. The theory of democratization in the framework of political economy frequently focuses on the pure redistributive model, for instance, Therborn (1977), Rueschemeyer et al. (1992) and Acemoglu and Robinson (2000, 2001). However, they don't distinguish between different dictators in the sense of growth-enhancing policies. Zak and Feng (2003) are more closely related to the current paper because they study also the relationship between economic growth and political transition. However, they emphasize the acceleration of democratization in different regimes' policies. In contrast to their work, we focus on the conditions under which different regimes (good or bad) exist. On the other hand, the literature of the new growth theory studies the impact of democracy on economic growth, e.g. Barro (1997, 1999) Kurzman et al. (2002), or the impact of redistribution policy on growth, e.g. Persson and Tabellini (1994), Benabou (1996, 2002), but few consider that the most growth-enhancing policies are implemented by dictators in non-democratic societies.

The paper proceeds as follows. In section 2.2, I will present the set-up of the model. In section 2.3, we study the exogenous growth case without the positive external effect of investments. Then we introduce the democratization process in section 2.4. In section 2.5, the external effect is investigated, in order to establish the relationship between political transition and economic growth. Moreover, we study the behavior of dictators who face the pressure of political transition. In section 2.6, the main results are summarized.

2.2 The Set-up of the Model

There are two types of political states: dictatorship and democracy, and two kinds of agents: the ruler and citizens. Citizens invest in a project which can increase their productive ability and produce output using this ability, whereas the ruler taxes the output after production in dictatorship. The dictator can choose to be good or bad. The good dictator shares a part of the tax income with some citizens, whereas the bad dictator consumes all tax revenue by herself. Democracy is characterized by equality: every citizen has the same political power to determine the tax rate and receives the same amount of transfers.

2.2.1 The Economic Environment

We consider an infinite horizon economy with two types of agents: a ruler and a continuum Λ of citizens. Citizens live infinitely long, but the ruler lives only if not killed in a revolution. Citizens are born with different ability levels which are invariant over time. These ability levels are assumed to be independent realizations of a continuous random variable taking values in the unit interval $[0,1]$. In that case we can also take $\Lambda = [0,1]$ as

indexing the set of citizens. This allows us to refer to citizens by their ability level: citizen $a \in \Lambda$ is that citizen with ability level a. For convenience, we assume a uniform distribution of abilities on [0,1].

The production function of citizen a in period t is given by:

$$y_{at} = A_t N a \lambda^{I_{at}}, \lambda > 1 \qquad (2.1)$$

where A_t represents the aggregate technology level, N is natural resources per capita, and I_{at} is an indicator function of investment in period t. $I_{at} = 1$ means that individual a invests at t, whereas $I_{at} = 0$ means that he doesn't invest. The investment cost is eA_t, $e > 0$, and it enables the investor to increase his productivity by the factor λ. Hence, the return rate of private investment for individual a is $\dfrac{N(\lambda - 1)}{e} a$. If his return rate is greater than 1, then he invests. This assumption implies that the private investment decision of individual a depends on his own productivity a, but not the aggregate technology level. The investment fully depreciates within one period. Hence, a citizen needs to invest in each period if investment is valuable to him. There exists a threshold, which is denoted by $\hat{a}$, i.e., individuals with ability lower than $\hat{a}$ do not invest, while others with ability higher than $\hat{a}$ invest. Hence, the investment ratio is $1 - \hat{a}$. In section 2.3, investment has no effect on A_t, because economic growth is assumed to be exogenous. In section 2.5, we assume that investment has a positive external effect on the aggregate technology level. As a result, long-run economic growth is endogenous.

2.2.2 The Political Environment

We assume that the ruler does not produce anything. However, she can tax the output of citizens. This is the crucial assumption of this paper. According to political economy literature, e.g. Benabou (1996), Persson and Tabellini (1994, 2000), non-democracy means that the rich, who are more productive, have more political power. We argue that this assumption describes an imperfect democracy well, but not dictators. This aspect does not apply to dictators such as Mobutu in the Democratic Republic of the Congo, and the dictators in Chinese history. They became dictators, not because they had higher productivity, but because of their military power in most cases. Our assumption is similar to that of McGuire and Olson (1996), where dictatorship impedes the growth of productivity due to taxation. The political institution is defined by the vector (τ, s)[7]. The tax rate τ lies between $[0, \bar{\tau}]$, $\bar{\tau} < 1$ and the social transfer s is financed through taxation. $\bar{\tau} < 1$ reflects the sustenance level. If taxed above that level, all citizens would revolt, because they have nothing left after taxation. For simplicity two extreme cases are considered: dictatorship and democracy. We assume that the initial political state is dictatorship, where the ruler can choose the tax rate and decide how to distribute the tax revenue. The bad dictator consumes the entire tax income alone, i.e., $s_a = 0 \ \forall a$. However, the good dictator shares the benefit with some citizens through social transfer, i.e., $s_a \geq 0$, for some $a, s_a > 0$, which is named the group-specific social transfer. The dictatorship is characterized by taxation. Both the good and

---

[7] This assumption comes directly from Lee (2003). To describe the difference between dictatorship and democracy he uses two variables, i.e., participation bias and redistribution bias. However, he does not consider the commitment problem. Hence, both of them are determined simultaneously in his model. This assumption is also consistent with Persson and Tabellini (2000). In their book (chapter 14), the taxation of capital income and the public investment in infrastructures are two policy instruments, which naturally affect private rates of return on investment, and in turn, economic performance. However, they study their effects in different models and do not consider group-specific public investment.

bad dictators tax citizens. The dictator is good in the sense that her redistribution policy ($s_a \geq 0$, for some $a, s_a > 0$) is growth-enhancing.

Dictator knows the distribution of private ability and the investment decision of citizens. However, she can not distinguish between individuals by ability. Hence, her group-specific social transfer policy can be set on the base of investment, but not upon private ability. For instance, Chinese government invests a lot in universities and sets an entrance examination. Only the candidates who pass the examination can study in these public universities. Here, studying in public universities is a human capital investment for the citizens as well as a social transfer from the government to the citizens who invest. Such social transfer is sunk and irreversible, e.g., universities are built before citizens decide to undergo education. It also is group-specific, e.g., only individuals with higher ability will undergo education. An entrance examination is not the single form of selection mechanism. For instance, the free infrastructure in industry zones is a social transfer from government to the individual who invests in this area.

In a democracy, there is no ruler and the tax rate is determined by all citizens through a "one-person-one-vote" majority voting system, where every agent gets the same transfer $s_a = s^{dem}, \forall a$. We assume that social transfer in a democracy is not group-specific, not because in reality there is no group-specific social transfer in the democratic society (in general, all social transfers are considered to be group-specific), but because the nature of democracy is such that everybody is treated equally. Hence, although the individual project, which is financed by the democratic government, could be group-specific, in the aggregate, the democratic government concerns itself with the interests of all citizens, and the social transfer is more equally distributed among individuals than under a dictator. Furthermore, allowing

group-specific social transfers in democracy would complicate our analysis of democracy, whereas the current article focuses on the non-democracy. Allowing group-specific social transfer in democracy does not qualitatively change our results concerning dictatorial behavior.[8] In fact, different majorities of citizens could support different group-specific social transfer schemes in democracy. Finally, everybody obtains the same *a priori*.

In order to attempt to change the political state through revolution, a citizen pays P for a weapon, and encounters a cost c if a revolution occurs. This cost of revolution could be either considered as the destroyed income in turmoil (Acemoglu 2001), or reflect the cooperation and/or coordination problem among a large scale of citizens. The cooperation problem among citizens has been modeled in details in some papers, e.g., Acemoglu et al. (2004a). The ruler acts by herself. Hence, she has no such problem. If the revolution is successful, the ruler dies. As a result, the ruler always tries to prevent the revolution. She also buys the weapon in order to repress the revolution. For simplicity, we assume the price of weapon to be fixed and the same for all. Whether the revolution can succeed depends on who possesses more weapons. This political transition is modeled by a sequential game. The citizens move first, the ruler then reacts. We assume that the ruler moves later, in order to reflect the advantage of holding political power. She can adjust the expenditure on weapons according to the revolution decision of citizens. However, this order of decisions reflects the actual weapon expenditures of citizens and the ruler but does not change the timing of the revolution. The current model focuses on the behavior of the ruler in dictatorship. Hence, the time of revolution, in turn, the life time of

---

[8] Appendix 2.2 shows that democracy is even better, if we allow that in democracy, social transfer is only given to the individual who invests. Then the citizens have higher incentive to revolt. Our result that the dictator faces a trade-off when she implements a positive social transfer policy has no qualitative change.

the dictator is the key issue. The amount spent on weapons by the citizens will not affect the social transfer policy of the dictator.

2.2.3 The Timing

Upon birth all citizens realize their abilities, and other exogenous parameters ($\bar{\tau}, \lambda, e$) are revealed. It is a finite repeated game between the ruler and a continuum of citizens until revolution succeeds. Within every period they play a sequential game, whose timing of events can be summarized as follows:

1. At the beginning of period t, the technology level A_t is determined either by the exogenous factor (section 2.3), or by the endogenous variables in time $t-1$ (section 2.5).

2. Citizens determine whether or not to undertake a revolution.

3. If there is no revolution, or if the revolution is repressed, the ruler can keep her political power. Then she decides whether to be a good dictator or not, i.e., to choose the scheme of the social transfer (s_a). The ruler can not observe the individuals' ability, but she can see whether the citizen invests or not.

4. If the revolution is successful, the ruler is killed and citizens establish the democratic political system.

5. After watching the political state and the behavior of the dictator, citizens decide whether to invest, i.e., $\hat{a}$ is determined.

6. Citizens produce output.

7. The tax rate τ is determined either by the ruler in dictatorship, or by the one-person-one-vote majority voting system in democracy. The tax revenue is collected and citizens receive the remaining output.

We assume that the tax rate is determined after production in order to reflect the idea that taxation is the key property of the dictatorship. The dictator has to tax the citizens because she holds all political power. Any promise to reduce the tax rate is incredible with regard to the citizens. This concept constitutes the basis of the democratization theory of Acemoglu (2000). However, we assume that the social transfer is paid to citizens before production, hence, it is credible. Thus, the prepaid social transfer gives the dictator an opportunity to become good.

We assume a perfect capital market with zero interest rate to finance all possible expenditures before production. With this crucial assumption the democratization process in the current model depends on the expected future income. The more the taxed income in dictatorship compared to that in democracy is, the greater the incentive to democratize is. Thus, the current model is consistent with the Lipset/Aristotle hypothesis. For simplicity, we assume that all debts should be cleared at the end of each period. The rest of income is eaten, thus, there is no saving.

All agents are risk neutral. Hence, utility can be measured by net income, which is totally consumed by agents within the period. Without taking the weapon expenditure into consideration, the net income of citizen a at the end of period t is:

$$Y_{at} = y_{at}(1-\tau_t) + s_{at} - I_{at}eA_t \qquad (2.2)$$

And the ruler's net income is:

$$Y_{ruler,t} = \int (\tau_t y_{at} - s_{at}) da \qquad (2.3)$$

In the following sections, we solve for the sub-game perfect equilibrium and analyze behavioral determinants of the dictator. According to backward induction, we first discuss the economic decision of agents and solve for the income of individuals in each political institution. In section 2.4, we study the political decision whether or not to revolt.

2.3 The Exogenous Growth Model

2.3.1 Dictatorship

The initial political institution is dictatorship. We assume that the technology level A_t grows exogenously. In section 2.4 we will know that the life time of the ruler depends on A_t (see equation 2.20 and 2.23), which is the single state variable in this simple model. Since this is a repeated game with finite periods and the tax rate is set after production, the dictator chooses $\tau^{dic} = \bar{\tau}$ regardless of whether she is good or bad. Although the good dictator is willing to encourage the citizen to invest, she cannot use the tax rate as the policy tool. As long as she holds all political power in her hand and taxation takes place after production, citizens are never convinced to invest by the promise of tax reduction.

The dictator selects a subset of the citizens and each will receive an amount s_t at time t. The ruler will make the social transfer if and only if the citizen will invest and thereby increase their taxable output. Clearly the dictator will not make a social transfer to non-investing citizens. Now define $s_t = A_t S$, where S is the steady state ratio of social transfer to technology level. We may treat S, rather than s_t, as the ruler's decision variable. Denote the ability level by $\hat{a}(S)$ for the citizen who is indifferent between

investing and not investing when the ruler uses the social transfer ratio S. For a citizen with ability a, $a > \hat{a}(S)$, then he will have income $A_t N \lambda (1 - \bar{\tau}) a + A_t S - e A_t$, after investing. If the citizen has ability a, $a < \hat{a}(S)$, he receives no social transfer and does not invest so has income $A_t N (1 - \bar{\tau}) a$. For the citizen with ability a, $a = \hat{a}(S)$, these two incomes must be equal: $A_t N \lambda (1 - \bar{\tau}) \hat{a}(S) + A_t S - e A_t = A_t N (1 - \bar{\tau}) \hat{a}(S)$. Solving for $\hat{a}(S)$ leads to:

$$\hat{a}(S) = \frac{e - S}{N(\lambda - 1)(1 - \bar{\tau})} \qquad (2.4)$$

After (2.4) is derived, suppose no transfer takes place ($S = 0$) to anyone. We want to assume that citizens with sufficiently high ability (near 1) will still invest. In other words, that means $\hat{a}(0) < 1$. Now we can see that $\hat{a}(0) < 1$ from (2.4) becomes inequality (A.2.1):

$$\frac{e}{N(\lambda - 1)} < 1 - \bar{\tau} \qquad (A.2.1)$$

This assumption states that the net benefit of investment for the individual with $a = 1$ ($N(\lambda - 1)(1 - \bar{\tau})$) is greater than the cost (e), even if he gets no transfer from the dictator. I.e., his net return rate of private investment $\frac{N(\lambda - 1)(1 - \bar{\tau})}{e}$ is greater than 1. With this assumption we avoid the corner solution, i.e., we always have $\hat{a}(S) < 1$ for all $S \geq 0$. The ruler chooses the optimal transfer S in order to maximize her income:

$$\max_{S} Y_{ruler,t} = \overline{\tau}\int_0^1 y_{at} da - \left(1 - \hat{a}(S)\right)A_t S \equiv A_t \widetilde{Y}_{ruler}$$

$$s.t. \quad \widetilde{Y}_{ruler} = \frac{1}{2}\overline{\tau}N\left[\lambda + (1-\lambda)\hat{a}(S)^2\right] - (1 - \hat{a}(S))S$$

Substitute (2.4) and recall the assumption that the social transfer is non-negative, we get S^{exg} from the first-order condition. The second-order condition for a maximum is satisfied.

$$S^{exg} = \begin{cases} S^{bad} = 0 & \text{if} \quad (1-\overline{\tau})^2 \geq \dfrac{e}{N(\lambda - 1)} \\[3ex] S^{good} = \dfrac{e - N(\lambda - 1)(1 - \overline{\tau})^2}{2 - \overline{\tau}} & \text{if} \quad (1-\overline{\tau})^2 < \dfrac{e}{N(\lambda - 1)} \end{cases} \qquad (2.5)$$

Proposition 2.1

If assumption (A.2.1) holds and A_t grows exogenously, the dictator will be bad if $(1-\overline{\tau})^2 \geq \dfrac{e}{N(\lambda - 1)}$; she will be good if $(1-\overline{\tau})^2 < \dfrac{e}{N(\lambda - 1)}$. The dictator is better the higher $\overline{\tau}$, the lower the level of natural resources and the lower the return rate of private investment.

As we assumed previously, the bad ruler consumes all tax income and sets the social transfer at $s_a = 0 \ \forall a$. $\hat{a}^{bad}$, reflecting this threshold in a bad dictatorship, equals to $\dfrac{e}{N(\lambda - 1)(1 - \overline{\tau})}$. Rearranging the condition of a good dictator $(1-\overline{\tau})^2 < \dfrac{e}{N(\lambda - 1)}$ and substituting from $\hat{a}^{bad}$, we have $1 - \hat{a}^{bad} < \overline{\tau}$. $1 - \hat{a}^{bad}$ is the investment ratio in the bad dictatorship, and $\overline{\tau}$ represents the highest tax rate. If private investment is not attractive to citizens, i.e., $1 - \hat{a}^{bad}$ is very low, the ruler has the incentive to be good thus

encouraging citizens to invest. As expected, if the highest tax rate declines, the ruler is less likely to be good. Because $1 - \hat{a}^{bad}$ strictly decreases in $\bar{\tau}$, we have a unique $\bar{\tau}^*$, so that $1 - \hat{a}^{bad}(\bar{\tau}^*) = \bar{\tau}^*$. For all $\bar{\tau} \leq \bar{\tau}^*$, the dictator is bad, and for all $\bar{\tau} > \bar{\tau}^*$ she is good. This result implies that the dictator wants to encourage private investment, if $\bar{\tau}$ is high enough, i.e., she can tax enough after production. For the dictator, the social transfer is similar to an investment, using the tax rate as her rate of return.

If the condition for being good is satisfied, the good dictator pays a positive social transfer S^{good} to the citizen who will invest. Substituting positive S^{good} in (2.4), we obtain:

$$\hat{a}^{good} = \frac{e + N(\lambda - 1)(1 - \bar{\tau})}{N(\lambda - 1)(2 - \bar{\tau})} \qquad (\,2.6\,)$$

The good dictator has two effects for citizens: first, the individual who invests can earn more due to the positive social transfer; second, the positive social transfer decreases the entry barrier investment, hence, more citizens will invest. (It is easy to see that $\hat{a}^{good} < \hat{a}^{bad}$) Of course the citizen who does not invest can not increase his income in the good dictatorship.

Proposition 2.2:

If condition $(1 - \bar{\tau})^2 < \dfrac{e}{N(\lambda - 1)}$ holds, the transition from the bad to the good dictatorship is a Pareto-improving process. More citizens invest, aggregate output increases and all agents obtain a higher or at least the same income.

The proposition is easy to prove, since S^{good} is the optimal choice for the ruler given $(1-\bar{\tau})^2 < \dfrac{e}{N(\lambda-1)}$, and citizens receive a positive social transfer from the ruler. The Pareto-improving process is achieved, because the transition ensures the income of the good dictator to exceed that of the bad dictator. The incomes of the ruler and citizens in the bad and good dictatorship in period t are given as follows, respectively:

$$Y^{bad}_{ruler} = \frac{1}{2}\bar{\tau}A_t N\left(\lambda - \frac{e^2}{N^2(\lambda-1)(1-\bar{\tau})^2}\right) \qquad (2.7)$$

$$Y^{bad}_a = \begin{cases} A_t N a(1-\bar{\tau}) & a < \hat{a}^{bad} \\ A_t N\lambda a(1-\bar{\tau}) - eA_t & a \geq \hat{a}^{bad} \end{cases} \qquad (2.8)$$

$$Y^{good}_{ruler} = \frac{A_t N\bar{\tau}}{2} + \frac{A_t(N(\lambda-1)-e)^2}{2N(\lambda-1)(2-\bar{\tau})} \qquad (2.9)$$

$$Y^{good}_a = \begin{cases} A_t N a(1-\bar{\tau}) & a < \hat{a}^{good} \\ A_t N\lambda a(1-\bar{\tau}) - eA_t + s^{good} & a \geq \hat{a}^{good} \end{cases} \qquad (2.10)$$

From (2.7) to (2.10) we can easily see that $Y^{good}_{ruler} > Y^{bad}_{ruler}$ and $Y^{good}_a \geq Y^{bad}_a$.

2.3.2 Democracy

In a democratic society, the tax rate is determined by all citizens through a "one-person-one-vote" majority voting system. The tax revenue is equally distributed to every citizen.[9] Hence, the median voter is the deciding person. He maximizes his income $Y_{0.5,t}$, subject to the budget constraint of redistribution:

---

[9] In Appendix 2.2, I will show the case where the tax revenue is only given to the individual who invests.

$$\underset{\tau}{Max}\ Y_{0.5,t} = 0.5 A_t N \lambda^{I_{0.5}}(1-\tau) + s - I_{0.5} e A_t$$

$$s.t. \quad s = \tau \int_0^1 y_{at} da = 0.5\tau A_t N(\hat{a}^2 + \lambda - \lambda\hat{a}^2)$$

There are two cases:

1. If $\hat{a} > 0.5$, i.e., the median voter doesn't invest. Hence, his maximization problem reduces to:

$$\underset{\tau}{Max}\ Y_{0.5,t} = 0.5 A_t N(1-\tau) + 0.5\tau A_t N(\hat{a}^2 + \lambda - \lambda\hat{a}^2)$$

The first order condition is:

$$\frac{\partial Y_{0.5,t}}{\partial \tau} = -0.5 A_t N + 0.5 A_t N(\hat{a}^2 + \lambda - \lambda\hat{a}^2) = 0.5 A_t N(\lambda-1)(1-\hat{a}^2) > 0$$

Hence $\tau^{dem,1} = \bar{\tau}$. In order to solve $\hat{a}^{dem,1}$, we have:

$$Y_{at}(invest) = Y_{at}(no\ invest) \Leftrightarrow A_t N \lambda \hat{a}(1-\bar{\tau}) - e A_t + s = A_t N \hat{a}(1-\bar{\tau}) + s$$

We get:
$$\hat{a}^{dem,1} = \frac{e}{N(\lambda-1)(1-\bar{\tau})} \qquad (\,2.11\,)$$

And:
$$s^{dem,1} = \frac{1}{2}\bar{\tau} A_t N\left(\lambda - \frac{e^2}{N^2(\lambda-1)(1-\bar{\tau})^2}\right) \qquad (\,2.12\,)$$

Hence,
$$Y_{at}^{dem,1} = \begin{cases} A_t N a(1-\bar{\tau}) + s^{dem,1} & a < \hat{a}^{dem,1} \\ A_t N \lambda a(1-\bar{\tau}) - e A_t + s^{dem,1} & a \geq \hat{a}^{dem,1} \end{cases} \qquad (\,2.13\,)$$

If condition $\hat{a}^{dem,1} = \dfrac{e}{N(\lambda-1)(1-\bar{\tau})} > \dfrac{1}{2}$ holds, democracy decreases inequality (comparing (2.13) and (2.10)). However, the aggregate output is the same as that of the bad dictatorship.

2. If $\hat{a} \le 0.5$, the median voter invests. His maximization problem then becomes:

$$\underset{\tau}{Max}\, Y_{0.5,t} = 0.5A_t N\lambda(1-\tau) + 0.5\tau A_t N(\hat{a}^2 + \lambda - \lambda\hat{a}^2) - eA_t$$

The first order condition is:

$$\dfrac{\partial Y_{0.5}}{\partial \tau} = -0.5A_t N\lambda + 0.5A_t N(\hat{a}^2 + \lambda - \lambda\hat{a}^2) = 0.5A_t N(1-\lambda)\hat{a}^2 < 0$$

Hence, $\tau^{dem,2} = 0$ and $s^{dem,2} = 0$.

We get: $\qquad \hat{a}^{dem,2} = \dfrac{e}{N(\lambda-1)}$ $\qquad\qquad\qquad$ (2.14)

$$Y_{at}^{dem,2} = \begin{cases} A_t Na & a < \hat{a}^{dem,2} \\ A_t N\lambda a - eA_t & a \ge \hat{a}^{dem,2} \end{cases}$$ $\qquad$ (2.15)

The aggregate output is as follows: $Y_t^{dem,2} = A_t\left(\dfrac{N\lambda}{2} - e + \dfrac{e^2}{2N(\lambda-1)}\right)$ (2.16)

If condition $\hat{a}^{dem,2} = \dfrac{e}{N(\lambda-1)} \le \dfrac{1}{2}$ holds, democracy is capable of increasing aggregate output. This is because the tax rate is set at the lowest level. Individuals are encouraged to invest.

The tax rate and the investment ratio in the democratic society depend on the behavior of the median voter. If he finds that it is not worth investing (this is the case $\frac{e}{N(\lambda-1)} > \frac{1}{2}$), then he supports a higher tax rate (here, $\tau^{dem,1} = \bar{\tau}$). Therefore, the democratic society suffers also from a lower investment ratio, which is as same as in the bad dictatorship. If $\frac{e}{N(\lambda-1)} \leq \frac{1-\bar{\tau}}{2}$, i.e., the investment is attractive enough to the median voter, then he would support a lower tax rate (here, $\tau = 0$). Consequently, the economy enjoys a higher output level due to a higher investment ratio. If $\frac{1-\bar{\tau}}{2} < \frac{e}{N(\lambda-1)} \leq \frac{1}{2}$, the median voter has two choices. Whether the investment is worthwhile to implement depends on his choice of the tax rate. If he decides to support a higher tax rate after production, he also knows that the investment is worthless to undertake. Hence, he chooses not to invest before production. All other citizens observe his investment choice and expect that he will support a higher tax rate after production. Hence, the investment ratio is at the lower level. Vice versa, if he would like to invest, then he must choose a lower tax rate after investment. Thus, two possible investment ratios and redistribution schemes could be achieved: $(\hat{a}^{dem,1}, \bar{\tau}, s^{dem,1})$, $(\hat{a}^{dem,2}, 0, 0)$. Which one is actually chosen by the median voter depends on the parameter constellation.

Proposition 2.3:

1. *If $\frac{e}{N(\lambda-1)} \leq \frac{1-\bar{\tau}}{2}$, democracy can increase aggregate output, and if*

   $\frac{e}{N(\lambda-1)} > \frac{1}{2}$ democracy can only decrease inequality, but cannot

   increase aggregate output.

2. *In the moderate case of* $\frac{1-\bar{\tau}}{2} < \frac{e}{N(\lambda-1)} \le \frac{1}{2}$, *the impact of democracy is ambiguous where two possibilities exist:* $(\hat{a}^{dem,1}, \bar{\tau}, s^{dem,1})$ *and* $(\hat{a}^{dem,2}, 0, 0)$. *The median voter is willing to choose* $(\hat{a}^{dem,2}, 0, 0)$, *if* $\bar{\tau} \ge \frac{1}{2}$.

Proof: The first part is already clear. The second part is easy to see, if we compare the incomes of the median voter in two cases. He will choose $(\hat{a}^{dem,2}, 0, 0)$, if it generates higher income for him. I.e.,

$$Y_{0.5}^{dem,2} - Y_{0.5}^{dem,1} \ge 0 \Leftrightarrow \bar{\tau} \ge \frac{2\hat{a}^{dem,1} - 1}{(1+\hat{a}^{dem,1})^2 - 2}.$$ Unfortunately, $\frac{2\hat{a}^{dem,1} - 1}{(1+\hat{a}^{dem,1})^2 - 2}$

depends on $\bar{\tau}$. Hence, the economic meaning of this condition is not very intuitive. However, notice that $\frac{2\hat{a}^{dem,1} - 1}{(1+\hat{a}^{dem,1})^2 - 2} < \frac{1}{2}$. Thus, the sufficient

condition is $\bar{\tau} \ge \frac{1}{2}$, i.e., the median voter will choose $(\hat{a}^{dem,2}, 0, 0)$ if the highest tax rate is high enough. Q.E.D.

The existence of multiple effects is consistent with the literature of political economy, which emphasizes the different effects of democracy on the economic performance. By assuming that the majority in a democracy is poor, this literature often argues that democracy hinders economic investment due to a higher level of redistribution. On the other hand, democracy also protects against taxation through the strong rule of law, which is good for economic performance. In the current model we argue that both could occur under different circumstances. The case of $(\hat{a}^{dem,2}, 0, 0)$ indicates the positive impact of democracy on economic performance, because democracy protects private investments from expropriative

taxation. On the other hand, the case of $(\hat{a}^{dem,1}, \bar{\tau}, s^{dem,1})$ reflects the negative impact of democracy on economic growth owing to the high tax rate. However, this negative effect has a different economic meaning compared to that of the bad dictatorship $(\hat{a}^{bad}, \bar{\tau}, 0)$. Whereas the former is pure redistribution, the latter is pure expropriation. Proposition 3 shows that which case occurs in the moderate case depends on parameters, in particular, the highest tax rate, $\bar{\tau}$. It reflects to what extent the political power is able to influence economic performance. If it is large enough $(\bar{\tau} \geq \frac{1}{2})$, individuals try to avoid redistribution and choose the lower tax rate. Hence, democracy has an aggregate effect on economic performance. For our purpose, it is more interesting to restrict attention to this case, i.e. $(\hat{a}^{dem,2}, 0, 0)$. Hence, we assume $\frac{e}{N(\lambda-1)} \leq \frac{1}{2}$ and $\bar{\tau} \geq \frac{1}{2}$ for simplicity.

Combining the above (A.2.1), we need to make the following assumption:

$$\frac{e}{N(\lambda-1)} \leq 1 - \bar{\tau} \leq \frac{1}{2} \tag{A.2.2}$$

We focus on the case where democracy has an aggregate effect on economic performance, because only in this case democratization is possible. The pure redistributive democracy $(\hat{a}^{dem,1}, \bar{\tau}, s^{dem,1})$ means that the expenditure of the ruler on weapons is more than that of the citizen net of the democratization cost. (For more details, see section 2.4) Hence, such "democratization" is impossible.

Combining the condition $(1-\bar{\tau})^2 < \frac{e}{N(\lambda-1)}$ and Assumption (A.2.2), we have:

The "goodness" ASSUMPTION: $\quad (1-\bar{\tau})^2 < \dfrac{e}{N(\lambda-1)} \le 1-\bar{\tau} \le \dfrac{1}{2} \quad$ (A.2.3)

This assumption is the sufficient condition of a good democracy in the sense that it has the aggregate effect on economic performance, and it also constitutes the condition of a good dictatorship. That is why we call it the "goodness" assumption. Since $\hat{a}^{dem,2} < \hat{a}^{good}$, the good democracy leads to a better economic performance than the good dictatorship. However, democratization is a social conflict, while the transition from the bad dictatorship to the good one is Pareto-improving.

2.4   Democratization

In the present paper the process of democratization is modeled as a two stage sequential game with perfect information. First the citizen decides whether to revolt, then the ruler decides whether to repress. Both revolution and repression require weapons. The citizen attempts to undertake a revolution, if he expects a higher level of income could be earned in a democratic society. Hence, if necessary, the citizen will offer the difference of his income in two political states as the highest payment for the weapon. Similarly, the dictator is willing to use her whole income to prevent the possible political transition, because she will lose all in the democratic society. Although both are willing to offer the whole life-time income, they cannot do so because we assume that the perfect financial market acts well only within one period. This assumption simplifies the analysis without loss of generality. Moreover, there is a revolution cost c for citizens. Hence, citizens don't invest in weapons if they expect their ruler to invest more than their highest payments in weapons net of the revolutionary cost. If they find that the ruler's income is lower than their highest payment for weapons net

of the revolution cost, their best choice is to invest in weapons a little more than the ruler's income. Thus we only need to compare the highest payments of both players for weapons, which are named the incentive of political transition. Revolution is the best choice for the citizen if and only if the citizens' incentive to revolt net of the revolutionary cost is higher than the incentive of the ruler to repress. For simplicity, we assume that the citizen will choose revolution when both are equal.

We will consider two possible democratization processes: from the bad dictatorship directly to democracy, and from the bad dictatorship indirectly to democracy via the good dictatorship.

2.4.1 The Incentive of Political Transition in the Bad Dictatorship

The highest payment of citizen a in period t is the difference between incomes in the bad dictatorship and the democratic society within t:

$$P_{at}^{bad} = \begin{cases} A_t N \lambda a \bar{\tau} & a \geq \hat{a}^{bad} \\ A_t N(\lambda-1)a - eA_t + A_t N a \bar{\tau} & a \in (\hat{a}^{dem}, \hat{a}^{bad}) \\ A_t N a \bar{\tau} & a \leq \hat{a}^{dem} \end{cases} \quad (2.17)$$

The first part ($A_t N \lambda a \bar{\tau}$) is the taxed income of the citizen who invests in both political states. The second difference of incomes ($A_t N(\lambda-1)a - eA_t + A_t N a \bar{\tau}$) comes from the citizen who invests in democracy but not in the bad dictatorship. The benefit of democracy for this group of citizens comes from two sides: the release of the expropriating taxation ($A_t N a \bar{\tau}$), and the investment return ($A_t N(\lambda-1)a - eA_t$). Finally,

the citizen, who invests neither in democracy nor the bad dictatorship, saves the tax in democracy ($A_t N a \bar{\tau}$). The sum of individual offers net of the revolutionary cost is the citizens' highest net expenditure on weapons.

$$P_{citizen,t}^{bad} = \int_0^1 P_{at}^{bad} da - c = \frac{A_t N \lambda \bar{\tau}}{2} - \frac{A_t e^2 \bar{\tau}}{2N(\lambda-1)(1-\bar{\tau})} - c \qquad (\,2.18\,)$$

For the ruler: $\qquad P_{ruler,t}^{bad} = Y_{ruler,t}^{bad} = \frac{1}{2}\bar{\tau} A_t N \left(\lambda - \frac{e^2}{N^2(\lambda-1)(1-\bar{\tau})^2} \right) \qquad (\,2.19\,)$

The difference of payments between the citizen and the dictator determines whether the revolution will succeed:

$$\Delta_t^{bad} = P_{citizen,t}^{bad} - P_{ruler,t}^{bad} = \frac{A_t e^2 \bar{\tau}^2}{2N(\lambda-1)(1-\bar{\tau})^2} - c \qquad (\,2.20\,)$$

If $\Delta_t^{bad} \geq 0$, the aggregate highest payment of citizens exceeds that of the ruler. Hence, citizens choose revolution and expend a little more on weapons than the highest payment of the ruler. The ruler knows the repression will not be successful, thus, the actual repression does not occur. If $\Delta_t^{bad} < 0$, citizens know that the revolution will be repressed, hence, they don't choose to revolt. We assume the society begins from the non-democracy. Hence, at the beginning period ($t = 1$), Δ_1^{bad} is negative. We have the following assumption:

The "status quo" ASSUMPTION: $\qquad \frac{A_1 e^2 \bar{\tau}^2}{2N(\lambda-1)(1-\bar{\tau})^2} < c \qquad (\,A.2.4\,)$

Equation (2.20) has the following important indications. First, the higher $\bar{\tau}$, the greater the incentive for citizens to seek democratization. Second, as most of the political economy literature argues, e.g. the Lipset/Aristotle hypothesis, democracy follows the good economic performance. Here, the economic growth rate is given by the exogenous growth rate of the aggregate technology level A_t. With A_t growing, the benefit from revolution increases. Third, the effects of the investment project on the incentive of democratization is demonstrated by the parameters N and λ, e. The more beneficial the project (i.e. the lower e and/or the higher λ and N), the lower the incentive to democratize. The first part of equation (2.20) is from the investment return of the "middle class", who invests in democracy but not in dictatorship, i.e., $A_t \int_{\hat{a}^{dem}}^{\hat{a}^{bad}} (N(\lambda-1)a - e)da$. The citizen of "middle class" has a higher incentive to revolt, if λ and N increases and/or e declines. However, the size ($\hat{a}^{bad} - \hat{a}^{dem}$) of this group decreases in N and λ. The more beneficial the investment project, the smaller the aggregate effect of democracy. Hence, the net social incentive (Δ_t^{bad}) decreases. This relationship between economic performance and political transition is possibly supported by the fact that oil impedes democratization (e.g., Ross 2001). In this framework, we can argue that a country's oil wealth increases the average return rate of the private investment ($\frac{N(\lambda-1)}{2e}$). Hence, the size of middle class shrinks. Such societies have a lower incentive to democratize.

Proposition 2.4:

In the bad dictatorship, the incentive of democratization increases in the technology level A_t, and decreases in the natural resource N. The higher the taxation level $\bar{\tau}$, the greater is the incentive of revolution. The net social

incentive of democratization decreases in the return of the investment project and increases in its cost e.

Comparing this result to Proposition 2.1, we find that the impact of $\bar{\tau}$, N and the return rate of the private investment on democratization is similar to that on the behavior of the bad dictator. If the highest tax rate increases, the bad dictator faces an increasing risk of revolution according to Proposition 2.4, and intuitively, she also has a larger incentive to become good according to Proposition 2.1.

2.4.2 The Incentive of Political Transition in the Good Dictatorship

For the good dictator the positive social transfer increases tax revenues. Hence, she also has more incentives to prevent the revolution than the bad dictator:

$$P_{ruler,t}^{good} = \frac{A_t N\bar{\tau}}{2} + \frac{A_t(N(\lambda-1)-e)^2}{2N(\lambda-1)(2-\bar{\tau})} \qquad (2.21)$$

The democratization incentive of citizens is as follows:

$$P_{at}^{good} = \begin{cases} A_t N\lambda a\bar{\tau} - s^{good} & a \geq \hat{a}^{good} \\ A_t N(\lambda-1)a - eA_t + A_t Na\bar{\tau} & a \in (\hat{a}^{dem}, \hat{a}^{good}) \\ A_t Na\bar{\tau} & a \leq \hat{a}^{dem} \end{cases} \qquad (2.22)$$

The poor who don't invest in both political states suffer the expropriative taxation in the dictatorship. Hence, he prefers to undertake revolution. Here, we model this as a positive payment $A_t Na\bar{\tau}$ for weapons. For the middle

class $a \in (\hat{a}^{dem}, \hat{a}^{good})$ who invest in democracy but not in the good dictatorship, their payment for weapons is positive, i.e., $P_{at}^{good} = A_t[N(\lambda - 1 + \bar{\tau})a - e] > A_t[N(\lambda - 1 + \bar{\tau})\hat{a}^{dem} - e] > 0$. They support democratization, because they can earn more in democratic society.

However, *a priori,* it is unclear whether the rich, who invest both in the good dictatorship and democracy, support democracy or not. If their payment for political transition $A_t N \lambda a \bar{\tau} - s^{good}$ is negative, they can earn more in the good dictatorship and become the supporter of this political institution.

Proposition 2.5:

The citizen with the highest ability 1 always supports democracy, whereas some of the rich, who invest both in the good dictatorship and democracy, could support the dictatorship under certain conditions.

Proof: see Appendix 2.1.

This Proposition indicates that the dictator can extend the social support of the regime by means of a positive social transfer. Surprisingly, the group which possibly supports the regime is not the one with the highest ability, but a group with a relatively lower ability, although their ability great enough to let them invest in both dictatorship and democracy. In this sense, the "top rich" do not sympathize with the good dictator.

Again, $P_{citizen,t}^{good} = \int_0^1 P_{at}^{good} da - c$ and the net social incentive of democratization of the whole society is:

$$\Delta_t^{good} = P_{citizen,t}^{good} - P_{ruler,t}^{good}$$

$$= A_t \frac{(N(\lambda-1)-e)^2(1-\bar{\tau})^2}{2N(\lambda-1)(2-\bar{\tau})^2} - c \qquad (2.23)$$

Proposition 2.6:

1) In the good dictatorship, the incentive of democratization increases in the aggregate technology level. The higher the taxation level, the less the incentive of revolution is. The net social incentive of democratization increases in natural resources and the return of the investment project and decreases in its cost.

2) Because of Pareto-improving social transfer the incentive of democratization in the good dictatorship is lower than in the bad one.

Proof: 1) It is clear that $\frac{\partial \Delta_t^{good}}{\partial A_t} > 0$, $\frac{\partial \Delta_t^{good}}{\partial \bar{\tau}} < 0$, $\frac{\partial \Delta_t^{good}}{\partial N} > 0$, $\frac{\partial \Delta_t^{good}}{\partial \lambda} > 0$,

$\frac{\partial \Delta_t^{good}}{\partial e} < 0$.

2) $\Delta_t^{bad} = P_{citizen,t}^{bad} - P_{ruler,t}^{bad} = \int_0^1 (Y_{at}^{dem} - Y_{at}^{bad})da - c - Y_{ruler,t}^{bad}$

$\Delta_t^{good} = P_{citizen,t}^{good} - P_{ruler,t}^{good} = \int_0^1 (Y_{at}^{dem} - Y_{at}^{good})da - c - Y_{ruler,t}^{good}$

$\Delta_t^{bad} - \Delta_t^{good} = (\int_0^1 Y_{at}^{good}da - \int_0^1 Y_{at}^{bad}da) + \left[Y_{ruler,t}^{good} - Y_{ruler,t}^{bad} \right] > 0$ Q.E.D.

Comparing to Proposition 2.4, it is of interest to see that the effects of investment and the tax rate on the incentive to revolt differ between the bad and good dictatorship. Analogously, the first term of (2.23) is also from the investment return of the "middle class", i.e., $A_t \int_{\hat{a}^{dem}}^{\hat{a}^{good}} (N(\lambda-1)a-e)da$. The size ($\hat{a}^{good} - \hat{a}^{dem}$) of this group increases, if N and λ increases and/or e declines. Hence, the net social incentive (Δ_t^{good}) increases. In other words,

this model predicts that natural resources accelerate democratization in the good dictatorship. This, however, requires future empirical evidence. In the good dictatorship, taxation is the mixture of redistribution and expropriation. The increase of the highest tax rate implies that social support of the dictatorship could widen. Hence, the incentive for democratization declines.

Proposition 2.6 strengthens Proposition 2.1. An increase in the highest tax rate gives rise to a higher incentive for the dictator to be good, because she can tax more. Furthermore, the good dictator faces a smaller danger of revolution if the highest tax rate increases. Analogously, if the private investment is more profitable, the dictator has less incentive to be good, and the good dictator faces a larger possibility of revolution.

Improvement of the citizen's income due to the positive social transfer decreases their incentive to change the political state, whereas the good dictator resists the democratization more than the bad one because of the higher economic benefit. Hence, given the technology level, we argue that the opportunity of democratization decreases in the economic performance during the transition from the bad dictatorship to the good. However, it does not directly contradict the Lipset/Aristotle hypothesis. As we have seen, if the economy grows with the technology level, the society has a higher incentive to become a democracy. In the following section, we consider the external effect of the individual's investment on the aggregate technology level and demonstrate that the technology progress enlarges the income difference between dictatorship and democracy.

2.5   External Effect and Endogenous Growth

So far we have assumed that the aggregate technology level, as well as the long-run economic growth, is given exogenously. The dictator is good if she finds that the positive social transfer can increase her instantaneous income. In other words, we have assumed that the behavior of the dictator can affect short run economic performance, but not long-run economic growth. Now we introduce endogenous technological progress to our simple model. As is standard in endogenous growth theory,[10] the aggregate technology level and, in turn, the economic growth rate, increases in the investment ratio $1-\hat{a}$. We assume for simplicity that private investment has a positive externality on the aggregate technology level, i.e., $A_t = A_{t-1}(1+G(\hat{a}_{t-1}))$, where $G(\hat{a}_{t-1})$ is the growth rate of the aggregate technology level, $G'(\hat{a}_{t-1}) < 0$. Because of (2.4), we know $\hat{a}_t = \dfrac{e-S_t}{N(\lambda-1)(1-\overline{\tau})}$. Hence, the growth rate of A_t is the increasing function of the social transfer in period $t-1$, denoted by $G(\hat{a}_{t-1}(S_{t-1})) \equiv g(S_{t-1})$, where $g'(S_{t-1}) > 0$. This is the single linkage across periods. According to the assumption that financial markets are perfect only within a period, no income can be transferred across periods. From equations (2.20) and (2.23) we know that the higher growth rate of technology level leads to a sooner political transition. Hence, there could be a tradeoff for the ruler between a greater benefit in the short run and relatively faster democratization in the long-run. From now on, we standardize $N=1$ for simplicity.

---

[10] There are two main approaches to model the role of human capital in economic growth: Lucas (1988) emphasizes the externality of human capital in production; Nelson and Phelps (1966), Romer (1990), Grossman and Helpman (1991), and Aghion and Howitt (1992) argue that the human capital will induce more innovation or let the economy accept new technology.

As the growth rate is endogenous, all individuals know the life-time of the dictator, which is the first period with a non-negative Δ_t. The dictator sets the tax rate on $\bar{\tau}$. As we know from (2.20) and (2.23),

$$\Delta_t = A_t \int_{\hat{a}^{dem}}^{\hat{a}^{dic}} ((\lambda - 1)a - e)da - c \qquad (2.24)$$

In order to make the analysis tractable, we consider a three-period model in this section. We assume that the revolution takes place in the third period. According to (2.24), it implies $\Delta_3 > 0$ for any S_1 and S_2. Hence, the sufficient and necessary condition for $\Delta_3 > 0$ is:

$$(1 + g(0))^2 \int_{\hat{a}^{dem}}^{\hat{a}^{dic}} ((\lambda - 1)a - e)da > \frac{c}{A_1} \qquad (A.2.5)$$

where $\hat{a}^{dic} = \dfrac{e - S^{exg}}{(\lambda - 1)(1 - \bar{\tau})}$ and S^{exg} is the optimal social transfer in the exogenous growth model, as shown in (2.5). This condition means that revolution will take place in the third period, even if the ruler sets the social transfer at the lowest level (i.e., $S = 0$) in the first two periods. Hence, the dictator knows that the second period is her last period. Then she acts the same as in the exogenous growth model, i.e., she maximizes her instantaneous income. Thus, $S_2 = S^{exg}$. What we want to show here is the social transfer in the first period S_1. This is the social transfer in the endogenous growth model. In period 1, the dictator is aware of two effects of her social transfer policy. First, her transfer can encourage more citizens to invest, and in turn, increase her income in period 1. Secondly, more investment implies the higher technology level in the second period, and in turn, will render the revolution more likely in period 2. If the revolution

takes place in the second period, then the first period is the last period for the dictator. Hence, the life-time income of the ruler is given as follows:

$$V_1 = \begin{cases} Y_{ruler}(S_1) + \rho Y_{ruler}(A_2(S_1), S^{exg}) & if \quad \Delta_2(S_1) \le 0 \\ Y_{ruler}(S_1) & if \quad \Delta_2(S_1) > 0 \end{cases} \qquad (\,2.25\,)$$

where $0 \le \rho \le 1$ is the discount factor. We define a threshold value S_1^r so that $\Delta_2(S_1^r) = 0$, i.e.,

$$(1 + g(S_1^r)) \int_{\hat{a}^{dem}}^{\hat{a}^{dic}} ((\lambda - 1)a - e)da = \frac{c}{A_1} \qquad (\,2.26\,)$$

Hence, for all $S_1 > S_1^r$, revolution occur in the second period. The ruler knows that. Hence, she chooses S^{exg} in the first period. For all $S_1 \le S_1^r$, the dictator can live for two periods. Hence, she chooses $S_1^* = \arg\max V_1 = \arg\max Y_{ruler,1}(S_1) + \rho Y_{ruler,2}(A_2(S_1), S^{exg})$, subject to $S_1 \le S_1^r$. We define $\hat{S}_1$ as the unconstrained optimal social transfer, so that $\frac{d\tilde{Y}_{ruler}}{dS_1}\Big|_{S_1 = \hat{S}_1} + \frac{dg}{dS_1}\Big|_{S_1 = \hat{S}_1} \rho \tilde{Y}_{ruler}(S^{exg}) = 0$. Sum up, $S_1^* = \min\{S_1^r, \hat{S}_1\}$. Because S^{exg} is the optimal social transfer in the exogenous growth model, we have $\frac{d\tilde{Y}_{ruler}}{dS_1}\Big|_{S_1 = S^{exg}} = 0$. Hence, $\hat{S}_1 > S^{exg}$. We define $\tilde{S}_1$ so that $\tilde{Y}(\tilde{S}_1) = (1 - \rho - \rho g(\tilde{S}_1))\tilde{Y}(S^{exg})$. Hence, if $S_1^r < \tilde{S}_1$, then the dictator sets $S_1 = S^{exg}$ and lives for one period. This social transfer decision of the dictator in the endogenous growth model is shown in Figure 2.1 and summarized in Proposition 2.7.

$$V_1$$

$$Y_{ruler}(S_1) + \rho Y_{ruler}(A_2(S_1), S^{exg})$$

$$Y_{ruler}(S_1)$$

$$\widetilde{S}_1 \quad S_1^r \quad S^{exg} \quad \hat{S}_1 \qquad\qquad S_1$$
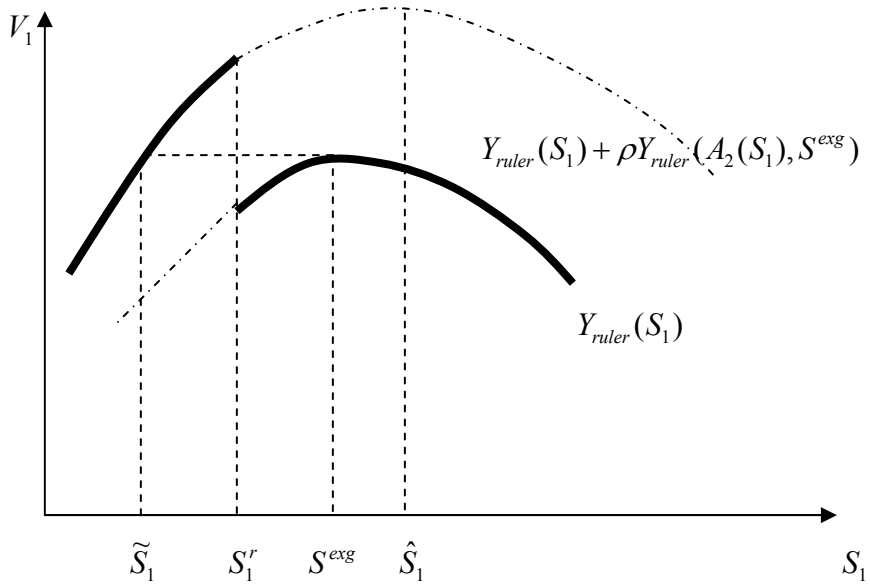
Figure 2.1 Social transfer in relation to the life-time income

of the dictator

Proposition 2.7

In the endogenous growth model, the dictator chooses the social transfer as follows:

1. *In the last period of her life-time, the dictator acts the same as in the exogenous growth model.*

2. *In the period before, the dictator sets $S^* = \min\{S^r, \hat{S}\}$. S^r increases in the revolution cost c and decreases in the initial technology level A_1. $S^* = \min\{S^r, \hat{S}\}$ could be smaller than S^{exg}.*

The effect of $\dfrac{c}{A_1}$ on the social transfer S_1 is non-linear (Figure 2.2).

Assuming a sufficiently small value of $\dfrac{c}{A_1}$, i.e.,

$\dfrac{c}{A_1} < (1 + g(0)) \int_{\hat{a}^{dem}}^{\hat{a}^{dic}} ((\lambda - 1)a - e)\,da$, the first period is the last period for the

46

dictator, thus, $S_1 = S^{exg}$. When $\dfrac{c}{A_1}$ exceeds this threshold value, the dictator could live for two periods. However, she isn't willing to live such a long time as long as $S_1^r < \widetilde{S}_1$, because living for two periods implies that she has to set the social transfer so low that her life-time income of two periods is even smaller than that of one period. In this case, the dictator would like to transfer more to citizens, although it leads to a sooner revolution. When $\dfrac{c}{A_1}$ increases further so that $S_1^r \geq \widetilde{S}_1$, the ruler can and is willing to live for two periods. Then she chooses $S^* = \min\{S^r, \hat{S}\}$. However, it does not mean directly that her social transfer is greater than S^{exg}. Whether S^* is greater than S^{exg} depends on $\dfrac{c}{A_1}$. If $\dfrac{c}{A_1}$ is not so big, S^r could be smaller than S^{exg}.
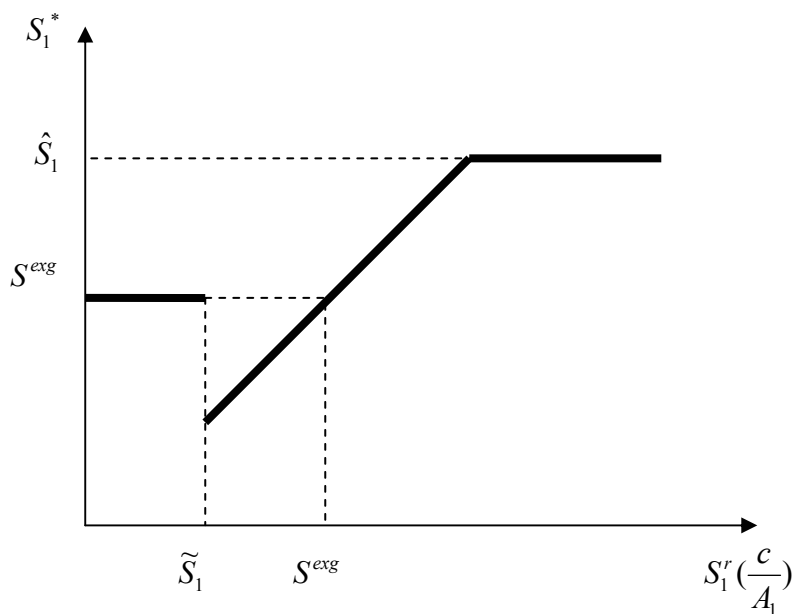
Figure 2.2: Effect of S_1^r, $\dfrac{c}{A_1}$ on optimal social transfers

According to McGuire and Olson (1996), the longer the ruler's life-time, the higher is her incentive to be good. Here, we show that it is also possible for the ruler to be worse, when her life-time increases. The reason is that she wants to keep her longer life-time, and is concerned more with the negative effect of her social transfer policy in the long-run. This relationship is shown in the Figure 2.2.

2.6 Summary

In the current paper we discussed the determinants of the dictator's incentive to be good in the sense that she would like to share the tax income with certain citizens. We emphasized two important effects of private investment in production: the individual effect which improves private output, and the positive externality on the aggregate technological level. We find that the dictator is more likely to be good if the individual faces a less profitable investment project. The dictator's incentive to be good is to tax more through encouraging citizens to invest more. Possible evidence is the gradual process of Chinese reform, in which regions and sectors reform one after another. For the local government, the investment from the central government in its region could be seen as a "natural resource", because the local government can use it free of charge. The less it is, the lower is the return rate of private investment, in turn, the lower the investment ratio. Hence, the local government, who is far away from the economic center in the old system, has a higher incentive to encourage private investment. Chinese reform began from the agricultural sector, where the central government invested nothing in the command economy. Moreover, the agriculture reform began from the poor province, Anhui. The nowadays fast growing provinces, e.g., Zhejiang, Fujian, Guangdong, are all less developed areas in the old system. Northeast China, where is the economic

center of the old system and attracted the most investment from the central government, is in recession now. Our finding does not directly contradict the study of Laffont and Qian (1999), where they argued that the necessary condition of reform in one sector is that the private return of investment in this sector is large enough to compensate the rent of government in the old system. We argue that the ruler would prefer to choose the sector with a lower private rate of return, if there are several sectors satisfying the necessary condition.

After endogenizing the growth rate, we find two different effects of economic performance on democratization. The good dictatorship is capable of reducing the incentive of a revolution through increasing the citizens' investment ratio and their income, but it is also possible to lead to an earlier democratization given higher economic growth rates. The effect of the revolutionary costs on the behavior of the ruler is non-linear. As a consequence long life-time does not always lead to a good dictator.

Chapter 3

Education, Income Distribution and Innovation

3.1 Introduction

The relationship between a country's income distribution and its economic growth is a permanent topic which sparks debates not only among economists but also policy-makers. In the last fifteen years, most cross-country studies (e.g., Berg and Sachs 1987, Persson and Tabellini 1994, Alesina and Rodrik 1994, Clarke 1995) show that if there is a relationship at all, inequality has a negative impact on long-run growth rates. Nonetheless, there also is evidence that inequality has a positive impact on short or medium run growth rates (Forbes 2000), or that the relationship between the income distribution and the long-run growth rate is non-linear (Chen 2003, Banerjee and Duflo 2003). In this article we provide a theoretical model to shed light on the ambiguous relationship between income distribution and the economic growth.[11] We argue that inequality, which is measured by the

---

[11] There are many different theoretical models to explain these different empirical results. E.g., Bénabou (1996) summarized three points of view to explain the negative impact of inequality on growth. Bénabou (2002) provides a model to illustrate the non-linear relationship between redistribution and growth.

Gini-coefficient, includes many variables, which may have a different impact on the economic growth.

For simplicity, we assume that there are two types of consumers, poor people and rich people. The income distribution can be measured by two variables, the income of the poor relative to the average income, and the population share of the poor. Both an increase in the relative income of the poor and a decline in the population share of the poor indicate a decrease in inequality. The minimal wage level, social insurance and so on could be considered as policies to improve the income of the poor, whereas mandatory education is easily understood as one to reduce the population share of the poor. If they have a different impact on growth, above cross-country evidence, which is based on the simple regression of the Gini-coefficient on the economic growth rate, could be ambiguous. In particular, we may be unable to draw from such simple empirical studies recommendations on redistribution policies for achieving a higher economic growth rate as well as a more equal income distribution.

We discuss the impact of the income distribution on the firms' profits in a vertically differentiated goods market in a model originally introduced by Shaked and Sutton (1982, 1983). However, our analysis focuses on the general equilibrium, whereas those papers are interested in issues of competition in a partial equilibrium framework. Rich consumers can afford more high quality goods than the poor and are willing to pay more for them. Hence, firms supply different qualities to different consumers in order to reduce the competitive pressure on prices. Therefore, the firms' profit depends on the income distribution.

Second, economic growth is achieved through innovation. The high quality good is firstly invented, and then produced by oligopolists. Innovation is

assumed to follow a Poisson process. An inventor can increase the Poisson arrival rate through employing more workers. The inventor's incentive to innovate depends on the profit of production after taking the cost of innovating into account. Hence, the income distribution can affect innovation through profits. If we consider the pooling case, where the oligopolistic market reduces to a monopolistic one, we are back to the case of Aghion and Howitt (1992).

Inequality may give rise to quality differentiation and a higher incentive for firms to innovate because rich consumers can pay more for high quality goods than the poor. But on the other side, the relatively small market share of high quality goods implied by inequality impedes the spread of better quality goods. Hence, the effect of the income distribution on innovation is *a priori* unclear. We consider an overlapping-generations economy, where individuals live for two periods: young and old. They decide whether to undergo education when young, and how to consume when old. We assume that individuals can become rich through education. If we increase the relative income of the poor, individuals have less of an incentive to undergo education. Hence, the population share of the poor increases. It reflects the idea of "social mobility" in political economy, which describes the movement of individuals among different income classes.

Assuming interdependence between relative income and population share seems realistic. The improvement of relative income of the rich may increase the incentive of the poor to become the rich. Of course, we can also argue that this incentive will decrease if the poor find that the rich are so rich that they can't catch up. If there are suitable channels in our society through which the poor are able to become the rich, for example, through education, immigration, or winning a lottery, we may find that the

population share of the poor is an endogenous variable given exogenous relative income of the poor.

This paper shows that interdependence between relative income and population share is crucial in the study of the impact of inequality on innovation. The main results of this paper are that there is a separating equilibrium, in which the high quality good is sold only to the rich and the low quality good only to the poor. In this equilibrium, a lower relative income of the poor is good for innovation, and a larger population share of the poor is bad for innovation. This result is consistent with Foellmi and Zweimüller (2002). But there they introduce hierarchic preferences[12], and the innovation induces new goods but not the improvement of quality. The result of Zweimüller and Brunner (2005), that the redistribution from the rich to the poor raises the innovation rate, does not hold under the assumption of interdependence.

Since we focus on the impact of the income distribution on the demand for quality goods, the labour market and production are assumed as simple as possible: the labour force is the single production factor, which is allocated among the production sectors and the research activity. Everybody inelastically supplies one unit of labour. This assumption generates another novel result that education enrollment is always positively associated with the innovation rate, although we have not assumed that education can increase productivity. Some might doubt why the educated individual can earn more than the non-educated, although they supply the same labour unit. In literature (e.g., Glomm et al. 1992, Croix et al. 2003), economists focus on the impact of education on economic growth through increasing productivity and/or human capital. However, there also are arguments that

---

[12] "A hierarchy of wants implies that goods can be ranked according to their priority in consumption" (Foellmi and Zweimüller 2002)

education does not necessarily produce human capital, e.g., the signalling education. Bils et al. (2000) show the empirical evidence that the positive relationship between education enrolment rate and the following economic growth rate can not be explained by the human capital theory. The current paper concentrates on the demand side effect of income distribution. I.e., we investigate the effect of income distribution on the consumption expenditure. Hence, it is not necessary to consider the effect of initial income distribution on the supply of production factor in this paper. In other words, we shed light on how rich consumers spend their income, but neglect why they are rich. We argue that education generates not only higher productivity, but also richer consumers. The signaling education is able to influence economic growth through the demand for better quality. This is almost neglected by most economic studies.

The paper is organized as follows. In section 3.2, the set-up of the model is introduced. In section 3.3 and 3.4, we study equilibrium and simulate the model to show the impact of the relative income of the poor on innovation given the endogenous population share of the poor. In section 3.5, we discuss the case, where the relative income of the poor is endogenous and the population share of the poor is exogenous. Section 3.6 presents the dynamics of the model. The main results are summarized in section 3.7.

3.2   The Model

We have two types of agents, consumers and firms. Consumers live for two periods: young and old. The young people decide whether to undergo education, and the old people decide how to consume. Firms produce two kinds of goods, the standard good and the quality good. In order to produce quality goods, they must first invent them.

3.2.1 The Environment

We consider an overlapping-generations economy populated by a continuum of consumers, who live for two periods: young and old. The population size is constant. We normalize the population size of young people to measure one. Then the population of the old also has measure one at any time. In the first period, young individuals receive an amount of money from old individuals which can cover their education cost denoted by e. This transfer is exogenous. This assumption implies that everybody is equal at birth. The difference in the old period is due to the education decision in the young period. This transfer e is the single source of income for the young.

The old individual works and owns firms. Our model focuses on the demand for consumption in the period "old". Hence, we assume a simplistic view regarding the production of consumption goods. Labour is the single productive factor, and every old individual inelastically supplies one unit of labour to the competitive labour market. The income of the old individual i consists of two parts: w is the basic income, and θA_i is the bonus, with θ as the constant dividend rate and A_i taking the value of firms owned by the individual i, we call it wealth. Hence, the total income of the old individual i is $y_i = w + \theta A_i$. For simplicity, we assume that there are only two groups of old individuals, the poor (p) and the rich (r), distinguished by wealth, $A_r > A_p$, and, consequently, by income $y_r > y_p$. Their consumption expenditure is the total income net of the education cost e. (In Table 3.1 below, $y_p - e$ and $y_r - e$, respectively.) We assume $A_p = dV$, where d ($0 < d < 1$) measures the wealth of the old poor relative to the average level of old people, V is the average wealth per capita in period "old". Hence,

$V = \beta A_p + (1-\beta)A_r$, where β $(0 < \beta < 1)$ denotes the population share of the poor in period "old". We get $A_r = \frac{1-d\beta}{1-\beta}V$. Hence, the more young people undergo education, the lower the relative wealth of the old rich people ($\frac{1-d\beta}{1-\beta}$), holding d constant. The average wealth V is accumulated by the profit of firms after netting research costs and interest payment. For the definition of profits see section 3.2.3 and for that of research costs see section 3.2.4.

Young individuals can decide whether or not to go to school. If they go, they pay e as tuition. Thus, they have nothing left to spend on consumption (in Table 1 below, the consumption expenditure of a student is 0). Otherwise they can consume with their budget e (in Table 3.1 below, there is consumption expenditure e for non-students). Without education, a young person is confined to a poor position in society upon reaching old age and gets wealth A_p, otherwise they can have the wealth of a rich person A_r. In other words, the population share of students in period "young" is $1-\beta$.

Table 3.1: Expenditure and population size of different individuals

	Young		Old	
	Student	Non-student	Poor	Rich
Consumption expenditure	0	e	$y_p - e$	$y_r - e$
Population size	$1-\beta$	β	β	$1-\beta$

This assumption implies that the educated individual can supply a higher qualified labour unit than the non-educated. Hence, the rich can earn more than the poor, although both work for the same time. The skill premium is presented by the difference of the bonus $\theta(A_r - A_p)$. In reality, we often

find that the manager (highly educated individual, or the rich) have more shares of firms than workers (lowly educated one, or the poor). Here, we do not need to justify exactly why educated labors can acquires more wealth. Our purpose is to show the impact of income distribution on the demand for the consumption good.

There are two kinds of goods, referred to as standard good and quality good, respectively. Let x be the quantity of the standard good, which has a constant quality (normalized to 1) and is traded in a competitive market. Hence, the price P_x is equal to its marginal cost, which is also normalized to 1. The marginal cost of the standard good can be expressed as wb, where w is the basic income of the worker and the unit labour demand is b, which measures how many units of labour are needed to produce one unit of the standard good. We get $P_x = wb = 1$.

The quality goods are traded in an oligopolistic market. At any time there are many qualities q_j, $j = 0,-1,-2,...$ available in the market, the high quality good is k times better than the next lower one: $q_j = kq_{j-1}$. But marginal costs are same, denoted by wa, where $a < 1$ is again the unit labour demand. Every quality good is first invented through research and then produced by one firm. After a successful innovation, the new inventor can produce a k-times better quality good than the existing best one in the next period. Innovation is a random process, which will be introduced in section 3.2.4. Hence, the life-time of the oligopolistic firm is uncertain. The firm which sells the highest quality good q_0 can keep its position until the successful inventor enters, after which its good becomes the second best good q_{-1} until the next new inventor enters and so on. Since in the original model of Zweimüller and Brunner (2005) only two firms can exist in their

vertical differentiation competition, we take their results to mean that the third best quality is driven out of the market. This third best quality supplier can be considered as the potential competitor, who sets the price at the marginal cost. Nonetheless its demand is still zero.

3.2.2 The Household's Decision Problem

Every individual faces a two-stage decision problem. At the beginning of young period she decides whether or not to go to school, i.e., she allocates the consumption expenditure over time. When she is old, the individual decides how to allocate her instantaneous consumption expenditure between standard good and quality good. It doesn't mean that young people have no consumption decision. They consume with budget e if they do not go to school. But we assume for simplicity that they can only consume the standard good.[13] Hence, they simply spend all of their income e on the standard good in order to maximize their instantaneous utility.

We begin our analysis from the second stage of the individual's decision problem. There is no saving. All income of old people except for the education lump-sum tax is spent either on the consumption of the standard good or the quality good. Every old individual can consume one and only one unit of the quality good q_j. There is no limitation to the consumption of the standard good x_i except for the budget constraint, i.e., $y_i - e = 1 \cdot x_i + P_j \cdot 1 \quad j = 0, -1$, where the price of standard goods is 1, the

---

[13] This assumption is made to ensure that the potential consumers of quality goods have only two types (old poor and old rich). It guarantees that there are only two qualities in the market (see Zweimüller and Brunner2005). This assumption also is reasonable. For example, we can imagine the quality goods to be automobile, alcohol and/or cigarette, which are prohibited for young people.

price of the best quality is denoted by P_0 and the second best one's price is P_{-1}. The preference for consumption of the standard good and the quality good is given by the following instantaneous utility function:

$$u_{old,i}(x_i, q_j) = \ln x_i + \ln q_j \qquad i = p, r \text{ and } j = 0, -1 \qquad (3.1)$$

which can also be expressed as: $u_{old,i} = \ln(y_i - e - P_j) + \ln q_j$.

Analogously, we assume the instantaneous utility of young individuals as follows:

$$u_{young,i} = \begin{cases} \ln e & i = p \\ 0 & i = r \end{cases} \qquad (3.2)$$

Young individuals maximize their life-time utility U_i at the beginning of their young period:

$$U_i = u_{young,i} + \rho u_{old,i} \qquad (3.3)$$

where ρ is the discount factor. If they decide not to undergo education, their life-time utility is $U_p = \ln e + \rho u_{old,p}$. Otherwise, they have $U_r = \rho u_{old,r}$. Suppose $U_p > U_r$, all young people are not willing to undergo education. Hence, $\beta \to 1$. Recall that the relative wealth of the rich ($\frac{1 - d\beta}{1 - \beta}$) increases in β, $\beta \to 1$ implies $y_r \to \infty$. We assume that the instantaneous

utility increases in the income[14], hence, $U_r \to \infty$, which contradicts $U_p > U_r$. It implies that $U_p > U_r$ cannot be an equilibrium. Analogy, $U_p < U_r$ cannot be an equilibrium, either. Hence, in equilibrium we must have $U_p = U_r$, i.e.:

$$\ln e = \rho(u_{old,r} - u_{old,p}) \qquad (3.4)$$

The left hand side of (3.4) is the cost of education, while the right hand side is the benefit. In equilibrium both should be equal. Hence, the heterogeneity among old consumers (poor and rich) comes from the indifference between education and non-education for the young. In this heterogeneous steady state, β is determined by exogenous parameters, although individuals are randomly divided between the poor and the rich.

3.2.3 The Pricing Decision of Oligopolists

Firms have all the above information but they are unable to distinguish between individuals by income. The strategy which firms can pursue is to choose a price while quality is fixed. We concentrate only on the steady state where prices are constant over time. The whole market size of oligopolists is 1 while only the old individuals can buy quality goods. We differentiate between rich and poor consumers of quality goods respectively, dropping "old" below. For simplicity, we assume that the consumer prefers better quality goods if both quality goods yield the same utility.

---

[14] Since we don't know the price of firms, we can only assume the monotonous relationship between the utility and the income. In next section, we analyze price decision of firms, the monotony expressed through equations (3.18) and (3.19) can then be proven.

Lemma 3.1: There are only two kinds of pricing equilibria: In the pooling equilibrium the best quality good is sold to both the rich and the poor, the second best quality good can't be sold; in the separating case the best quality good is consumed by the rich and the second best quality good is consumed by the poor.

Proof: see Appendix 3.1.

The second best quality supplier considers only how to attract the poor to purchase her good while the rich never buy the second best good in equilibrium, according to Lemma 3.1. Because of the existence of potential competitors which offer the price at marginal cost, the highest price which the second best firm offers satisfies:

$$\ln(y_p - e - \overline{P}_{-1}) + \ln q_{-1} = \ln(y_p - e - wa) + \ln q_{-2} \qquad (3.5)$$

The left hand side of this equation is the utility when poor individuals buy the second best quality good q_{-1} and the right hand side is the utility when they consume the third best quality good q_{-2}. Only if the second best quality good can yield at least the same utility as the third best quality good to consumers, the consumer prefers buying it. Substituting $q_{-1} = kq_{-2}$ and rearranging the equation, we get the highest price of the second best quality good:

$$\overline{P}_{-1} = (1 - \frac{1}{k})(y_p - e) + \frac{wa}{k} \qquad (3.6)$$

The lowest price which the second best quality firm can offer is at marginal cost wa. Analogously, the best firm can set its highest price satisfying:

$$\ln(y_i - e - \overline{P_0}) + \ln q_0 = \ln(y_i - e - P_{-1}) + \ln q_{-1} \qquad (3.7)$$

leading to
$$\overline{P_0} = (1 - \frac{1}{k})(y_i - e) + \frac{P_{-1}}{k}, \quad i = p, r \qquad (3.8)$$

These two reaction functions are depicted in Figure 3.1. In order to attract the poor to buy its products the best firm sets its price as high as $(1 - \frac{1}{k})(y_p - e) + \frac{P_{-1}}{k}$ (it is the line CD in Figure 1). Because the rich can afford more good quality goods than the poor, they are willing to buy the best good too if the poor prefer the best good to the second best good. Hence, the area below CD (including CD) is the pooling strategy case, where the best quality good captures the entire market and the second best quality good is not sold. Above CD the poor don't purchase the best quality good. The line AB in Figure 3.1 expresses the highest price of the best good, given P_{-1}, if the best firm wants to attract only the rich. Hence, the area ABCD excluding the line CD is the separating strategy.
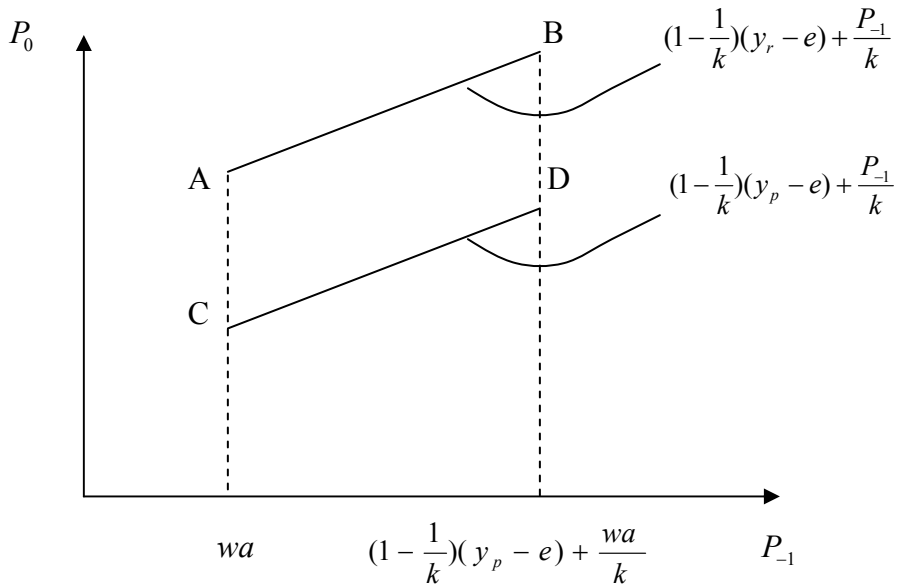
Figure 3.1: Pricing decision of quality goods firms

We define the profit of firms as $\pi_j = (P_j - wa)(\textit{market share of firm j})$, $j = 0, -1$. The firms set their prices as high as possible given the market share of their goods. The two possible equilibria in price competition are summarized below:

1. Pooling: q_0 is sold to all consumers. In this case the second best firm becomes the potential competitor and its price and profit are respectively:

$$P_{-1} = wa \qquad (3.9)$$

$$\pi_{-1} = 0 \qquad (3.10)$$

The best firm can set its price at:

$$P_0 = (1 - \frac{1}{k})(y_p - e) + \frac{wa}{k} \qquad (3.11)$$

and earn profit $\quad \pi_0 = (1 - \frac{1}{k})(y_p - e - wa) \qquad (3.12)$

2. Separating: q_0 is sold to the rich and q_{-1} to the poor. Because this is a repeated game until a new inventor comes in, many possible equilibria could exist. In order to get a unique result, we assume that no player is punished if she changes her price without affecting the other player's profit. Then the single separating equilibrium is point B:

$$P_{-1} = (1 - \frac{1}{k})(y_p - e) + \frac{wa}{k} \qquad (3.13)$$

$$P_0 = (1 - \frac{1}{k})(y_r - e) + (\frac{k-1}{k^2})(y_p - e) + \frac{wa}{k^2} \qquad (3.14)$$

$$\pi_{-1} = \beta(P_{-1} - wa) \tag{3.15}$$

$$\pi_0 = (1 - \beta)(P_0 - wa) \tag{3.16}$$

Proof: see Appendix 3.2.

3.2.4 Innovation

As mentioned before, the new entrant of this oligopolistic market should do research before production. Only after the successful innovation it can produce a quality k-times better than the currently best. Following the work by Aghion and Howitt (1992), we assume that the innovation is random and arrives according to a Poisson process with parameter ϕ. The researcher can employ n workers to reach the Poisson arrival rate ϕ, i.e., $\phi = \lambda n$, where λ is the productivity of workers in research, which is given by the technology of research. This assumption of innovation means that the success of research depends only on current input, not upon past research. The flow of research cost is wn. And the flow of research benefit is ϕB, where B is the present value of future profits when innovation takes place:

$$B = \sum_{t=0}^{\infty} \left(\frac{\pi_0}{(1+\theta)^t} prob\{0 \text{ innovation before } t\} + \frac{\pi_{-1}}{(1+\theta)^t} prob\{1 \text{ innovation before } t\} \right)$$

$$= \sum_{t=1}^{\infty} \left(\frac{\pi_0 (1-\phi^e)^{t-1}}{(1+\theta)^t} + \frac{\pi_{-1}(t-1)(1-\phi^e)^{t-2}\phi^e}{(1+\theta)^t} \right)$$

leads to
$$B = \frac{\pi_0}{\phi^e + \theta} + \frac{\phi^e \pi_{-1}}{(\phi^e + \theta)^2} \tag{3.17}$$

where t is a time index, $\phi^e = \lambda n^e$ is the expected future arrival rate of innovation, and n^e is the expected future number of workers in the research sector.

We are now in a position to define V. We assume that the firms' profits net of interest payments and research costs consist of average wealth, i.e., $\Delta V = \pi_0 + \pi_{-1} - wn - \theta V$, where ΔV presents the difference of the average wealth between two subsequent periods. V can be interpreted as the aggregate value of firms. According to our assumption this wealth is distributed among old individuals according to their education level.

3.3　Equilibria

The general equilibrium, which consists of three conditions, is presented in this section. Substituting the price decisions of firms into equilibrium conditions, we obtain two possible equilibria: the pooling and the separating.

3.3.1　Equilibrium Conditions

The equilibrium condition of the education decision is given by equation (3.4). Substituting the pooling price (3.11) and the separating prices (3.13), (3.14) in (3.4), respectively, we get the same form as follows:

$$\ln e = \rho \ln \left(\frac{k(y_r - y_p)}{y_p - e - wa} + 1 \right) \tag{3.18}$$

leads to

$$\ln e = \rho \ln \left(\frac{k\theta V \dfrac{1-d}{1-\beta}}{w + \theta dV - e - wa} + 1 \right)$$

(3.19)

From (3.19) we know the interdependence between β and d. The left hand side of (3.19) is the education cost and the right hand side is the benefit from education. If we improve the relative wealth of the poor (d rises), y_p increases and y_r decrease. In other words, the benefit of education declines. Therefore, young people have less of an incentive to undergo education. This means *ceteris paribus* a higher population share of the poor. We assume now that d is exogenous and β is endogenous. In section 3.5, we discuss the impact of an exogenous β on the innovation rate, given that d is endogenous.

Lower time preference ρ indicates more impatience. Hence, fewer individuals invest in education, which means a higher β. The effect of e on β is not so obvious. At first, e is the education cost. The increase in the education cost decreases the incentive of education for the young. Hence, β increases. This is the effect of e on the left hand side of equation (3.19). On the other hand, e is also a social transfer from the old to the young. The old becomes poorer if e increases. Hence, the marginal utility of education increases, i.e., $y_r - y_p$ yields more utility if y is lower. This induces a lower β. If e is not so large, the latter effect is dominated by the former. We can proof that β increases in e if it satisfies the following sufficient condition:

Assumption 3.1 $\qquad e \le \dfrac{w - wa}{1 + \rho}$ $\hspace{3cm}$ (3.20)

Proof: see Appendix 3.3.

For the research sector we assume free entry and perfect foresight for firms in equilibrium, which is analogous to Aghion and Howitt (1992). Hence, the innovation equilibrium condition means that the flow of research costs should be equal to the flow of expected profits, i.e. $wn = \phi B = \lambda n B$ and $\phi = \phi^e$ (or, $n = n^e$) due to perfect foresight. This leads to:

$$\frac{w}{\lambda} = \frac{\pi_0}{\phi + \theta} + \frac{\phi \pi_{-1}}{(\phi + \theta)^2}$$ $\hspace{2cm}$ (3.21)

The underlying intuition is similar to Aghion and Howitt (1992). The left hand side of equation (3.21) represents the flow cost of research per efficient worker, which decreases in the productivity of research workers λ. The effect of λ on ϕ is positive, because the researcher employs more workers to do research, if the productivity of workers increases. The interest rate affects the innovation rate via two channels: first it is a discount factor, hence, the higher θ, the lower is the benefit of research. Therefore, the innovation rate decreases in the interest rate. On the other hand, higher θ means more interest income of individuals, which increases the profit of firms. Hence, the benefit of research increases. This implies the positive impact on the innovation rate. The main difference between our model and that of Aghion and Howitt (1992) lies in the market structure. They assume a monopoly market. Hence, firms can survive only until a successful inventor comes in. Hence, there is no π_{-1}. In our oligopoly model firms can exist in two stages: best quality supplier and second best supplier. Note that

in the pooling case $\pi_{-1} = 0$, which is equivalent to the case of the monopoly in Aghion and Howitt (1992).

The income of the education sector $(1-\beta)e$ originates from students, and the education sector employs workers to supply courses. We assume the labour demand of the education sector is S, thus, the cost of the education sector is wS. In equilibrium the budget of the education sector should be balanced: $(1-\beta)e = wS$.

The labour market equilibrium condition is standard, i.e., at any point in time the labour supply should be equal to the labour demand:

$$1 = n + a + b(e\beta + x_p\beta + x_r(1-\beta)) + S \qquad (3.22)$$

which implies:
$$\frac{w\phi}{\lambda} = \pi_0 + \pi_{-1} - \theta V \qquad (3.23)$$

By assumption, each old individual supplies one unit of labour, so the total labour supply is 1. The total labour demand is illustrated by the right hand side of equation (3.22): n is the labour demand in the research sector. The quality good sector needs labour a, because the total demand for quality goods is 1 and the unit labour demand of quality goods is a. $e\beta + x_p\beta + x_r(1-\beta)$ is the total demand for standard goods, which consists of three parts: the non-students' demand, and the demand of the poor and the rich. Recall that b is the unit labour demand of standard goods, hence, the third item of the right hand side of equation (3.22) measures the demand for labour in the standard goods sector. Finally S is the labour demand of the education sector. Recall that $\Delta V = \pi_0 + \pi_{-1} - \theta V - \frac{w\phi}{\lambda}$. Hence, (3.23) implies that in a stationary equilibrium, the average wealth V remains

constant ($\Delta V = 0$). Now we have three equations (3.19) (3.21) (3.23) in the three variables β, ϕ and V. We omit the discussion of equilibrium existence condition, because it is similar to that in Zweimüller and Brunner (2005).

3.3.2 The Pooling Equilibrium

Substituting price and profit equations (3.9), (3.10), (3.11), (3.12) and (3.13), (3.14), (3.15), (3.16), respectively, in the above equilibria conditions (3.21), (3.23) leads to two different equilibria, namely "Pooling" and "Separating". We discuss the simple case first.

Pooling equilibrium:

$$\ln e = \rho \ln \left(\frac{k\theta V \dfrac{1-d}{1-\beta}}{w+\theta dV - e - wa} + 1 \right) \qquad (3.19)$$

$$\frac{w}{\lambda} = \frac{(w+\theta dV - e - wa)(1-\dfrac{1}{k})}{\phi + \theta} \qquad (3.24)$$

$$\frac{w\phi}{\lambda} = (w+\theta dV - e - wa)(1-\frac{1}{k}) - \theta V \qquad (3.25)$$

It is of interest to see the impact of e, ρ, d on the equilibrium value of variables β, ϕ and V. If we substitute (3.24) in (3.25), then $V^* = \dfrac{w}{\lambda}$. Hence, e, ρ, d have no impact on V^*. We depict these equations below in Figure 3.2. In the pooling equilibrium (3.24) and (3.25) are independent of β. This simplifies the analysis. The right hand side of Figure 3.2 shows just

the model with an exogenous β, which is the special case in my model without (3.19).
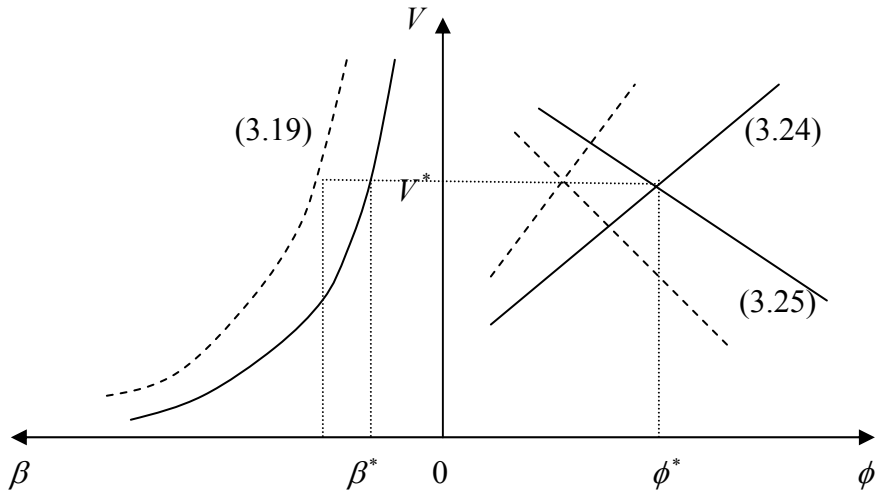
Figure 3.2: The pooling equilibrium

When e increases, the opportunity cost of education increases (dotted line in Figure 3.2). This means that fewer young people undergo education. Hence, β^* increases under the assumption 1. On the other hand, e is a transfer from the old to the young, viz. to those who cannot buy quality goods. It is equivalent to say that the consumer of the quality good becomes poorer. A lower willingness to pay translates into a reduced price and less profit from quality goods. This, in turn, leads to a lower incentive to innovate.

If ρ increases, then β^* decreases, because the young are more patient and thus more of them invest in education. However, it does not necessarily imply a higher innovation rate, because in the pooling equilibrium the market of the best quality goods is made up of the whole population of old people, the change of β^* does not change the market share and profits of

the best quality good supplier. Hence, there is no impact on the incentive to invent.

From equation (3.24) and (3.25) we get $\phi^* = (\lambda - a\lambda - e\lambda / w + \theta d)(1 - 1/k) - \theta$. When d increases, the benefit of education decreases. This leads to fewer students, in addition to fewer rich consumers of quality goods. But in the pooling equilibrium the market of quality goods is the whole population of the old. The profit of firms is independent on the population share of the poor. Moreover, the decisive consumer in the price decision is the poor. The improvement of their budgets means that they can pay more for quality goods. The firm can charge a higher price and earn a larger profit, which increases the incentive to innovate. In this sense we can say that redistribution (increase in d) is good for the long-run growth rate (here, a higher innovation rate ϕ).

We summarise the relationship between inequality and innovation as follows:

Proposition 3.1:

In the pooling equilibrium, the relative wealth of the poor has a positive impact on the innovation rate; and the innovation rate is independent on the population share of the poor. A higher education cost leads to a lower innovation rate. The time preference ρ has no impact on the innovation rate.

3.3.3 The Separating Equilibrium

Now we turn to the separating equilibrium. According to the vertical differentiation model of Shaked and Sutton (1982, 1983), maximization of quality differentiation is the optimal choice for firms in order to reduce price competition. The target of quality differentiation is to separate the whole market into several parts. The firms can then play in some sense a monopolistic role. The pooling equilibrium appears only because we assume for simplicity that there are two types of consumers. If income is continuously distributed among individuals, the pooling equilibrium cannot exist any more. Hence, the separating equilibrium is more general and more important than the pooling case. We will concentrate on the separating equilibrium in following discussion.

In the separating equilibrium q_0 is sold only to the rich and q_{-1} is sold only to the poor. Hence, β enters the profit function of firms and the equilibrium equations of innovation and labour market. So the analysis is more involved. In order to show the impact, we simulate the model in the next section.

$$\ln e = \rho \ln \left(\frac{k\theta V \dfrac{1-d}{1-\beta}}{w + \theta dV - e - wa} + 1 \right) \qquad (3.19)$$

$$\frac{w}{\lambda} = \frac{\left(\left(w - e + \dfrac{1-d\beta}{1-\beta}\theta V \right)\left(1 - \dfrac{1}{k}\right) + \left(w - e + d\theta V\right)\left(\dfrac{k-1}{k^2}\right) + \dfrac{wa}{k^2} - wa \right)(1-\beta)}{\phi + \theta}$$

$$+ \frac{\phi\left(\left(w - e + d\theta V\right)\left(1 - \dfrac{1}{k}\right) + \dfrac{wa}{k} - wa \right)\beta}{(\phi + \theta)^2} \qquad (3.26)$$

$$\frac{w\phi}{\lambda} = \left(\left(w - e + \frac{1 - d\beta}{1 - \beta}\theta V\right)\left(1 - \frac{1}{k}\right) + \left(w - e + d\theta V\right)\left(\frac{k-1}{k^2}\right) + \frac{wa}{k^2} - wa\right)(1 - \beta)$$
$$+ \left(\left(w - e + d\theta V\right)\left(1 - \frac{1}{k}\right) + \frac{wa}{k} - wa\right)\beta - \theta V \qquad (3.27)$$

3.4   Simulation

As equations (3.19), (3.26), (3.27) show, the separating equilibrium is not as easy to analyze as the pooling case, because the population share of the poor is able to influence the firms' profits. Our main purpose is to study the impact of the income distribution (here measured by d) on the innovation rate ϕ and the population share β, as well as the impact of the education cost e and the time preference ρ on ϕ and β. For the purpose of simplification, we show the impact by numerically simulating the model.

3.4.1 Simulation Procedure

First, we assume d to be exogenous and β to be endogenous. (The reverse case is considered in the next section.) In order to show the impact of parameter changes and exogenous variables on the endogenous variables, we analyze them *ceteris paribus*. E.g., in order to show the impact of d on ϕ, β, V, we let d move away from the benchmark value 0.4 to 0.2 and 0.6, respectively, holding the other parameters at the fixed benchmark value (the value of d from 0.2 to 0.6 and the according values of endogenous variables are shown in Table 3.2, first part).

We set all parameters and exogenous variables at the following values as a benchmark. $\theta = 0.5, \rho = 0.5, w = 10, e = 2, a = 0.3, \lambda = 1, k = 4, d = 0.4$. They are chosen for the following reasons. 50% is the suitable interest rate per

period, because the period in the current paper reflects the generation. In reality it is about 20-30 years. We also assume the subjective discount rate $\rho = 0.5$, which reflects the time preference between young and old periods. The basic income w and the education cost e are set to satisfy Assumption 1. We choose $a = 0.3$ so that total labour supply is almost equally allocated among research, quality good production and the standard good sector. The other two parameters characterizing research and innovation -- λ, k -- are chosen only for simplicity, because we know very little about such characteristics in this pure theoretical model.

3.4.2 Simulation results

The simulation result is summarized as below Proposition 3.2 and Table 3.2:

Proposition 3.2:

In the separating equilibrium, redistribution from the rich to the poor (i.e., d increases) has a negative impact on the innovation rate; the population share of the poor increases with the relative wealth of the poor. A higher education cost and a lower time preference leads to a lower innovation rate.

Table 3.2: Simulation results of the separating equilibrium

with $\theta = 0.5, \rho = 0.5, w = 10, e = 2, a = 0.3, \lambda = 1, k = 4, d = 0.4$ as benchmark

d	0.2	0.3	0.4	0.5	0.55	0.6
ϕ	0.35	0.33	0.31	0.29	0.28	0.26
β	0.08	0.21	0.34	0.46	0.52	0.57
V	10.43	11.28	12.34	13.67	14.49	15.45
$\dfrac{\pi_0^{po}(V^{se})}{\phi^{se} + \theta}$	5.33	6.05	6.91	7.99	8.64	9.51

ρ	0.25	0.3	0.4	0.5	0.6	0.65
ϕ	0.18	0.21	0.27	0.31	0.35	0.37
β	0.83	0.74	0.54	0.34	0.16	0.07
V	17.91	16.46	14.09	12.34	11.0	10.44
$\dfrac{\pi_0^{po}(V^{se})}{\phi^{se}+\theta}$	9.47	8.76	7.61	6.91	6.35	6.11

e	1.75	2	2.25	2.5	2.75	3
ϕ	0.38	0.31	0.26	0.22	0.18	0.15
β	0.15	0.34	0.47	0.56	0.62	0.67
V	10.92	12.34	13.41	14.26	14.95	15.51
$\dfrac{\pi_0^{po}(V^{se})}{\phi^{se}+\theta}$	6.34	6.91	7.33	7.66	7.99	8.19

We first check whether the above solution of separating price strategy can yield at least as much benefit as under pooling. Otherwise the firm would switch to pooling. The benefit from separating is $\dfrac{\pi_0^{se}}{\phi^{se}+\theta}+\dfrac{\phi^{se}\pi_{-1}^{se}}{(\phi^{se}+\theta)^2}$, which is equal to $\dfrac{w}{\lambda}$ in equilibrium. The benefit from switching to pooling given the average wealth level V in separating equilibrium is $\dfrac{\pi_0^{po}(V^{se})}{\phi^{se}+\theta}$, where indices "se" and "po" refer to separating and pooling, respectively. The necessary condition for separating to occur is: $\dfrac{\pi_0^{po}(V^{se})}{\phi^{se}+\theta}\le\dfrac{w}{\lambda}=10$. Parameter values satisfy this condition, see above Table.

As opposed to the pooling case we have two different results in a separating equilibrium:

1) An increase in d still implies less education, but no longer large incentives to innovate, because the consumer determining the profit of the best quality good is the rich. The lower education, the fewer rich individuals exist. d has a direct negative impact on the relative wealth of the rich $A_r = \frac{1-d\beta}{1-\beta}V$, and an indirect positive impact on it via β and V. According to this simulation, the net effect is positive. Hence, the best quality supplier faces a smaller market share and a higher willingness to pay. However, the positive effect on the consumers' willingness to pay is dominated by the negative effect on the market share. Thus, the profit of the best quality supplier decreases in d. On the other hand, the second best quality supplier has a higher demand. The population share of the poor increases and their willingness to pay has been improved. Intuitively, the profit of the best quality is more important, because it has a bigger weight than that of second best quality in the equation of present value calculation (3.21). Hence, the relative wealth of the poor has a net negative impact on the innovation rate. For the simulation results of the impact of d on A_r, π_0, π_{-1} and their weights see Appendix 4. The increase in d delays the realization of profits (π_0 decreases and π_{-1} increases). Hence, less research is undergone. It implies the lower research cost and the higher V, according to (3.23).

2) An increase in ρ results in more education in the young generation, and therefore increases the share of the rich in the old generation. Hence, the profit of the best quality increases and the profit of the second best quality declines. Again because the profit of the best quality good has a large weight in the valuation of innovation, firms have thus more incentives to innovate when they face an increasing

population share of the rich. This impact of the population share on innovation does not appear in the pooling case, where the best quality good is sold to the entire population of the old generation. More incentive to innovate implies more labour in research sector. Hence, the research cost increases and the value of oligopolists (V) decreases.

The effect of the education cost e is similar as in the case of pooling, but the reason is different. An increase in e impedes individuals to accept education, so that the population share of the rich decreases. Moreover, the consumer of the quality good becomes poorer when she has to pay higher transfers to young people. Both of them decrease the profit from selling the best quality good. But the impact on the second best quality good is *a priori* unclear, because the market share increases and the consumers' willingness to pay decreases. The simulation results suggest that the net effect of education cost on innovation is negative.

In order to show that it is crucial to assume interdependence between the population share of the poor and the relative income of the poor, we show the simulation results of the model given β exogenously. Then we lose an equilibrium condition of education (3.19). For the sake of comparison we set $\beta = 0.34$, which is the equilibrium value in our simulation benchmark:

Table 3.3: The impact of d on ϕ given $\beta = 0.34$

d	0.2	0.3	0.4	0.5	0.55	0.6
ϕ	0.302	0.308	0.313	0.319	0.322	0.325

This simulation implies that redistribution has a positive effect on innovation although it is very weak. An increase in d improves the willingness to pay for the second best quality by the poor. So π_{-1} is

increasing in d. This in turn allows the best quality producer to charge a higher price. On the other hand, the rich becomes poorer with an increase of d. If the population share of the rich is big enough, their wealth doesn't reduce strongly when d increases. The negative effect on π_0 is dominated by the positive effect on π_{-1}. In sum, less inequality in the sense of a higher d is good for innovation. But this effect is weak because π_{-1} is lower weighted than π_0 in (3.21). From the supply side, the sum of π_{-1} and π_0 is increasing in d ($\frac{\partial(\pi_0 + \pi_{-1})}{\partial d} = (1-\beta)\theta V\left(\frac{1}{k} - \frac{1}{k^2}\right) > 0$). There is less expenditure for the standard goods. Hence, more labour units can be allocated in the research sector, which induces a higher innovation rate.

Population shares in our model are not constant. The redistribution (d increases) cannot only improve the budget constraint of the poor, but also decrease the incentive to education, thus decrease the population share of the rich. Hence, a higher d has two effects on π_0: the direct effect is that it decreases the wealth of the rich; the indirect effect is that it reduces the market share of the best quality. In other words, the negative effect on π_0 is strengthened by the endogenous population share. Although the positive effect on π_{-1} is also strengthened, its weight in (3.21) is smaller than that of best quality good. Hence, as we have seen in the simulation results, the net effect of redistribution on innovation is negative when we assume the population share is endogenous through education.

Another interesting result from our model is that the education enrollment $1 - \beta$ and the innovation rate ϕ are positively correlated although we don't assume that education can increase the productivity. In this sense, education looks much more like a tool of distribution, but not like a production factor in this paper. However, it can still increase the innovation rate (or growth

rate in some sense, see section 3.5) because it produces richer consumers, who induce society to allocate more resource in the research sector. Normally economists discuss the impact of education only through the supply side, i.e., education increases the productivity. Hence, the supply increases and the economy will grow. However, Bils and Klenow (2000) supply empirical evidence that the human capital (as a production factor) which is produced by a higher education level cannot explain the higher long-run growth rate which is associated with this higher education level. Their explanation is that education is much more like consumption than productive investment. Hence, individuals will increase education in the young period, if they expect that in the future they will have a higher income level. In other words, it is the expected higher long-run growth rate that leads to a higher education level today, but not vice versa. Our model supplies another explanation for the empirical results of Bils and Klenow (2000). If individuals can become richer through education, a higher growth rate can be achieved regardless of whether education increases productivity. The higher education level is associated with the higher growth rate through the demand for better quality goods.

3.5 Exogenous Population Shares and Endogenous Relative Income

The above analysis assumes that the income of the poor relative to the average level is an exogenous variable and the population share of the poor is endogenous. We can also assume that the population share of the poor is exogenous and the relative income of the poor is endogenous. These different assumptions could imply different policies. We can interpret it as follows. If the government wants to decrease the income inequality by setting a higher minimum wage or by giving more social transfers, it is

similarly to say, d increases exogenously with respect to our model's analysis. Then we should consider its impact on the innovation rate not only through its direct effects on the willingness to pay, but also indirect effects through the population share. On the other hand, if government sets up a mandatory education law to improve the population share of the student, it decreases β exogenously. Now we discuss what the impact of the exogenous β is on the innovation rate by assuming an endogenous d. The simulation results are as follows:

Proposition 3.3:

In separating equilibrium, a higher population share of the poor has a negative impact on the innovation rate. The relative wealth of the poor is positively associated with the population share of the poor.

Table 3.4: The impact of the exogenous β on ϕ, d, and V

where $\theta = 0.5, \rho = 0.5, w = 10, e = 2, a = 0.3, \lambda = 1, k = 4$

β	0.2	0.3	0.34	0.4	0.5	0.6
ϕ	0.33	0.32	0.31	0.30	0.28	0.25
d	0.29	0.37	0.40	0.45	0.54	0.63
V	11.19	11.98	12.34	12.96	14.24	16.0
$\dfrac{\pi_0^{po}(V^{se})}{\phi^{se} + \theta}$	5.98	6.60	6.91	7.42	8.19	10

In this example, if the government increases the education opportunity (β decreases), and in equilibrium all such education opportunities are used by individuals, the relative wealth of the poor has to decline in order to push individuals to enter school. Because of the increasing population share of the rich, the profit of best quality good increases. This in turn raises the incentive to invent.

What is the impact of inequality (through β or d, respectively) on utility? From (3.1) we have:

$$\Delta u = \frac{\Delta x}{x} + \frac{\Delta q}{q} \qquad\qquad (\,3.28\,)$$

In a steady state, the consumption of standard goods is constant ($\Delta x = 0$), and $\Delta q = \phi(k-1)q$. Hence, we have $\Delta u = \phi(k-1)$. The higher the innovation rate, the larger is the increase in the utility. Redistribution from the rich to the poor (d increases exogenously, β is endogenous) increases the price and profit of the quality good, and the average wealth, too. Hence, consumers become richer through redistribution and consume more standard goods. The production resource, labor, is shifted from the research sector to the standard good sector. Consumers enjoy a higher utility level in the short run, but the long-run growth rate of the utility is lower than before because of a lower innovation rate. In contrast to redistribution, the decrease of the population share of the poor can induce a higher innovation rate. Hence, in a new equilibrium consumers have lower consumption of standard goods, but the long-run growth rate of the utility becomes higher.

3.6   The Model Dynamics

So far, we have only discussed the steady state. The comparative statics show us the long-run impact of parameters on the equilibrium. In this section we discuss the effect of parameters on the endogenous variables in the short run and the medium run, i.e., the path to equilibrium. We focus on the case where d is exogenous and β is endogenous (the case in section 3.3 and 3.4). From section 3.2.4, we know that the accumulation function of average wealth is:

$$\Delta V = \pi_0 + \pi_{-1} - \theta V - \frac{w\phi}{\lambda} \qquad (3.29)$$

The total profits net of the interests payment $(\pi_0 + \pi_{-1} - \theta V)$ is the net income of firms. $\frac{w\phi}{\lambda}$ is the research cost. If the net income of firms is higher than the research cost, the economy can accumulate the average wealth.

From equation (3.19) we know:

$$\Delta \beta = -\frac{(1-d\beta)(1-\beta)(w-e-wa)}{(1-d)(w+\theta dV - e - wa)V}\Delta V \qquad (3.30)$$

According to Assumption 3.1, $w - e - wa > 0$. Hence, the accumulation of the average wealth ($\Delta V > 0$) enlarges the income gap between the poor and the rich. More young people are attracted to education. Hence, the population share of the poor declines ($\Delta \beta < 0$).

The innovation rate $\phi = \lambda n$ is determined by the number of workers in the research sector (n). From (3.22) we have:

$$n = 1 - a - eb - b(x_p\beta + x_r(1-\beta)) \qquad (3.31)$$

where $x_i, i \in \{p, r\}$ is the consumption of the standard good by the poor and the rich, respectively. Because $1 - a - eb$ is constant, n depends only on the aggregate consumption of standard goods, i.e., $\Delta n = -b\Delta(x_p\beta + x_r(1-\beta))$.

We concentrate on the impact of d on the dynamics of V, β, ϕ. Suppose there is a shock (d increases). We define the short run as the time just after the shock and before any other endogenous change of variables, and the medium run as the period when all endogenous variables move simultaneously to the long-run equilibrium values. Distinguishing the short run from the medium run enables us to study the different effects of the shock. The direct effect of d on the endogenous variables can be observed in the short run. The indirect effect of d through the interaction among the endogenous variables takes place in the medium run.

In the short run, when d increases and V is still unchanged, β increases (from equation 3.19) because the benefit of education decreases.

The total profits of firms are as follows:

$$\sum_i \pi_i = \beta(P_{-1} - wa) + (1 - \beta)(P_0 - wa)$$

$$= \beta[(1 - 1/k)(w + \theta dV - e) + \frac{wa}{k} - wa] + (1 - \beta)[(1 - 1/k)(w + \theta \frac{1 - d\beta}{1 - \beta} V - e)$$

$$+ (\frac{k-1}{k^2})(w + \theta dV - e) + \frac{wa}{k^2} - wa]$$

Since $\dfrac{\partial \sum \pi_i}{\partial d} = \dfrac{k-1}{k^2}(1 - \beta)\theta V > 0$, the total profit of firms increases in d. The intuition is as follows: if d increases, the poor become richer. Thus the firm producing the second-best quality goods is able to raise the price without losing consumers, which enables the supplier of the best quality good to increase her price, too. Hence, an increase in d raises the profit from quality goods. Therefore, the net income of firms is above the total cost of research, and $\Delta V > 0$ (from (3.29)).

83

From (3.31) we get $\dfrac{\partial n}{\partial d} = b\dfrac{k-1}{k^2}(1-\beta)\theta V > 0$, i.e., in the short run, the effect of d on the innovation rate is positive. Because V is kept unchanged in the short run, the total income of all consumers doesn't change. However, we know from above that the price of the quality goods increases in d. Hence, consumers have to decrease the consumption of the standard goods. This leads to more labor input in the research sector and a higher innovation rate.

Proposition 3.4:

In short run, the relative wealth of the poor (d) has a positive impact on the population share of the poor, the accumulation of the average wealth, and the innovation rate.

In the medium run, d reaches the new level. But the accumulation of the average wealth has just started. From (3.29) we get:

$$\frac{\partial \Delta V}{\partial V} = \frac{\partial \sum \pi_i}{\partial V} - \theta = (1-\frac{1}{k})[1+(1-\beta)\frac{d}{k}]\theta - \theta = \frac{\theta}{k}[(1-\beta)(1-\frac{1}{k})d - 1] < 0$$

It indicates that the average wealth does increase, but the change of the average wealth is diminishing. According to (3.30), β decreases in the average wealth. Hence, after the immediate jump in the short run the population share of the poor declines with the accumulation of the average wealth.

From (3.29) we know also:

$$\frac{\partial \sum \pi_i}{\partial \beta} = (1 - 1/k)(w + \theta dV - e) + wa/k - wa - [(1 - 1/k)(w - e)$$

$$+ (\frac{k-1}{k^2})(w + \theta dV - e) + \frac{wa}{K^2} - wa] - (1 - 1/k)\theta dV$$

$$= (\frac{k-1}{k^2})(wa - P_{-1}) < 0$$

Hence, at first ΔV decreases because of the immediate jump of β in the short run, then increases because β declines in the medium run. The average wealth V increases before it reaches the equilibrium value in long term.

From (3.31) we get $\frac{\partial n}{\partial V} = -\frac{b\theta}{k}[1 - d(1 - \beta)(1 - \frac{1}{k})] < 0$. The intuition is as follows: Because of the accumulation of the average wealth all consumers become richer than before, and the demand for standard goods is increasing. Hence, more labor units have to be allocated into the standard sector and less into research. The innovation rate begins to decrease after the immediate increase in the short run. Additionally, we have $\frac{\partial n}{\partial \beta} = -\frac{b(k-1)}{k^2}(w + \theta dV - e - wa) < 0$. This implies that the decrease of the innovation rate in the medium run is accelerated by the immediate jump of β in the short run.

To sum up, after the shock the average wealth increases slowly until it reaches a new higher equilibrium value. The population share of the poor increases in the short run but then decreases in the accumulation of the average wealth. However, our simulation results show that the long-run equilibrium value is still higher than before. Because the demand for the standard good sinks in the short run, the innovation rate achieves a higher

level. In the medium run, the innovation rate decreases because both the average wealth and the population share of the poor increase.

3.7 Conclusions

In this paper we distinguish between two measures of income inequality, the population share of the poor and the relative wealth of the poor. We discuss their different impact on the rate of innovation. Our results are established on the basis of a model by Zweimüller and Brunner (2005), but we do not assume the independence between the population share and the relative income. The relaxation of this assumption leads to the novel result that in separating equilibrium, the improvement of the relative income of the poor impedes the innovation rate, and a decrease of the population share of the poor accelerates the rate of innovation.

There are some important implications regarding our result. First, since the Gini-coefficient does not differentiate between the relative income of the poor and the population share of the poor, it is not suitable for policy recommendations. Each different measure of inequality has a different impact on economic growth. Second, the interdependent relationship between relative income and the population share is very important when considering the impact of inequality on growth. Finally, the effect of education on growth or innovation is not only due to an increase of productivity, which is discussed by most economists, but also due to an increase of the demand for better quality. The latter is almost neglected by most economists.

We believe that future research should be directed to empirical work bringing to focus the relationship between the relative income of the poor

and the education enrollment rate. We also need evidence to support the argument that education could produce rich consumers, through which the education enrollment is positively associated with the growth rate. Moreover, the current paper points out that there are possible different polices with different effects on the economic growth, e.g., the redistribution from the rich to the poor, and the public school. However, the more important question is under what conditions society would choose the one, which can achieve a higher economic growth rate, as well as a fair income distribution.

Chapter 4

Inequality and Growth: a Joint Analysis of Demand and Supply

4.1 Introduction

The relationship between a country's wealth inequality and its economic growth has been a major concern of economists for more than a century. Yet it is far from being well understood. In theoretical modelling, the distribution of wealth is the relevant inequality source. However, most empirical studies use income inequality data as a proxy for wealth inequality because of the scarcity of available data on the distribution of wealth.[15] "It is generally argued that this is unlikely to be a major problem since both measures of inequality generally vary together in cross-sections." (Aghion et al. 1999). In the current paper, initial wealth inequality coincides with income inequality through human capital investment.

The empirical evidence on this link is ambiguous. Some cross-country studies (e.g., Berg and Sachs 1988, Persson and Tabellini 1994, Alesina and

---

[15] There also are studies using other proxies. For instance, Alesina and Rodrik (1994) and Deininger and Squire (1998) include land inequality along with income inequality, Castelló and Doménech (2002) investigate human capital inequality.

Rodrik 1994, Clarke 1995) show that income inequality, as a proxy for wealth inequality, negatively impacts long-run growth rates. Nonetheless, there also is evidence that income inequality has a positive impact on short or medium run growth rates (Forbes 2000), and that the relationship between income distribution and the long-run growth rate is non-linear (Chen 2003, Banerjee and Duflo 2003). The ambiguous empirical results imply that there is not a clear relationship between income inequality and economic growth (Barro 2000b). Hence, it is important for economists to develop models which illustrate the possible different effects of inequality on economic growth under different circumstances. The existing theoretical wisdom has proposed either a negative or a positive relationship between initial wealth inequality and economic growth. Here it will be shown that both are extreme cases in an integrating simple model. We further the analysis of the relationship between wealth inequality and economic growth in two directions.

First, in a simple model with two types of individuals, the poor and the rich, the distribution of wealth comprises two variables, namely the relative wealth of the poor and the population share of the poor. We argue these variables may have different, even opposite effects on economic growth under certain conditions. Hence, cross-country evidence which is based on the simple regression of the Gini-coefficient on the economic growth rate can be ambiguous. In particular, we may be unable to obtain from such empirical studies recommendations on redistribution policies for achieving a higher economic growth rate as well as a more equal distribution.

Second, we combine the supply of production factors and the demand for the new quality goods in a general equilibrium model. Thus, wealth inequality in two areas can affect the economic performance: the supply side and the demand side. Most of the literature maintains that wealth inequality

reduces the aggregate human capital investment, given a neoclassical production function of investment and imperfect capital market. Consequently, inequality has a negative effect on the supply of consumption goods. We name this effect "the supply-side effect". The main arguments of the supply-side effect are included in the survey of Benabou 1996. On the other hand, following the literature on endogenous growth with quality-improving innovation (Aghion and Howitt 1998, Zweimüller et al. 2000) we argue that innovation is the engine that drives economic growth. This can improve the quality of goods and, in turn, increase the utility of consumers. The innovation cost is compensated by the monopolistic profit after successful innovation. Thus, the incentive of innovation is the monopolistic profit. Wealth distribution can affect the demand for the newly invented goods, and subsequently the price and profit of monopolist. We name this "the demand-side effect".

As we assume that there are only two types of individuals, the monopolistic supplier of newly invented goods can set the price either at the separating level, i.e. only the rich are able to buy it, or at the pooling level that even the poor can afford. Because wealth distribution has different effects on the profit in both cases, the relationship between inequality and economic growth is non-linear. Inequality may give rise to a higher incentive for firms to innovate because rich consumers can pay more than the poor for high quality goods. However, on the other hand, the relatively small market share of high quality goods implied by inequality impedes the spread of better quality goods.

This paper shows that in a separating equilibrium, a lower relative wealth of the poor is good for innovation, and a larger population share of the poor is bad for innovation. This result is consistent with Foellmi et al. (2002) and

the 3rd chapter. In Foellmi et al. (2002) hierarchic preferences[16] are introduced, and innovation induces new goods but does not improve quality. The 3rd chapter considers the interdependent relationship between the relative wealth of the poor and the population share of the poor. In the pooling equilibrium, the lower relative wealth of the poor is bad for innovation, and the population of the poor has no effect on innovation. The threshold value which distinguishes between these two equilibria depends on the strength of the supply-side effect. These findings imply that two nations with the same Gini-coefficient could have different economic growth rates if their wealth inequality is reached for different reasons (e.g., low relative wealth of the poor or large population share of the poor). Hence, it is important to decompose the Gini-coefficient in empirical research.

This paper integrates two main streams of theory relating growth and inequality. Recent surveys of the supply-side effect are by Benabou (1996) and Aghion et al. (1999), where three broad categories corresponding to the main feature are stressed: imperfect financial market, political economy and social unrest. The demand-side effect is illustrated Murphy et al. (1989), Foellmi et al. (2002) and Zweimüller et al. (2000 and 2005).

The rest of this paper is organized as follows: Section 4.2 discusses briefly the measurement of inequality. Section 4.3 lays out the basic framework. In section 4.4 we analyze the equilibrium and in Section 4.5 we give an example and present some empirical implications with section 4.6 concluding.

---

[16] "A hierarchy of wants implies that goods can be ranked according to their priority in consumption" (Foellmi et al. 2002)

4.2   The Measurement of Inequality

Since Corrado Gini, the Italian statistician, published his paper "Variabilità e mutabilità" in 1912, the Gini coefficient is widely used as a measurement of inequality. It is a number between 0 and 1, where 0 corresponds with perfect equality (everyone has the same wealth) and 1 means perfect inequality (one person has all the wealth; everyone else has nothing). The Gini index is the Gini coefficient expressed in percentage form, and is equal to the Gini coefficient multiplied by 100.

The Gini coefficient is calculated as a ratio of the areas on the Lorenz curve diagram. (see Figure 1(a)). If the area between the line of perfect equality and the Lorenz curve is A, and the area beneath the Lorenz curve is B, then the Gini coefficient is $\dfrac{A}{A+B}$.   The advantages of using the Gini coefficient are clear: It is both scale and population-independent, hence, it can be compared across countries and is easily interpreted; by retaining anonymity it doesn't matter who the high and low earners are; last but not least, it is simple. However, economies with similar wealth and Gini coefficients can still have very different distributions. This is because the Lorenz curves may have different shapes and yet yield the same Gini coefficient. As an extreme example, an economy where half the households have no wealth, and half share the wealth equally has a Gini coefficient of 0.5 (Lorenz curve *abd* in Figure 4.1(b)); but an economy with complete wealth equality except for one wealthy household that has half the total wealth also has a Gini coefficient of 0.5 (Lorenz curve *acd* in Figure 4.1(b)). In this paper, we address the question: Does the shape of Lorenz curve having the same Gini coefficient matter?
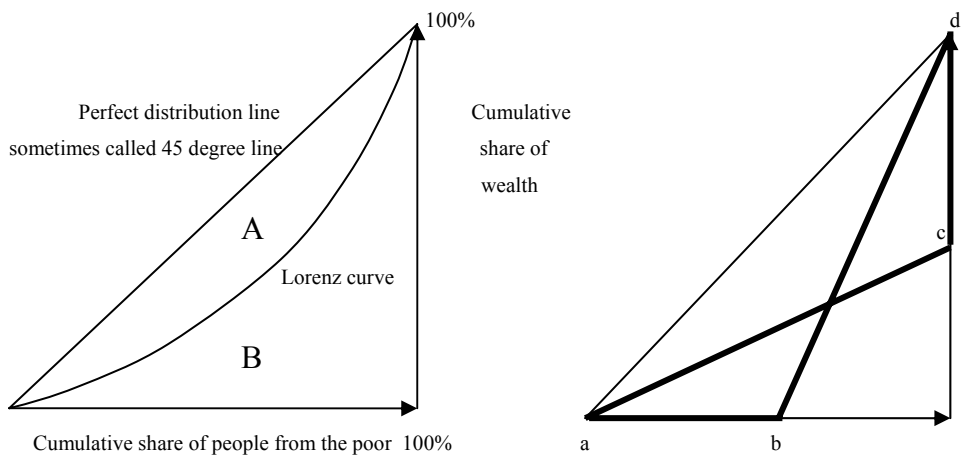
(a) Lorenz curve   (b) different shapes of Lorenz

curve with the same Gini coefficient

Figure 4.1 Lorenz curve and the Gini coefficient

4.3   The Model

We consider a closed economy with two types of individuals: the poor and the rich. They work for firms and consume products of firms. There are two kinds of goods: standard goods and quality goods. The quality improves over time due to innovation. Hence, the innovation rate represents the growth rate of quality, but not the growth rate of quantity. In turn, the economic growth is the growth of the consumers' utility, not the output.

4.3.1 The Environment

There is a continuum of individuals at each point in time, who live for two periods, young and old. Time is discrete, indexed by $t = 1,2,...,$. The population size of each generation is constant over time and normalized to 1. Individuals, within as well as across generations, are identical in their

preferences. However, they may differ in their family wealth and thus, due to the absence of perfect financial markets, in their capacity to invest in human capital. For simplicity, we assume that there are two kinds of individuals: the poor and the rich, their population shares being β ($0 < \beta < 1$) and $1 - \beta$, respectively. The average wealth of the whole society is denoted by V, which is the value of firms. Firms earn a flow profit and the value of firms equals the present value of this flow profit.[17] The poor individual has wealth $A_p = dV$, where d ($0 < d < 1$) means the wealth of the poor relative to the average level V. As a result the rich have $A_r = \dfrac{1 - d\beta}{1 - \beta} V$.[18] For simplicity, we assume that the wealth should not be eaten and can be transferred from generation to generation. Thus, there is no social mobility in this simple model. At birth, a young individual i receives an amount of dividend θA_i, where θ is a constant dividend rate. Therefore, the wealth distribution is equivalent to the distribution of the initial income of the young people.

Figure 4.2 shows the resulting Lorenz-curve. Given our assumptions, the Lorenz-curve is piecewise linear. The Gini-coefficient of the wealth, as well as that of the income of the young is $(1 - d)\beta$, (see Appendix 4.1). Both an increase in the population share of the poor and a decrease in relative wealth of the poor can increase the inequality level of the wealth. However, we claim that they have different effects on the economic growth.

---

[17] See section 4.3.3 and 4,3.4.

[18] According to the definition of the average wealth, $V = A_p\beta + A_r(1 - \beta)$. After substituting $A_p = dV$ and rearranging, we have $A_r = \dfrac{1 - d\beta}{1 - \beta} V$.

There are four sectors in the economy. The education sector produces skilled labor which is the only production factor and is expressed by the efficient labor unit denoted by L. The education sector is run as a non-profit organization. It collects an education fee H from young individuals and hires S efficient labor units from old individuals to teach. The more that young individuals invest in the education sector, the more efficient labor units of old generation will be employed as teachers. As a result, more efficient labor units can be produced for the next period. Following education, young individuals have L units of efficient labor, which can be used in four sectors when individuals are old.
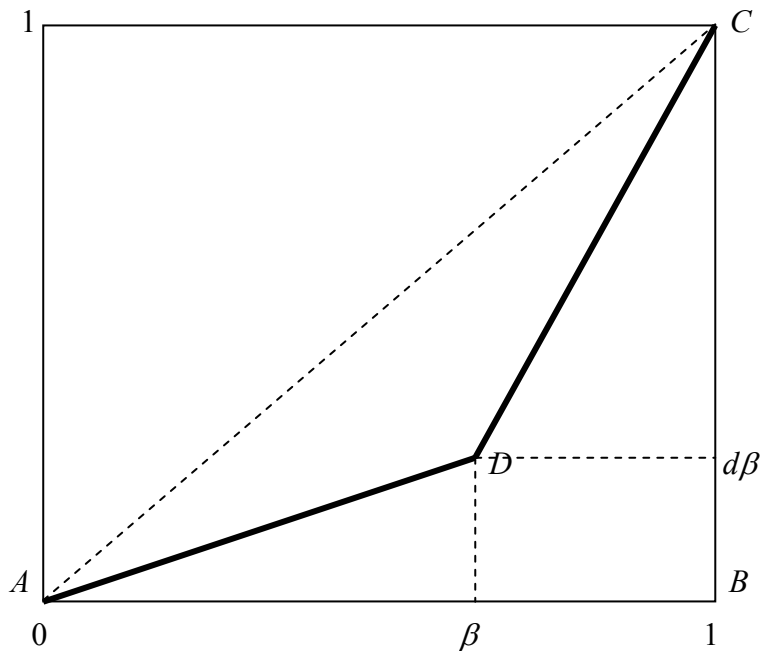
Figure 4.2 The wealth distribution

Two production sectors produce two kinds of goods, referred to as standard goods and quality goods, respectively. Let x be the quantity of the standard goods, which has a constant quality (normalized to 1) and is traded in a competitive market. Hence, the price P_x is equal to its marginal cost which

is also normalized to 1. The marginal cost of the standard goods can be expressed as wb, where the unit labor demand is b. This determines how many units of efficient labor are needed to produce one unit of standard goods. w is the wage rate of the efficient labor unit. We get $P_x = wb = 1$.

In the quality goods sector, one monopolist produces the best quality goods. Anyone is allowed to produce competitively any other quality goods. We denote the quality level as $q_j \quad j = 0, -1, -2, ...$, where q_0 is the best quality, q_{-1} is the second best one and so on. Furthermore, we assume for simplicity: $q_j = kq_{j-1} \quad k > 1$. Despite the different qualities, the quality goods have the same marginal cost wa, where a is again the unit labor demand. Since all quality goods except the best quality are sold on a competitive market, they have the same price $P_j = wa, \quad j = -1, -2, ...$, and the monopolist sets P_0 to maximize her profit. For convenience, we assume that every consumer can consume one and only one unit of quality goods.

The new quality good is invented by the research sector. The research sector needs n units of efficient labor to achieve the innovation rate, ϕ, which is the probability of success. Each successful innovation introduces a k-times better quality good than the existing best one in the next period. The authority to produce this best quality will be sold to one monopolist. After successful innovation the current best quality becomes the second best quality in the next period. Hence, any competitor can produce it. Since in equilibrium the amount of consumption is constant, the economic growth throughout this model is not the growth of quantity, but of quality. Here the innovation rate ϕ represents this growth rate of quality. Later on, we will see that the growth rate of quality also is that of the consumers' utility.

The assumption of two kinds of goods, one with constant quality and the other with constant quantity, is an abstract of two dimensions of consumption. In reality the quality of each of the goods can be improved and there is no limit that consumers can only consume one unit. However, we can still find goods whose quality consumers readily appraise: automobiles for instance. Normally we have one car. However, we sometimes buy a new, better quality car, and sell the old one in order to improve our utility. There are other goods about whose quantity consumers also readily appraise, for example, leisure.

4.3.2 The Household's Decision Problem

As we assumed in last section, a young individual i has initial income θA_i at birth which can be used to buy standard goods[19] x_i^1 at the price 1 and invest in education H_i. The production function of efficient labor is $L(H_i) = l + H_i^\alpha, \quad (\alpha < 1)$[20], where l ($l > 0$) is constant and represents the basic supply of labor without any education. The H_i^α are the efficient labor units produced by the human capital investment, H_i, which is the choice variable of the young individual i. This production function is a strictly concave increasing function satisfying the neoclassical boundary conditions, and $L(0) = l$. For simplicity, we assume that l is equal to the unit labor

---

[19] This assumption ensures only two kinds of consumers in the quality goods market. Allowing young people to be able to buy quality goods will not change our result qualitatively, but complicate the analysis, because then there are four types of consumers in the quality goods market. This assumption also is reasonable. For example, we can imagine the quality goods to be automobile, alcohol and/or cigarette, which are prohibited for young people.

[20] It is a closed form of $L = l + l'S^\alpha$, where l' is a constant parameter and S the labor units of teachers. Since the education sector has no profit, H is totally paid for teaching. Hence, $H = Sw$. For simplicity, we assume $l' = w^\alpha$. Thus, the closed form for the production function of labor is $L = l + H^\alpha$.

demand of quality good: $l = a$, which simplifies the calculation without loss of generality. Hence,

$$L(H_i) = a + H_i^{\alpha}, \quad (\alpha < 1) \qquad (4.1)$$

In the second period, old individual i has L_i units of efficient labor. We assume a simplistic view regarding the production of consumption goods. Efficient labor is the single productive factor, and every individual inelastically supplies all of her efficient labor units to the competitive labour market. As a result, incomes of the poor and the rich in the second period are respectively: $y_i = wL_i$, $i = p,r$. In section 4.4, we will show that the poor invest less than the rich. Hence, $L_p < L_r$, in turn, $y_p < y_r$. It means that there is no social mobility in our simple model.

We assume the instantaneous utility function in the first period to be $u_i^1 = \ln x_i^1$. Because the standard good is the single good which young people can consume. Substituting the budget constraint in the first period $\theta A_i = x_i^1 + H_i$, we have:

$$u_i^1 = \ln(\theta A_i - H_i) \qquad (4.2)$$

There is no saving for the old individual. All income is spent both on the consumption of the standard good and the quality good. Every individual can consume one and only one unit of the quality good q_j. There is no limit to the consumption of the standard good x_i^2 except for the budget constraint, i.e., $y_i = 1 \cdot x_i^2 + P_j \cdot 1 \quad j = 0,-1,...$, where the price of standard goods is 1, the quantity of standard goods is x_i^2 and the price of the quality j is

denoted by P_j. The preference for consumption of the standard good and the quality good is given by the following utility function:

$$u_i^2(x_i^2, q_j) = \ln x_i^2 + \ln q_j \qquad i = p, r \text{ and } j = 0, -1 \qquad (4.3)$$

By substituting the budget constraint in the second period, (4.3) can be expressed as:

$$u_i^2 = \ln(y_i - P_j) + \ln q_j \qquad (4.4)$$

The life-time utility function of individual i is assumed to be:

$$U_i = u_i^1 + \rho u_i^2 \qquad (4.5)$$

where ρ is the subjective discount factor. It can, but need not necessarily, be equal to $\dfrac{1}{1+\theta}$, where θ is the dividend rate. The old individual i chooses the quality level q_j to maximize u_i^2, given income y_i being constant. By backward induction, when the subject is young she chooses H_i to maximize her life-time utility (4.5) with the rational expectation that q_j will be optimally determined in the second period. Hence, in order to solve the household's decision problem, we need to know the price of the quality good.

4.3.3 The Pricing Decision of the Monopolist

Firms have all the above information but are unable to distinguish between individuals based on income. The strategy which firms can pursue is to

choose a price while quality is fixed. We concentrate only on the steady state where prices are constant over time. First of all, only the most recent old quality good (q_{-1}) can be sold at the price wa in the competitive market of quality goods q_j $j < 0$. Hence, the price that the monopolist can offer has to satisfy the condition:

$$\ln(y_i - P_0) + \ln q_0 \geq \ln(y_i - wa) + \ln q_{-1} \qquad (4.6)$$

The left hand side of (4.6) is the utility when individuals buy the best quality good q_0 and the right hand side is the utility when they consume the second best quality good q_{-1}. Further, we assume that the consumer prefers better quality goods if both quality goods yield the same utility. Substituting $q_0 = kq_{-1}$ and rearranging (4.6), we get the highest price $\overline{P_0}$ of the best quality good:

$$\overline{P_0} = (1 - \frac{1}{k})y_i + \frac{wa}{k} \quad i = p, r \qquad (4.7)$$

The monopolist thereby has two possible price strategies -- either to set the price high, to attract only the rich consumers (separating price), or, low to occupy the entire market (pooling price). The instantaneous profits are as follows:

$$\pi^{sep} = (1 - \beta)(1 - \frac{1}{k})(y_r^{sep} - wa) \qquad (4.8)$$

$$\pi^{pool} = (1 - \frac{1}{k})(y_p^{pool} - wa) \qquad (4.9)$$

The monopolist sets the separating price in steady state, if 1) given the separating strategy before, she has no incentive to deviate, which means:

$$(1-\beta)(H_r^{sep})^\alpha \geq (H_p^{sep})^\alpha \qquad (4.10)$$

2) the profit of separating strategy in steady state is larger than that of the pooling, viz.:

$$(1-\beta)(H_r^{sep})^\alpha \geq (H_p^{pool})^\alpha \qquad (4.11)$$

Since the supplier of the best quality goods is monopolistic, it has a positive flow profit. All other firms have zero profit and their value also is zero. All firms are owned by individuals. Hence, the value of this monopolistic firm is equal to V.

4.3.4 Innovation

As mentioned earlier, the quality improves over time due to innovation. Following the work by Aghion and Howitt (1992), we assume that the innovation is random and arrives according to a Poisson process with parameter ϕ. The researcher can employ n units of efficient labor to reach the Poisson arrival rate ϕ, i.e., $\phi = \lambda n$, where λ is the productivity of efficient labor in research. Hence, the flow of research cost is wn. This assumption of innovation means that the success of research depends only on current input, not upon past research.

Once innovation succeeds, a new quality good is invented. This newly invented good is k-times better than the current best quality good, and can

be produced by one monopolist in the next period. The authority to produce this new best quality good is sold to the monopolist via a simple auction. We assume that researchers prefer to sell the authority to the incumbent as long as its offer is at least the same as that of others. In order to keep this priority, the incumbent has to buy the new innovation from researchers at a price which is equal to the present value of the future monopolistic profit. Thus, we have a single monopolist who produces the best quality in every period. The price paid by the monopolist to the research sector is the flow of research benefit, ϕB, where ϕ is the probability of success and B is the present value of the future monopolistic profit:

$$B = \sum_{t=1}^{\infty} \left(\frac{\pi}{(1+\theta)^t} \, prob\{no\ innovation\ before\ t\} \right) = \sum_{t=1}^{\infty} \left(\frac{\pi(1-\phi^e)^{t-1}}{(1+\theta)^t} \right)$$

Leading to
$$B = \frac{\pi}{\phi^e + \theta} \qquad\qquad (4.12)$$

where t is a time index, $\phi^e = \lambda n^e$ is the expected future arrival rate of innovation, n^e is the expected future number of efficient labor units in the research sector, and θ is interest rate. The sum of the interest rate and the innovation rate is the discount factor of the monopolistic profit. In steady state, all agents have perfect foresight. Consequently, $\phi = \phi^e$ (or, $n = n^e$).

Now we are in a position to define the average wealth of the whole society V. As we mentioned before, the average wealth is the value of monopolistic firm, which can generate dividends θV in each period. Hence, the per period increase in the average wealth is the monopolistic profit net of the dividend and the payment to the researcher.

$$\Delta V = \pi - \theta V - \phi B \qquad\qquad (4.13)$$

4.4 Equilibrium

According to section 4.3.3 there are two possible equilibria, namely separating and pooling respectively. If the monopolist chooses the separating strategy, then the poor buy q_{-1} and the rich consume q_0. Hence, the rich young people maximize their life-time utility as follows:

$$\max_{H_r} U_r = \ln(\theta A_r - H_r) + \rho(\ln(y_r(H_r) - P_0) + \ln q_0)$$

substituting (4.7) in this equation and solving the first order condition, we have:

$$H_r^{sep} = \frac{\alpha\rho}{1+\alpha\rho}\theta A_r^{sep} \qquad\qquad (4.14)$$

Similarly, $\qquad\qquad H_p^{sep} = \frac{\alpha\rho}{1+\alpha\rho}\theta A_p^{sep} \qquad\qquad (4.15)$

If the monopolist chooses the pooling price, then the poor set the optimal investment at:

$$H_p^{pool} = \frac{\alpha\rho}{1+\alpha\rho}\theta A_p^{pool} \qquad\qquad (4.16)$$

103

$\dfrac{\alpha\rho}{1+\alpha\rho}$ is the saving rate of the young people. The results (4.14) - (4.16), consistent with Bénabou (1996), reflect the fact that the poor invest less in human capital than the rich. Due to the neoclassical production function of human capital investment, (4.1), the marginal productivity of the human capital investment of the poor is higher than that of the rich. Hence, the inequality of initial wealth reduces the aggregate supply of efficient labor units. This is the negative supply-side effect.

Substituting (4.14) and (4.16) in (4.8) and (4.9), we have

$$\pi^{sep} = (1-\beta)(1-\frac{1}{k})w(\frac{\alpha\rho}{1+\alpha\rho}\theta A_r^{sep})^{\alpha} \qquad (4.17)$$

$$\pi^{pool} = (1-\frac{1}{k})w(\frac{\alpha\rho}{1+\alpha\rho}\theta A_p^{pool})^{\alpha} \qquad (4.18)$$

The instantaneous profit of the monopolist in the separating equilibrium depends on the initial wealth of the rich young people. Analogously, the profit in the pooling equilibrium depends on the wealth of the poor.

Furthermore, we assume free entry in the research sector, which is the traditional assumption of the quality-improving model, to obtain the research arbitrage equation (Aghion and Howitt 2004). Hence, $wn = \phi B$, where wn is the flow cost of the research sector and ϕB is the flow benefit (see section 4.3.4). This leads to:

$$\frac{w}{\lambda} = \frac{\pi}{\phi+\theta} \qquad (4.19)$$

The underlying intuition is similar to Aghion and Howitt (1992). The left hand side of equation (19) represents the flow cost of research in order to achieve a successful innovation, which decreases in the productivity of research workers λ. The effect of λ on ϕ is positive because the researcher is able to achieve a higher innovation rate with the same number of efficient labor units if their productivity increases. The effect of the interest rate is ambiguous. First, it is a discount factor. Hence, the higher θ, the lower the research benefit. Therefore, the innovation rate decreases in the interest rate. The other way through in which the interest rate can affect the innovation rate is, by raising the initial income of individuals. Hence, the higher θ, the larger the human capital and consequently, the larger the monopolistic profit. It has a positive effect on the innovation rate.

As the single production factor, the supply of efficient labor units should be equal to the demand for efficient labor units in equilibrium. The total efficient labor supply is $\beta L_p + (1-\beta)L_r$, which is equal to $a + \beta H_p{}^\alpha + (1-\beta)H_r{}^\alpha$. The demand for efficient labor consists of four parts. First, the research sector needs n. Second, the quality goods sector needs a because every consumer consumes one unit of quality good. Third, the standard goods sector needs $b(x^1 + \beta x_p^2 + (1-\beta)x_r^2)$. And finally, the education sector needs S. Hence, the total demand for efficient labor units is $n + a + b(\beta x_p + (1-\beta)x_r) + S$. In equilibrium, the labor market clearing condition is as follows:

$$\beta L_p + (1-\beta)L_r = n + a + b(\beta x_p + (1-\beta)x_r) + S \qquad (4.20)$$

Solving (4.20) yields $\qquad \pi = wn + \theta V \qquad\qquad (4.21)$

Proof: see Appendix 4.2.

From (4.19), (4.21) and $\phi = \lambda n$ we know that the average wealth $V^* = \dfrac{w}{\lambda}$ in equilibrium regardless of the price strategy of the monopolist. The higher the wage rate, the greater is the average wealth. This is because the high wage rate involves the rich consumer (recalling that the old people are the consumer of quality goods, their income is given by $y_i = wL_i$ $i = p,r$, which depends on the wage rate). Then the monopolist can set a high price and earn more profit (see equations 4.17 and 4.18). The larger λ, the higher is the innovation rate. Thus, the value of the monopolistic firm is less.

After substituting (4.14), (4.15) and (4.16) into (4.10) and (4.11), and using $V^* = \dfrac{w}{\lambda}$, we get the unique condition of the separating price in equilibrium:

$$(1-\beta)^{1-\alpha}(\frac{1}{d}-\beta)^\alpha \geq 1 \qquad (4.22)$$

This condition shows that the larger the population share of the poor, and/or the richer the poor, the less probable will the monopolist choose the separating price strategy. The larger the α, the bigger the difference of income of old individuals. Hence, more probably will the separating price be chosen.

Rearranging (4.19) and substituting (4.17) and (4.18), we have two possible innovation rates in the separating and the pooling equilibria respectively:

$$\phi^{sep} = \lambda(1-\frac{1}{k})(1-\beta)(\frac{\alpha\rho}{1+\alpha\rho}\theta A_r)^\alpha - \theta \qquad (4.23)$$

106

$$\phi^{pool} = \lambda(1-\frac{1}{k})(\frac{\alpha\rho}{1+\alpha\rho}\theta A_p)^\alpha - \theta \qquad\qquad (4.24)$$

where $A_r = \dfrac{(1-d\beta)w}{(1-\beta)\lambda}$, $A_p = \dfrac{dw}{\lambda}$.

Proposition 4.1

The effect of wealth inequality on the innovation rate is non-linear and ambiguous:

1) *Given β constant, the effect of d on ϕ is negative for $d \in [0, d^*]$ and positive for $d \in [d^*, 1]$. The threshold value $d^* \in (0,1)$ satisfies*

$$(1-\beta)^{1-\alpha}(\frac{1}{d^*}-\beta)^\alpha = 1.$$

2) *Given d constant, the effect of β on ϕ is negative for $\beta \in [0, \beta^*]$. In the pooling case $\beta \in [\beta^*, 1]$, β has no effect on ϕ. The threshold value $\beta^* \in (0,1)$ satisfies $(1-\beta^*)^{1-\alpha}(\frac{1}{d}-\beta^*)^\alpha = 1$.*

The non-linear relationship between initial income inequality and economic growth has two interpretations in the current model: For one, d and β have different effects on the innovation rate. For the other, both the effect of d and that of β on ϕ are non-linear. Inequality can affect the innovation rate not only through the supply of the production factor (here, labor) but also the demand for the new better quality. The supply-side effect is discussed by most economists. Here, we assume the strictly concave increasing production function of labor and an imperfect capital market as in the literature; hence, the negative effect of inequality on growth is not surprising (see Appendix 4.3). The parameter α is a measure of the strength of the supply-side effect.

Figure 4.3 shows different effects of the relative wealth of the poor on the innovation rate in two extreme cases. Both are the examples where the supply-side effect disappears. Suppose $\alpha \to 0$, then the saving rate of the young people ($\frac{\alpha\rho}{1+\alpha\rho}$) approaches zero. Both the poor and the rich young people have little incentive to invest in human capital. Hence, the difference in income for the old people approaches zero. The threshold value $d^* \to 0$. The monopolist faces a more equally distributed society and thus sets the pooling price. Consequently, the income of the poor is crucial for the price of the quality good. In this case, if the poor have more income, then the price of the quality good increases. Finally, the innovation rate increases. The effect of d on ϕ is overall positive in the case of (a).

(a) $\alpha \to 0$ (b) $\alpha \to 1$

Figure 4.3 The pure demand-side effect of d on ϕ

The picture is reversed, if $\alpha \to 1$, $d^* \to \frac{1}{1+\beta}$, i.e., $Gini^* \to \frac{\beta^2}{1+\beta}$. Since

$\frac{1}{1+\beta} > \frac{1}{2}$, we can argue that d has negative effect on the innovation rate

over the most range through the demand side. If the condition (4.22) holds, then the poor are too poor and/or the population of the poor is too small. Hence, the monopolist sets the separating price to sell the best quality good only to the rich. In this case, if the rich become poorer and the poor become richer (d increases), i.e., if the Gini-coefficient decreases given constant β, then this inequality brings about less incentive for the researcher to innovate because of falling profits. If d increases further and exceeds the threshold value $\dfrac{1}{1+\beta}$, the monopolist sets the pooling price and then d has a positive effect on the innovation rate ϕ. This is case (b).

Contrary to d, the population of the poor β has a different effect on the innovation rate. In the case of the separating price, if the Gini-coefficient increases because the population of the poor β increases given constant d, then the inequality leads to a small market size for the quality good. Hence, the monopolistic firm has less profit, and the innovation rate decreases. If a country has a relatively even initial income distribution ($\beta \in [\beta^*, 1]$) then the monopolist sets the pooling price. Since the market of the quality good is the whole society, the population share of the poor does not affect the innovation rate.

What is the impact of wealth inequality (β or d, respectively) on utility? From (4.2) and (4.3) we have:

$$
\begin{aligned}
\Delta u^1 &= \frac{\Delta x^1}{x^1} \\
\Delta u^2 &= \frac{\Delta x^2}{x^2} + \frac{\Delta q}{q}
\end{aligned}
\qquad (4.25)
$$

In a steady state, the consumption of standard goods is constant ($\Delta x = 0$), and $\Delta q = \phi(k-1)q$. Hence, we have $\Delta U = \rho\phi(k-1)$. The higher the innovation rate, the larger the increase in the utility is. Redistribution from the rich to the poor (d increases) decreases the wealth inequality, hence, the aggregate supply of efficient labor increases. This is the supply-side effect. What is the demand-side effect of this redistribution? If $d < d^*$, the monopolist sets separating price. Redistribution leads to a decrease in the initial wealth of the rich, in turn, a less monopolistic profit. Consequently, the research sector employs less efficient labor units. Recalling that the quality good sector needs always a units of efficient labor and the education sector needs same efficient labor as long as the aggregate investment of human capital keeps constant, more efficient labor units are shifted from the research sector to the standard good sector. This reallocation of efficient labor among different sectors is the demand-side effect. Sum up, consumers enjoy a higher utility level in the short run, but the long run growth rate of the utility is lower than before because of a lower innovation rate. If $d > d^*$, we have a pooling equilibrium. In contrast to the separating case, redistribution from the rich to the poor can induce a higher price of quality goods and more monopolistic profits. Consequently, the research sector has a higher incentive to employ more efficient labor units and a higher innovation rate will be achieved. It is not a priori clear whether consumers have more or less consumption of standard goods in a new pooling equilibrium. It depends on which effect is dominant, the supply-side effect or the demand-side effect.

4.5 An example: China

Because d and β have different effects on the innovation rate, in particular, their effects offset each other in the separating equilibrium, the

Gini-coefficient has no overall effect on economic growth. In this sense, it is important for us to decompose the Gini-coefficient in the empirical research. The different effects of the relative poorness and the population share of the poor imply the different policy recommendation. In a country where the separating equilibrium is overwhelming and the goal of government policy is to achieve both an increase in economic growth and a decrease in inequality, one should consider decreasing the population share of the poor but not redistributing from the rich to the poor.

Chinese experience in the last two decades bears witness to this prediction. In China, the disparity between urban and rural residents is assured by the Chinese household registration (Hukou) system, (Yang and Zhou 1999). Lacking free migration between urban and rural areas, the Chinese government has invested more in public goods such as education, social insurance and infrastructure, in the cities than in the rural areas since 1949. This can be stylized by assuming V to be the public social wealth.[21] The government implements an urban-biased redistribution policy, (Yang 1999). Hence, the urban resident is rich and the rural resident is poor. The goal of Chinese reform above all is to have a high economic growth rate. Government can control both the population share of the poor through the Hukou system and the relative poorness of the poor through the redistribution policy.

After the 1980s, this Hukou system was relaxed. However, it has never been abandoned. As a result the urban population (the rich) increased dramatically from 21% in 1982 to 36% in 2000. At the same time, the

---

[21] It reflects the central planning economy in China before reform, at which time almost all firms were state-owned. Hence, the Chinese government did have the power to distribute V between urban and rural. Since 1980, the power of this distribution diminishes as more and more firms went private. However, many state-owned firms remain, particularly, in the monopolistic branches and capital intensive industries.

relative income of rural residents (y_p/y) decreased from 0.76 (1980) to 0.61 (2000), (China Statistical Yearbook 2002). Combining these results, Chinese firms have a great incentive to invent better quality goods and set prices at the separating level. The evidence for separating price strategy lies in the fact that the most new and better quality goods are sold in Chinese cities. According to the China Statistical Yearbook 2002, Chinese average growth rate of GDP per capita was approximately 9% over the last 20 years. Although there are many reasons for the rapid growth, we cannot deny that one of them is the high demand for the better quality goods.

4.6   Conclusion

This paper investigates the ambiguous relationship between wealth inequality and economic growth in a framework of a quality-improving growth model. Our contribution is to enhance the analysis of this relationship in two ways. First, we argue that the Gini-coefficient, used by most empirical research in this branch, can include too many variables which have diverse effects on economic growth. Therefore, we need to decompose the Gini-coefficient into different variables. The current model supplies an example that divides the Gini-coefficient into the relative wealth of the poor and the population share of the poor. We have shown that they induce a contradictory effect under certain conditions. This result indicates that we need to investigate not only the Gini-coefficient, but also the shape of wealth distribution. The empirical research on the base of the Gini-coefficient cannot generate a clear relationship between wealth distribution and economic growth. In particular, we may be unable to draw from such simple empirical studies recommendations on redistribution policies for

achieving a higher economic growth rate as well as a more equal income distribution.

Additionally, we have combined two sides of the market within one simple model: the supply of production factors and the demand for the consumption goods. Thus, in this model, there are two different channels, by which wealth distribution can affect economic performance. Whereas the supply-side effect of wealth inequality is negative on the economic performance, the demand-side effect could be positive under certain condition. Hence, there is non-linear relationship between the wealth inequality and economic growth. This result is partly consistent with the empirical evidence (Chen 2003), although he uses the Gini coefficient, but not other variables which we investigate.

Chapter 5

Final Remarks

This thesis emphasizes two important sources of economic growth: political institutions and inequality. Since each essay has been summarized in the introductory chapter, this final chapter briefly addresses some points for future research.

Although national leadership plays a crucial role in economic development (Jones et al. 2005), we know little about the precise mechanisms. Chapter 2 provides a possible way to think about the behavior of dictators in developing economies. However, the model is quite simple. We have not distinguished between human capital and physical capital investments. Hence, it is impossible to know from chapter 2 why the Chinese government is willing to invest so much in infrastructure, but not in education. It also is a very simple assumption that dictators can influence the economy only via social transfers. Barro (2000a) points out the importance of the rule of law to economic development. While social transfers can be interpreted as public investments and/or subsidy to private investments, the rule of law is a commitment to ensure the return of private investments. Although the good dictators in East Asia implemented the rule of law, economists have little theory to explain why dictators can keep their promise (Acemoglu et al. 2000).

Another way to understand the behavior of dictators is to incorporate more variables into their utility function. As a national leader, a dictator does have some political interest, e.g., independence of the nation, her status in history. The question is why she is concerned about these and how these variables influence economic performance. In this sense modern behavioral economics might be helpful in understanding the functionality of a dictatorship.

It also is of interest to compare economic performance in both dictatorial and democratic countries. With a fast growing economy, China, as the biggest dictatorial developing country, showed somewhat of an advantage compared to the greatest democratic developing country, India, in the 1980s and 1990s. Does this imply that a dictatorship is better off than a democracy when considering developing economies? Since India began to grow fast in the last 5 years, the question arises again: does democracy help India to catch up with China? Can democratic India grow under more stable conditions and thereupon benefit the poor more than China? In order to understand the fast growth in China and India, we need more theories with regard to political institutions and economic growth.

Wealth inequality is another important source of economic growth. While traditional economic research focuses on the supply-side effect of inequality, we stress the demand-side effect. However, this is not the whole picture. Inequality can also affect political institutions, and in turn, economic performance. The interaction between political institutions and inequality will be helpful to understand the political transition and economic growth.

Last but not least, we need more empirical evidence to support our predictions. Some important findings are testable. For instance, both the

population share of the poor and the relative wealth of the poor negatively affect economic growth. Another interesting point is to test in an econometric model whether dictatorial institutions induce a higher variance in growth rates. These aspects are all important for future research.

Appendix

Appendix 2.1:

The payments of citizens whose ability over $\hat{a}^{good}$ are:

$$P_{at}^{good} = \mathrm{A}_t N\lambda\bar{\tau}a - s^{good} = A_t(N\lambda\bar{\tau}a - \frac{e - N(\lambda-1)(1-\bar{\tau})^2}{2-\bar{\tau}})$$

In order to determine the political attitude of this group of citizens, we should check whether $\mathrm{A}_t N\lambda\bar{\tau}a - s^{good}$ is positive or not.

For the person with ability 1, the payment is:

$$P_{1,t}^{good} = A_t(N\lambda\bar{\tau} - \frac{e - N(\lambda-1)(1-\bar{\tau})^2}{2-\bar{\tau}}) \geq A_t(N\lambda\bar{\tau} - \frac{N(\lambda-1) - N(\lambda-1)(1-\bar{\tau})^2}{2-\bar{\tau}}) = A_t N\bar{\tau} > 0$$

Hence, the citizen with ability 1 always supports democracy.

For the person with ability $\hat{a}^{good}$, the payment is:

$$P_{\hat{a}^{good},t}^{good} = A_t(N\lambda\bar{\tau}\hat{a}^{good} - \frac{e - N(\lambda-1)(1-\bar{\tau})^2}{2-\bar{\tau}}) = A_t \frac{e(\lambda\bar{\tau} - \lambda + 1) + N(\lambda-1)(1-\bar{\tau})(\lambda + \bar{\tau} - 1)}{(\lambda-1)(2-\bar{\tau})}$$

If conditions $\lambda > \dfrac{1}{1-\bar{\tau}}$ and $e > \dfrac{N(\lambda-1)(1-\bar{\tau})(\lambda+\bar{\tau}-1)}{\lambda-\lambda\bar{\tau}-1}$ are satisfied, $P_{\hat{a}^{good},t}^{good}$

is negative. It implies that there is $a^* \in (\hat{a}^{good},1)$, $\forall a \in (\hat{a}^{good},a^*)$ $P_{at}^{good} < 0$

and $\forall a \in [a^*,1]$ $P_{at}^{good} \geq 0$. Hence, the citizen $\forall a \in (\hat{a}^{good},a^*)$ becomes the

supporter of the good dictatorship under conditions $\lambda > \dfrac{1}{1-\bar{\tau}}$ and

$e > \dfrac{N(\lambda-1)(1-\bar{\tau})(\lambda+\bar{\tau}-1)}{\lambda-\lambda\bar{\tau}-1}$.

Appendix 2.2:

One reason to assume an equally distributed social transfer in democracy is that, we cannot know *a priori* who constitutes the majority. Theoretically, 50% of the population plus one individual could make up the majority, who support the social transfer policy only benefiting them. However, some may doubt whether the result of this paper is sensitive to the assumption of equally distributed social transfer in a democracy. This appendix shows us that results do not change qualitatively, if we assume that the social transfer policy in a democracy is same as that in dictatorship, i.e., the individual who invests gets social transfer.

Since my model is based on a trade-off between the short run benefit and long-run costs for the dictator, we need to show that citizens, in the aggregate, still have incentives to revolt under the new group-specific social transfer policy. I.e., democracy is still better than a dictatorship.

Analogously, the median vote maximizes his income.

$$\underset{\tau}{Max}\ Y_{0.5,t} = 0.5 A_t N \lambda^{I_{0.5}}(1-\tau) + I_{0.5}(s - eA_t)$$

$$s.t. \quad s = \frac{\tau \int_0^1 y_{at} da}{1 - \hat{a}} = \frac{0.5\tau A_t N(\hat{a}^2 + \lambda - \lambda \hat{a}^2)}{1 - \hat{a}}$$

If $\hat{a} > 0.5$, i.e., the median voter doesn't invest, his maximization problem reduces to:

$$\underset{\tau}{Max} Y_{0.5,t} = 0.5 A_t N(1 - \tau)$$

The first order condition is: $\quad \dfrac{\partial Y_{0.5,t}}{\partial \tau} = -0.5 A_t N < 0$

Hence $\tau^{dem,1} = 0$, $s^{dem,1} = 0$ and $\hat{a}^{dem,1} = \dfrac{e}{N(\lambda - 1)} < \hat{a}^{good}$.

If $\hat{a} \leq 0.5$, the median voter invests. His maximization problem is then:

$$\underset{\tau}{Max} Y_{0.5,t} = 0.5 A_t N\lambda(1 - \tau) + \frac{\tau A_t N(\hat{a}^2 + \lambda - \lambda \hat{a}^2)}{2(1 - \hat{a})} - e A_t$$

The first order condition is: $\quad \dfrac{\partial Y_{0.5}}{\partial \tau} > 0$

Hence, $\tau^{dem,2} = \bar{\tau}$ and $s^{dem,2} > 0$. Because there is no taxation, the individual who invests gets more social transfer in a democracy than in a dictatorship. Thus, $\hat{a}^{dem,2} < \hat{a}^{good}$.

Summarizing, in a democracy the investment ratio is always greater than in a dictatorship. I.e., the citizens in aggregate can earn more in democracy. Hence, they are willing to revolt if possible. The dictator must face the

trade-off between the short run economic benefit and the earlier democratization in the long-run. Our result would not change qualitatively.

Appendix 3.1

First, it is easy to see that both qualities cannot be sold to the same consumer because of the assumption that consumers will choose the better quality if both generate the same utility. There are two possible cases: the best quality good is sold either to the rich or to the poor. In the first case the second best good cannot be sold to the rich but to the poor. This is the separating equilibrium. The other case is the pooling equilibrium, in which the best quality good is accepted by the poor, i.e., the utility of the poor from consuming the best quality good is larger than that of the second best quality:

$$\ln(y_p - e - P_0) + \ln q_0 \geq \ln(y_p - e - P_{-1}) + \ln q_{-1}$$

$$\Leftrightarrow (k-1)(y_p - e) - P_0 k + P_{-1} \geq 0 \qquad \text{and} \quad y_r > y_p$$

$$\Rightarrow (k-1)(y_r - e) - P_0 k + P_{-1} > 0$$

$$\Leftrightarrow \ln(y_r - e - P_0) + \ln q_0 > \ln(y_r - e - P_{-1}) + \ln q_{-1}$$

The rich prefer the best quality good to the second best one, too. Hence, in the pooling equilibrium the second best quality good is not sold.

Appendix 3.2

1. Pooling: Given the price of q_{-1} the firm of q_0 will charge his price as high as possible. Hence, the possible equilibria lie on the line CD of Figure

3.1. Suppose P_0 is higher than $\left(1-\dfrac{1}{k}\right)(y_p - e) + \dfrac{wa}{k}$, then firm q_{-1} can charge a price higher than marginal cost and attract all poor consumers. It is the separating case. This contradicts the pooling assumption. Hence, the single stage pooling equilibrium is $P_0 = \left(1-\dfrac{1}{k}\right)(y_p - e) + \dfrac{wa}{k}, P_{-1} = wa$. We should also consider if other possible equilibria can be sustained through any punishment in a repeated game. Because here the lowest profit which firm q_{-1} can earn is zero, it is impossible to punish him because what the firm has in equilibrium is also zero. Hence, above stage equilibrium is also the equilibrium for the whole repeated game.

2. Separating: The best-reply function of the best quality firm is $\overline{P}_0 = \left(1-\dfrac{1}{k}\right)(y_r - e) + \dfrac{P_{-1}}{k}$, which is the line AB in Figure 3.1. For the poor the utility if he consumes q_{-1} is strictly greater than that if he consumes q_0 given above best-reply function. It implies that firm q_{-1} has an incentive to increase its price without losing its consumers. Hence, the single stage equilibrium is point B in Figure 1. However, for the whole repeated game, other points on AB can also be sustained as equilibria because the firm q_0 can punish the other to set the pooling price (then firm q_{-1} can earn only zero profit) in future if firm q_{-1} increases its price in this stage. Hence, theoretically there are many possible separating equilibria. But such punishment is in some sense unrealistic because the deviation in P_{-1} is not able to decrease the profit of the firm q_0. Hence, under the assumption that the deviation will not be punished if such deviation does not affect other's profit, we have a single separating equilibrium B.

Appendix 3.3

Denote ED as follows: $ED = \rho \ln\left(\dfrac{k\theta V \dfrac{1-d}{1-\beta}}{w + \theta dV - e - wa} + 1\right) - \ln e$.

Hence, $\left.\dfrac{\partial \beta}{\partial e}\right|_{ED=0} = -\dfrac{\partial ED/\partial e}{\partial ED/\partial \beta}$.

We have:

$$\left.\frac{\partial \beta}{\partial e}\right|_{ED=0} = (1-\beta)\left[\frac{(w + \theta dV - e - wa - e\rho)}{e\rho(w + \theta dV - e - wa)} + \frac{(1-\beta)(w + \theta dV - e - wa)}{e\rho k\theta V(1-d)}\right]$$

The sufficient condition of $\left.\dfrac{\partial \beta}{\partial e}\right|_{ED=0} = -\dfrac{\partial ED/\partial e}{\partial ED/\partial \beta} \geq 0$ is that

$w - e - wa - e\rho \geq 0$.

Appendix 3.4

Table 3.5: The impact of d on $A_r, \pi_0, \pi_{-1},$ and their weights

d	0.2	0.3	0.4	0.5	0.55	0.6
A_r	11.16	13.38	16.15	19.49	21.55	23.64
π_0	8.34	7.92	7.40	6.82	6.49	6.20
$1/(\phi+\theta)$	1.18	1.20	1.23	1.27	1.28	1.32
π_{-1}	0.36	1.05	1.90	2.90	3.50	4.12
$\phi/(\phi+\theta)$	0.48	0.48	0.47	0.46	0.46	0.45

When d increases, the relative wealth of the rich increases. However, the profit of the best quality good decreases because of the decreasing market

share $1-\beta$. π_{-1} increases because both the market share and the income of the poor increase. But in the present value of innovation π_0 has a higher weight factor $\frac{1}{\phi+\theta}$ than π_{-1} ($\frac{\phi}{(\phi+\theta)^2}$). Hence, the net effect of d on the present value of innovation is negative, which impedes the innovation rate.

Appendix 4.1

According to the definition of the Gini-coefficient, it is equal to the ratio of the areas ACD and ABC. As we normalized AB and BC to 1, we have:

$$Gini = 2 \bullet the\ area\ of\ ACD = 1 - 2 \bullet the\ area\ of\ ABCD$$
$$= 1 - \beta \bullet d\beta - (d\beta+1) \bullet (1-\beta)$$
$$= (1-d)\beta$$

Appendix 4.2

The labor market clearing condition is

$$\beta L_p + (1-\beta)L_r = n + a + b(\beta x_p + (1-\beta)x_r) + S$$

Substituting (4.1) and budget constraint equations of both periods, we have two possible cases:

1) if the monopolist sets the price at the pooling level:

$$a + \beta H_p^\alpha + (1-\beta)H_r^\alpha = n + a + b(\beta(\theta A_p - H_p) + (1-\beta)(\theta A_r - H_r) + \beta(y_p - \overline{P_0}) +$$
$$(1-\beta)(y_r - \overline{P_0})) + S$$

recall $wb = 1$, $wS = \beta H_p + (1-\beta)H_r$ and $\pi^{pool} = \overline{P_0} - wa$:

$$w(\beta H_p^\alpha + (1-\beta)H_r^\alpha) = wn + \beta(\theta A_p - H_p) + (1-\beta)(\theta A_r - H_r) + \beta(y_p - \overline{P_0}) + (1-\beta)(y_r - \overline{P_0})$$
$$+ \beta H_p + (1-\beta)H_r$$
$$\Leftrightarrow 0 = wn + \theta V - \pi^{pool}$$

2) if the monopolist sets the price at the separating level:

$$a + \beta H_p^\alpha + (1-\beta)H_r^\alpha = n + a + b(\beta(\theta A_p - H_p) + (1-\beta)(\theta A_r - H_r) + \beta(y_p - wa) +$$
$$(1-\beta)(y_r - \overline{P_0})) + S$$
$$\Leftrightarrow 0 = wn + \theta V - \pi^{sep}$$

Summing, we have $\pi = wn + \theta V$.

Appendix 4.3

Here we show that the effect of d on L is positive. Hence, the redistribution from the rich to the poor can increase the supply of labor; in turn, the innovation rate increases.

$$L = \beta L_p + (1-\beta)L_r = a + \beta H_p^\alpha + (1-\beta)H_r^\alpha$$

From (4.14) (4.15) (4.16) and $V^* = \dfrac{w}{\lambda}$, we know

$$H_p = \frac{\alpha \rho}{1 + \alpha \rho} d \frac{w}{\lambda}, \quad H_r = \frac{\alpha \rho}{1 + \alpha \rho} \frac{1 - d\beta}{1 - \beta} \frac{w}{\lambda}$$

Hence,

$$\frac{\partial L}{\partial d} = \frac{\alpha \rho w}{(1+\alpha \rho)\lambda} \left[\beta \alpha d^{\alpha-1} + (1-\beta)\alpha \left(\frac{1-d\beta}{1-\beta} \right)^{\alpha-1} \left(\frac{-\beta}{1-\beta} \right) \right]$$

$$= \frac{\beta \alpha^2 \rho w}{(1+\alpha \rho)\lambda} \left[d^{\alpha-1} - \left(\frac{1-d\beta}{1-\beta} \right)^{\alpha-1} \right] > 0$$

.

In both extreme cases $\alpha \to 0$ and $\alpha \to 1$, the supply-side effect of inequality on growth disappears, i.e., $\frac{\partial L}{\partial d} \to 0$. α reveals the strength of this supply-side effect.

References

Acemoglu, Daron and Robinson, A. James (2000) "Why did the West Extend the Franchise? Democracy, Inequality and Growth in Historical Perspective" *Quarterly Journal of Economics* 115(4), pp. 1167-1199.

_____. **(**2001) **"**A Theory of Political Transitions" *American Economic Review* 91(4), pp. 938-963.

Acemoglu, Daron, Robinson, A. James and Verdier Thierry (2004a) "Kleptocracy and Divide-And-Rule: a Model of Personal Rule" *Journal of European Economic Association*, 2(2/3), pp.162-192.

Acemoglu, Daron, Johnson, Simon and Robinson, James (2004b) "Institutions as the Fundamental Cause of Long Run Growth" *NBER Working Paper* No: w10481.

Aghion, Philippe, Caroli, Eve and Garcia-Penalosa, Cecilia. (1999) "Inequality and Economic Growth: The Perspective of the New Growth Theories", *Journal of Economic Literature*, 37(4), pp. 1615-1660.

Aghion, Philippe and Howitt, Peter. (1992) "A Model of Growth through Creative Destruction", *Econometrica*, 60(2), pp. 323-351.

_____. (1998) Endogenous Growth Theory, MIT Press, Cambridge, MA.

_____. (2004) "Growth with Quality-Improving Innovations: An Integrated Framework", Forthcoming in *Handbook of Economic Growth*, edited by Philippe Aghion and Steven N. Durlauf. Amsterdam, North-Holland.

Alesina, Alberto and Rodrik, Dani. (1994) "Distributive Politics and Economic Growth.", *Quarterly Journal of Economics*, 109(2), pp. 465-490.

Banerjee, Abhijit V. and Duflo, Esther. (2003) "Inequality and Growth: What can the Data say?", *Journal of Economic Growth*, 8(3), pp. 267-299.

Barro, Robert J. (1997) *Determinants of Economic Growth: a cross-country empirical study,* Cambridge, MA: MIT press.

＿＿＿＿＿＿＿＿. (1999) "Determinants of Democracy": *Journal of Political Economy*, 107(6), pp.158-183.

＿＿＿＿＿＿＿＿. **(**2000a) "Democracy & the Rule of Law" in Fraser Forum, http://oldfraser.lexi.net/publications/forum/2000/06/section_04_full.html

＿＿＿＿＿＿＿＿. (2000b) "Inequality and Growth in a Panel of Countries", *Journal of Economic Growth*, 5(1), pp. 5-32.

Barro, Robert J. and Lee Jong-wha (1994) "Data Set for a Panel of 138 Countries", Homepage of Barro, http://post.economics.harvard.edu/faculty/barro/data.html.

Bénabou, Roland. (1996b) "Inequality and Growth" in Ben S. Bernanke and Julio J. Rotemberg, eds., NBER macroeconomics annual 1996. Cambridge, MA: MIT Press, pp. 11-74.

＿＿＿＿＿＿＿＿. (2002) "Tax and Education Policy in a Heterogeneous-Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?" *Econometrica* 70(2), pp.481-517.

Bénabou, Roland and Ok, A. Efe. (2001) "Social Mobility and the Demand for Redistribution: the POUM Hypothesis.", *Quarterly Journal of Economics*, pp. 447-487.

Berg, A. and Sachs, J. (1988) "The Debt Crisis: Structural Explanations of Country Performance.", *Journal of Development Economics*, 29, pp. 271-306.

Bils, Mark and Klenow, J. Peter. (2000) "Does Schooling Cause Growth?", *American Economic Review*, 90(5), pp. 1160-1183.

Boix, Carles and Stokes, Susan C. (2003) "Endogenous Democratization" *World Politics* 53(4), pp.517-547.

Bueno de Mesquita, B., Smith, A., Siverson, R., Morrow, J. (2003) *The Logic of Political Survival.* MIT Press, Cambridge.

Cass, David (1965) "Optimum Growth in an Aggregate Model of Capital Accumulation" *Review of Economic Studies*, 32, pp.233-240.

Castelló, A. and Doménech, R. (2002) "Human Capital Inequality and Economic Growth: Some New Evidence", *the Economic Journal*, 112(478), pp. 187-200.

Chen, Been-Lon. (2003) "An Inverted-U Relationship between Inequality and Long-Run Growth", *Economics Letters*, 78, pp. 205-212.

China Statistical Yearbook 2002. China Statistics Press.

Clarke, George R. (1995) "More Evidence on Income Distribution and Growth.", *Journal of Development Economics*, 47(2), pp. 403-427.

Collins, S.M. and Bosworth, B.P. (1996) "Economic Growth in East Asia: Accumulation versus Assimilation", *Brookings Papers on Economic Activity* 2, pp.135-203.

Collier, P. (2001) "Implications of Ethnic Diversity", *Economic Policy* 32, pp. 127-166.

Cox, Gary W. (1997) *Making Votes Count*, New York: Cambridge University Press.

Croix, David de la and Doepke, Matthias. (2003) "Inequality and Growth: Why Differential Fertility Matters", *American Economic Review*, 93(4), pp. 1091-1113.

Deininger, K and Squire L. (1996) "New Data Set Measuring Income Inequality", *World Bank Economic Review*, 10(3), pp. 565-591.

_____. (1998) "New Ways of Looking at Old Issues: Inequality and Growth.", *Journal of Development Economics*, 57(2), pp. 259-287.

Easterly, W. and Levine, R. (1997) "Africa's Growth Tragedy: Policies and Ethnic Divisions", *Quarterly Journal of Economics* 112, pp.1203-1250.

Foellmi, Reto and Zweimüller, Josef. (2002) "Heterogeneous Mark-ups, Demand Composition, and the Inequality-Growth Relation.", Royal

Economic Society Annual Conference 2002, 76, Royal Economic Society, http://ideas.repec.org/s/ecj/ac2002.html

Forbes, J. Kristin. (2000) "A Reassessment of the Relationship between Inequality and Growth", *American Economic Review*, 90(4), pp. 869-887.

Gini, Corrado. "Variabilità e mutabilità", (1912) Reprinted in Memorie di metodologica statistica (Ed. Pizetti E, Salvemini, T). Rome: Libreria Eredi Virgilio Veschi, 1955.

Glaeser, L. Edward, Porta, La Rafael, Lopez-de-Silnes, Florencio and Shleifer, Andrei (2004) "Do Institutions Cause Growth?" *Journal of Economic Growth* 9(3), pp. 271-303.

Glomm, Gerhard and Ravikumar, B. (1992) "Public Versus Private Investment in Human Capital: Endogenous Growth and Income Inequality." *Journal of Political Economy*, 100(4), pp. 818-834.

Grossman, G. and Helpman, E. (1991) "Quality Ladders in the Theory of Growth" *Review of Economic Studies* 58, pp.43-61.

Huntington, Samuel P. (1991) *The Third Wave: Democratization in the Late Twentieth Century.* Norman: Univ. Oklahoma press.

Jones, Benjamin F. and Olken, Benjamin A. (2005) "Do Leaders Matter? National Leadership and Growth since World War II", *The Quarterly Journal of Economics* 120(3), pp.835-864.

Konrad, Kai A. (2002) "Investment in the Absence of Property Rights: the Role of Incumbency Advantages", *European Economic Review* 46(8), pp. 1521-1537.

Koopmens, Tjalling C. (1965) "On the Concept of Optimal Economic Growth" in *The Economic Approach to Development Planning*, Amsterdam: North-Holland.

Kurzman, C., Werun R. and Burkhart RE (2002) "Democracy's Effect on Economic Growth: A pooled time-series analysis, 1951-1980" *Studies in Comparative International Development* 37(1), pp.3-33.

Laffont, Jean-Jacques and Qian, Yingyi (1999) "The Dynamics of Reform and Development in China: a political economy perspective" *European Economic Review, Papers and Proceedings*, 43, pp. 1085-1094.

Lee, Woojin (2003) "Is Democracy More Expropriative than Dictatorship? Tocquevillian wisdom revisited" *Journal of Development Economics* 71(1), pp. 155-198.

Lucas, Robert E. (1988) "On the Mechanics of Economic Development", *Journal of Monetary Economics* 22(1), pp.3-42.

_____. (1993) "Making a Miracle" *Econometrica*, 61(2), pp.251-272.

McGuire, M.C. and Olson, M. (1996) "The Economics of Autocracy and Majority Rule", *Journal of Economic Literature,* 34, pp.72-96.

Murphy, Kevin M., Shleifer, Andrei and Vishny, Robert. (1989) "Income Distribution, Market Size and Industrialization", *The Quarterly Journal of Economic,* 104(3), pp. 537-564.

Nelson, R. and Phelps, E. (1966) "Investment in Humans, Technological Diffusion, and Economic Growth", *American Economic Review: Papers and Proceedings* 51(2), pp.69-75.

North, Douglass C., and Robert P. Thomas (1973) *The Rise of the Western World: A new Economic History,* Cambridge University Press.

Persson, Torsten and Tabellini, Guido. (1994) "Is Inequality Harmful for Growth?" *American Economic Review* 84(3), pp. 149-187.

_____. (2000) *Political Economics: Explaining Economic Policy.* Cambridge: MIT Press

_____. (2003) *The Economic Effects of Constitutions: What Do the Data Say?* Cambridge: MIT Press

Romer, P. (1990) "Endogenous Technological Change" *Journal of Political Economy* 89(5), pp.71-102.

Ross, L. Michael (2001) "Does Oil Hinder Democracy?" *World Politics* 53, pp. 325-361.

Rueschemeyer, Dietrich, Stephens, Evelyn H. and Stephens, John D. (1992) *Capitalist Development and Democracy.* Chicago: University of Chicago Press.

Sah, Raaj K. (1991) "Fallibility in Human Organizations and Political Systems" *Journal of Economic Perspectives* 5, pp.67-88.

Schumpeter, Joseph A. (1947) *Capitalism, Socialism and Democracy*. 2d ed. New York: Harper.

Shaked, Avner and Sutton, John. (1982) "Relaxing Price Competition through Product Differentiation", *Review of Economic Studies*, 49, pp. 3-13.

_____. (1983) "Natural Oligopolies", *Econometrica*, 51, pp. 1469-1483.

Shepsle, Kenneth A. and Weingast, Barry R. (1995) *Positive Theories of Congressional Institutions*, Ann Arbor: University of Michigan Press

Solow, Robert M. (1956) "A Contribution to the Theory of Economic Growth", *Quarterly Journal of Economics*, 70, pp. 65-94.

Taylor, C. and Hudson, M. (1972) "*World Handbook of Political and Social Indicators*", New Haven and London, Yale University Press.

Therborn, Goran. (1977) "The Rule of Capital and the Rise of Democracy", *New Left Review,* 103(1), pp.3-41.

Tornell, Aaron and Lane, Philip R. (1999) "The Voracity Effect", *American Economic Review*, 89(1), pp.22-46.

Wintrobe, Ronald (1998) *The Political Economy of Dictatorship.* Cambridge: Cambridge University Press.

Yang, Dennis Tao. (1999) "Urban-biased Policies and Rising Income Inequality in China", *American Economic Review*, 89(2), pp. 306-310.

Yang, Dennis Tao and Zhou, Hao. (1999) "Rural-urban Disparity and Sectoral Labor Allocation in China", *Journal of Development Studies,* 35(3), pp. 105-133.

Zak, Paul J. and Feng, Yi (2003) "A Dynamic Theory of the Transition to Democracy" *Journal of Economic Behavior & Organization* 52, pp.1-25.

Zweimüller, Josef. (2000) "Schumpeterian Entrepreneurs Meet Engel's Law: The Impact of Inequality on Innovation-Driven Growth", *Journal of Economic Growth*, 5(2), pp. 185-206.

Zweimüller, Josef and Brunner, K. Johann. (2005) "Innovation and Growth with Rich and Poor Consumers", *Metroeconomica,* 56(2), pp.233-262.