Five Statistical Analyses on Efficiency and Productivity. The Data Envelopment Analysis in Economic History.

Inaugural-Dissertation

zur Erlangung des Doktorgrades

der Wirtschafts- und Sozialwissenschaftlichen Fakultät

der Eberhard-Karls-Universität Tübingen

vorgelegt von

Dipl.-Kaufmann Dominic Behle

Spaichingen

Tübingen

2010

Tag der mündlichen Prüfung:	18. Mai 2011
Dekan:	Prof. Dr. rer. soc. Josef Schmid
Erstkorrektor:	Prof. Dr. Jörg Baten
Zweitkorrektor:	Prof. Dr. Heinz Gert Preuße

Acknowledgments

First, I would like to thank my supervisor, Prof. Dr. Jörg Baten, who gave me the opportunity to work in the research-oriented and stimulating environment of his research group in Tuebingen. He supported my choice to work on efficiency, provided ideas, valuable criticism and shared his impressive knowledge about a large variety of topics in economic history and econometrics. He also provided the opportunities to present my work on numerous international conferences, where I benefited from comments and ideas from scholars working on different fields of economics and economic history.

I also thank Prof. Dr. Heinz Gert Preuße for his willingness to be my second advisor. His focus on economic policy balanced the more methodological parts of the thesis.

My colleagues at the research group in economic history and seminar participants at the University of Tuebingen provided many ideas, criticism and valuable other support. I want to express special gratitude to Dr. Kirsten Jäger, Dr. Daniel Schwekendiek, Dr. Dorothee Crayen, Dr. Normann Müller, Dr. Kerstin Manzel, Linda Twrdek, Mojgan Stegl, Matthias Blum, Ralph Hippe, Valeria Prayon and Yvonne Stolz. Numerous discussions with Dr. Normann Müller taught me much about economics and scientific thinking. From Dr. Luis Huergo I learned a lot about econometrics.

Many researchers in economic history provided comments and ideas. The exchange of ideas and knowledge on conferences and workshops proved invaluable for this thesis.

Without Jeremy Atack, Fred Bateman and Thomas Weiss this thesis would not exist. Their samples from the American Manufacturing Census were the starting point for most of this work. Jeremy Atack also shared his expertise on nineteenth century industrial development with me.

Last but not least I want to thank my family, which has provided invaluable support and encouragement. They endured the moments of joy and frustration associated with writing a dissertation and listened to stories of methods they never thought did exist.

Table of Contents

1 Introduction.....1
1.1 Outline.....5
References.....9
2 Industrial Efficiency in the Antebellum USA and its Implications for Industrial Structure and Development.....10
2.1 Introduction.....11
2.2 Methodology and Data.....13
2.3 Factory vs. Artisanal Shop.....16
2.4 The Northern and Southern Industry before the Civil War.....20
2.5 The Hypotheses.....21
2.6 Earlier Research on the Hypothesis of Southern backwardness.....23
2.7 Results.....25
2.8 Efficiency and long-run Economic Growth.....32
2.9 Conclusion.....33
References.....35
Appendix 1: Figures and Tables.....37
Appendix 2: The Constant Returns to Scale specification.....43
3 The Impact of the American Civil War on Manufacturing Efficiency.....45
3.1 Introduction.....46
3.2 The Civil War and its Economic Impact on American Industry.....48
3.3 The Data Envelopment Analysis.....50
3.4 The Civil War and Manufacturing Efficiency.....54
3.5 The War and Efficiency: Empirical Results.....56
3.6 Conclusion.....71
References.....73
Appendix 1: Figures.....78
Appendix 2: Tables.....81
4 Do Numeracy and Health determine Labor Productivity in Tsarist Russia?.....85
4.1 Introduction.....86
4.2 Literature Review.....87
4.3 Historical Background.....88
4.4 Data and Variables.....91
4.5 Regression Results.....97
4.6 Conclusion.....104
References.....106
Appendix.....110
5 Regional Determinants of Firm Creation in the Russian Empire. Evidence from the 1870 Industrial Exhibition. 111
5.1 Introduction.....112
5.2 Data 113
5.3 Determinants of Firm Creation.....115

5.4 Results 119
5.5 Conclusion.....121
References.....123
Appendix.....125
6 On the Determinants of a successful State: Good Governance between the 1850s and the 1980s.....126
6.1 Introduction.....127
6.2 Methodology.....129
6.3 Data.....133
6.4 Results – the Development of Government Performance 137
6.5 How did we distinguish Inputs, Outputs and explanatory Variables?.....143
6.6 Data used in the Second Step – Determinants of Efficiency.....143
6.7 Regression Results.....150
6.8 Conclusion.....153
References.....155
Appendix A: The Bootstrap.....162
Appendix B: The Data Set.....164
Appendix C: Figures.....176
7 Summary and Research Outlook.....183
References.....187

Abbreviations

DEA	Data Envelopment Analysis
DMU	Decision Making Unit
ELF	Index of Ethno-Linguistic Fractionalization
EMS	Efficiency Measurement System
EU	European Union
GDP	Gross Domestic Product
N	Number of observations
OECD	Organization for Economic Cooperation and Development
OLS	Ordinary Least Squares
pte	pure technical efficiency
UK	United Kingdom
US	United States
WHO	World Health Organization
WW II	World War Two

1 Introduction

In the last decades, the study of efficiency has gained much attention in the scientific literature. More efficient industries can produce more goods, so that higher levels of welfare are possible. This makes the study of the development and determinants of efficiency a topic which is highly relevant for politics and academics. Empirical studies about the efficiency of industrial firms, storage shops and even nations have emerged. Productivity is related to efficiency; however efficiency is a relative concept that does not reflect technical progress. Nevertheless, both are important aspects of industrial development. Research on efficiency and productivity has a long and fruitful tradition. An early example is Adam Smiths (1776) analysis of the needle industry.

For a profound understanding of economic history, knowledge on efficiency is necessary. Efficiency gains release resources in the production of goods, so that more production and welfare can be realized. As rising efficiency enables states to spend more resources on sanitary systems, health of the population increases. Rising efficiency in agriculture allowed mankind a society in which some people could focus on science and culture. Moreover, a comparably high productivity level allows countries to dominate world market, finance large armies and achieve cultural breakthroughs.

This thesis consists of five separate papers, which empirically investigate the development of productivity and efficiency in economic history using state-of-the art methods.[1] Four of them are concerned with the industry of the nineteenth century, a time when the industrial revolution led to large increases in productivity and fundamentally changed many economies around the globe. At the same time governments and adverse events, such as wars, had a strong impact on firms, which makes it interesting to study their impact on efficiency.

In this thesis productivity and efficiency are defined as follows: Productivity refers to the amount of output per unit of an input a firm produces. It is therefore an absolute measure that allows for inter-temporal comparisons. Computation is conceptually easy, as, in the case of labor-productivity, it is measured as output divided by the number of workers. If workers differ considerably in quality, which is for example the case when many children are employed in an industry, full worker equivalents should be used. These account for the quality differences in the workforce (see for example Sokoloff and Tchakerian 1997). Productivity might depend on the industrial structure; when some industries are more competitive than others, profit rates and revenue will depend on the considered industry. Actual technology may also have an influence, as technological progress (e.g. the personal computer) has increased the output a single worker can produce.

Efficiency as defined here measures relative differences in productivity between firms. A firm is called inefficient if it could produce more output with the given inputs and the technology available, which means its productivity is lower than that of another comparable firm. It is possible that a firm produces more output than last year with the same inputs, which

[1] All five chapters were written as independent articles and are intended for separate publication or already published. I therefore refer to them as “article”, “paper” or “study”.

means rising productivity, but when other comparable firms increase their output to a larger extent, then its efficiency has decreased according to this definition. Normally only firms at the same point in time and environment should be compared. Inter-temporal comparisons are possible, but they require careful handling. Chapter three presents such an application of an inter-temporal comparison. Efficiency shows us how successful an industry or a firm is in using the resources that are available with the technology at hand. However, as defined here, efficiency does not necessarily imply a value judgment. If a firm is called inefficient, it just means that with the available technology it would have been able to produce more with the available inputs, but switching to the other production possibilities would not necessarily be increasing welfare. It is for example possible that using the full technological possibilities is associated with more demanding work conditions or more pressure and might therefore decrease the welfare of the entrepreneur (Saraydar 1989).

To study the structure and development of an industry, abundant quantitative descriptive material from the past is available, for example through census manuscripts or surveys from industrial or public institutions. Census inquiries also produce enough data to study productivity and its development. Revenue and labor statistics which can be found in archives make the computation of labor-productivity indices possible. One chapter of this dissertation will use labor productivity indices. Studying efficiency is more difficult. Comparable firms have to be found and efficiency measures have to be computed. While it is relatively easy to study revenue efficiency for comparable firms, as revenues from different firms can be directly compared, this is critical when outputs are produced for which no prices are available or outputs differ (high versus low quality boots might serve as an example).

One main contribution of my thesis is that I propose a method that allows efficiency comparisons, while it has not yet been widely used in research on economic history. This method – the so called Data Envelopment Analysis (DEA) – is commonly used in the modern management literature and allows evaluating the efficiency of a large number of firms. Though the technique needs comparable data sets that are reliable, they offer new possibilities to study industry history. In chapters 2, 3 and 6, I show some examples how it can be used in economic history. I also incorporate actual developments from the statistical literature, like a bootstrapping algorithm, which allow computation of reliable standard errors. I am therefore able to provide more reliable results compared to existing research.

The DEA is widely used today for managerial and scientific purposes. One managerial application of it is to compare stores or production plants to see which units could be better managed and also to show where optimization is possible. This is done by showing for example which input a certain plant uses in an excessive way. In science, it is often used to study organizations where conventional production function approaches are difficult. This is the case when prices for inputs or outputs are missing, as is for instance the case when studying universities. How should graduates, the first output, be weighted against research, the second output? The DEA circumvents this problem by estimating optimal weights from the data (see Cooper, Seiford and Tone 2006 for a detailed description of the technique).

The DEA also enables us to look deep into firms, deeper than it is possible with indicators like revenue or capital. Revenue does not tell us how the firm looks like – whether the firm is fully equipped, or whether the production processes are well organized, etc. To assess those facets of a firm we can first do case studies of one or a few firms in an industry using in-

ternal documents. This approach has the disadvantage that it first takes a huge effort to study only few enterprises. Second, the researcher is restricted to the opinions of managers or owners of the firms at the respective time. This is especially problematic as not all managers know enough about competing firms to classify their enterprise. The second approach is to use statistical techniques able to handle many firms and to produce comparable results. The DEA allows such insights, as a firm that is considered highly inefficient compared to its competitors likely has internal problems that limit its efficiency.

To summarize, it is important and equally difficult to study the distribution and development of efficiency in history. The methods that are used in this paper allow a deeper and broader examination of efficiency than was previously possible and might therefore be able to answer some questions where no satisfying answers exist today.

1.1 Outline

This thesis begins with two studies that use the DEA to study nineteenth century American industry. Though this industry has been intensively studied in the literature, the DEA can still shed light on topics that could not be answered hitherto. I begin by studying economies of scale and I show that substantial economies of scale existed before the Civil War. Afterwards, I examine efficiency differences between the Southern and Northern states before the Civil War. The literature could not establish whether a Southern efficiency lag might explain the differences in the economic structure between the North and the South. These efficiency dif-

ferences might be the results of less efficient management in the Southern states. My results show that, though the South is among the less efficient regions, one can not definitely state efficiency deficits for the South. Additionally the existing differences in efficiency are small and likely not able to explain the differences in the economic development after the war.

The American Civil War is the background of the second chapter. Its political, economic and social effects have been discussed in a numerous literature. The focus of my work is on the effect of the war on efficiency. While the literature examines the impact of the war on the location or structure of the industry, I ask what the war changed in the way the economy works. With a large data set sampled from the American Manufacturing Census I estimate efficiency and find that the war was indeed used to restructure the production processes in the Southern territories, which gives the Southern industry a slight advantage in technical efficiency after the war. In other efficiency categories I show a larger effect of the war, though the effects all in all seem small compared to the conventional historiography that reports large and terrible destructions. But my results are more in line with a newer reinterpretation of the destruction reports (Neely 2007) which documents that commanders in the war were well aware that they were fighting against brothers and often tried to limit destruction.

Next, I present two studies, both co-authored with Prof. Dr. Jörg Baten, which are concerned with the industry of Tsarist Russia. These do not use the DEA but other approaches to give a balanced methodological approach to efficiency and productivity measurement. The first article, “Do numeracy and health determine labor productivity in Tsarist Russia?”, uses OLS regressions to asses the determinants of labor productivity. The main focus of this study

is the impact of human capital and heights on labor productivity, and we find a large positive effect for both. Political factors, especially taxation, also show an influence on productivity.

The second study looks at the determinants of firm creation rates in Tsarist Russia. This chapter is based on a study published in the Russian Economic History Yearbook, which is the reason why the structure of that chapter differs slightly from the other chapters. Though it is not concerned with efficiency, firm creation is a central part of successful industrial development, which is also a topic discussed in the other papers. Next, new firms add competitive pressure to existing firms and thereby may increase the efficiency of the industry as a whole (see Hermalin 1992). The most important finding is that human capital has a strong positive impact on firm creation rates. Educated people are ceteris paribus more likely to create firms (for evidence in modern times see Acs and Armington 2002), thereby increasing economic growth.

The sixth chapter leaves the direct focus on industry to use the methods introduced in this thesis in another branch of the historic literature. Together with Matthias Blum and Luis Huergo I analyze government performance from the 1870s to the 1980s. We view governments as a form of producers, as each country has endowments, such as land, capital etc. and uses those to produce welfare. This transformation of the available inputs into outputs is viewed as a kind of production process and we use the DEA to assess the efficiency of this process. The chapter expands the existing literature, which only considers the period after 1965, and we are the first to study the interesting world war two and inter-war period. Our results show among others that fractionalized societies had markedly less success in creating welfare than more unified nations.

The last section concludes this thesis by summarizing the most important findings and showing possible areas for future research, where the methods used in this thesis might be fruitfully employed to further enhance our understanding of history.

References

Acs, Z. J. and C. Armington (2002). The Determinants of Regional Variation in New Firm Foundation. Regional Studies 36 (1), 33-45.

Cooper, W. W., Seiford, L. M. and K. Tone (2006). Introduction to data envelopment analysis and its uses. With DEA-solver software and references. New York: Springer.

Hermalin B. E. (1992). The Effects of Competition on Executive Behavior. The RAND Journal of Economics 23 (3), 350-365.

Neely, M. E. (2007). The Civil War and the limits of destruction. Cambridge, Mass.: Harvard University Press.

Saraydar, E. (1989). The Conflation of Productivity and Efficiency in Economics and Economic History. Economics and Philosophy 5 (1), 55-67.

Smith, A. (1776). An Inquiry into the Nature and Causes of the Wealth of Nations.Chicago: University Of Chicago Press.

Sokoloff, K. L. and V. Tchakerian (1997). Manufacturing where Agriculture predominates: Evidence from the South and Midwest in 1860. Explorations in Economic History 24 (3), 243-264.

2 Industrial Efficiency in the Antebellum USA and its Implications for Industrial Structure and Development.

Abstract:

Empirical evidence on the industrial efficiency of the antebellum economy is scarce. This paper presents efficiency estimates and uses these to shed light on two unresolved questions. First, it is controversial whether factories enjoyed efficiency advantages over artisanal shops. The estimates here confirm the existence of such economies of scale. The second question is whether efficiency differences between Northern and Southern states existed before the Civil War and whether these were responsible for the diverging development patterns found after the Civil War. The findings here show that there is no clear Southern efficiency disadvantage that was responsible for the slower industrial development.

This chapter is an extended version of a working paper called “Industrial Efficiency in the antebellum USA. A Southern disadvantage?”.

2.1 Introduction

Efficiency in the antebellum US is a widely discussed topic in economic history. However, quantitative evidence is not abundant. The efficiency of the Southern agricultural system which was based on slavery has been subject of major discussions. Though a topic in the historical literature since the Civil War, the empirical debate surged with Fogel and Engerman's Book "Time on the Cross" and went on for more than two decades (Fogel and Engerman 1974, Field 1988, Grabowski and Pasurka 1988 and 1989, Field-Hendrey 1995, Craig and Field-Hendrey 1996). Most papers used production functions, few the Data Envelopment Analysis as does this paper. The result of the debate is that the gang system made large slave farms (more than 15 slaves) more efficient than smaller slave farms, which in turn were more efficient than the free (which means they used no slave labor) Southern farms.

If we look at industrial efficiency, the literature, especially the empirical works, is far less numerous. This paper contributes to the literature by estimating efficiency for a large sample of firms from 1850 and 1860. The estimates are also used to shed light on two questions that are discussed in the literature.

First, we analyze scale economies and their impact on industrial development. It is controversial in the literature whether scale effects existed before the Civil War and especially whether the transition from the artisanal shop to the factory led to rising efficiency and therefore strengthened American industrial development. We show the existence of substantial economies of scale.

In the second step, we look at the connection of industrial efficiency and development. Many scholars described the antebellum South as being industrially „backward" compared to

the North. Just to give one example, Starobin claims that "… Southern industrialization lagged behind that of the North and of Great Britain." (1970, p. 163). Though the South was far less industrialized than the North at the eve of the War, this was not unique, as the West was also largely agricultural and in many ways similar to the South (Bateman and Weiss 1981, p.55).

Nevertheless, the fact that the South developed less industry than the North needs to be explained. Scholars have discussed different hypotheses to explain this phenomenon. These include the idea that a Southern competitive advantage in cotton production delayed manufacturing development, that limited market size in the South restrained manufacturing, that certain characteristics of slavery were unfavorable for industry and that noneconomic barriers or entrepreneurial inability prevented manufacturing from developing its full capacity (See Bateman and Weiss 1976 and 1981). Several of the mentioned hypotheses can be rejected by quantitative research, others could not be tested till now. But it seems that one of the major reasons was investors' irrationality, which means that planters and other people who had funds for investment available did not invest in profitable manufacturing, perhaps because of prejudices and the wish to retain the plantation system.

In this paper two of these hypotheses, which center around efficiency, are investigated. The literature has used indirect evidence, like profit rates, to study these hypotheses. Until now, no direct evidence has been put forward, and this paper is the first to close this gap. The first hypothesis states that Southern firms were not as efficient as their Northern counterparts and could therefore not develop as rapidly as firms in the North could. The second hypothesis is similar, but it argued that Southern firms were not as efficient as their Northern counterparts because they were too small. In other words, they were scale inefficient.

The Data Envelopment Analysis (DEA) is used to examine these hypotheses. This methodology is an alternative to the traditional production function approach and has several advantages. Most importantly, it measures efficiency directly. Nevertheless, it has not been used intensively in the debate on American economic history. Several papers discussed issues concerning the efficiency of the plantation system using the DEA (Grabowski and Pasurka 1988, 1989, Field-Henry 1995). Behle (2009) used the technique to estimate the impact of the American Civil War on industry.

The paper develops as follows. First the methodology and data are discussed. Section three deals with the question whether scale economies existed and whether the introduction of the factory system therefore increased industrial efficiency. We will then give a short overview over industrial conditions in the North and South before the Civil War in section 4. The two hypotheses concerning the differences between the Southern and Northern industry are discussed in section 5. Earlier research on this question will be discussed next, and afterwards the new estimates are presented, which allow for a direct test of these hypotheses for the first time. Section 8 analyzes the implications of the estimations for Southern industrial development. The last section concludes.

2.2 Methodology and Data

The previous studies on efficiency or scale economies operated with a Cobb-Douglas production function or other regression-based tests (see Schäfer and Schmitz 1982 and the literature cited there). Normally, output of firms is regressed on capital, labor, and sometimes land. This approach has the disadvantage that the researcher has to specify the functional form of the production function. In this article we will employ the Data Envelopment Analysis, which is a non-parametric technique, to estimate efficiency. With this method, we are able to circumvent the question whether the chosen production function is appropriate for the industry under study. There is also no need for estimating different production functions for different regions, or to make other potentially arbitrary choices.

The Data Envelopment Analysis is able to differentiate between efficient and inefficient firms. The DEA constructs a production frontier from the data and evaluates each firm against this frontier, measuring by how much the input quantities of a given firm have to be reduced so that this firm would lie on the frontier. This proportional decrease in inputs is the measure of the firms' inefficiency. All estimated efficiency scores have values between 0 and 1.

There are several models available which differ in their assumptions, especially concerning the returns to scale. It is possible to use these different models to decompose the efficiency score into three different categories: pure technical efficiency (pte), scale efficiency and mix efficiency. Pure technical efficiency describes the technical side of production, which means how much the firm produces with the available inputs. Imagine 2 firms which have the same inputs available, but achieve different levels of output, for example because they produce different amounts of waste or have different machines so that one firm can produce more in the same time with the same amount of workers. The firm which has a lower output will be

shown as not technically efficient, as with the available technology it is possible to produce more from a given quantity of inputs. If firms operate under non-constant returns to scale they are called scale inefficient. Consider for example a firm that produces under increasing returns to scale. If this firm would double its inputs, the output would, due to the increasing returns, increase to more than double. The ratio of output to input would therefore increase and the firm be called more efficient. Firms are called mix inefficient if they are using some inputs in an excessive way, which means they could decrease the usage of one input without thereby altering their production. For example imagine two establishments that use the same amount of capital, but one uses 10 workers while the other uses 12 workers, to produce the same output. The second establishment could obviously decrease their labor-force by 2 compared with the first one and still produce the same output, so this second firm is mix-inefficient.

To evaluate the hypotheses a regression approach is used. After estimating the efficiency scores of all firms in the data set with the DEA, truncated regressions are used to estimate the determinants of the efficiency scores. Though a regression approach after a DEA estimation needs to be handled cautiously, as it is consistent but the rate of convergence is slow, this is no point in this estimation as the DEA models use only few variables and the number of observations is about 6000 (see Simar and Wilson 2007, who propose the use of truncated regressions. In that article the rate of convergence is given in formula (11). Here we have enough observations to achieve convergence). The independent variables include dummies for different regions, which show whether some regions were really more efficient than others. Outliers were excluded in this setting. An outlier is defined as an observation that has an efficiency score of over 200% in a superefficiency model. In this model, the establishment which is evaluated is excluded from the estimation of the frontier, so that its efficiency-score

can exceed 100 %, which makes identification of outliers possible. Values of over 200 %, which mean that the firm could double the used input quantities to produce the same outputs and still be considered efficient, are seen as implausible in a competitive market and therefore likely come from data errors.[2]

The data come from various American Manufacturing Censuses, namely the ones from 1850 and 1860.[3] During every Census firms were asked to give detailed information about their used inputs, the produced outputs, labor force and many more.

To estimate efficiency, we use the stated amount of capital, the effective workforce, which is estimated as $L_T = L_{male} + 0{,}5L_{female} + 0{,}33L_{children} + 1$ (Golding and Sokoloff 1982), and the dollar value of all inputs the firms used. The stated dollar value of outputs is used as the output here. With this information, efficiency values for every firm are estimated.[4] To adjust for price changes, the data for capital, inputs and outputs were deflated with price series. The price series and sources are given in the appendix (table A1).

2.3 Factory vs. Artisanal Shop

In this chapter the existence of economies of scale in the antebellum industry is discussed. This question is of importance as it is connected with the question whether the change from artisanal shops to the factory systems led to efficiency increases and if yes, what were the

[2] In modern competitive markets values of nearly 200 % are surely implausible, so one could argue that the threshold should be somewhat lower than 200 %. But the exact value of the threshold is unknown and therefore a high value is used as the antebellum industry was surely not completely competitive.
[3] The data were made available by Jeremy Atack.
[4] For every firm the other firms of the same industry are used to construct the efficiency frontier.

reasons for this increase. For a long time, historians have diverged in their opinion on this question. The main empirical contributions to this question are Sokoloff (1984) and Robinson and Briggs (1991). Sokoloff first showed that the size of establishments increased significantly from 1820 to 1850. He then estimated various production functions, in which he found evidence for increasing returns to scale for firms until they had between 5 and 16 employees, which indicates that the increasing size of the firms in the antebellum industry led to increases in efficiency. His results were questioned by Robinson and Briggs (1991) who found no evidence for returns to scale in a small sample of firms from Indianapolis. They also argued that many factories were only larger shops, in which the same production techniques were used with a larger workforce. This does not lead to increased efficiency.

One problem with estimations of production functions is that one has to explicitly specify the functional form. Though the specifications Sokoloff and Robinson and Briggs used were very simple and widely used ones,[5] the possibility of specification errors and resulting biases remains. Therefore, our efficiency values are used to test the earlier results.

The Scale efficiency values are regressed on a factory dummy, which equals one for all firms with between 6 and 15 laborers (following the literature; see Sokoloff 1984), and a large factory dummy for firms with more than 15 laborers. As in the other regressions in this chapter industry, time and region dummies are also included.

Table 1: Truncated regressions of scale efficiency on the factory dummies.

	Coefficient	P-Value
Factory	0.58722***	0.000
Large Factory	-0.28925***	0.000
Number of Obs.	0	0.511

[5] They used Cobb-Douglas and Translog-specifications.

Industry Dummies	included	
N	12389	

Notes: */**/***: significant on the 10/5/1 percent level. A crop service establishment with less than 6 workers located in New England is the reference category.

The results for scale efficiency reinforce the findings of Sokoloff. Factories are significantly more scale efficient than small firms, and the coefficient is large and highly significant. Large factories are less efficient than small factories, so the economies of scale vanish after a certain threshold.

Next, a second model is presented, which assumes constant returns to scale. This specification ensures that the results do not report the spread of efficiency in certain size classes, which is possible when variable returns to scale are used and therefore firms are only compared to firms from their respective size class.[6] Table 2 shows the results from the constant returns to scale specification.

Table 2: Truncated regressions of constant returns to scale efficiency values on the factory dummies.

	Coefficient	P-Value
Factory	0.07163***	0.000
Large Factory	0.02798***	0.001
Number of Obs.	0.00017***	0.000
Industry Dummies	included	
N	12408	

Notes: */**/***: significant on the 10/5/1 percent level. A crop service establishment with less than 6 workers located in New England is the reference category.

[6] See Appendix 2 for a detailed examination of the technical reasons for the constant-returns-to-scale specification.

This model shows similar results as the first, with factories being significantly more efficient than shops. It is also confirmed that scale efficiencies vanish somewhere, although this specification shows large factories to be still more efficient than shops.

As Sokoloff remarked, some historians believe in the efficiency advantage of factories only in mechanized industries, as textiles and iron, but not in other industries. To test this, the regressions shown above were repeated by industry. Though, due to limited observations, not all facets of textile and iron manufacturing are in our dataset, the regressions showed that in 13 out of 19 cases the pattern shown above was found, namely higher efficiency in factories, while large factories had somewhat lower values than the normal factories. Rising efficiency values even for the group of larger factories accounts for four of the remaining six cases, while only two cases showed higher efficiency in artisanal shops than in factories. This suggests that the efficiency advantage of factories, though varying in size over the different industries, was a universal facet of the antebellum industry and not only limited to some early mechanizing industries.

Sokoloff estimated the amount of value added when scale economies were exhausted to be 9.500 $ (in prices of 1860) for non-mechanized factories. Our estimates point towards a values of 6.500 $ of value added for the turning point of scale economies. Both the scale and constant returns to scale model lead to similar values here. This means that scale economies existed, but were exhausted early. Though our estimates suggest that scale economies were exhausted earlier than we believed until now, they still show that scale economies were important for a large number of establishments.

2.4 The Northern and Southern Industry before the Civil War

After showing the existence of scale economies, we move on to the differences between Northern and Southern industry. Before we discuss the hypotheses, the situation that has to be explained is described. North and South were, though living under the same constitution, very different in their economic structure. The Northeast was a heavily industrialized region even before the Civil War, while the Southern economy was relying on cotton and some small manufacturing establishments. The most significant institutional difference was slavery. Of the 9 million people living in the Confederacy, 3.5 Million were slaves (Faulkner, Schreiber and Vatter 1976). The low income of the slaves together with the mass of poor white farmers left the South with a small consumer market. The higher urbanization of the North with its large labor force and available capital made the Northern territories a far more suitable place for manufacturing than the South, where most laborers were working in agriculture. But profit rates were not especially low in the South; they were comparable to the returns in agriculture, which means that profitability was not the reason for the slow Southern industrial development (Bateman, Foust and Weiss 1975). The reason for the slow industrialization of the South might be in the preferences of the planter class. They did not like to invest in manufacturing, for example because they did not want to loose political influence or because they believed that the lifestyle of a planter was superior to that of an entrepreneur. As migration was small the influx of new ideas and values was not significant. Investors from outside the South were discouraged by many obstacles, including capital market imperfections (Atack and Passell 1994).

In the literature most authors only discuss differences between South and North, assuming the South to be exceptional. To judge whether the South was really exceptional, one should compare the South with all regions. Comparing South and West, Bateman and Weiss (1975) show that the industrial structure of the two regions looks quite similar before the Civil War. Even though settlement history in the South was longer than in the West, industrial development before 1860 was minor, so the regions are comparable in this regard. Some important industrial facilities were located in the South, such as the Tredegar Iron Works in Richmond. This means that, although the South and the Northeast had different economic structures, the South was not unique. If there were efficiency disadvantages of Southern industry those should also be visible for Western establishments.

2.5 The Hypotheses

In their summary of reasons for the relative backwardness of the American South, Bateman and Weiss discuss several hypotheses (Bateman and Weiss 1976). These range from slavery and a Southern comparative advantage in cotton production to the limited market size hypothesis. On page 47 they say that that "*... organizational and managerial inefficiencies prevented industry from developing in the South as well as it had in other regions"* (The Entrepreneurial Inability Hypothesis).

This means that because of lacking managerial skills or because entrepreneurs did not use modern organizational options Southern firms were not producing as efficiently as it would be possible compared to the state of the art for that time. Examples could include new

accounting techniques, communication possibilities or machinery. Managerial inefficiencies could be attributed to a lacking tradition of managerial qualities in the South, as most rich men were focused on cotton, either because they behaved irrationally, or because they did not accept profit as an important goal, or due to lacking human capital (Bateman and Weiss 1981). Inefficiencies will increase consumer prices, therefore slowing economic development. They may also prevent investors from investing in manufacturing, as the likelihood of failures increases. On the other hand, inefficiencies may attract investors who believe that they can increase the efficiency and profitability of the firms they are investing into. For the South it is also possible that inefficient slave labor made the industry less efficient than their Northern counterpart. We will discuss this question below. Bateman and Weiss argue that those inefficiencies will be visible "*in investment patterns, in the relative efficiency of Southern firms, or in the comparative failure rate of enterprises.*" (Bateman and Weiss 1981, p. 36). In this paper we focus on the second point by analyzing relative efficiency.

A second hypothesis discussed by Bateman and Weiss is "*That effective market size in the South kept manufacturing too small to reap the benefits of internal economies of scale. Southern manufacturing consequently were unable to compete with those in the North, creating a situation that stifled Southern industrialization.*" (The Limited-Market Size Hypothesis) The core of this hypothesis is that economies of scale allow large producers to work at comparably low cost per unit, which gives them an advantage in the competition. If markets in the South were smaller than in the North, for example because of shortcomings in the transportation system, firms will remain smaller and therefore produce on higher costs per unit, ceteris paribus. This can also be interpreted as inefficiency, as larger firms can, due to the lower per unit costs, produce more output per unit of input.

Inefficient producers could dampen economic development, as the price per good is higher, which will decrease the number of goods available in an economy, and additionally tying money into manufacturing that could be invested otherwise. The inefficient firms may not be able to survive the competition from Northern firms and therefore manufacturing firms in the South go bankrupt. This might lead to a smaller industrial sector in the South compared to the situation when all firms are producing efficiently.

2.6 Earlier Research on the Hypothesis of Southern backwardness

Many hypotheses related to Southern industrial development were thoroughly investigated by Bateman and Weiss (1981). The two hypotheses outlined above were also investigated. However, efficiency was not directly measured by the authors. Several other indices of industrial development and success were investigated which give indirect evidence on efficiency. Their findings will be described here so we can compare them with the new estimates below.

As inefficiencies in industry would deter development, we should see differences in industrial indices between the South and other regions. Bateman and Weiss show, for example, that Southern firms were smaller than firms in other parts of the United States, but if we differentiate for industry, the smallness of Southern firms is less clear-cut, for in some industries the Southern firms are the largest in the country. Especially when compared with the West, the size-distribution of the two regions looks quite similar.

The major contribution of Bateman and Weiss is their estimation of profit rates in manufacturing, measured as earnings over invested capital. Estimates of the profitability of

slavery in agriculture are available, but they were the first to present representative estimates for manufacturing. Their estimates, which are similar to the methods related to the slave economy to allow comparability, show profit rates for Southern firms often in excess of 20 percent, up to 40 percent. If production methods were inefficient we would expect a lower profitability, as resources are wasted in the production process and more resources needed to achieve a specific level of output, which lowers profits. The profit rates that could be achieved in Southern manufacturing are surely impressive and higher than the ones available in the plantation system. This means that, although industry still could be inefficient, the inefficiency cannot be very pronounced, because then these high profit rates would not be achievable. This result is strengthened by a comparison of industry profit rates in the different regions of the antebellum United States. The results from Bateman and Weiss show that Southern profit rates are if anything at the top of the profit rates of this time. Even if these high profit rates are the result of monopoly conditions and cover inefficiencies, these inefficiencies are more than counterbalanced by the monopoly profits and are therefore no reason for the manufacturing sector developing at an unusually slow pace.

Regarding the hypothesis of returns to scale which could not be exploited in the South because of small markets Bateman and Weiss find scale parameters that are equal to or smaller than one. They use a Cobb-Douglas production function, which differs from our method below. The coefficients in the South were not different from the coefficients in the other regions, which means that "… for manufacturing as a whole, Southern producers did as well in reaping returns to scale as did firms in every other region." (Bateman and Weiss 1981, p. 59). Though markets in the antebellum South were small, in most cases smaller than the state, usu-

ally only as large as a county, Bateman and Weiss find that market size did not restrain Southern producers from exploiting the returns to scale.

2.7 Results

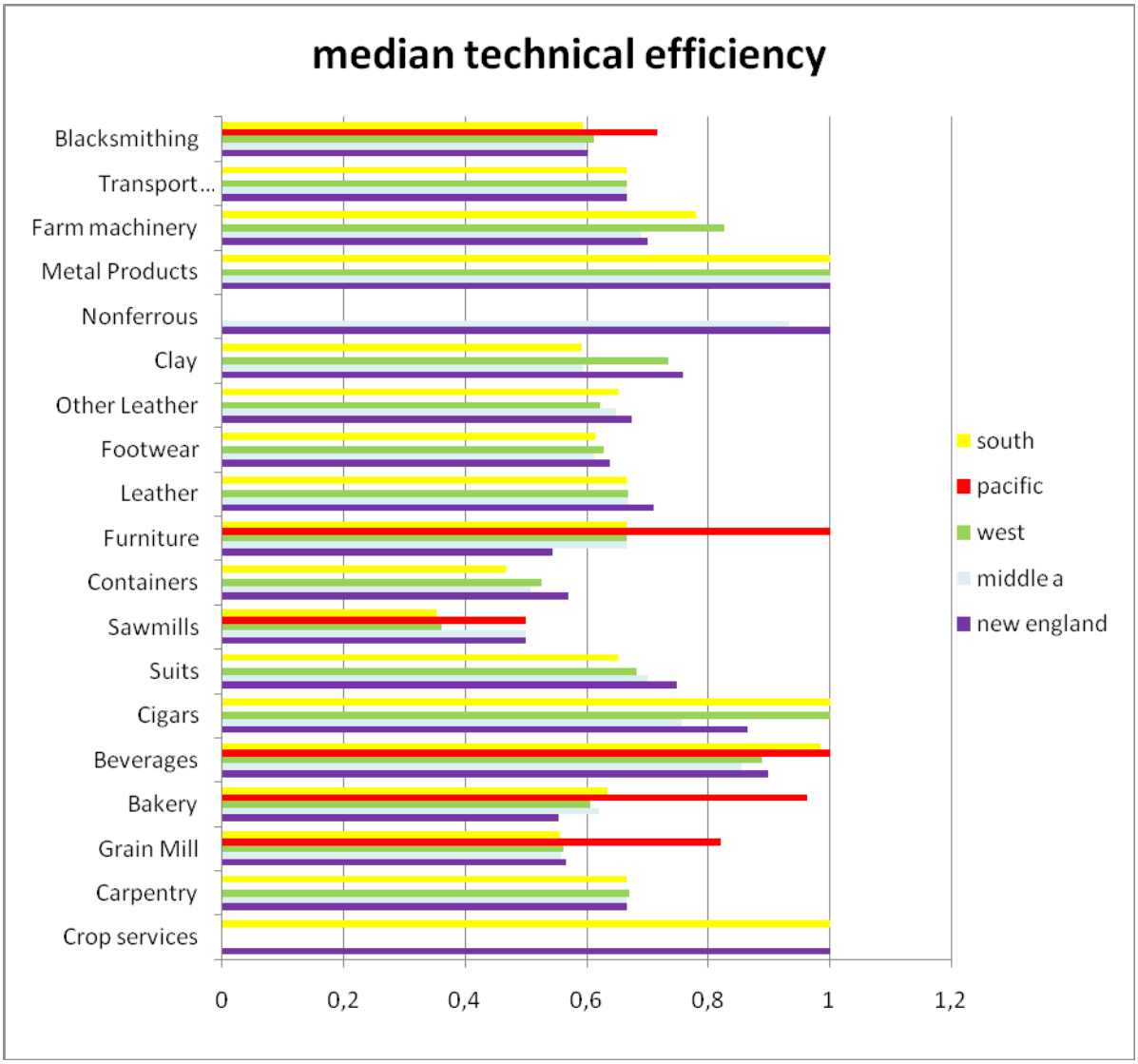
In this section the efficiency estimates are presented. At first we show the median efficiency of each industry and region. We use the 5 regions New England, Middle Atlantic, West, Pacific and South. The definition of the regions follows Atack (1985).[7]

Figure 1 shows median technical efficiency scores for every industry in 1850.

No clear pattern emerges from the figure. No region can claim a definite advantage in technical efficiency over the other regions. Though some industries were not at all located in certain regions which might be due to efficiency disadvantages, this is not the case for the South. In some industries the South is among the least efficient regions, but there is no large and consistent efficiency disadvantage visible. In some industries, like in beverages and cigars, the South is even the most efficient region.

Figure 1: median technical efficiency

[7] New England: Connecticut, Maine, Massachusetts, New Hampshire, Vermont
Middle Atlantic: Delaware, District of Columbia, Maryland, New Jersey, New York, Pennsylvania
Pacific: California, Oregon, Washington
South: Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Texas, Virginia, West Virginia
West: Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, Ohio, Wisconsin.

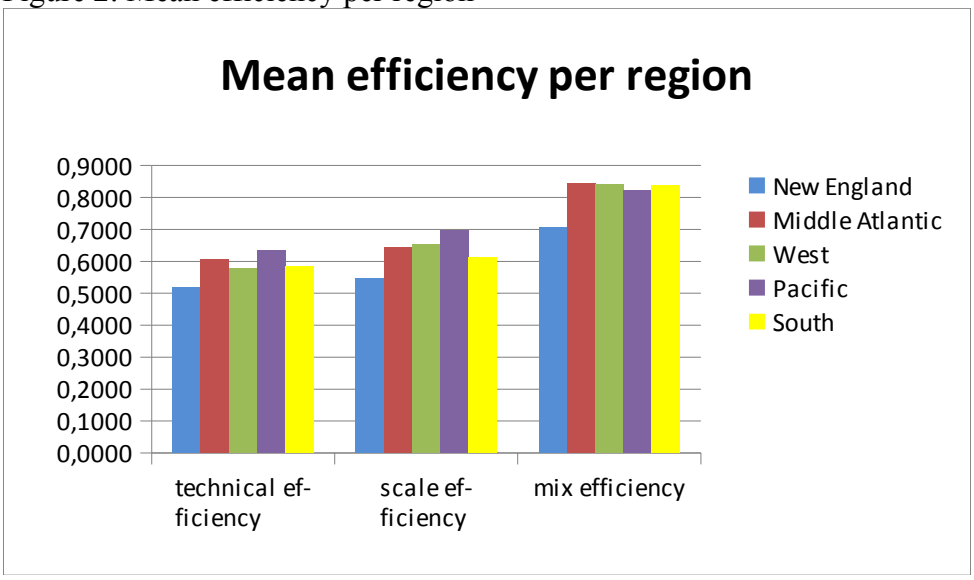


Source: own calculations. For the data sources see text.

The results for scale and mix efficiency and those for technical efficiency in 1860 can be found in the appendix (figures A1 to A5). None of these figures reveals large and persistent differences between the regions, strengthening the conclusions drawn above.

Mean efficiency per region is shown in figure 2. The underlying numbers can be found in table A2 in the appendix.

Figure 2: Mean efficiency per region



Source: own calculations. For the data sources see text.

All regions show similar efficiency values, the South has a small disadvantage in scale efficiency, while in technical and mix efficiency it is in the middle of all regions. The visible differences here are mostly around 5 to 10 percentage points, which is an effect of medium size, as the distance from the 25 % quintile to the median is about 12 percentage points and from the median to the 75 % quintile it is 15 percentage points. As the South was economically similar to the West, this result seems reasonable. So the consideration of industry or regional differences does not show large differences.

In a further step we will have a closer look on the efficiency scores. By using truncated regressions (see Simar and Wilson 2007 who propose the truncated regression approach when using DEA estimates) we can have a more detailed look and access the statistical significance of the differences. It would also be possible that the values found above might come from differences in the industrial composition of the regions, and in this case the regressions would reveal these hidden differences. The efficiency scores are regressed on industry dummies, re-

gional dummies and the number of firms used in the estimations.[8] We use two separate models for 1850 and 1860; the reference category is a crop service firm in the South.

Table 3: Dependent Variable: Technical efficiency scores.

	1850	1860
West	0.00667 (0.287)	-0.00745 (0.259)
Pacific	0.12148*** (0.000)	0.05084*** (0.000)
New England	0.03074** (0.000)	0.01276 (0.159)
Middle Atlantic	0.01184*** (0.004)	0.01514** (0.023)
Industry dummies	included	included
N	5369	6220

Notes: P-Values in parentheses. */**/***: significant on the 10/5/1 percent level. For 1860 the industries crop services and nonferrous industries had to be skipped. A Southern carpentry firm is the reference category. Robust standard errors are used.

The South is significantly inferior in technical efficiency to the Pacific, Middle Atlantic and New England region in 1850 (see table 3). In 1860, the Pacific and Middle Atlantic region are more efficient than the South, while now the West is slightly inferior to the South. Both the coefficient for the West and New England are not statistically significant. The differences are not large, except for the pacific dummy. Apart from this, the estimated differences between the regions are smaller than 3 percentage points in efficiency. This means a weak to medium support for the *entrepreneurial inability hypothesis*. One possible explanation next to the ones discussed above is that Southern entrepreneurs had less experience than their Northern competitors, although we know that some firms hired Northern managers (Starobin 1970). Another possibility is that slaves, who were employed extensively in industry, were not as efficient as white workers. But Starobin's calculations show that slaves were cheaper than white

[8] See Behle (2009) for more details on the variables.

workers, at least by 20 percent, and he gives evidence of businessmen claiming that slaves would work harder than white workers, one observer even claiming them to be 50 percent more productive than whites. Some voices also claimed slaves to be more reliable than whites, as they were not striking or bargaining for higher wages. Together this may mean that if anything, slave labor in industry was more efficient than free labor.

Another reason for the inefficiency of Southern manufacturing might be technology diffusion. Important inventions that were made in Northern cities needed time to diffuse to firms in remote parts of the United States. Sokoloff (1988) shows patenting rates per 1 million residents of the US per region for 1846 and finds values from 20 to 65 for Northern counties, while for the rest of the US he states a figure of 9.9 (the national average being 27.3). This shows clearly that Northern firms had cheap access to many patents, even those who did not patent themselves, while Southern firms were far away from the centers of inventive activity. He also argues that waterways, because they lower the transportation costs to markets, led to higher patenting activity in their proximity. So it might be that newer machines available to Northern firms are responsible for the efficiency disadvantage of Southern industry. Immigration might have also increased this problem. As David Jeremy (1981) concluded, much of the technology transfer between Britain and the United States occurred in the form of immigration. As most immigrants stayed in the healthier climate of the North (Ferrie 1996), this imported knowledge was available in Southern states to a smaller extent only, giving the North an advantage. A second possibility is the effect of learning-by-doing, which discriminates against the younger industries of the South and West. Paul David for example showed the existence of learning-by-doing effects in Northern textile mills (David 1970). These effects were substantial and can also explain part of the efficiency differences.

Table 4: Dependent Variable: Scale efficiency scores.

	1850	1860
West	0.10168*** (0.000)	0.07416*** (0.000)
Pacific	0.62751*** (0.000)	0.21886*** (0.000)
New England	0.02941* (0.090)	-0.01028 (0.674)
Middle Atlantic	0.00544 (0.669)	-0.01896 (0.307)
Industry dummies	included	included
N	5871	6518

Notes: P-Values in parentheses. */**/***: significant on the 10/5/1 percent level. For 1860 the industries crop services and nonferrous industries had to be skipped. A Southern carpentry firm is the reference category. Robust standard errors are used.

Table 4 shows that in regard to scale efficiency, the South in 1850 is inferior to all regions except the Middle Atlantic states, where the difference is not significantly different from zero. In 1860, the New England and Middle Atlantic states are inferior to the South, though these differences are not significant. Southern firms were generally smaller than their counterparts in the Northeast (Bateman and Weiss 1981, p. 52), and my estimations show that there was some disadvantage for Southern firms.

Table 5: Dependent Variable: Mix efficiency scores.

	1850	1860
West	0.01718*** (0.000)	-0.00064 (0.880)
Pacific	0.00743 (0.514)	-0.03061*** (0.000)
New england	-0.00120 (0.781)	0.00241 (0.676)
Middle Atlantic	0.01045*** (0.001)	0.00560 (0.159)
Industry dummies	included	included
N	5884	6528

Notes: P-Values in parentheses. */**/***: significant on the 10/5/1 percent level. For 1860 the industries crop services and nonferrous industries had to be skipped. A Southern carpentry firm is the reference category. Robust standard errors are used.

Mix efficiency is significantly worse in the South than in the West and the Middle Atlantic region, while the New England states are slightly inferior (table 5). In the next decade the South is catching up heavily so that in 1860 the New England and Middle Atlantic region have a slight advantage, which is all but significant. The West and Pacific have lower mix efficiency scores than the South, for the pacific region the difference is highly significant. This shows that the South might have had a disadvantage in mix efficiency first but it was eliminated until 1860.

One interesting feature of the estimations above is the excellent achievement of the Pacific region in scale and technical efficiency. The three states that are included in this region, Washington, Oregon and California, were only in the beginning of colonization at the time of the Census of 1850. Oregon for example had only a resident population of around 12 000 people in 1850, Washington just about 1 200.[9] Only California was already part of the Union, Oregon (1859) and Washington (1889) joined the Union later. Though most people came due to the gold rush, it is important to note that in the estimations above only manufacturing establishments and no gold mines were included to maintain comparability. In this early stage of development the few existing firms could be built at the best places with the cheapest and easiest access to water-power or raw materials. This might have enhanced efficiency as here only the best places were occupied while in other regions places of minor quality had to be used for the erection of a mill. This explanation is also consistent with the low mix efficiency

[9] Data come from the US Census Office. http://www.census.gov/dmd/www/resapport/states/washington.pdf and http://www.census.gov/dmd/www/resapport/states/oregon.pdf. Access date April, 9th, 2009.

scores in the Pacific region, as the far better infrastructure in the older regions favored the industry there.

2.8 Efficiency and long-run Economic Growth

The efficiency differences documented above might have contributed to the slow industrial development of the Southern states. This section tests the hypothesis more formally. If lacking efficiency was a major barrier to Southern industrial development, then we should observe that regions with higher efficiency scores grew more quickly in their industrial sector than other regions. To check this hypothesis correlations between the coefficients from the regressions above and the development of manufacturing output per region are shown in table 6. The manufacturing output data come from the Manufacturing Censuses[10] and were aggregated by the regions defined previously. From them yearly growth rates of manufacturing output between 1850 and 1860 were estimated. Growth rates for longer time periods are difficult to estimate because of the war, which changed the efficiency rankings somewhat.

Table 6: Correlations between efficiency estimates and the growth of manufacturing output.

	PTE	SCALE	MIX
Correlation	-0.41	0.92	0.79

Source: own calculations.

The correlation coefficients do not reveal a clear pattern. Scale and mix efficiency are strongly and positively correlated with subsequent growth as expected. In contrast, technical efficiency is negatively correlated with growth. This might indicate catch-up growth. The idea

[10] Accessed from http://mapserver.lib.virginia.edu/php/newlong2.php. Access date September, 9th, 2010.

here is that regions that are in early stages of development, which have small industrial sectors with low efficiency, will grow faster than the already developed ones where the prospects for manufacturing growth were already exploited.

To test this idea, the growth rate between 1850 and 1860 is regressed on the manufacturing output per capita in 1850 and the efficiency scores. If catch-up growth existed, regions with high per capita manufacturing output will grow more slowly than the ones with small industrial sectors. Therefore, a negative impact of initial values on growth is expected.

Table 7: Determinants of the growth of manufacturing output per person from 1850 to 1860.

	Coefficient	P-Value
Output1850	-0.00372**	0.043
PTE1850	-7.17	0.325
R^2	0.24	
N	32	

Notes: */**/***: significant on the 10/5/1 percent level.

The results in table 7 show evidence of catching-up growth, the coefficient of the initial level of output being negative and statistically significant. The coefficient for technical efficiency in turn is negative, but not significant. This result is the same for scale efficiency. The regressions therefore point to catching-up growth in manufacturing, but suggest that efficiency differences were not responsible for differences in manufacturing output growth.

To sum up, our results lend little support to the hypothesis that regional efficiency differences were responsible for differences in economic development.

2.9 Conclusion

For a long time historians have brought forward industrial efficiency as an explanation for the antebellum industrial development in the US. Despite its importance in the historical literature, quantitative evidence on efficiency trends is scarce. This paper provides direct evidence for the first time. The Data Envelopment Analysis is used to measure efficiency and truncated regressions are used in a second step to link efficiency with its correlates. With this technique, the existence of economies of scale is shown, which are available for small shops and factories. This result is in line with Sokoloff (1984) and indicates that the transition from the artisanal shop to the early factory was associated with efficiency gains.

After showing the existence of economies of scale, the efficiency measures were used to test the hypothesis that efficiency disadvantages of Southern firms were responsible for the different economic development of the Southern and Northern industry during the nineteenth century. The resulting picture is somewhat mixed. If we look at technical efficiency we see that the South is one of the least efficient regions, but there is no consistent disadvantage visible. Looking at scale efficiency we see the South in an average position, so the hypothesis above cannot be confirmed. Also for mix-efficiency we find this average achievement of the Southern industry.

The resulting differences do not suggest that regional efficiency differences were the crucial factor for the differences in the developing industrial structures in these regions. But neither do our results refute the idea that efficiency was one factor. It seems therefore that the South was not as efficient as the North and that this was one factor that determined the path of development.

References

Atack, J. (1985). The estimation of economies of scale in nineteenth century United States manufacturing. New York: Garland.

Atack, J. and P. Passel (1994). A New Economic View of American History from Colonial Times to 1940. Second Edition. London and New York: W. W. Norton & Company.

Bateman, F., Foust, J. and T. Weiss (1975). Profitability in Southern Manufacturing: Estimates for 1860. Explorations in Economic History 12 (3), 211-231.

Bateman, F. and T. Weiss (1976). Manufacturing in the Antebellum South, in Lindert, Peter et. al. (eds.): Research in Economic History: An Annual Compilation of Research, 1. Greenwich, Conn: JAI Press.

Bateman, F. and T. Weiss (1981). A deplorable scarcity: the failure of industrialization in the slave economy. Chapel Hill: University of North Carolina Press.

Behle, D. (2009). The impact of the American Civil War on efficiency in manufacturing. Working paper. University of Tuebingen.

Carter, S., Gartner, S. S., Haines, M. R., Olmstead, A. L., Sutch, R. and G. Wright (eds.). Historical Statistics of the United States: Earliest Times to the Present. Millennial Edition. New York: Cambridge University Press.

Craig, L. and E. Field-Hendrey (1996). Were free Southern farmers "driven to indolence" by slavery? A stochastic production frontier approach. NBER working paper. Historical Paper 82.

David, P. (1970). Learning By Doing and Tariff Protection: A Reconsideration of the Case of the Ante-Bellum United States Cotton Textile Industry. The Journal of Economic History, 30 (3), 521-601.

Faulkner, H. U., Schreiber, H. N. and H. G. Vatter (1976). American Economic History, 9. Edition. New York: Harper & Row Publishers Inc.

Ferrie, J. P. (1996). The Impact of Immigration on Natives in the Antebellum U.S. Labor Market, 1850-60. Working Paper, NorthWestern University.

Field, E. (1988). The Relative Efficiency of Slavery Revisited: A Translog Production Function Approach. American Economic Review 78 (3), 543-549.

Field-Hendrey, E. (1995). Application of a stochastic production frontier to slave agriculture: an extension. Applied Economics 27 (4), 363-368.

Fogel, R. W. and S. L. Engerman (1974). Time on the Cross: The Economics of American Negro Slavery. Boston: Little Brown.

Goldin, C. and K. L. Sokoloff (1982). Women, Children and Industrialization in the early Republic: Evidence from the Manufacturing Censuses. The Journal of Economic History 42 (4), 741-744.

Grabowski, R. and C. Pasurka (1988). The relative technical efficiency of Northern and Southern U. S. farms in 1860. The Southern Economic Journal 54, 598 – 614.

Grabowski, R. and C. Pasurka (1989). The relative efficiency of slave agriculture: An application of a stochastic production frontier. Applied Economics 21 (5), 587 – 595.

Hofler, A. R. and S. T. Folland (1991). The relative efficiency of slave agriculture: a comment. Applied Economics, 23 (5): 861-868.

Jeremy, D. J. (1981). Transatlantic industrial revolution: The diffusion of textile technologies between Britain and America, 1790-1830. Cambridge, Mass.: MIT Press.

Robinson, R. V. and C. M. Briggs (1991). The Rise of Factories in Nineteenth-Century Indianapolis. The American Journal of Sociology 97 (3), 622-656.

Schäfer, D. and M. Schmitz (1982). Efficiency in Antebellum Southern Agriculture: A Covariance Approach. Southern Economic Journal 49 (1), 88-98.

Simar, L. and P. W. Wilson (2007). Estimation and inference in two-stage, semi-parametric models of production processes. Journal of Econometrics 136 (1), 31-64.

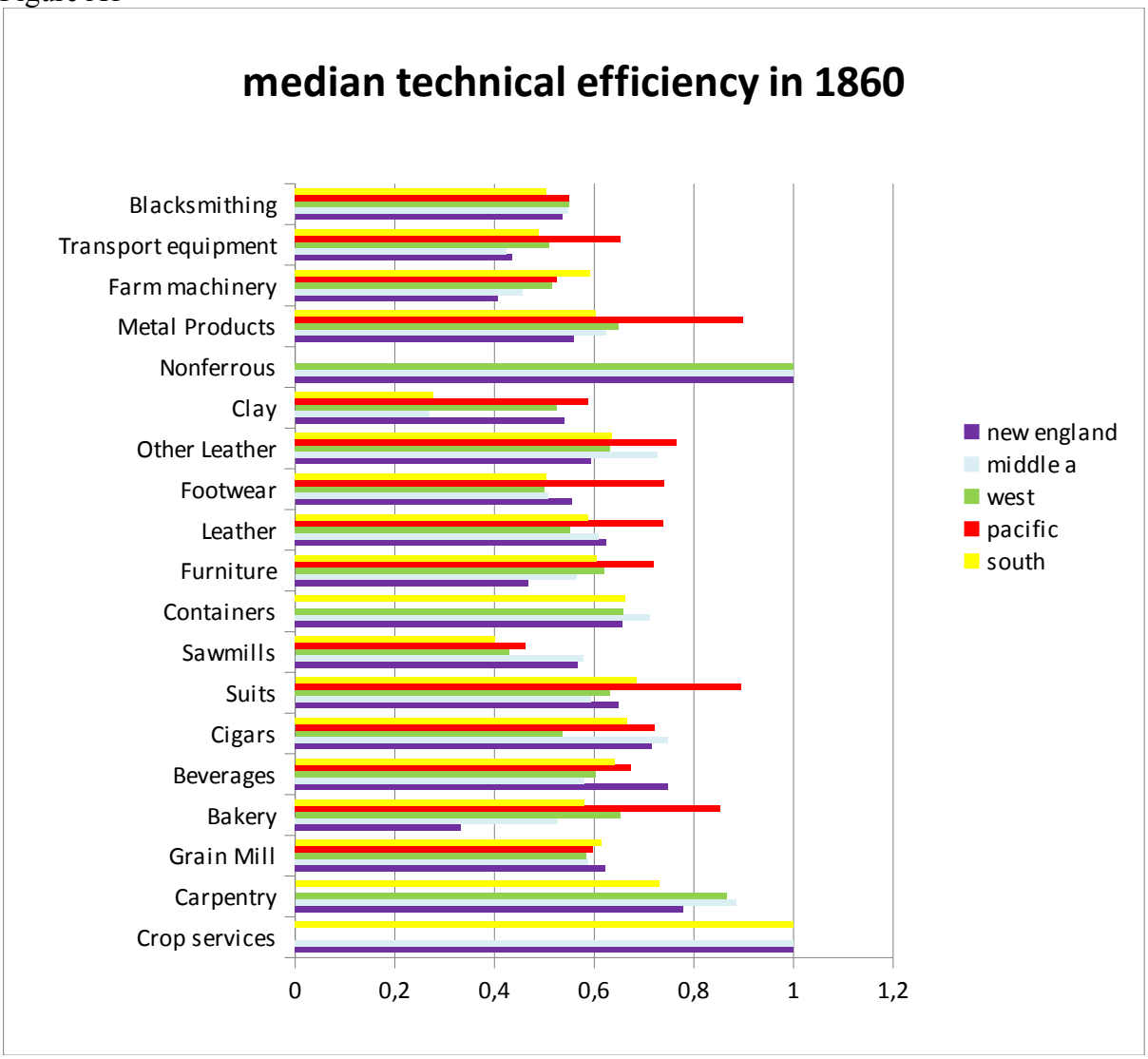
Sokoloff, K. L. (1984). Was the transition from the artisanal shop to the Nonmechanized Factory Associated with Gains in Efficiency? Evidence from the U.S. Manufacturing Censuses of 1820 and 1850. Explorations in Economic History 21 (4), 351-382.

Sokoloff, K. L. (1988). Inventive Activity in Early Industrial America: Evidence From Patent Records, 1790-1846. The Journal of Economic History 48 (4), 813-850.

Starobin, R. S. (1970): Industrial Slavery in the Old South. London, Oxford, New York: Oxford University Press.

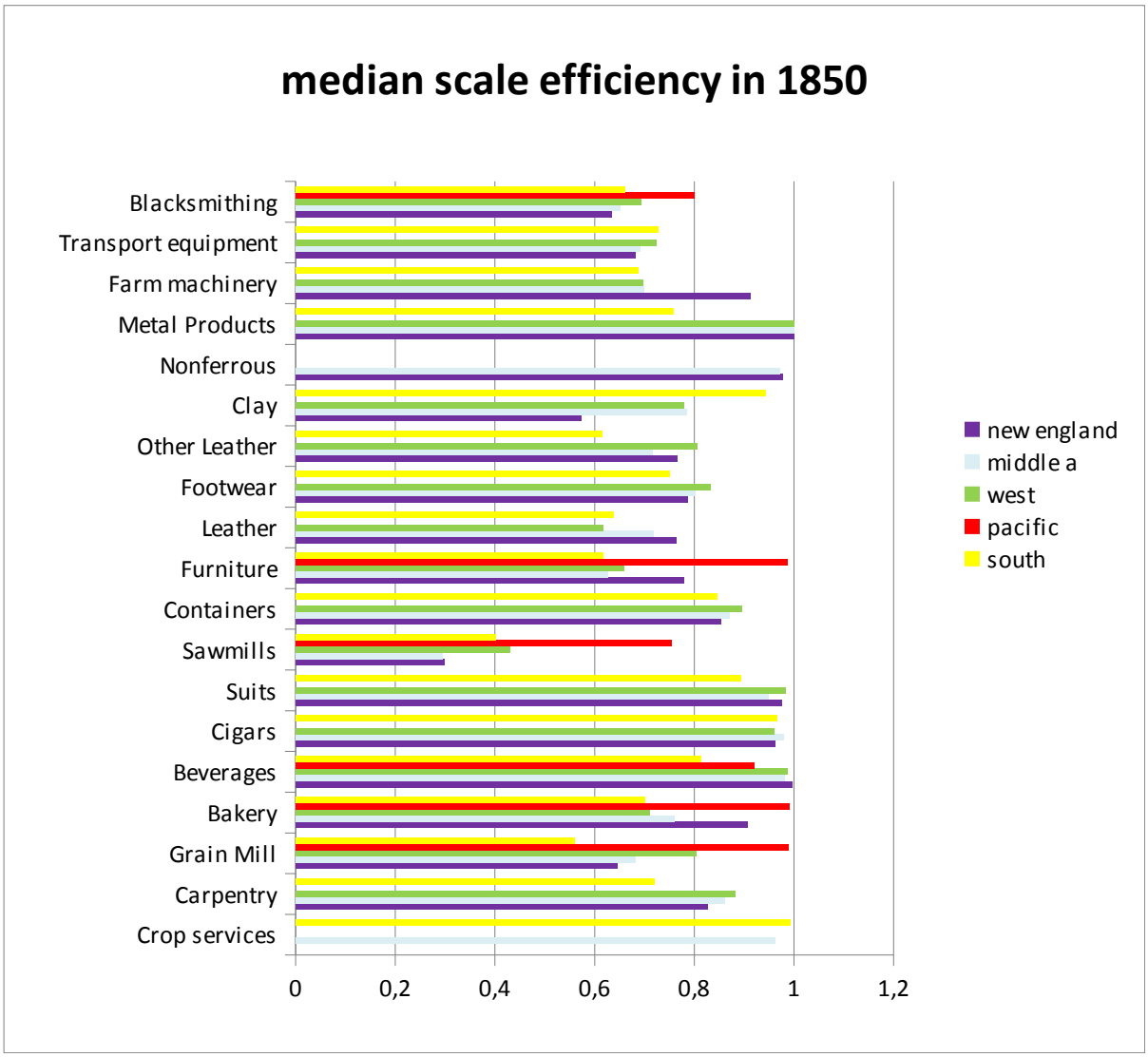
Appendix 1: Figures and Tables

Figure A1



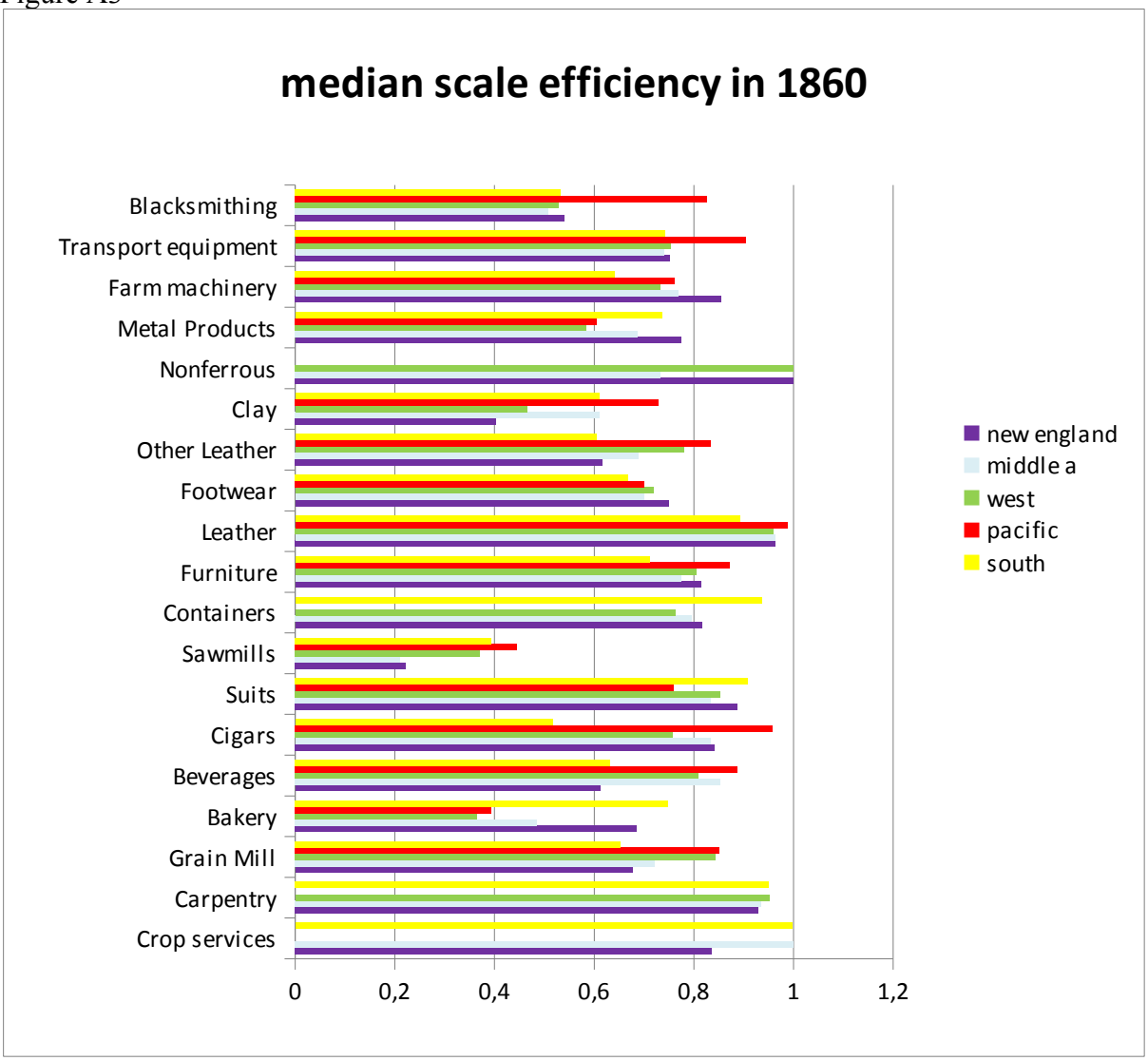
Source: own calculations. For the data sources see text.

Figure A2



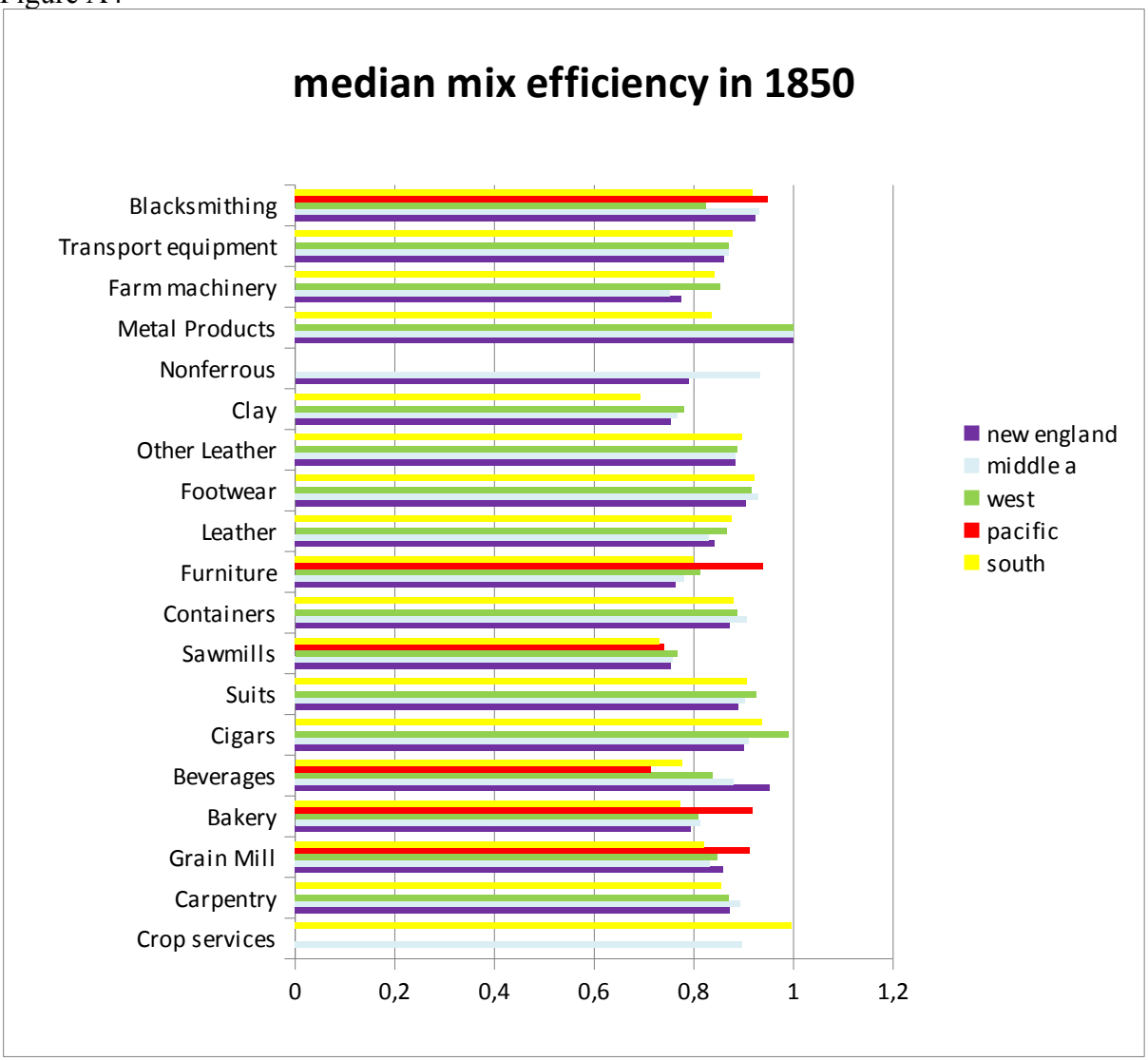
Source: own calculations. For the data sources see text.

Figure A3



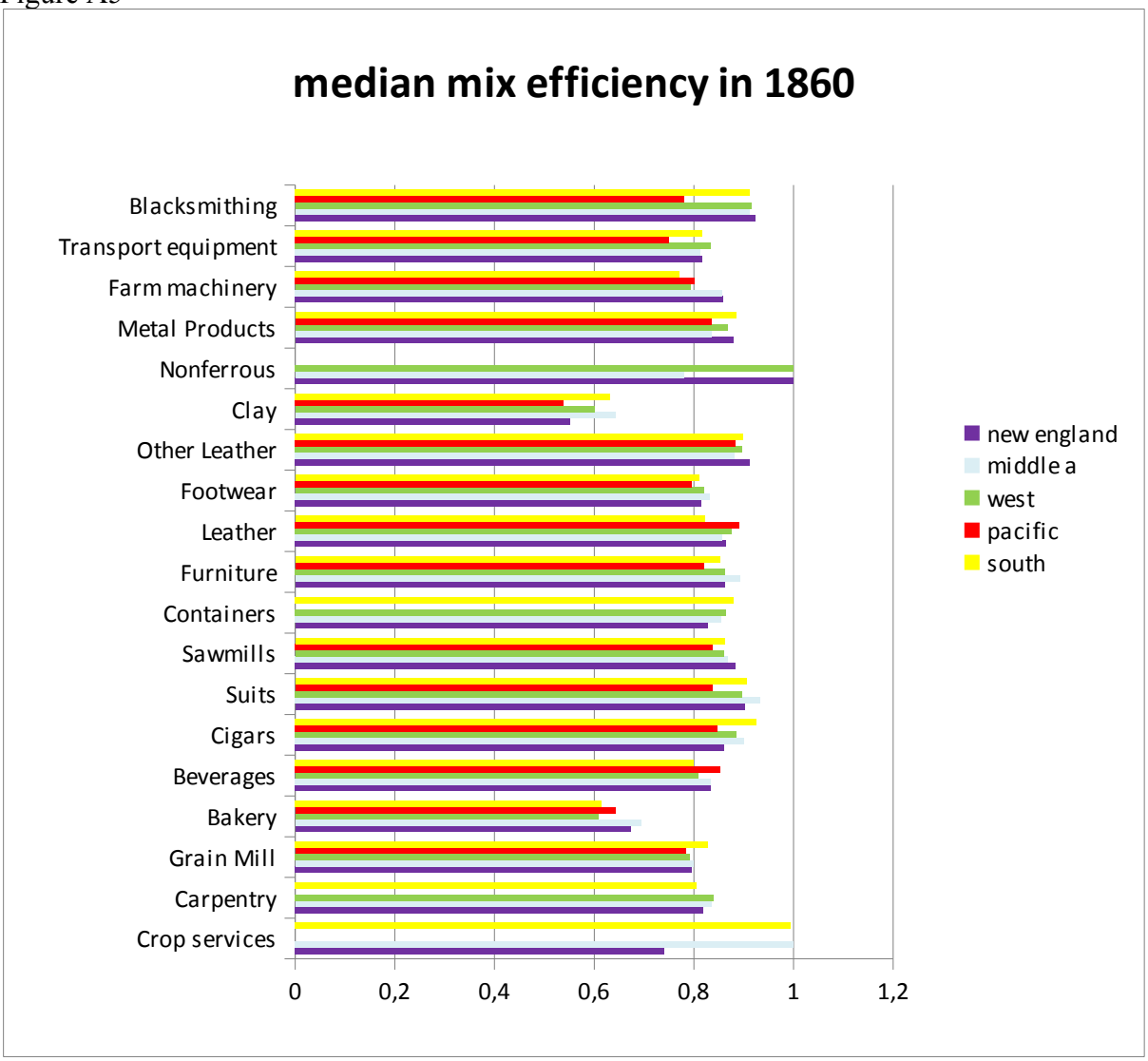
Source: own calculations. For the data sources see text.

Figure A4



Source: own calculations. For the data sources see text.

Figure A5



Source: own calculations. For the data sources see text.

Table A1: Price series.

	South	Pacific	Rest
1850	110	236	84
1860	112	113	93

Notes: Prices come from Carter et. al. (2006).

Table A2: Mean efficiency values per region

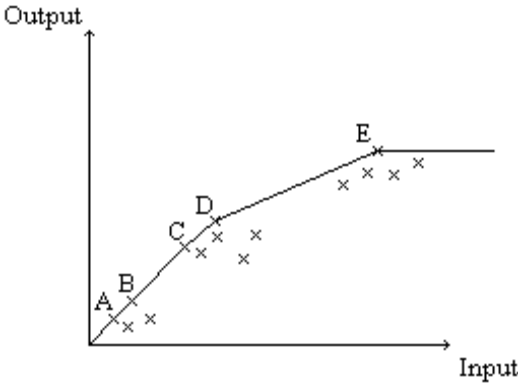
	Technical Efficiency	Scale Efficiency	Mix Efficiency
New England	0.5193	0.5458	0.7072
Middle Atlantic	0.6070	0.6430	0.8437
West	0.5764	0.6525	0.8410
Pacific	0.6344	0.6979	0.8228
South	0.5856	0.6126	0.8377

Source: own calculations. For the data sources see text. The numbers show what share of the output of a firm on the frontier the actual firm achieved.

Appendix 2: The Constant Returns to Scale specification

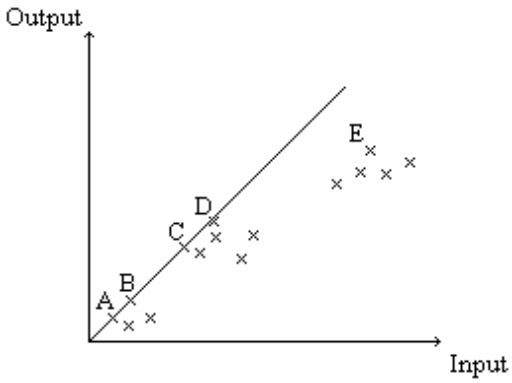
As was already pointed out in the introduction, efficiency is a relative concept, and the DEA is the mathematical representation of this concept as it estimates efficiency of an observation relative to the other observations in the dataset. Using variable returns to scale leads to an efficiency frontier as depicted in figure A 6:

Figure A 6: The Variable Returns to Scale Model.



In the picture above the small establishments will be evaluated against the frontier that is constituted by the firms A, B, C, D and E and linear connections between them. Therefore, although the technology inhibits significant decreasing returns to scale and the ratio of output to input is much higher for the small establishments, all three groups will have high efficiency values, as they are all close to the efficiency frontier. This means that efficiency values show in this case more about how homogeneous the three groups are than about their respective performances. If we used a constant returns to scale model, the resulting frontier would look like in figure A 7.

Figure A 7: The Constant Returns to Scale Model.



Now, Firms D and E would receive efficiency values of less than 100 percent, they are identified as inefficient. Additionally, as E is now no longer on the frontier, the large firms would all receive lower efficiency values. This indicates that they cannot reach the ratio of output to input that the smaller ones achieve. This example shows that the constant returns to scale model is more appropriate to evaluate the efficiency of different size classes compared to each other.

3 The Impact of the American Civil War on Manufacturing Efficiency.

Abstract:

This paper uses the Data Envelopment Analysis to measure the development of efficiency around the American Civil War and to compare the counties that were affected by battles with those that were not. The results show that the battleground counties developed better in technical and scale efficiency, and had a disadvantage in mix efficiency relative to the counties that were not affected. Comparing the impact of the abolition of slavery and the war destruction we find the former to have had a stronger impact on manufacturing.

This chapter is based on a working paper with the same title currently under review at the journal of business history.

3.1 Introduction

The American Civil War is an event that has been intensively studied. With regard to manufacturing, scholars have asked what impact the war had on the firms and whether this can explain Southern backwardness. In looking at the consequences of the war for industry one can take two different views. One is to look at the structure of the manufacturing sector and see how many establishments were destroyed and how output expanded or contracted. A second view is to look inside the industry and look at the production processes. Did the war introduce new production methods? How strong was the loss of capital, e.g. how many machines were missing after the war? These questions relate to the efficiency of firms, which will be altered through damages and will have an impact on the development of industry. This topic has not been studied quantitatively before, and in order to accomplish this task we use the Data Envelopment Analysis, which has not been used widely in the war literature, to gain direct evidence on efficiency changes through the war.

War destruction has a negative influence on industry, as capital is destroyed, workers are killed and communication lines are disrupted. However, destruction may also have positive consequences. If more advanced technology is available at the time of reconstruction, it will give the newly developed industry a technological advantage. Olson (1982) argued that this cannot be beneficial to the economy, as better technology would have been implemented even without the war (see Arbetman and Kugler 1989), but if capital markets are imperfect, war can lead to the use of superior technology that otherwise would not have been used. This may increase efficiency in industry.

In this study, we provide direct evidence of efficiency changes by a war while we do not have to rely on conclusions drawn from growth studies. Specifically, we consider the effect of the American Civil War on manufacturing efficiency. The Civil War has some features that are favorable for our analysis. First, we have data shortly before and after the war, which minimizes the effect of third factors. Second, we can compare regions that suffered from the war and regions that were not directly affected by the fighting which have a similar institutional context and are therefore comparable. Third, we have a large data set from the American Manufacturing Census at hand which decreases the likelihood of finding spurious results that are due to the small sample sizes and are not related to the war.

The discussion about the effects of the Civil War on industry has a long tradition in the economic history literature, and we will discuss this literature below. A systematic discussion of the effect of the Civil War on efficiency is though still missing in the literature. The impact of the War on industry is often treated in an aggregated way, such as in the literature discussed below, or in small samples like Gallman's (1990) study on Philadelphia. Contrary to that we use a large firm-level data set.

We use the Data Envelopment Analysis technique to estimate manufacturing efficiency and to separate it into scale, mix, and technical efficiency.[11] Afterwards, regressions are used to separate the effects of the war from third factors, especially the abolition of slavery. As this event freed black slaves in manufacturing and as it is possible that slaves and free workers are not equally productive, it could influence efficiency and therefore our regressions separate the effects of the two events.

[11] Technical efficiency refers to the amount of waste produced in the production process, scale efficiency means how successful scale economies were exploited while mix efficiency looks at the overutilization of production factors. Below we provide a more detailed treatment of the concepts.

There are some studies that come close to the method and period used in the present paper (Grabowski and Pasurka 1988 and 1989, Field-Hendrey 1995, Craig and Field-Hendrey 1996). These articles are concerned with agricultural efficiency and the effect of slavery. None of these studies is concerned with the effect of the Civil War on manufacturing as is the case here. Hutchinson and Margo (2006) study capital intensity and labor productivity with a production function approach. As a by-product they look at the development of economies of scale as we do, but our approach differs as we also study both technical and mix efficiency.

This work develops as follows: The next section describes which economic influences the war had on industry. Then the Data Envelopment Analysis is briefly explained and how it can be used to decompose measured efficiency into technical-, mix-, and scale-efficiency. From some theoretical considerations, testable hypotheses are derived about the effect of the war on efficiency. We then describe the data, present our estimates and discuss the results. The last section concludes.

3.2 The Civil War and its Economic Impact on American Industry

The American Civil War was the most devastating event in the history of the United States as the direct costs alone can be estimated at about seven billion dollars (in 1860 US-$) while the GDP of the United States in 1869 was just about four billion dollars (O'Brien 1988, Ransom 1998). From a total population of around 30 million people, 620,000 were killed during the war. Destruction in the South was severe and the old plantation system based on slavery was dismantled by the emancipation of the slaves. In the major cotton producing states in the

South output per person of the rural population failed to reach its pre-war level for the rest of the nineteenth century (O'Brien 1988). William Russel attempted to pin down the impact of the war and the end of slavery on the South and concludes that the negative effect from the abolishment of slavery affected the incomes of southerners more strongly than the war itself (Russel 2001).

Destruction in the South was unequal. Georgia and the Shenandoah Valley for example were severely devastated as Union troops tried to destroy the crop supplies of the confederate armies. Other states experienced less direct fighting. In Georgia, General Sherman actually used the census of 1860 to help identify the rich counties that had many valuable goods to seize (Hughes 1994).

There was a lively debate in the literature concerning the impact of the Civil War on manufacturing (Engelbourg 1979). Charles and Mary Beard (1927) and Louis Hacker (1940) claimed that the Civil War was a decisive event that strengthened economic growth. This view was later tested intensively and largely rejected (Engerman 1966, Cochran 1961, Salsbury 1962). It seems as if the Civil War had no boosting effect on manufacturing at all, and although the period of the Civil War was a period that led to the increased distribution of agricultural machinery (Schmidt 1930), few new machines were invented (Wright 1943). In some industries however, such as small arms and textiles, the production processes were modified by the Civil War. But other important industries were not at all changed by the war (Clark 1962).

3.3 The Data Envelopment Analysis

Data Envelopment Analysis (DEA) is a technique that allows the estimation of efficiency values of decision making units (DMU). In this paper every establishment is a decision making unit. It is a non-parametric technique that uses a given set of DMUs to estimate a production frontier so that every DMU lies on or below the frontier. Every DMU is then compared with the frontier. An establishment that produces with the same amount of input only 70% of the output of a second establishment on the frontier will receive an efficiency score of 70%. No assumption regarding the underlying production function or market structure is necessary, only an assumption concerning the returns to scale. It can deal with multiple inputs and outputs, and can be seen as a multivariate generalization of the simple output/input measure.[12] No weights have to be assigned to the inputs or outputs a priori, as those are estimated from the data. A basic DEA-model for n DMUs can be represented by the following linear program (Cooper, Seiford and Tone 2006).

$$\max_{\mu,\nu} \theta = \mu_1 y_{1o} + \ldots + \mu_s y_{so}$$

subject to

$$\nu_{1o} x_{1o} + \ldots + \nu_m x_{mo} = 1$$

$$\mu_1 y_{1j} + \ldots + \mu_s y_{sj} \leq \nu_1 x_{1j} + \ldots + \nu_m x_{mj} \, (j = 1,\ldots,n)$$

$$\nu_1, \nu_2, \ldots, \nu_m \geq 0$$

$$\mu_1, \mu_2, \ldots, \mu_s \geq 0$$

[12] It is the fraction of a weighted sum of outputs over a weighted sum of inputs.

Here θ denotes the efficiency score, μ denotes the weights for the s outputs, y the corresponding output, ν denotes the weights for the m inputs, while x denotes the input. All weights have to be non-negative, and the weighted output must be smaller or equal to the weighted input, so that the efficiency score is between 0 and 1.

The DEA has its origin in a paper by Farrell (1957) and is common in the literature since the 1980s. Today it is widely used to measure the efficiency of different institutions like US Army bases or hospitals (See Taveres 2003 for a literature review).

There are a number of different models available which use different assumptions, in particular concerning the nature of returns to scale in the production function. These different models can be used to decompose the efficiency of DMUs into different efficiency-categories, namely pure technical efficiency, scale efficiency and mix efficiency (Cooper, Seiford and Tone 2006, chapter 5, Diacon, Starkey and O'Brien 2002). Pure technical efficiency (pte) means how efficient the DMU used the available inputs, so that two DMUs using the same inputs but with different levels of output will have different levels of pure technical efficiency. For example, if a new machine or organizational invention enables a firm to produce more goods with the same amount of labor and capital then its pte scores will increase. Also a reduction in the amount of waste that is produced increases technical efficiency. Contrary to that a DMU will be called mix-inefficient if it can reduce input of one factor without reducing output. This DMU could increase the produced output by using the inputs in a different mix, thereby gaining efficiency. Scale efficiency refers to the case when a DMU does not produce in the most productive scale size. This means for example that a DMU which is operating under increasing returns to scale will be called scale-inefficient. When this DMU increases the

size of its operations, it would, due to the increasing returns to scale, achieve a higher ratio of outputs to inputs. Of course, all three sources of inefficiency may appear simultaneous.

In the following we will first report the development of the median efficiency of an industrial sector, which means the median of all efficiency scores from that sector obtained from input-oriented models. Looking at the median efficiency gives us information about the distribution of the firms with respect to efficiency. If the median efficiency is, for example, around 0.9 then most firms produce efficiently. If the estimated median efficiency is around 0.5, then the distribution of efficiency scores is much wider and hence in this sector some firms do not make proper use of their inputs.

Looking only at the median efficiency would not allow direct comparisons of industries over time. If the median efficiency decreases from one point in time to a later date it is still possible that the technology in all firms improved, but some firms improved more than others. What is observed here is only how similar the firms are.

We avoid these problems by estimating the frontier using all firms in the data set, but the median efficiency is estimated separately for those firms that were located in the states which were directly impacted by the war and firms that were in the other states. If there was no war impact, the development of the industry in different parts of the country would have been equal with respect to the efficiency measures. So the level of the median efficiency of the Southern firms should develop similarly to the median efficiency of the Northern firms. If the development of efficiency was different in the South compared to the North (evaluated against all enterprises), this indicates a real effect of the war. Because of this we will mainly

discuss the differences in efficiency development between the North and the South, and not whether efficiency actually improved or not.

The North and the South had a different economic structure. The North was more industrialized than the South, which relied heavily on cotton production. This does not influence our results for several reasons. First, we only compare the industries available and differences in agriculture will not influence this comparison. Second, if there are differences between the regions, such as different institutions, these differences will mostly influence the level of efficiency and also the development in the decades around the Civil War. We focus solely on differences in the Civil War decade that are not present in the other decades, hence the influence of institutions is minimized. It was argued in the literature (Pessen 1980) that the institutional situation in the North and the South was roughly equal. What about institutional differences between the United States and the Confederacy? The data point in 1870 could be affected, but the Confederacy lasted only four years and had ended 5 years before the next census. Moreover, there was no peace-time Confederacy, and the war-time effects are exactly what we want to measure. We conclude therefore that institutional differences clearly influence efficiency, but they pose no problem in the setting we use.

For identifying outliers a super-efficiency model is employed. In this model the analyzed firm is excluded from the estimation of the frontier, which means that the efficiency score may exceed 100% if the firm constitutes a part of the frontier in a normal model. An estimated score of 200% for example would mean that the firm could double the quantities of all their inputs, and, producing the same output, would still be defined efficient. DMUs that receive a very high score in a super-efficiency model seem to be much more efficient than the

rest of this industry, which is unlikely in a competitive market. Therefore, those highly efficient establishments are regarded as data errors.

3.4 The Civil War and Manufacturing Efficiency

Regarding the technical side of production, a differentiation between two points in time is necessary. Directly after the war, when most factory-equipment was destroyed, technical efficiency should have lower in the battlefield-regions because fewer machines were available and older equipment was used, which was not as efficient as its more modern counterpart. So in the years directly following the war technical efficiency will be lower in the damaged regions. But as the economy recovers from the war many new machines will be bought and the old equipment will be replaced. The new equipment will be technologically more advanced than the older one because of the technology developed during the time of the war. Wars also lead to many inventions for the military that can be used afterwards for civilian purposes. It may also be the case that governments intervene in the industry to strengthen important military suppliers, which will lead to efficiency gains in those factories. Some years after the war technical efficiency of the firms impacted by the war will be higher than in the other firms, as they use new equipment, while in the other firms not all of the old equipment will be replaced, for example because of financial constraints. But then over time also firms from regions which saw no fighting replace their old machines with the new equipment, so the technical efficiency scores should converge in the long-run (Modernization hypothesis).

During battles and campaigns factories will be destroyed which have to be rebuilt after the war. If the economy before the war produced under constant returns to scale, scale inefficiency after the war will be observed, as the firms are smaller than before (Scale hypothesis).[13] Atack (1985) estimated returns to scale in manufacturing and found that the industry generally worked in the region of constant returns to scale before the war, though some industries realized increasing or decreasing returns.[14] Overall, the assumption of constant returns seems reasonable. Destruction of firms as well as reduced labor supply by freed slaves (Foner 1988; Ransom and Sutch 1977) may have reduced the scale of operations and therefore lead to scale inefficiency in the South. This effect will vanish in the course of industrial development as firms are increasingly able to attract labor and capital and increase their size.

A war normally leads to an unequal destruction of production resources. In some border regions, the destruction of physical capital may be particularly large. The losses of life are more equally shared by all regions, as all regions send soldiers into battle, which can lead to a capital-scarcity in the regions that suffered much from fighting (Prichard 1939; Hutchinson and Margo 2006). Regions behind the front lines, on the contrary, suffer less loss in capital as no direct fighting takes place there. So in these regions one would expect a relative scarcity in labor compared to capital. These considerations lead to the expectation that after a war, inputs in production may not be used in their optimal relation (Disruption hypothesis). Also the emancipation of the slaves leads many former slaves to reduce their work effort (Ransom and Sutch 1977), which reduces labor supply. This works against the capital-scarcity in the war

[13] The assumption here is that small firms produce under increasing returns to scale, while they produce under constant returns to scale when they are of middle size and above a certain size class firms produce under decreasing returns to scale (see figures A1 and A2 in the appendix).
[14] See Tables 5.2-7.3 in Atack (1985).

torn regions, so that the sign of the scarcity here cannot be predicted. For the regions not affected by the battles, the two effects run in the same direction, so the labor-scarcity will be even more pronounced. This might lead to declining mix efficiency, as destroyed infrastructure impedes the reallocation of resources and production processes will take time to adjust.

In the case of the American Civil War one has to keep in mind that American industry recovered rapidly after the war (Lerner 1959). The industry in the South was already better off in 1870 than it had been ten years ago, while agriculture was below its pre-war level. This might limit the measurable effects. For the modernization hypothesis this means that we expect a better pure technical efficiency development in the war-torn regions than in the rest of the county. In the next section the hypotheses of a technological advantage and a disadvantage in scale and mix efficiency in the war regions will be tested.

3.5 The War and Efficiency: Empirical Results

We now turn to the results of the Data Envelopment Analysis. The data comes from the American manufacturing census from the 19th century and the years 1850, 1860, 1870 and 1880 are used here.[15] The samples include data on capital, labor-force, used inputs and produced outputs of between eight and fourteen thousand manufacturing firms. Unfortunately, the manuscripts from Georgia and Louisiana before 1880 are lost, which would be interesting for this study, as destruction there was quite strong. Between six and ten thousand establish-

[15] The data are available from the homepage of Jeremy Atack. See Atack and Bateman (1999) for an overview over the data set and Bateman and Weiss (1981) for additional information on the censuses. For 1865 some census data are available but none for Southern states, so those information are useless for the purpose of this study.

ments are used in the estimations, which come from 19 different industries.[16] At first we differentiate between war regions and non-war regions. The Confederacy and the states of the Union where fighting took place are classified as war regions, the other states as non-war regions. Roughly one third of our observations come from a war region.

The firms were asked which four inputs they used the most, what quantities of those inputs they used and what value each input had. The same question was asked regarding the outputs. As firms in the same industry often reported different inputs and outputs and the units of measurement of the inputs were quite different, we decided to use capital, labor and the sum of the values of the inputs in our estimation. The reason for this is that by using different inputs we would have to drop every firm that did not state this particular input. The total labor force is estimated as $L_T = L_{male} + 0{,}5\, L_{female} + 0{,}33\, L_{children} + 1$ (Golding and Sokoloff 1982).[17]

We follow the literature and add one to the stated number for labor to account for the owner-operator (Sokoloff 1984).[18]

As output, the value of all outputs is used. To correct for price changes we use price series gathered from the historical statistics of the United States and Somer (1965). The price series can be found in the appendix.[19] We estimate efficiency scores separately for every in-

[16] These are 072 Crop Services, 175 Carpentry and Floor Work, 204 Grain Mill Products, 205 Bakery Products, 208 Beverages, 212 Cigars, 231 Men's and Boys' Suits, Coats, and Overcoats, 242 Sawmills and Planing Mills, 244 Wood Containers, 251 Household Furniture, 311 Leather Tanning and Finishing, 314 Footwear, Except Rubber, 319 Leather Goods, not Elsewhere Classified, 325 Structural Clay Products, 335 Rolling, Drawing, and Extruding of Nonferrous, 349 Miscellaneous Fabricated Metal Products, 352 Farm and Garden Machinery and Equipment, 379 Miscellaneous Transportation Equipment, 769 Blacksmithing.

[17] Using the formula $L_T = L_{male} + L_{female} + 0{,}5 L_{children} + 1$ does not change the results.

[18] It is difficult to judge where the correction is necessary, as all regions in all time periods have establishments with 0 workers. We therefore follow the practice of the literature in applying the correction for every firm.

[19] Most of our results are fairly robust to the price series used. Where large differences occur through the chosen price series, we will note that in the text.

dustry for every census-year and report mean differences in median efficiency for each region and decade below.

It would be good to correct the capital and labor data to account for the effect of part-year operation (Inwood and Keay 2006), as firms operating only few months per year will seem to be less efficient than they really were compared to full-year operating establishments. Although the share of part-year operating establishments in 1880 was substantial (around 40% according to Atack, Bateman and Margo 2002), this is not possible, as this information was only reported for 1870 and 1880. The comparison between 1860 and 1870 would not be robust if we corrected for part-year-operations only in 1870.[20]

As was stated in the literature (Bartels and Zhang 1998) the median efficiency of a sample of firms decreases, when the number of firms increases. In our data set the number of firms in some industries varies substantially over time and we correct for this bias by adjusting the median efficiency using ratios estimated from simulations done by Bartels and Zhang.[21]

Even in normal years median efficiency will change over time and evolve differently in different geographical regions, as the development of median efficiency in the decades between 1850 and 1880 shows. If the numbers in the Civil War decade differ significantly from the other decades, this shows an effect of the Civil War.

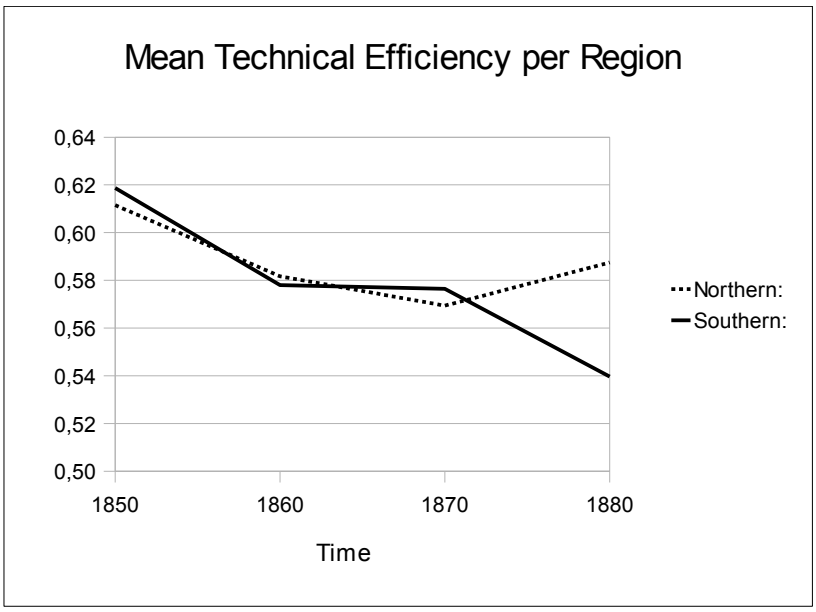
We exclude outliers in the calculations using the following method. All firms whose efficiency scores exceeded 200% in a super-efficiency model are deleted. On average three

[20] Moreover, comparison between a sample corrected for part-year operation and without the correction revealed only minor differences in median and mean efficiency. It is not useful to look for industries which have a high fraction of part-year operation in 1870 or 1880 and correct all firms in those industries as if they were only operational for 6 months per year. If we would correct all firms for the same number of months of operations, the efficiency scores would not change.

[21] See tables 2 and 3 in Bartels and Zhang (1998).

percent of the firms are classified as outliers, only 5 out of 76 cases have more than 10 percent of the firms removed and the maximum is 17 percent. This happens in industries where only few observations are available.[22] At first we present mean efficiency values for Northern and Southern states from 1850 to 1880.

Figure 1: Development of mean technical efficiency.



Source: Own estimates based on the Atack/Bateman samples.[23]

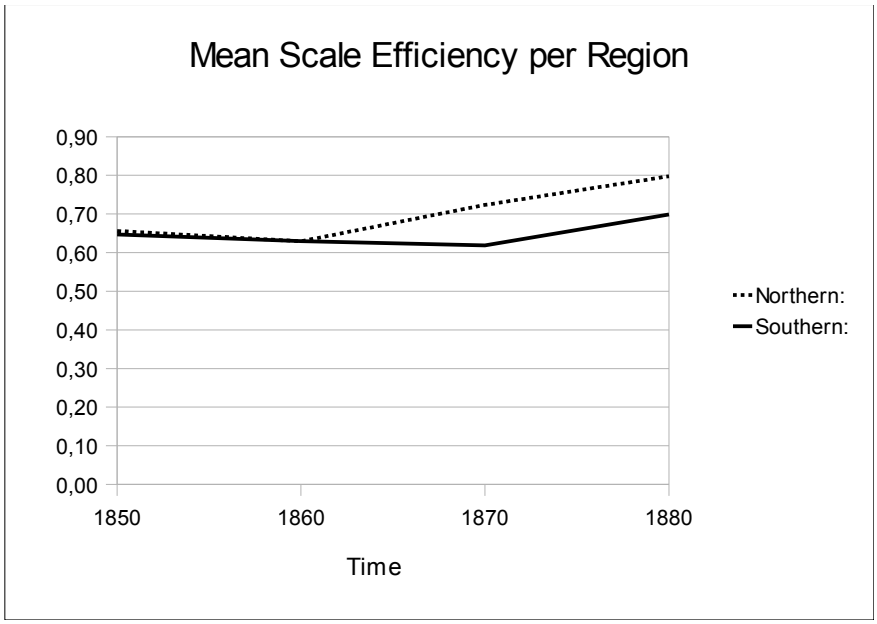
Before the war technical efficiency in both regions was similar (figure 1) and the differences in 1850 and in 1860 are neither statistically significant. In 1870 the South gained a technical advantage, which is still not significant (p-value 0.14). In the war region capital intensity fell from 591 in 1860 to 476 in 1870, while it increased in the non-war region from

[22] For 2 industries and years we did not exclude outliers, as here only three and five observations were available at all.

[23] These are unweighted means. For sensitivity testing we tried median efficiency and weighted means (labor, capital or value of outputs were used as weights). The general conclusion of a better Southern development in the 1860 and a loss afterwards is robust.

701 in 1860 to 725 in 1870. This means that production in the North became more machine-intensive, which should lead to higher efficiency. So the results indicate that firms in the battleground counties used the war to change production processes and were therefore even able to compensate the higher capital intensity. We have to keep in mind here that, although industry after the war was severely damaged and by 1870 many firms were still in ruins, human capital was available directly after the war, as veterans returned home (Ash 1988). In industries where small-scale construction prevailed and labor was more important than machines, this leads to smaller efficiency losses than the war-time destruction might suggest. This advantage reverses, and by 1880 the North has a large and highly significant advantage over the South.

Figure 2: Development of mean scale efficiency.



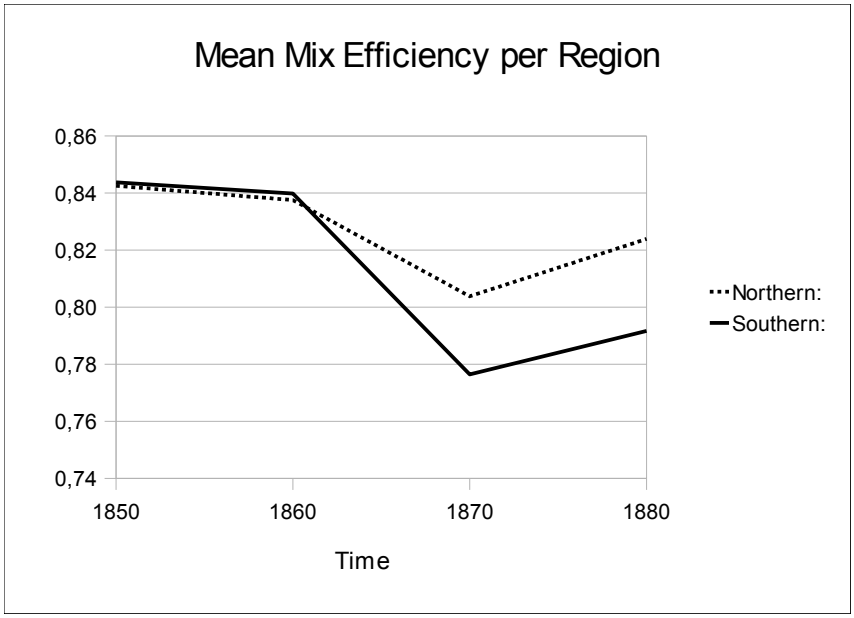
Source: Own estimates based on the Atack/Bateman samples[24].

[24] Sensitivity testing confirmed here the loss of the war decade, though the models differ in whether the South can catch up in the 1880s or not.

The differences in scale efficiency (see figure 2) are larger than the technical efficiency differences. Before the war median scale efficiency was similar in both regions (the difference is under one percentage point and not statistically significant). During the war a gap arises and in 1870 the Northern firms have an advantage of 10 percentage points, which is highly significant.

The size of this gap remains stable over the next decade. The war had a negative scale effect, according to the Scale hypothesis. This finding is also in line with Hutchinson and Margo (2006), who show that the decrease in labor productivity in the South cannot be fully attributed to a decrease in capital per worker; the residual is due to the loss of scale efficiencies, as their specification implies constant returns to scale.

Figure 3: Development of mean mix efficiency.



Source: Own estimates based on the Atack/Bateman samples[25].

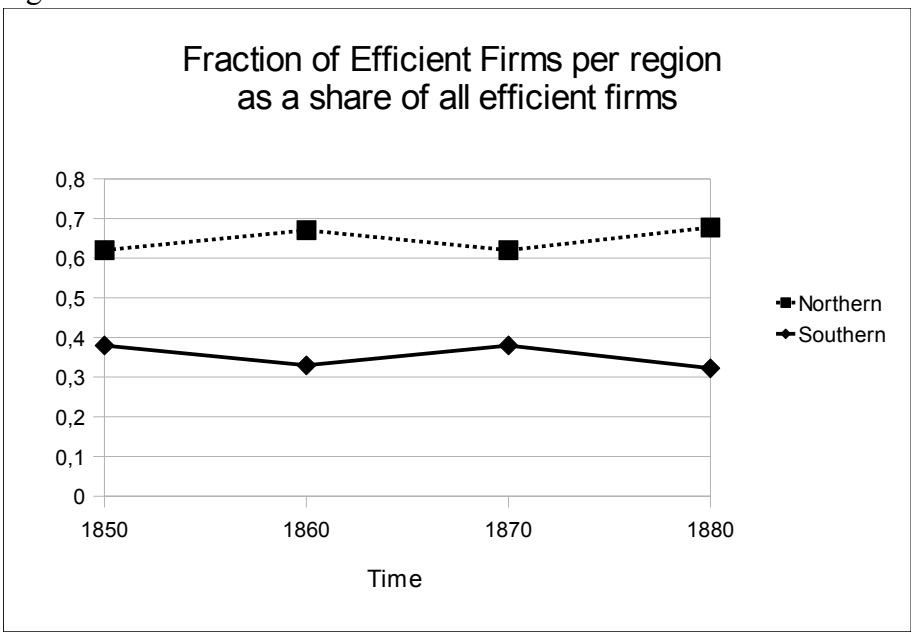
[25] The efficiency loss of the South in the war decade is common in all models, though the magnitude of the effect differs. The development of the 1870s is also sensitive to the model used.

With regard to the Disruption hypothesis the mix-efficiency difference in the war decade is very small, much smaller than the differences before and after that decade, so there does not seem to be a strong effect from the Civil War (figure 3). This finding is in line with the results obtained by Niemi (1974) who argues that capital and labor were unimportant in the decision on industrial location, because they were quite mobile. This mobility limits the effect of unequal destruction of inputs.

As the threshold for excluding outliers is somewhat arbitrarily chosen with 200% efficiency, we estimated a model without these outliers and one with all firms in the sample. Comparing these models reveals only minor differences. The overall picture and the conclusions remain valid, though the numbers are slightly different.

To sum up, the results suggest that the Civil War influenced efficiency, giving the war regions a slight advantage in technical efficiency, while they lost in scale and mix efficiency. As we argued above the sign for technical efficiency changes differs between the 1860s and the 1870s. If the firms in the war regions were first to replace their destroyed machinery during the war decade and the firms in non-battleground counties were to follow in the next decade, the figures show the results we expect. The firms in the war regions were the first to get technically more efficient and are the firms on the frontier which lowers scores for the non-war regions. Then, these firms make widespread use of the technology and push the frontier outward, which decreases efficiency for the war regions. Figure 4 below shows the fraction of technically efficient firms that came from Northern and Southern states.

Figure 4:



Source: own calculations based on the Atack/Bateman samples.

The Northern states increased their share of efficient firms during the 1860s, while the share declined during the war decade and rose again after 1870. The Southern states had more efficient firms after the war as predicted by the modernization hypothesis.

Comparable studies for other wars to compare the size of the effects are lacking. But differences from only a few percentage points (where efficiency scores are in the range of 0.5 to 0.8) seem small.

We have to keep in mind that the results may be underestimated for two reasons. First, our data are lacking firms from Georgia, where General Sherman's march to the sea occurred, which was one of the most destructive campaigns in the war, and Louisiana. Second, during the five years after the Civil War (we have data only every ten years), reconstruction was pronounced, so the effects from the war will be underestimated. This might be especially true as

many new firms were created, as the number of firms in the South increased sharply during the Civil War period (Wright 1982).

During the Civil War not only the destruction of stocks and factories changed the industry, but also the emancipation of the slaves had an important impact on manufacturing. The freedmen lowered their labor supply and demanded wages, which changed optimal input mixes in the industry. This effect is surely not negligible, so we try to separate the effect of physical destruction from that of emancipation. To accomplish this we divide the United States into three regions. First we define battleground counties, which are likely to have experienced large physical destruction.[26] We then classify all states that allowed slavery at the beginning of the Civil War as slave states.[27] Around one third of our observations come from a state in which slavery was allowed, and 16 to 25 percent of all firms come from a battleground county.

We employ a regression approach to separate the effects of emancipation and physical destruction. For doing this, we regress the efficiency scores of our 30000 firms on industry-dummies, dummies for the time the census was taken, a dummy for firms located in cities (more than 2500 inhabitants), capital and labor.[28] We then use interaction variables between the region- and time dummies to assess the effects. As nearly all of the battleground counties are also regions in which slavery was legal the battleground-variables give direct evidence on the effects of physical destruction. In the second stage after the DEA, OLS, tobit

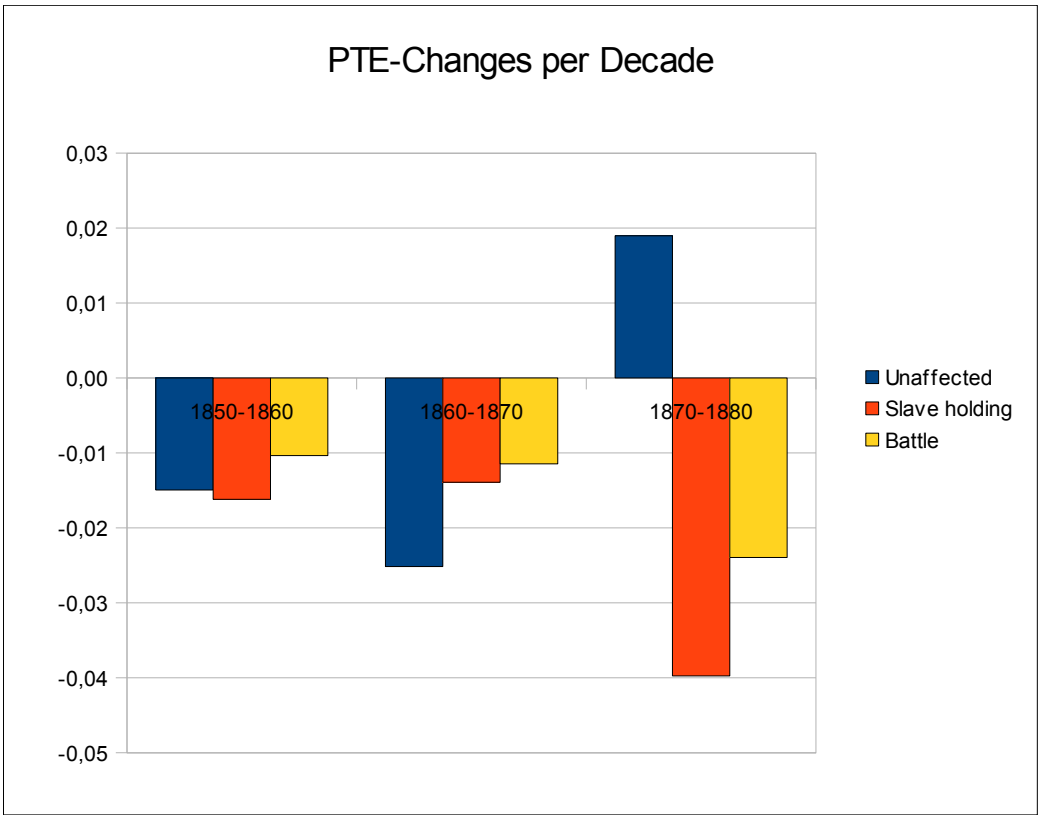
[26] These counties include all those counties where a battle is recorded, the counties that Sherman passed during his campaign in the Carolinas, the counties from the Shenandoah Valley, and all counties from Arkansas, Kentucky, Louisiana, Mississippi, Missouri, Tennessee and Virginia, which are classified as major battlegrounds.
[27] These are: Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Texas, Virginia, Missouri, Delaware.
[28] The efficiency scores are not aggregated for the regression. Only the information on region (battleground, slaveregion or else) is aggregated by county.

and truncated regressions are commonly employed in the literature (see McDonald 2008 for a discussion). We follow Simar and Wilson (2007) in using truncated regressions, though we do not need bootstrapping, as we have sufficient observations.[29] We also control for the number of observations in the industry, as more observations decrease the individual efficiency scores, as explained above. Outliers were excluded from the regressions and the standard errors are corrected for clustering at the state level.[30] The regression results can be found in the appendix. As the interaction variables are difficult to interpret, we show here only the efficiency changes over time in the three regions[31].

As can be seen in figure 5, the 1850s show some change in technical efficiency, but the development is quite uniform in all regions. The development in unaffected counties was negative in the Civil War decade (2.5 percentage points). It was somewhat less negative in the slave counties (1.5 percentage points) and similar in the battlefield counties (1 percentage points). Note that it is not important here that efficiency declined in the unaffected counties, but only that the development in the battlefield counties was better than in the unaffected counties. This gives evidence in favor of our modernization hypothesis, although the differences are generally small. The largest change in efficiency values is about 2 percentage points, and only 8% of all firms are fewer than 2 percentage points below or above the median. In the decade following the war, the unaffected counties experience rising technical effi

[29] See Figure A3 in the appendix for the distribution of efficiency scores. Note that OLS regressions yield similar results.
[30] In some industries, which were enumerated by so called special agents, under-enumeration of some establishments occurred (Atack, Bateman and Margo 2002). It is not possible to correct our data for under-enumeration in these industries, however, from our sample only industry 208 is affected and exclusion of this industry does not change the results.
[31] The literature also argues that the standard errors may be biased (see Hahn 2005), so we use them with care and focus on the effects shown in the figure.

Figure 5:



Source: Table A5.

ciency values while the slave counties and the battlefield counties experienced decreasing efficiency. This is in line with the finding that more efficient firms came from battleground counties after the war (figure 4). Some firms were apparently quickly able to receive enough Northern capital to acquire new machinery (Somer 1965), while others could not achieve this. An example for a successful firm is the Tredegar Iron Works, which were able to sell property and receive some capital from Northern investors early (Dew 1966). Harold Wilson (2002) lists several other entrepreneurs who bought new machinery directly after the war; some of the machinery even came from England. Harold James reports that access to the national capital market was possible for Southern investors, and this is especially true for large firms

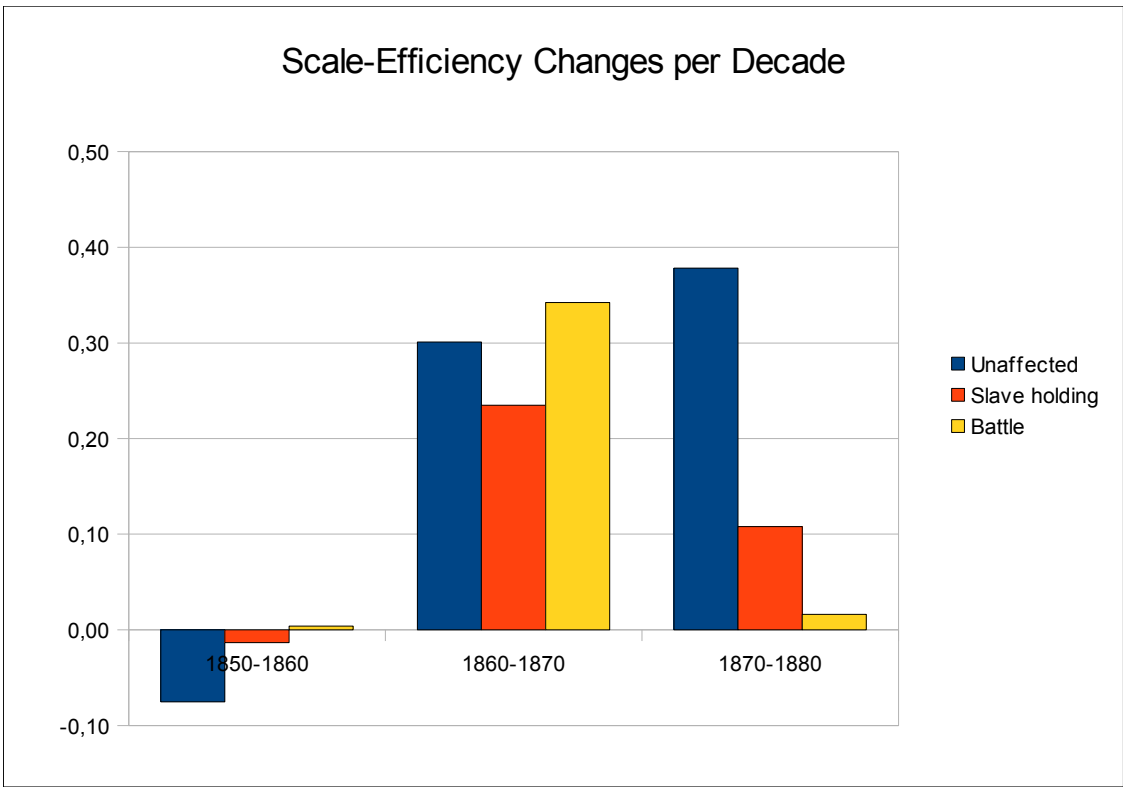
which promised great profits (James 1981). But generally the Southern financial system after the war is considered inadequate (Davis 1963). One other possible aspect needs to be mentioned here: it is possible that the shocks, like lost access to the Northern market, and destruction eliminated only inefficient firms, as only successful firms were rebuilt afterwards. This is clearly the case, but this is also an effect of war, hence it should be included here. All in all we find a small advantage for battleground firms compared to their competitors. In this aspect the South seems to be different to today's countries that experience a civil war, as for them capital flight and therefore a reduction in investment is a major obstacle to recovery (Imai and Weinstein 2000).

Another area in which the war had positive impacts is the restructuring of social life. Acemoglu et al. (2009) argued that the French conquests after the French revolution had a positive economic impact on the conquered German parts as the power of the old elites, who are generally opposed to economic change, were destroyed. This allows new industries to emerge. New legal and institutional systems which are better in line with the needs of industry may be introduced more easily than in peaceful times. In the US, former slaves after the war had more freedom to act independently and started businesses which was not possible before.

The results for scale efficiency (figure 6) show that physical destruction and emancipation had a rather strong negative impact on scale efficiency. Before the war the unaffected counties are doing worst. During the war scale efficiency increased in all regions, the battlefield counties gaining most, second are the unaffected counties while the slave counties do worst. In the 1870s, the Northern counties gain most, followed by the slave- and battleground

counties. The estimated effects here are much stronger than the ones for technical and mix efficiency.

Figure 6:



Source: Table A6.

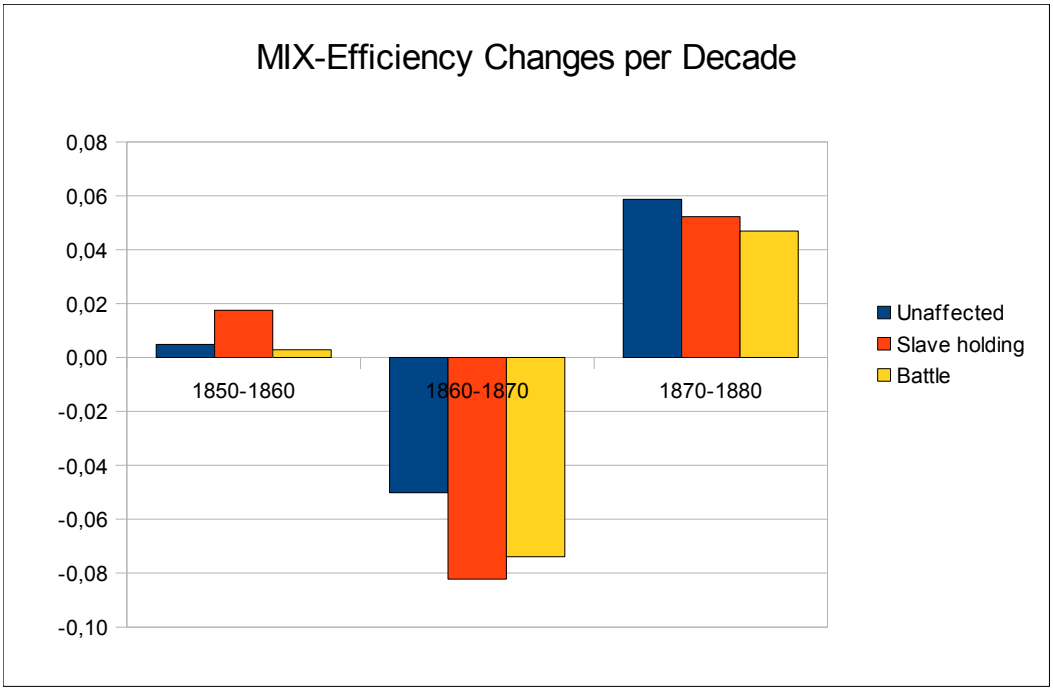
That the battlefield counties are doing best in the war decade is surprising, as the scale hypothesis predicts the opposite. One reason might be that the factories that were soon rebuilt were only those that were highly efficient before the war and were therefore very promising. Other more inefficient producers were eliminated by the war and were not able to gain enough capital soon after the war to start again. Therefore only few large and highly efficient firms were already in existence in 1870. The weak performance of the slave counties is surely con-

nected to emancipation. Labor supply was reduced as the freedmen preferred to start own farms or moved to the North and therefore firms had problems to gain enough workers. Former slaves who tried to start new enterprises had in 1870 firms that were too small to produce in the efficient scale size. Next, all Southern industry had lost markets in the North through the secession. The shrunken market made it difficult for new firms to reach an efficient size and the large landowners, who were an important source of demand before the war were now often impoverished, which reduced demand and with it firm size even further.

The results for scale efficiency are somewhat sensitive to the price series used. If we do not use the price deflation, the results are more in line with the scale hypothesis outlined above (see figure A4 in the appendix). We do not know the exact price development in the South, but the price series used here seems to be more on the upper limit of plausible values.

Regarding mix efficiency, which is depicted in figure 7, the 1850s show nearly no movement, but the slaveholding regions have a small advantage. In the war decade we see the expected pattern of a stronger loss in the battleground and slave holding regions, the latter doing a bit worse than the battleground regions. During the 1870s the unaffected counties are even able to push farther ahead, while the slaveholding states can regain some of their lost ground compared to the battleground counties. The war had a definite impact on mix efficiency, and emancipation seems to be very important, as it suddenly changed labor supply.

Figure 7:



Source: Table A7.

Clearly, market integration also plays a role in the results demonstrated above. Antebellum capital markets were fairly integrated (Bodenhorn 1992; La Croix and Grandy 1993), and product markets before the war show marked signs of integration (Slaughter 1995). After the Civil War Southern and Northern financial markets were less integrated than before the war (Davis 1965) and many firms lost former markets. At the same time mechanization was strongly on its way, but still many establishments were producing with low levels of mechanization. Human capital, in our context skills and knowledge about production processes, was surely more important in many industries than physical capital. Through the war the size of the markets was severely restrained, while many soldiers returned to their old businesses. As human capital was therefore available, technical efficiency is not strongly affected as the im-

portant human capital was still in existence, while the disintegration of markets played a role in the scale inefficiency.[32]

There are some possible biases in our analysis, as it is possible that the census data contain errors. For example, false numbers may have been written in the data sets or factory owners may have been unwilling to give detailed information, as this took time for them to calculate the correct data. This source of errors cannot be eliminated, but we want to point out that during the collection of the data efforts were taken to eliminate them (Atack and Bateman 1999). American Census data from the 19th century are among the most reliable for this period (Robison and Briggs 1991).

3.6 Conclusion

Looking directly at the impact of the Civil War on efficiency we offer a new way of studying the effects of wars. This impact is decomposed into changes in technical-, scale and mix efficiency.

Our results suggest that the war increased technical efficiency and also scale efficiency in the affected counties, while it decreased mix efficiency. This shows that the Civil War was used to rearrange production and that the influx of new machines compensated destruction (see Mill 1884). The differences are, except for scale efficiency, of medium size, but our findings could be downward-biased, as Georgia and Louisiana could not be included in the

[32] Please remember that the results for scale efficiency are sensitive to the price series used. As no reliable price series for the South from antebellum to postbellum times is available we cannot say for sure what exactly happened, but the results suggest that at least the loss of economies of scale was not very important for the Southern industry after the Civil War.

samples. The influences from the Civil War were largely eliminated during the 1870s. This finding and the result that the effects of the war were not large might be surprising given the heavy destruction and the large loss of lives, but we have to keep in mind two things. First, it was shown that destruction was largely concentrated in military industries, while the civilian industries were not affected so heavily. Secondly, reconstruction in the South was quite fast, as the government invested heavily to restore infrastructure, especially in railroads (Foner 1988).

Separating the effects of the abolition of slavery and the war our results suggest that the impact of the abolition of slavery was more important for efficiency than that of destruction.

The results strengthen the hypothesis that the Civil War was not crucial for the path of the industrialization in the second half of the nineteenth century. To achieve such an impact, larger and more persistent effects seem necessary.

References

Acemoglu, D., Cantoni, D., Johnson, S. and J. A. Robinson (2009). The Consequences of Radical Reform: The French Revolution. NBER Working Paper 14831.

Arbetman, M. and J. Kugler (1990). Exploring the "Phoenix Factor" with the Collective Goods Perspective. *The Journal of Conflict Resolution*, 33, 84-112.

Ash, S. (1988). *Middle Tennessee society transformed: 1860 – 1870; war and peace in the upper South. Baton Rouge: Louisiana State University Press.*

Atack, J. (1985). *The estimation of economies of scale in nineteenth century United States manufacturing*. New York: Garland.

Atack, J., Bateman, F. (1999). Nineteenth-Century U.S. Industrial Development through the Eyes of the Census of Manufacturers. A New Resource for Historical Research. Historical Methods 32 (4), 177-188.

Atack, J., Bateman, F., Margo, R. A. (2002). Part-Year Operation in Nineteenth-Century American Manufacturing: Evidence from the 1870 and 1880 Censuses. The Journal of Economic History 62 (3), 792-809.

Bartels, R., Zhang, Y. (1998). The Effect of Sample Size on the Mean Efficiency in DEA with an Application to Electricity Distribution in Australia, Sweden and New Zealand. Journal of Productivity Analysis 9 (3), 187–204.

Bateman, F., Weiss, T. (1981). A deplorable scarcity. The failure of industrialization in the slave economy. Chapel Hill: The University of North Carolina Press.

Beard, C. and M. Beard (1927). The Rise of American Civilization. New York: Macmillan.

Bodenhorn, H. (1992). Capital Mobility and Financial Integration in Antebellum America. The Journal of Economic History 52 (3), 585-610.

Carter, S., Gartner, S. S., Haines, M. R., Olmstead, A. L., Sutch, R. and G. Wright (2006). Historical Statistics of the United States: Earliest Times to the Present. Millennial Edition. New York: Cambridge University Press.

Clark, V. (1962). Manufacturing Development during the Civil War. In Andreano R (Ed.) The Economic Impact of the American Civil War. New York: Schenkman Publishing Co.

Cochran, T. (1962). Did the Civil War retard industrialization? The Mississippi Valley Historical Review 48, 197-210.

Cooper, W. W., Seiford, L. M. and K. Tone (2006). Introduction to data envelopment analysis and its uses. With DEA-solver software and references. New York: Springer.

Craig, L. and E. Field-Hendrey (1996). Were free Southern farmers "driven to indolence" by slavery? A stochastic production frontier approach. NBER working paper. Historical Paper 82.

Davis, L. (1963). Capital immobilities and financial capitalism: a study of economic evolution in the United States 1820-1920. Explorations in Entrepreneurial History 2 (1), 88-105.

Davis, L. (1965). The Investment Market, 1870-1914: The Evolution of a National Market. The Journal of Economic History 25, 355-99.

Diacon, S. R., Starkey, K. and C. O'Brien (2002). Size and Efficiency in European Long-term Insurance Companies: An International Comparison. The Geneva Papers on Risk and Insurance 27 (3), 444-466.

Dew, C. B. (1966). Ironmaker of the Confederacy. Joseph R. Anderson and the Tredegar Iron Works. New Haven and London: Yale University Press.

Engelbourg, S. (1979). The Economic Impact of the Civil War on Manufacturing. Business History 21 (2), 148-162. DOI: 10.1080/00076797900000024.

Engerman, S. (1966). The Economic Impact of the Civil War. Explorations in Economic History 3, 176-199.

Farrell, M. J. (1957). The Measurement of Productive Efficiency. Journal of the Royal Statistical Society, Series A, 120, 253-290.

Field-Hendrey, E. (1995). Application of a stochastic production frontier to slave agriculture: an extension. Applied Economics 27 (4), 363-368. DOI: 10.1080/00036849500000120.

Foner, E. (1988). Reconstruction: America's unfinished revolution, 1863-1877. New York: Harper & Row.

Gallman, M. (1990). Mastering wartime. A social history of Philadelphia during the Civil War. Cambridge: Cambridge University Press.

Goldin, C. and K. L. Sokoloff (1982). Women, Children and Industrialization in the early Republic: Evidence from the Manufacturing Censuses. The Journal of Economic History 42 (4), 741-744.

Grabowski, R. and C. Pasurka (1988). The relative technical efficiency of Northern and Southern U. S. farms in 1860. Southern Economic Journal 54, 598 – 614.

Grabowski, R. and C. Pasurka (1989). The relative efficiency of slave agriculture: An application of a stochastic production frontier. Applied Economics 21 (5), 587 – 595. DOI: 10.1080/758524892.

Hacker, L. (1940). The Triumph of American Capitalism. New York: Columbia University Press.

Hahn, F. (2005). Environmental Determinants of Banking Efficiency in Australia. Unpublished Manuscript, Österreichisches Institut für Wirtschaftsforschung.

Hughes, M. J. (1994). Sherman's 1864-65 campaign: strategic assessment and lessons for today. Naval War College, Newport.

Hutchinson, W. K. and R. A. Margo (2006). The impact of the Civil War on capital intensity and labor productivity in Southern manufacturing. Explorations in Economic History 43 (4), 689-704. doi:10.1016/j.eeh.2005.08.001.

Imai, K. and J. Weinstein (2000). Measuring the Economic Impact of Civil War. CID Working Paper No 51.

Inwood, K. and I. Keay (2006). Assessing economic performance among North American Manufacturing Establishments, 1870/71: Data, Methodology and Measurement Issues. Unpublished Manuscript, Economics Department, Queen's University.

James, J. (1981). Financial Underdevelopment in the Postbellum South. Journal of Interdisciplinary History 11 (3), 443-454.

La Croix, S. and C. Grandy (1993). Financial Integration in Antebellum America: Strengthening Bodenhorn's Results. The Journal of Economic History 53 (3), 653-658.

Lerner, E. (1959). Southern Output and Agricultural Income, 1860-1880. Agricultural History 33 (3), 117-125.

McDonald, J. (2008). Using Least Squares and Tobit in Second Stage DEA Efficiency Analyses. Unpublished Manuscript, Flinders University.

Mill, J. S. (1884). Principles of political economy. New York: Appleton and Company.

Niemi A. (1974). State and regional patterns in American manufacturing. Westport: Greenwood Press.

O'Brian, P. K. (1988). The Economic Effects of the American Civil War. Hampshire and London: Macmillan Education Ltd.

Olson, M. (1982). The Rise and the Decline of Nations. New Haven: Yale University Press.

Pessen, E. (1980). How different from each other were the antebellum North and South? The American Historical Review 85 (5), 1119-1149.

Prichard, R. (1939). The Effects of the Civil War on the Louisiana Sugar Industry. Journal of Southern History 5 (3), 315-332.

Ransom, R. L. (1998). The economic consequences of the American Civil War. In Wolfson M (Ed.) The Political Economy of War and Peace. Boston/Dordrecht/London: Kluwer Academic Publishers, 49-74.

Ransom, R. L. and R. Sutch (1977). One kind of freedom: the economic consequences of emancipation. Cambridge: Cambridge University Press.

Robinson, R. V. and C. M. Briggs (1991). The Rise of Factories in Nineteenth-Century Indianapolis. The American Journal of Sociology 97 (3), 622-656.

Russel, W. (2001). The economic impact of the American Civil War: An empirical examination (Doctoral dissertation). Available from Proquest database.

Salsbury, S. (1962). The effect of the Civil War on American industrial development. In Andreano R (Ed.) The Economic Impact of the American Civil War. New York: Schenkman Publishing Co.

Schmidt, L. B. (1930). The Agricultural Revolution in the United States, 1860-1930. Science 72, 585-594.

Simar, L. and P. W. Wilson (2007). Estimation and Inference in two-stage, semi-parametric models of production processes. Journal of Econometrics 136 (1), 31-64.

Slaughter, M. (1995): The Antebellum Transportation Revolution and Factor-Price Convergence. NBER Working Paper 5303.

Sokoloff, K. (1984). Was the Transition from the Artisanal Shop to the Nonmechanized Factory Associated with Gains in Efficiency? Evidence from U.S. Manufacturing Censuses of 1820 and 1850. Explorations in Economic History 21 (4), 351-382.

Somer, R. (1965). The Southern States since the war 1870-71. Alabama: University of Alabama Press.

Taveres, G. (2003). A Bibliography of Data Envelopment Analysis (1978-2001). Rutcor, Pascatuwy.

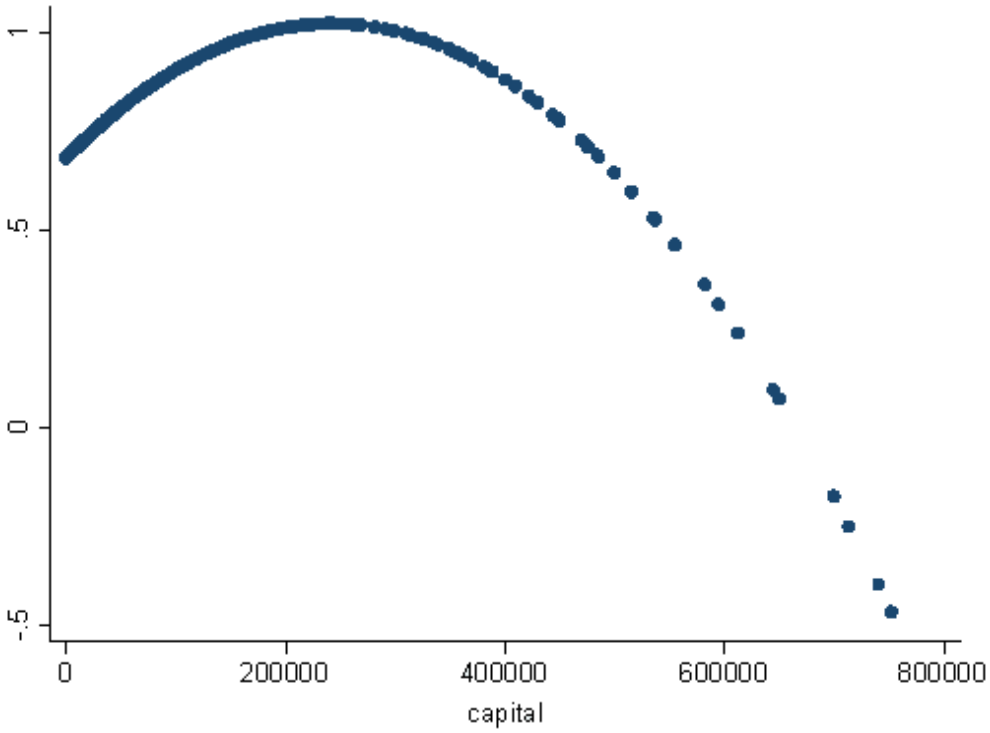
Wilson, H. S. (2002). Confederate industry: manufactures and quartermasters in the Civil War. Jackson Mississippi: University Press of Mississippi.

Wright, C. W. (1943). The more enduring economic consequences of America's Wars. The Journal of Economic History, 3 (Supplement: The Tasks of Economic History), 9-26.

Wright, G. (1982*)*. The strange career of the new Southern Economic History. Reviews in American History 10 (4), 164-180.

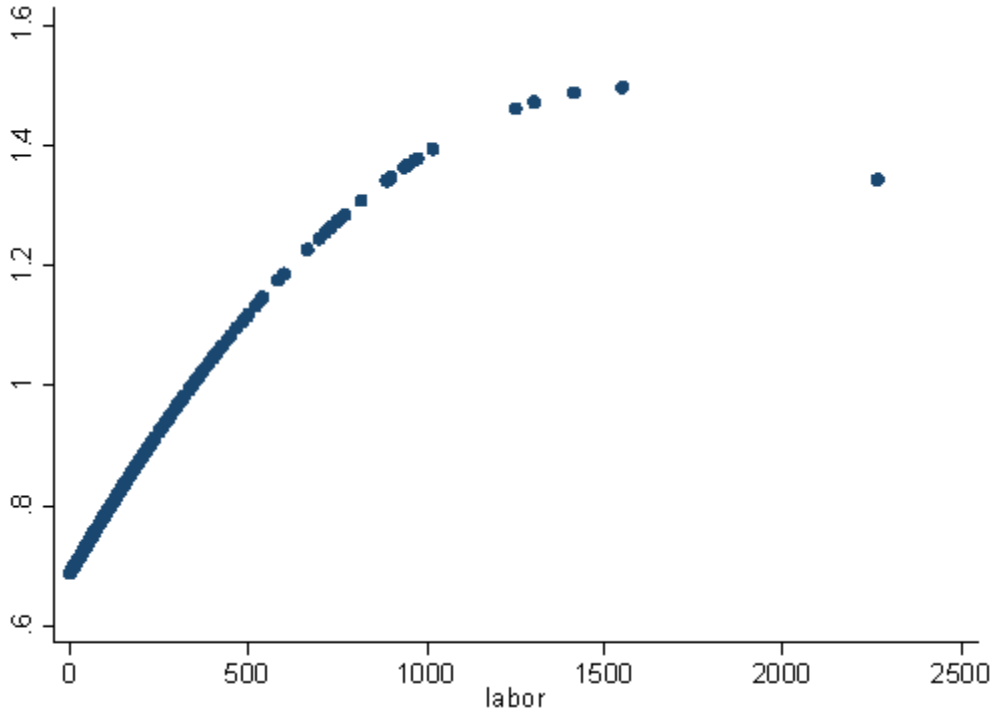
Appendix 1: Figures

Figure A1: Distribution of scale efficiency values.



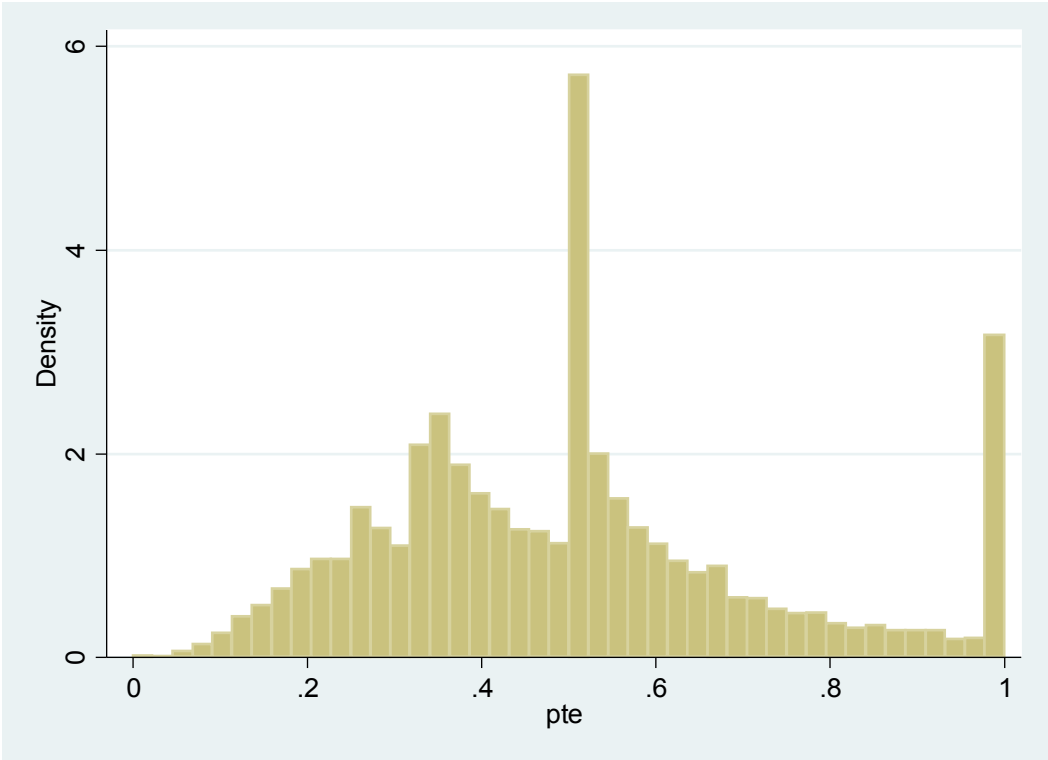
Notes: The Values are the predicted values of a regression of the scale efficiency scores on capital and its squared term.

Figure A2: Distribution of scale efficiency values.



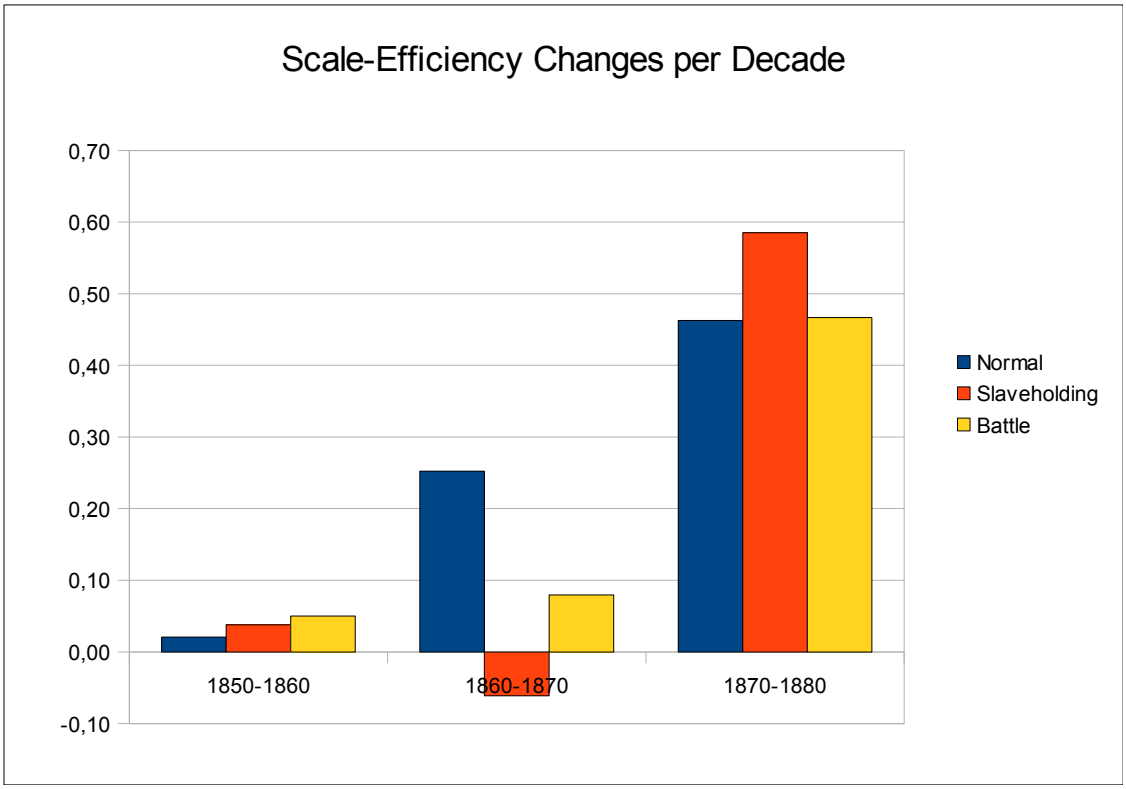
Notes: The Values are the predicted values of a regression of the scale efficiency scores on labor and its squared term.

Figure A3: Histogram of technical efficiency scores.



Source: Own calculations.

Figure A4: Scale efficiency changes without price deflation



Source: own calculations.

Appendix 2: Tables

Table A1: Price index series.

	South	Pacific	Rest
1850	110	236	84
1860	112	113	93
1870	300	118	135
1880	300	103	100

Notes: Prices come from Carter et. al (2006) and Somer (1965, p. 71).

Table A2: Regression of technical-efficiency scores.

Slave1850	0.0020321 (0.759)
Slave1860	0.0006326 (0.920)
Slave1870	0.0102025** (0.050)
Slave1880	-0.011334*** (0.008)
Battlecounty1850	-0.0099429 (0.202)
Battlecounty1860	-0.0156144** (0.033)
Battlecounty1870	-0.0014986 (0.790)
Battlecounty1880	-0.00274 (0.631)
Year1850	0.008341** (0.037)
Year1870	-0.0272974*** (0.000)
Year1880	-0.0146525*** (0.000)
City	0.0207985*** (0.000)
Number of observations	-0.0000339*** (0.000)
Time and Industry dummies	Included
N	27982

Note: p-values in parentheses. */**/*** means significant on the 10/5/1 percent level. The reference category is a crop service firm from 1860 located in a non-battleground- and non-slave county.

Table A3: Regression of scale-efficiency scores.

Slave1850	0.0538117** (0.036)
Slave1860	0.0706471*** (0.004)
Slave1870	-0.2429057*** (0.000)
Slave1880	-0.1206526*** (0.000)
Battlecounty1850	-0.0208553 (0.480)
Battlecounty1860	-0.0083224 (0.773)
Battlecounty1870	0.1323577*** (0.000)
Battlecounty1880	0.0140704 (0.652)
Year1850	-0.0208142 (0.197)
Year1870	0.2522698*** (0.000)
Year1880	0.7151821*** (0.000)
City	0.3536622*** (0.0000)
Number of observations	-0.0003371*** (0.000)
Time and Industry dummies	Included
N	29983

Note: p-values in parentheses. */**/*** means significant on the 10/5/1 percent level. The reference category is a crop industry firm from 1860 located in a non-battleground- and non-slave county.

Table A4: Regression of mix-efficiency scores.

Slave1850	-0.0170269*** (0.001)
Slave1860	0.0082508** (0.050)
Slave1870	0.0048173 (0.158)
Slave1880	0.0164297*** (0.000)
Battlecounty1850	0.0100957* (0.080)
Battlecounty1860	0.0010811 (0.824)
Battlecounty1870	-0.0023293 (0.532)
Battlecounty1880	-0.0000955 (0.978)
Year1850	0.0286753*** (0.000)
Year1870	-0.0388989*** (0.000)
Year1880	0.0087678*** (0.001)
City	0.0001174 (0.948)
Number of observations	-0.0000602*** (0.000)
Time and Industry dummies	Included
N	30028

Note: p-values in parentheses. */**/*** means significant on the 10/5/1 percent level. The reference category is a crop industry firm from 1860 located in a non-battleground- and non-slave county.

Table A5: Development of technical efficiency scores over time.

	Normal county	Slave county	Battlefield county
1850	0,0149	0.0178	-0.0005
1860	0.0000	0.0016	-0.0108
1870	-0.0251	-0.0123	-0.0223
1880	-0.0062	-0.0521	-0.0462

Note: The numbers are calculated from the results of table A1. Reference category is a crop service firm that was neither in a battleground county nor in a slave county in 1860.

Table A6: Development of scale efficiency scores over time.

	Normal county	Slave county	Battlefield county
1850	0.0754	0.0433	0.0127
1860	0.0000	0.0299	0.0167
1870	0,3011	0.2647	0.3590
1880	0.6793	0.3726	0.3750

Note: The numbers are calculated from the results of table A2. Reference category is a crop service firm that was neither in a battleground county nor in a slave county in 1860.

Table A7: Development of mix efficiency scores over time.

	Normal county	Slave county	Battlefield county
1850	-0.0048	-0.0068	0.0022
1860	0.0000	0.0107	0.0050
1870	-0.0502	-0.0715	-0.0689
1880	0.0085	-0.0193	-0.0220

Note: The numbers are calculated from the results of table A3. Reference category is a crop service firm that was neither in a battleground county nor in a slave county in 1860.

4 Do Numeracy and Health determine Labor Productivity in Tsarist Russia?

Abstract

Using data gathered around the industrial exhibition of 1870 in St. Petersburg, this study assesses the determinants of labor productivity in nineteenth century Russia and finds that, amongst others, human capital and health are important determinants of labor productivity. The analyses also reveal important differences between industries. We show that political influences, specifically the different policies between the inner Russian areas and the Russian parts of Poland, were important for labor productivity. The tax-system, which in Poland discriminated against landowners, made manufacturing an attractive investment alternative and this increased productivity compared to other Russian parts, where taxes favored agriculture.

This Chapter is based on a working paper with the same title that I wrote together with Jörg Baten. The idea was developed jointly. We both wrote equal shares of the paper.

4.1 Introduction

The question what determines the productivity of workers has a long tradition in the economic literature (See Bartelsman and Doms, 2000, for a review of a part of that literature). One robust result is that education is an important determinant of productivity (Black and Lynch 1996), as is capital intensity (Haltiwanger, Lane and Spetzler, 1999). Moreover, the literature links R&D to productivity, and it showed that firms engaging in R&D are more productive than others (Crépon, Duguet and Mairesse, 1998). Many of the studies have been performed on an aggregated level (See for example Englander and Gurney, 1994, and Brahmananda, 1982), and only few have been done on historical periods (for example Baten, 2001). The question remains whether the determinants of productivity change over time, so it is important to analyze historical samples.

Our contribution to the literature is that we analyze a sample of Russian firms from around 1870, a time when Russia was in its early phase of industrialization. The historical literature on productivity is mainly concerned with productivity development in whole industries or countries, while we use individual firm-level data for our analysis. We will assess the determinants of labor productivity at that time and also give quantitative evidence on the state of competition in various industries (Galenson 1955, p. 8). Another point is that we analyze how different political decisions in Poland and other Russian Provinces influenced productivity (Scherner, 2001).

Our main focus is on the impact of nutrition and education on productivity. As the Tsarist Empire was in an early phase of industrialization, it was still marked by a large agricultural sector and strong landlords who tried to keep their power. We will analyze whether nutritional conditions in Tsarist Russia had an impact on labor productivity. The theory of efficiency wages says that higher income and hence better food makes manual workers stronger and skilled workers more concentrated. This will increase their productivity (Bouis and Haddad, 1991). To approximate education, we use the concept of numeracy, which is explained in detail below (see A'Hearn, Baten and Crayen 2009).

The paper is organized as follows. First, we show the literature that deals with labor productivity in historical settings. We then sketch some relevant historical facts about Eastern Europe which help to understand our choice of variables and the results. A highly relevant point here is the comparison of Poland and the other parts of the Russian Empire. Next we explain the data and the variables that are used in the regression. The next section contains our empirical investigation of gross productivity determinants. The last section concludes.

4.2 Literature Review

The literature on historical labor productivity is mainly concerned with the evolution of aggregate productivity for an entire industry or country. Examples include the works of Broadberry (1997), which looks at Great Britain and a comparison of the British development with

other countries like Germany or the United States. Another study is Gadisseur (1983), who traces the evolution of Belgian labor productivity from 1846 to 1910.

Studies on an industry level include Walters (1975), who studies productivity in the coal mines of South Wales. His findings are that productivity varied during different time periods, and that voluntary absenteeism can explain parts of this variation.

Baten (2001) analyzes productivity in German firms from the late nineteenth century and provides evidence that smaller firms were more productive than larger ones.

For the US, Atack, Bateman and Margo (2006) use a data set of individual firms to study the impact and diffusion of steam machines in the American manufacturing from 1850 to 1880. Two studies by Sokoloff (1984 and 1985) are concerned with aggregate productivity development for whole American regions.

4.3 Historical Background

In the second half of the nineteenth century, the Tsarist Empire was in a phase of industrial development, although the country was still dominated by agriculture (Mironov, 2003). During the the nineteenth century it grew at a rate comparable to other European countries. Serfdom was abolished in 1861,[33] although it took some time till the peasants were really free. Railways were built and the railway companies were subsidized and forced to buy Russian material (Bovykin, 1975, p. 191). At the end of the nineteenth century Russia had the second

[33] In some parts of the Empire peasants were freed before this date, for example in Poland in 1807.

largest railway-network in the world, only exceeded by the USA. It had grown from 2 000 km 1861 to 60 000 km in 1905 (Thomas, 1992, p. 6).

Important industries were the textile industry, mining in the Urals, sugar processing in the Ukraine, and other manufacturing in St. Petersburg and Moscow. The industrial sectors developed quickly (Gregory 1972).

While foreign capital was flowing into the Russian Empire (Bovykin, 1975, p. 194ff) and some factories were able to produce goods of very high quality, the country remained an agricultural state and was hit by several regional famines in the 1880s and 1890s, especially in 1891. Nevertheless, agricultural output grew and Russia was one of the worlds leading grain producers (Gregory 1994).

Another important topic for the Tsarist Empire was schooling. Human capital accumulation was modest (Guseinov, 2000, p. 44) and started growing later than in other European countries. In the 1870s the younger parts of the population were well educated, while, older generations were still working in the factories with much less education. But we find a more rapid educational progress than in many other European countries in the Russian Empire.

The case of Poland deserves special attention. Since the Congress of Vienna in 1815 “Congress Poland” was part of the Tsarist Empire. While in Russia labor was abundant and the old elites were strengthened and interested in fostering agriculture from which they drew their wealth, the Tsarist state tried to disempower the Polish elites in order to punish them for participating in revolts (Rostworowski, 1896, p. 57f.). Therefore the Russian state freed peasants and set high taxes on agriculture, while the taxes on manufacturing were low, so that many land owners invested in industry. Another point that supported the Polish industry was a

relatively good access to the Russian market, which needed textiles and machines. The Empire built railway lines between Russia and Poland for strategic purposes (Westwood 1966, p. 38), which reduced transport costs. In Poland a flourishing manufacturing sector developed which expanded greatly between 1850 and 1870 (Luxemburg, 1898, p. 11). In the beginning of the nineteenth century western textile craftsmen and entrepreneurs came to Poland (Mühle, 1995, p. 42, Scherner, 2001, p. 164 and Luxemburg, 1898, p. 20) because in Prussia and Bohemia they were threatened by English competition after the removal of Napoleon's continental blockade, while in Poland they were protected from these competitors and had easy access to the Russian market. These entrepreneurs were the core of the later important textile sector in Polish industry. Another point is that because of the taxation and scarce labor in agriculture (Scherner, 2001, p. 136) in Poland many landowners invested their capital in the industry and not in agriculture, which means that the Polish factories had a better capital endowment than firms in the other parts of the Empire (Scherner, 2001, pp. 134ff.). The Polish factories were mainly active in textile and metal working (Scherner, 2001, pp. 173f.). Hence we need to control for substantial differences between Polish and Russian factories in the following.

4.4 Data and Variables

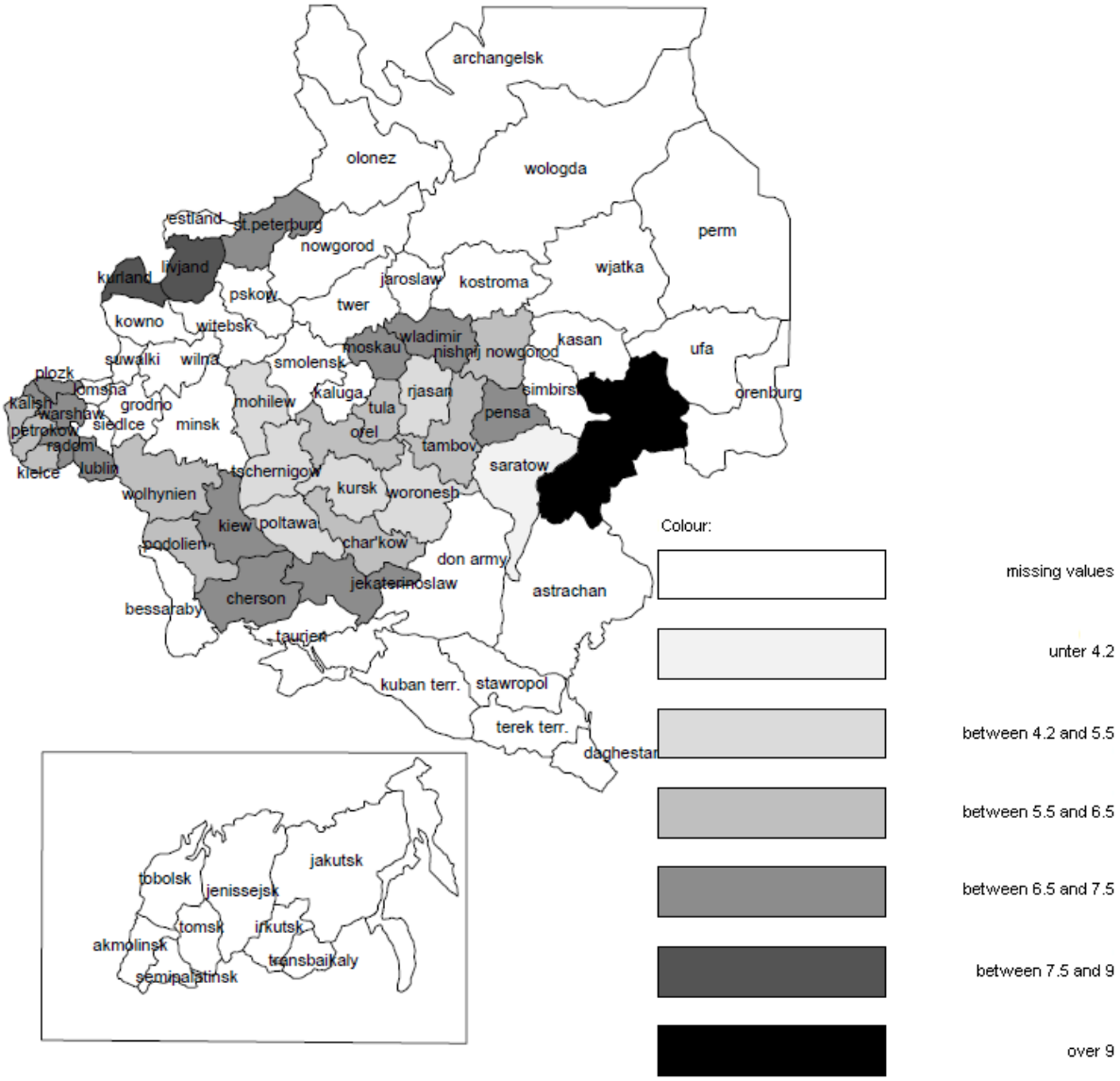
For the analysis we use data from a survey of Friedrich Matthäi (1873)[34] of the Russian Imperial industry and the exhibition of manufacturing in 1870. It contains information on the location of a factory, the value of the yearly production (in roubles), the number of employees and in most cases the use of steam machines for several hundred factories. The primary sources for his book are the official records of the ministry of finance. These are known to be not entirely reliable, but he also used additional sources like the exhibition catalog, which increases the reliability of the data. Our data are biased toward larger establishments, as Matthäi included detailed data mostly on those. But if the determinants of productivity do not vary with size, which is a reasonable assumption, the bias does not influence our estimations. We estimate yearly production per worker as a measure of productivity (see figure 1).

We have data about 57 factories active in mining, 313 which are processing food and liquors, 212 units from engineering, 157 factories from spinning and weaving and another 174 from various smaller industries like chemicals and leather processing, altogether making up 913 cases. Compared to all Russian industries, sugar-processing firms are slightly overrepresented, while leather is slightly underrepresented, but the overall picture represents the industry quite well. Of our 913 cases 351 firms are located in Poland, while the rest are firms from other parts of the Tsarist Empire. About 250 of our firms used steam machines. The mean number of workers in our sample is at 230, the median 76, which is quite close to the figures for

[34] The RUSCORP database from Thomas Owen could not be employed in this study due to unavailability of some variables.

the whole industry, although firms in our sample are a bit larger than the industry as a whole (Portal 1975).

Figure 1: mean log gross productivity

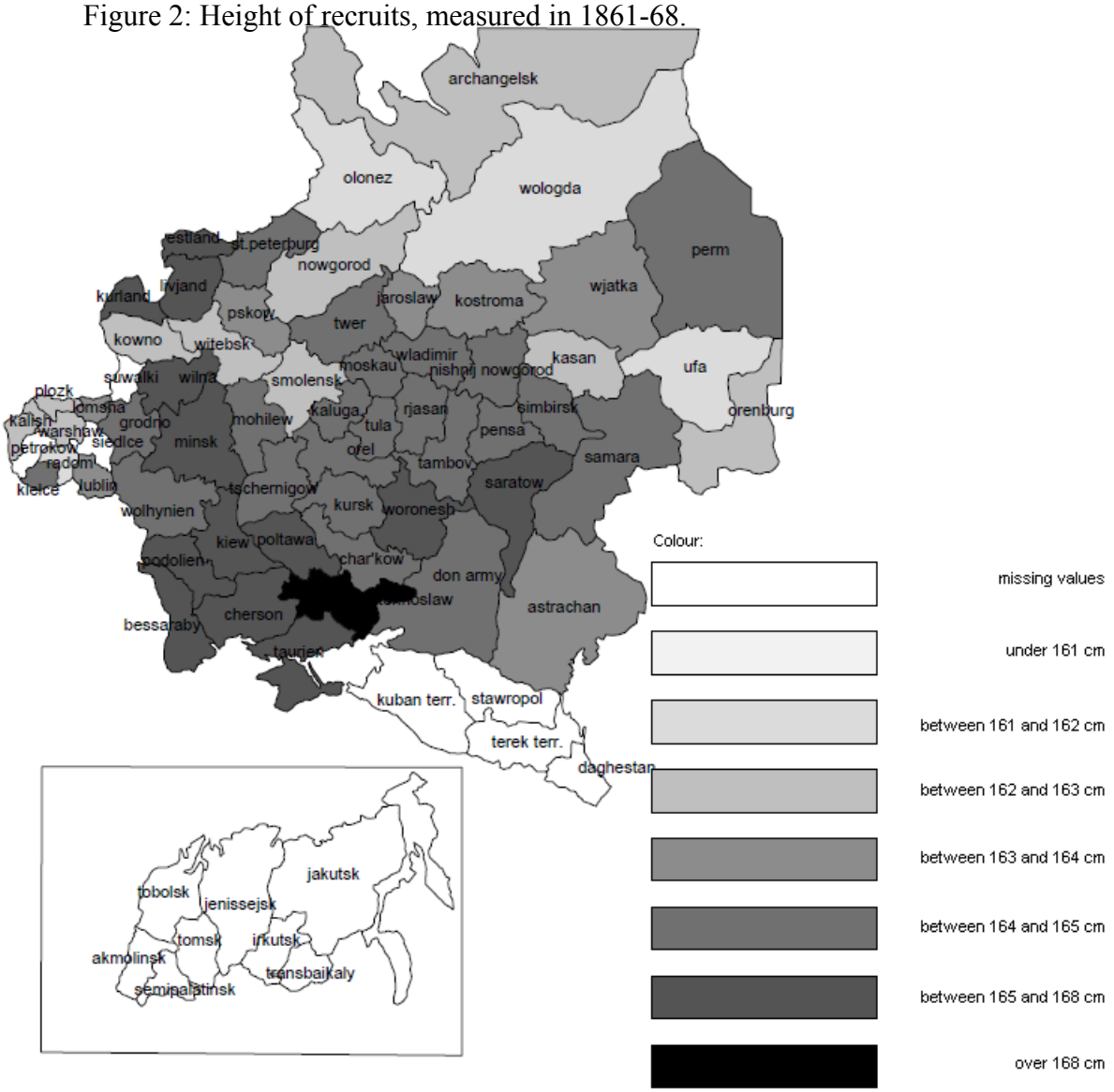


Source: own calculations based on Matthäi.

This firm data set has been connected to information about health (Strauss and Thomas, 1998) and human capital. In particular, we use height data derived from recruits

between 1863 and 1868. The height data were used on the level of gouvernements. We matched it to all firms which were located in that gouvernement. The average height in the gouvernements ranges from 161 cm in Radom and Kjeletz to 167.5 cm in Jekaterinoslav (see figure 2).

Figure 2: Height of recruits, measured in 1861-68.



Source: Woenno-Statistikscheskii Sbornik, 1871. The ranking is similar to Mironov's analysis of height differences, see Mironov (2004).

The reason for including heights in the analysis is that several scholars have found a correlation between nutrition/health and productivity for less developed countries (Strauss, 1986, p. 298, Schultz, 1961, Arora, 2001, Conlisk, Haas, Martinez, Martorell, Murdoch, Rivera, 1995). As the height of an adult person is heavily influenced by the food intake in the first years of life, height data are a proxy for the nutritional situation. Larger people may be more productive because they are stronger and healthier than the rest of the population and so will be better able to cope with the heavy work which has to be done in agriculture and factories.

Height might be endogenous here. Higher productivity leads to higher wages which enables the workers to consume more and better food, which increases their height. In our work this is not a huge problem as we have individual data for productivity and the effect of one firm on the height of a whole region should be limited, as roughly 90% of our firms have less than 500 workers. A second point is that the largest effect of food on height is during early childhood, so increasing productivity will increase the height of recruits with a lag of about 20 years. As Russia before 1850 was not very industrialized the impact of productivity on heights will be limited.

To control for numeracy,[35] one component of human capital, we construct ABCC indices[36] from the Russian Imperial census of 1897, also on the level of gouvernements. It refers to individuals born between 1825 and 1874. Numeracy is to a large extent determined by conditions around birth and during childhood. People who do not know their exact age often de-

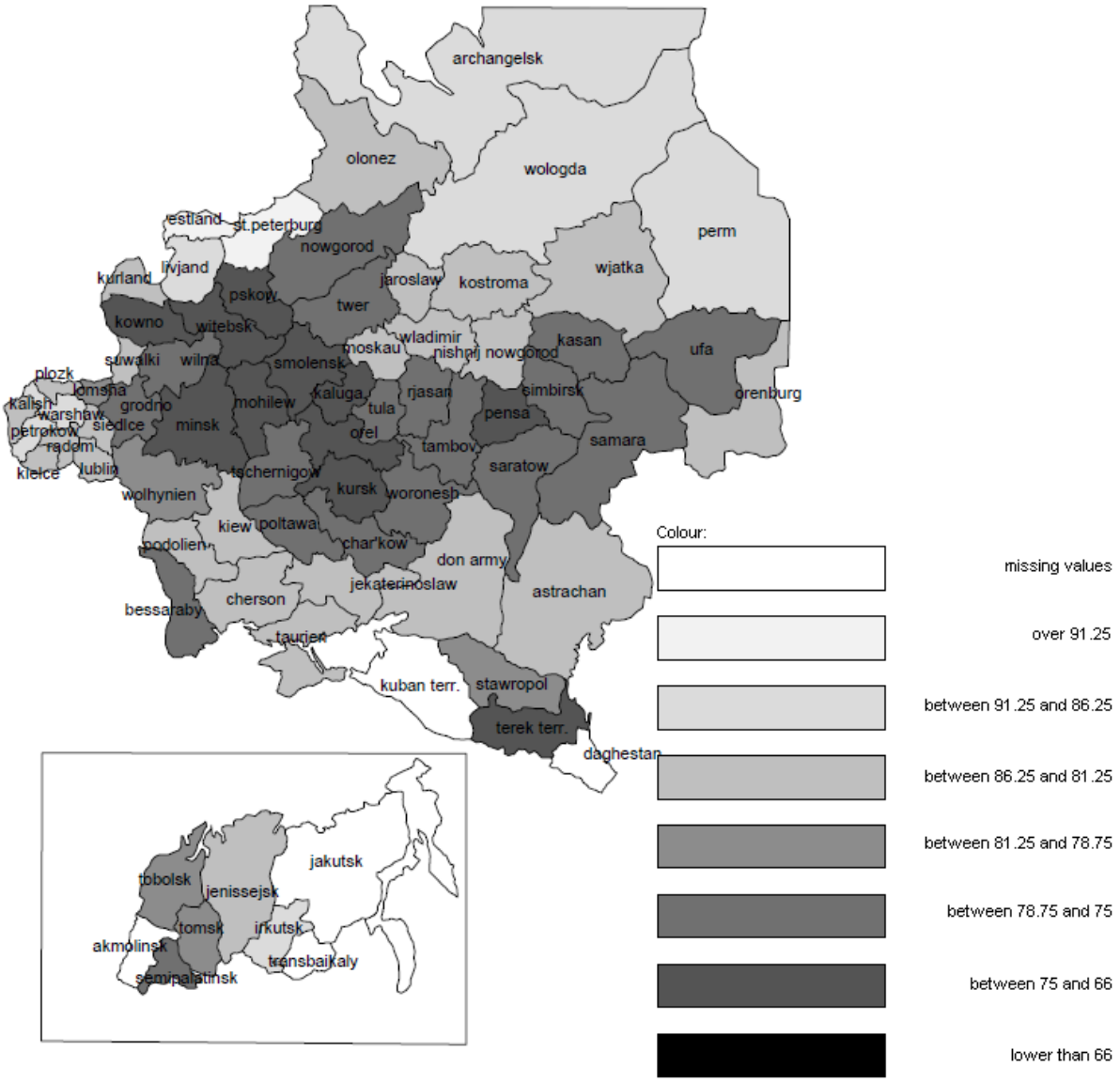
[35] Numeracy means the ability to operate with numbers and mathematical concepts.
[36] For the choice of the index see A'Hearn, Baten and Crayen 2009.

clare an age ending on 0 and 5.[37] The ABCC is a linear transformation of the Whipple index, which measures human capital by dividing the number of people who declared an age ending on 0 and 5 through the whole population between 23 and 72, then multiplying this number with 500. The ABCC is rescaled so that its range is from 0 (everybody is heaping) to 100 (the actual age distribution equals the theoretical distribution), which makes interpretation easier as a higher ABCC means higher numeracy (A'Hearn, Baten and Crayen, 2009, Crayen and Baten, 2006, 2010). The index therefore serves as a proxy for education. The ABCC index varies between 41 in Erivan and 97 in Estland (see figure 3).

We have to point to a potential problem of this variable here. Education could be endogenous, if for example more productive firms (for example because they use more complicated machines) need more highly educated workers and therefore begin training on the job. The government could also support more productive firms and build more schools in their locations. If this is the case, we see in our specification that education is a significant determinant of productivity although it is mainly the result of it. We believe that this is not a huge problem here because we use ABCC indices, and training on the job will more likely increase other skills apart from knowing one's age. We will check the endogeneity problem below by using another variant of the ABCC index, where we use just those people who were in 1897 43 or older, which means that they were born before 1855 and were mostly in the labor force in 1870. If more productive firms influenced the education in 1870, the educational level before that date should not be influenced. Because industrialization was advancing slowly in Eastern Europe in the early nineteenth century the effect of firms on

[37] The age-distribution in Russia shows a clear heaping on 0 and 5.

Figure 3: ABCC indices in Tsarist Russia



Source: Own calculations based on the Census of 1897. We thank Andrei Volodin and Leonid Borodkin for their help).

education was not as strong. The correlation between the two ABCC indices is around 0.96, which strengthens our assumption that endogeneity is no large problem here. The range of this new ABCC is from 28 in Erivan to 95 in Estland, which shows an improvement in education in the later nineteenth century.

We then control for the capital intensity of the factories. For this we created a dummy variable that equals 1 if the factory uses any steam machine, no matter how high the total number of machines (Crafts, 2004). If we use our variable for steam use in the firm our sample is reduced to one third of all cases due to data unavailability. We therefore constructed a proxy for the use of steam machines in the following way: The proxy was only assigned to firms from regions where we had information on steam use for more than 10% of the firms from that region (at least 3 firms). When over 50% of these firms used steam machines all firms in that gouvernement received a "1", otherwise a "0" was assigned to them.

The choice of our industry dummies is as follows: We control for sugar and cotton because these were very important industries in Tsarist Russia; the same is true for engineering and mining. We include the chemical industry because this was a new industry for which we assume high productivity and also for the alcohol producing industry. These dummy variables control for the difference between production and value added in the industries.

We also include a dummy variable which equals one if the firm has less than 50 workers, because it is controversial in the literature whether smaller or larger firms are more productive (Baten, 2001). Using the number of workers as an independent variable would introduce an endogeneity bias, so we use the dummy-variable.

4.5 Regression Results

To investigate the determinants of firm productivity in the Russian Empire we use OLS regressions. We use the following model to link productivity to our explanatory variables:

$$\lg arbp = \beta_0 + \beta_1 Height + \beta_2 indcenter + \beta_3 Poland + \beta_4 Whipple + \beta_5 Small + \beta_6 Steam$$

$$+ \sum_{i=7}^{n} \beta_i Industrydummy + u$$

The dependent variable lgarbp is the logarithm of gross productivity, which is calculated as the value of the yearly production of the firm divided through the number of employees of the firm (Walters, 1975, Gadisseur, 1983). This measure of productivity has the advantage that it does not depend on the financial situation of the firm, as interest payments or other financial activities do not influence our measure. As independent variables, we use the height of recruits in the 1860s, the ABCC index from the 1897 census to study effect of education[38] and several other dummy variables: First of all an industrial center dummy which equals one if the firm was located near or in St. Petersburg, (Yatsunsky, 1974, p. 131.) Moscow (Yatsunsky, 1974, p. 132) or the Vladimir gouvernement (v. Schulze-Gävernitz, 1899, p. 53). This variable allows for influences of these large industrial centers, for example through networking or educational externalities, when one firm can learn from the neighbor (Lall, Shalizi and Deichmann, 2004); next a variable for Polish factories to control for the particularities of that country as described above.

We control for economies of scale with a dummy for small firms. Next we use a set of industry dummies to control for different technologies in the different industries. We also use

[38] See Hall and Jones (1999) for a theoretical and empirical investigation of the effect of human capital on output per worker. They find that institutions explain most differences in output per worker, which does not apply for our dataset as all our observations come from one country, so differences in institutions are limited.

dummy variables for the utilization of steam machines or the proxy variable, as explained above. As infrastructure leads to economic growth and expands the market available for a firm, we control for the infrastructure in a region using a dummy-variable which equals 1 if the region had access to a railway or a major river.

Moulton (1990) has shown that merging micro-data with aggregated data leads to invalid standard errors. The t-values increase in that case by a factor of about 3 to 5 which yields highly statistically significant results. As our human capital and height variables are aggregated we use corrected standard errors.

We test several models to assess the robustness of our results and test for endogeneity of our human capital measure. We estimate the equation with and without the respective variable for the steam use and once with our ABCC index and then with the ABCC index constructed for the birth cohorts of 1825-1855. The results of our estimation are shown in table 1 and the appendix.

Our adjusted R^2 of 0.472 in Model I shows that our variables capture nearly one half of the productivity differences between factories, which is quite high for a cross-sectional data set, especially as we have no data on management or other firm-specific factors. Even if we drop our industry dummies the R^2 remains at 0.33, which is still very large (not shown).

Table 1: Determinants of gross productivity in the Russian Empire in 1870.

Variable	Model I	Model II	Model III
Height	0.136* (0.073)	0.247*** (0.009)	0.160* (0.004)

ABCC	0.035** (0.018)	0.041** (0.032)	0.035** (0.000)
Industry center	0.416** (0.048)	0.673** (0.014)	0.154** (0.001)
Poland	0.462* (0.087)	1.098** (0.012)	0.197* (0.008)
Small	0.440*** (0.000)	0.446*** (0.007)	0.188*** (0.000)
Proxy for steam power using region	0.125 (0.681)		0.0047 (0.284)
Sugar	-0.796** (0.010)	-0.679** (0.0112)	-0.313** (0.000)
Mining	-0.450** (0.010)		-0.095** (0.002)
Engineering	-0.561*** (0.000)	-0.625*** (0.001)	-0.205*** (0.000)
Cotton	-0.327 (0.187)		-0.108 (0.006)
Distillery	1.797*** (0.000)	1.939*** (0.000)	0.271*** (0.000)
Chemical	0.451*** (0.001)		0.063*** (0.018)
Steam use		0.886*** (0.000)	
Transportation	0.131 (0.453)	0.028 (0.884)	0.056 (0.113)
Constant	-18.789 (0.123)	-38.226** (0.013)	-18.789 (0.013)
R^2	0.473	0.704	0.473
No. of cases	913	334	913

Notes: ***/**/* p-value $\leq$ 0.01/0.05/0.1. P-Values in Parentheses. Dependent variable is the logarithm of revenue per worker.

We find a positive and significant influence of the heights of recruits on gross productivity (Model I). This means that health and nutritional differences in Russia influenced the productivity of workers in 1870.[39] The ABCC index is statistically highly significant and

[39] Because we use gross-productivity and not wages we can say that this positive influence is not the result of a selection process. It might for example be the case that employers pay taller workers more because they only expect them to be more productive.

has the expected positive sign. This shows that human capital had a positive influence on productivity (Broadberry and Ghosal, 2002).

Our industrial center variable, which controls for firms in St. Petersburg, Moscow and Vladimir, is significant, which shows that location mattered, perhaps due to network-effects. Our dummy variable for Poland is significant and positive; hence the different economic conditions in Poland and the rest of the Empire gave the Polish industry a productivity advantage. Of course, this could be due to other factors like the availability of raw materials. But comparing the steam data for Poland and other gouvernements, we find that the fraction of steam-using firms is higher in Poland than in Russia as a whole, 84% compared to 74%. The steam proxy variable is insignificant and positive. We show in later models that the insignificant effect is due to the rough specification of our variable.

Our regression also shows that in our sample small firms were significantly more productive than larger ones. This strengthens the results obtained by Baten (2001) who shows that small firms in Germany around 1900 were more productive than large ones. This result might be influenced by our productivity measure. We measure it as revenue per worker, and large textile firms tend to employ more female workers, who were discriminated against by education, and large sugar firms employ more peasant (seasonal) workers. As excluding those industries does not change the productivity advantage by small firms (not shown), we conclude that our result is not only a statistical artifact.

The industry dummies are partly significant, which reflects different technologies and raw material shares in different industries, but also industrial market structure, because more competitive industries will have lower prices which will reduce the value of output and hence

decrease our productivity measure. The workers in the distillery sector were the most productive workers in our sample, which might be due to weak competition.

An interesting case is the low productivity of the cotton industry (v. Schulze-Gävernitz, 1899, p.68). Yatsunsky states that around 1800 the cotton industry was the technologically most advanced sector in the Russian Empire (Yatsunsky, 1974, p. 114). This advance should be reflected in higher productivity. Factors explaining the low productivity include seasonal workers, other forms of part-time employment and the discrimination of women who worked in this industry. Moreover, the cotton industry was a strong part of the Polish industry; in 1879 it comprised around 30 % of all workers in manufacturing (Scherner, 2001, p. 165). It had also nearly unlimited access to the Russian market, where it was a strong competition for the domestic industry. Although after 1850 tariffs limited this access, many Polish textile workers emigrated over the border into other parts of the Empire (Scherner, 2001, p. 167). So the Russian cotton market was highly competitive, which lowers our production value measure. Additionally Yatsunsky states that the cotton industry was hit very hard by the crisis of 1867 (Yatsunsky, 1974, p. 118). So our low productivity measure for cotton could also be a temporal appearance, as our cross-sectional data represent the state of the industry around 1870.

The chemical industry seems to be a highly productive industry which is apparent through the positive and significant coefficient. Firms in the engineering sector achieve comparably lower output per worker (Kahan, 1989 p. 22).

In Model IV[40] we exclude the proxy for steam engines, because the variable could be misspecified. But the exclusion of the variable does not change much and no conclusion changes. This indicates that the variable does not suffer from a large bias.

In Model II we use the variable with the data on steam use instead of our regional steam proxy. Due to data unavailability our sample is reduced to 334 firms, which represent 37 % of the original sample. Our R^2 increases to over 0.7, the other coefficients do not change much. The increase of the R^2 is due to the smaller sample, as can be seen in model V, where we estimate the original equation with the small sample. There, our small firm dummy gets insignificant. This is probably caused by the fact that the use of machines was more common in large firms.

As we mentioned above we do not believe that endogeneity is a huge issue in our data but to test it, we ran additional regressions and replaced our original ABCC variable with the ABCC indices from 1855. The differences in the estimates are rather small.[41]

The last column of Table 1 shows the same regression as in Model I, but this time we computed standardized coefficients. These coefficients allow us to compare the strength of the effects, as they measure the effect a one standard deviation increase of the independent variable has on our dependent variable. The coefficients for the dummy variables are not of special interest, but we see that the effect of education as measured through the ABCC index is much smaller than that of heights. This shows that in countries where nourishment is poor, large productivity gains from better food can be expected.

[40] Models IV, V, VI and VII are included in the appendix in table A1.
[41] Model VI corresponds to model I and model VII to model II in table 1 except of the other ABCC-variable.

As an additional robustness-check we estimate the basic-regression (Model I), this time ignoring the industry-dummies, each regression using only firms from the same industry. For the sugar-processing industry with 255 observations we get the same results like in the full specification. For the 204 observations from the engineering industry we find the AB-CC-index being insignificant, the other results remain as before. Analyzing the cotton industry, we find the ABCC-index being significant, while Poland and height become insignificant, as these firms are nearly all from Poland where heights vary only little. All in all, the main results remain fairly stable.

4.6 Conclusion

The Russian Empire was in a phase of industrialization in the later parts of the nineteenth century. Nutritional status and health were still far worse than today, and we find large height differences between the Russian gouvernements. The government aimed at supporting industry while at the same time protecting the rights and privileges of the rich landowners. The special treatment of Poland, where the Russian authorities tried to harm large landowners, led to increased investment in industry and machines. This paper examines which influence these conditions have on productivity.

The analysis demonstrates the importance of human capital and health for worker productivity in countries that are at a comparable stage as the Russian Empire in the late nineteenth century, with a GDP per capita of 943 $ (in real $ of 1990) (Maddison, 2001, p. 264).

We therefore confirm the ideas of Schultz (1961) and others, who argued that education has to be seen as an investment that can produce large increases in productivity. According to Maddisons GDP estimates, Russia in the nineteenth century was comparable to modern developing countries (Somalia and Nepal have similar levels). Therefore, our results are interesting for developing countries which have, as Russia did back in 1860, a focus for industrialization although nourishment of the population is far from adequacy. Our results suggest that improving the supply of high-quality food and education can increase productivity.

Our results show that Polish firms were more productive than their competitors from other parts of the Empire, and we present evidence that this is due to either the better capital equipment of Polish firms. These results therefore confirm the results of Scherner (2001, pp. 181f.) who suggests that Poland benefited from the Russian foreign rule because in Poland the old elites were weakened while in Russia they kept their power which slowed down the industrial progress. The steps taken in Poland helped the peasants which provided a strong home market for the products of the own industry.

References

A'Hearn, B., Baten, J. and D. Crayen (2009). Quantifying quantitative literacy: age heaping and the history of human capital. The Journal of Economic History 69 (3), 783-808.

Atack, J., Bateman, F. and R. Margo (2006). Steam power, establishment size, and labor productivity growth in nineteenth century American manufacturing. NBER Working Paper 11931, http://www.nber.org/papers/w11931.

Arora, S. (2001). Health, human productivity, and long-term economic growth. The Journal of Economic History 61 (3), 699-749.

Bartelsmann, E. J. and M. Doms (2000). Understanding Productivity: Lessons from Longitudinal Microdata. Journal of Economic Literature 38 (3), p. 569-594.

Baten, J. (2001). Produktivitätsvorteil in kleinen und mittleren Unternehmen, Sicherheit in Großunternehmen? Die Gesamtfaktorproduktivität um 1900. Tübinger Diskussionsbeitrag Nr. 217.

Black, S. and L. Lynch (1996). Human-Capital Investments and Productivity. The American Economic Review, Papers and Proceedings, 86 (2), 263-267.

Bliss, C. and N. Stern (1978). Productivity, wages and nutrition. Part I: the theory. Journal of development economics 5, 331-362.

Bouis, H. and L. Haddad (1991). The impact of nutritional status on agricultural productivity: wage evidence from the Philippines. Oxford Bulletin of Economics & Statistics 53, (1), 45-68.

Brahmananda, P.R. (1982). Productivity in the Indian Economy. Bombay: Himalaya Publishing House.

Broadberry, S. (1997). The Productivity Race: British Manufacturing in International Perspective. Cambridge: Cambridge University Press.

Broadberry, S. and S. Ghosal (2002). From the counting house to the modern office: explaining Anglo-american productivity differences in services, 1870-1990. The Journal of Economic History 62 (4), 967-998.

Bovykin, V. (1975). Probleme der industriellen Entwicklung Russlands, in: Geyer, Dietrich (Editor): Wirtschaft und Gesellschaft im vorrevolutionären Russland, Köln: Kiepenheuer & Witsch.

Chandler, A. D. (1990). Scale and Scope: The Dynamics of Industrial Capitalism. Cambridge: Belknap Press of Harvard University Press.

Crafts, N. (2004). Steam as a general purpose technology: a growth accounting perspective, The Economic Journal 114 (1), 338-351.

Crayen, D. and J. Baten (2006). Numeracy, Inequality, Age Heaping, and Economic Growth: New Estimation Strategies for Western Europe and the U.S. (17th - 19th centuries), Economic History Review 63 (2), 452-478.

Crayen, D. and Baten, J. (2010). Global trends in numeracy 1820–1949 and its implications for long-term growth; Explorations in Economic History; Volume 47 (1), 82-99.

Crépon, B., Duguet, E. and J. Mairesse (1998). Research, Innovation and Productivity: an economic analysis at the firm level. Economics of Innovation and New Technology 7, 115-158.

Englander, S. and A. Gurney (1994). Medium-term determinants of oecd productivity, OECD Economic Studies.

Gadisseur, J. (1983). Output per Worker and its evolution in Belgian industry, 1846-1910, in: Fremdling, Rainer (editor): Productivity in the economies of Europe. Stuttgart: Klett-Kotta.

Galenson, W. (1955). Labor productivity in Soviet and American industry. New York: Columbia University Press.

Gregory, P. (1972): Economic Growth and Structural Change in Tsarist Russia: A Case of Modern Economic Growth? Soviet Studies (23) 3, 418-434.

Gregory, P. (1994): Before command. An economic history of Russia from emancipation to the first five-year plan. Princeton: Princeton University Press.

Guseinov, R. M. (2000). The phantom of capitalism in Russia (I). Problems of economic transition 42 (12), 37-52.

Conlisk, E., Haas, J., Martinez, E., Martorell, R., Murdoch, S. and J. Rivera (1995). Nutritional supplementation during the preschool years and physical work capacity in adolescent and young adult Guatemalans. The Journal of Nutrition 125: 1078-1089.

Hall, R. and J. Charles (1999). Why do some countries produce so much more output per worker than others? Quarterly Journal of Economics, 114, 83-116.

Haltiwanger, J., Lane, J. and J. Spetzler (1999). Productivity differences across employers: the roles of employer size, age, and human capital. The American Economic Review 89 (2) Papers & Proceedings of the one hundred Eleventh Annual Meeting of the American Economic Association, 94-98.

Kahan, A. (1989). Russian economic history. The nineteenth century. Edited by Roger Weiss. Chicago: The University of Chicago Press.

Lall, S., Zmarak S. And U. Deichmann (2004). Agglomeration economies and productivity in Indian industry. Journal of development economics 73, 643-673.

Luxemburg, R. (1898). Die Industrielle Entwicklung Polens. Leipzig: Duncker & Humblot.

Maddison, A. (2001). The World Economy. A Millenial Perspective. OECD.

Matthäi, F. (1873). Die Industrie Russlands in ihrer bisherigen Entwicklung und in ihrem gegenwärtigen Zustande mit besonderer Berücksichtigung der allgemeinen russischen Manufaktur-Ausstellung im Jahre 1870. Volume 2, Leipzig: Verlag von Hermann Fries.

Mironov, B. N. (2003). Russia: Modern Period, in: Mokyr, Joel (Editor): The Oxford encyclopedia of economic history. Oxford: Oxford University Press.

Mironov, B. N. "Zhiznennyi uroven' Sovetskoi Rossii pri Staline po antropometricheskim dannym," in Ekonomicheskaia istoria. Ezhegodnik 2004. Moskva: ROSSPEN. 565-588.

Moulton, B. (1990). An illustration of a pitfall in estimating the effects of aggregated variables on micro units. The Review of Economics and Statistics 72 (2), 334-338.

Mühle, R. (1995). Colonist Traditions and Nineteenth-Century Emigration from East Elbian Prussia, in: Hoerder, Dirk and Jörg Nagler (ed.): People in Transit: German Migrations in Comparative Perspective, 1820-1930, Cambridge: Cambridge University Press.

Portal, Roger (1975). Die russische Industrie am Vorabend der Bauernbefreiung, in: Geyer, Dietrich (Editor): *Wirtschaft und Gesellschaft im vorrevolutionären Russland*, Köln: Kiepenheuer & Witsch.

Rostworowski, A.J.F.C. v. (1896). Die Entwicklung der bäuerlichen Verhältnisse im Königreich Polen im 19. Jahrhundert, in: Conrad, J. (Editor): Sammlung nationalökonomischer und statistischer Abhandlungen des Staatswissenschaftlichen Seminars zu Halle a. d. S., Vol. 12, Jena.

Scherner, J. (2001). Eliten und wirtschaftliche Entwicklung – Kongreßpolen und Spanien im 19. Jahrhundert. Münster: Lit. Verlag.

Schultz, T. W. (1961). Investment in Human Capital. *The American Economic Review* 51 (1), 1-17.

v. Schulze-Gävernitz, G. (1899). Volkswirtschaftliche Studien aus Russland, Leipzig: Duncker & Humblot.

Sokoloff, K. L. (1984) Was the Transition from the Artisanal Shop to the Nonmechanized Factory Associated with Gains in Efficiency?: Evidence from the US. Manufacturing Censuses of 1820 and 1850. Explorations in Economic History, 21, 351-382.

Sokoloff, K. L. (1985) Productivity Growth in Manufacturing during Early Industrialization: Evidence form the American Northeast, 1820 to 1860. NBER Working Paper No. 1685.

Strauss, J. (1986). Does better Nutrition raise farm productivity? Journal of Political Economy 94 (2), 297-320.

Strauss, J. and D. Thomas (1998). Health, Nutrition and Economic Development. Journal of Economic Literature 36 (2), 766-817.

Thomas, L. (1992). Rußland im Jahre 1900. Die Gesellschaft vor der Revolution. Inaugural lecture, 13.10.1992. http://dochost.rz.hu-berlin.de/humboldt-vl/thomas-ludmila/PDF/Thomas.pdf, downloaded at 6.07.2007.

Walters, R: (1975). Labor productivity in the South West Steam-Coal Industry, 1870-1914. The Economic History Review, New Series, 28 (2), 280-303.

Westwood, J. N. (1966). Geschichte der russischen Eisenbahnen, Zürich: Orell Füssli.

Yatsunsky, V. K. (1974). The industrial revolution in Russia, in: Blackwell, William (Editor): Russian Economic Development from Peter the Great to Stalin, New York : New Viewpoints.

Appendix

Table A1: Regressions of gross productivity of Russian firms in 1870

Dependent variable: logarithm of labor productivity.				
Variable	Model IV	Model V	Model VI	Model VII
Height 1860	0.144* (0.056)	0.333*** (0.003)	0.148* (0.059)	0.268*** (0.006)
ABCC	0.034** (0.021)	0.056*** (0.007)		
Industry-center	0.395* (0.081)	0.816** (0.011)	0.491** (0.013)	0.760*** (0.004)
Poland	0.453* (0.099)	1.402*** (0.005)	0.542** (0.047)	1.189*** (0.006)
Small	0.439*** (0.000)	0.182 (0.175)	0.441*** (0.000)	0.447*** (0.007)
Proxy for steam power using region			0.140 (0.650)	
Sugar	-0.732*** (0.003)	-0.717** (0.032)	-0.789** (0.011)	-0.651** (0.020)
Mining	-0.455** (0.010)		-0.465*** (0.007)	
Engineering	-0.563*** (0.000)	-0.726*** (0.000)	-0.559*** (0.000)	-0.631*** (0.001)
Cotton	-0.325 (0.196)		-0.314 (0.219)	
Distillery	1.785*** (0.000)	1.673*** (0.000)	1.789*** (0.000)	1.922*** (0.000)
Chemical	0.454*** (0.000)		0.452*** (0.000)	
Steam use				0.902*** (0.000)
Transportation	0.128 (0.457)	-0.016 (0.944)	0.143 (0.427)	0.032 (0.867)
ABCC1855			0.023** (0.030)	0.029** (0.046)
constant	-19.933 (0.100)	-52.737** (0.004)	-19.576 (0.121)	-40.410** (0.012)
R^2	0.473	0.638	0.471	0.702
N	913	334	913	334

Notes: ***/**/* p-value $\leq$ 0.01/0.05/0.10. P-Values in Parentheses.

5 Regional Determinants of Firm Creation in the Russian Empire. Evidence from the 1870 Industrial Exhibition.

Abstract

Firm creation is a central aspect of industrial development. New firms create opportunities for employment, new products and add competitive pressure to existing firms. We assess firm creation rates in Russia around 1870 using cross-sectional data from the industrial exhibition in St. Petersburg. The Econometric results show that human capital was an important pre-requisite for firm creation.

This chapter is based on an article that was written together with Jörg Baten and that has been published in the Russian Economic Yearbook. The idea was developed jointly, the writing equally shared.

5.1 Introduction

In this paper we study the determinants of firm creation in Russia in the 1870s. Our main interest is the importance of human capital for firm creation. Firm destruction and creation are vital parts of our capitalist world. Disappearance of inefficient firms lowers average production costs and prices, which in turn increases welfare of a society. Creation of new firms increases competition and thereby increases efficiency in the industry.

A large literature has developed concerning the determinants and implications of firm creation. This literature has considered many factors regarding firm creation, among them human capital, state aid and competitiveness of the financial markets. Our interest in this study is to see whether the determinants of firm creation change in the course of development, which means whether the determinants today and in former times are identical. Now why could the determinants differ? It might be that factors that were important in the 19th century are no more important today, for example because state infrastructure replaced that factor. One example could be the location near large markets, for example cities, which might have lost importance as infrastructure is more developed today that it was a century ago. Human capital might be an example that is more important today than it was earlier, as now the requirements for firm creators are much stronger than they had been. Today's technology is far more advanced, so a new entrepreneur needs a much larger knowledge to create competitive products than before. Also, regulation and applications for state aid are sometimes quite difficult today, things an entrepreneur has to cope with.

The study develops as follows. We first show our data set and discuss the representativeness for the Russian economy. We then discuss potential determinants of firm creation which will be tested in the empirical part. Our regression results are shown and we conclude with a short summary of the main findings.

5.2 Data

Our data set comes from a survey of Russian industry, which uses government statistics and additional information, which the attending firms showed at the industrial exhibition in St. Petersburg in 1870, by Friedrich Matthäi, an Austrian military officer (Matthäi 1873).

He reported the firms from all industries that were located in each gouvernement. As the Russian Empire included during this period not only Russian territory, but also today's territories of Ukraine, Belarus, Poland, Lithuania, Estonia, Latvia, Moldavia and the Caucasian and Central Asian countries, we can study the firm development process in all those regions (however, we excluded the non-Russian Caucasus and the region east of the Ural, and Finland for lacking some of the variables). We do not know exactly when those firms were created, or when they were transformed from small craftsmen workshops to modern industrial enterprises. It is clear that they were created before the exhibition, and most of them probably underwent this process of creation or transformation in the decades before the exhibition, as we can guess from similar samples of German or American firms (Baten 2003). This is also confirmed by the fact that the Russian Industry expanded greatly in the 1860s. A transforma-

tion from a craftsman workshop to an industrial firm has many features of creating a new enterprise, so we will speak of "creation" in the following. We can relate this figure to the size of the population and study regional factors which might have increased the previous creation activity of firms, relative to the potential firm creators. We define the potential firm creators as the population born before 1850, i. e., those aged 20 and above at the time of the exhibition, assuming similar mortality in the regions.

To which extent are the exhibiting firms a selective sample of the total underlying firm population? Certainly, not all firms participated in the exhibition, and Matthäi (1873), who is our main source, might not have included all exhibiting firms in his lists of "excellent firms". We therefore study the creation of promising firms in Russia. Secondly, there was certainly some minimum criterion of visibility for demand, which might be higher, if the firm was addressing not only the regional market, but targeted national or international customers as well. Thirdly, firms with innovative products or new production methods might have been overrepresented among the exhibitors. Finally, a certain bias might result from proximity to the exhibition place, which reduced transport costs to the exhibition, and increased the likelihood that an entrepreneur was informed at all about the exhibition. On the other hand, the organizers of the exhibition were clearly interested in showing a wide range of Russian firms, and hence advertised broadly. They also wanted to create the impression of a booming and promising industrial base which had a large and geographically dispersed base. Therefore, the number of firms is quite impressive and is not dominated, say, by machinery firms of St. Petersburg or Moscow, but also many sugar processing firms from Ukraine and even firms from Perm near the Urals, among many others. Matthäi also aimed at presenting a complete picture

of the Russian industry, so research for his tables was thorough. We gain the impression that while there might have been some bias towards size and modernity, we can obtain in general a broad picture of Russian industry in its geographical dispersion.

5.3 Determinants of Firm Creation

What might have determined the number of "excellent" firms that were created in the regions of the Russian Empire during the beginning of industrialization?

1. Urban demand of goods and supply of industrial and organizational skills might be one major driving force. Many rural taxes went to the large urban centers, therefore creating additional demand. At the same time, urban handicrafts had developed the necessary skills to produce industrial goods, as is shown by the examples Matthäi names for Russian firms that won international awards for their products. It is quite evident that this variable will be important; the value-added of our study is rather to quantify the extent of this urban factor. We use the three large cities of St. Petersburg, Moscow and Kiev to study this effect.
2. Related to urban demand and supply, but less obvious is the proximity-to-city effect. It might be that a combination of cheaper land prices (for the factory site and worker's dwellings) and modest transport costs led to setting up firms in the vicinity of large cities. Gouvernements near big cities might also be affected by human capital externalities (See Baten et al. 2007). Marshall (1920), Arrow (1962), and Romer (1986) ar-

gued that industrial agglomerations (also termed "industrial districts") lead to advantages in the diffusion of knowledge and technology, as well as more effective labor markets, and this is even more likely for urban agglomerations. While those authors refer to agglomeration effects of similar industries, other effects have been proposed for the spread of knowledge across industries (Jacobs 1970). On the other hand, Christaller's (1933) theory might suggest that the areas near cities might rather specialize on producing agricultural products of limited transportability to feed the urban industrial producers. Moreover, there might have been industrial obstacles in the countryside of the Russian Empire, stemming from the traditional organization of the villages and estates. Hence we could expect both positive and negative signs from our "near city" dummies, which identify rural districts near big cities.

3. Special supply factors could play a large role, which is most obvious in the sugar processing industry, which relies heavily on a raw material that cannot be transported economically to distant locations. So the gouvernements where the raw material is available will have more firms than other comparable gouvernements.
4. In a similar vein, trade policy factors might influence industrial location. This point is quite influential in the case of Poland (the part of Poland which was, after the partition, included in the Russian Empire). The Russian protectionist policy attracted textile producers from Prussia and the Habsburg Empire, which were not competitive anymore after the removal of the Napoleonic continental blockade. Those producers were able to develop rapidly in Poland, protected against British productivity advantage, serving the Russian and Polish market (Wolf 2006). Moreover, Tsarist Imperial

policy of the time was aimed at weakening the Polish large landowners and nobility, creating incentives to invest in factories rather than agricultural production capital (Scherner 2001, p. 163). We will create a dummy variable for the Polish gouvernements.

5. Transport infrastructure might have encouraged the creation of firms, such as harbors, rivers, and railways (Holl 2004). We will distinguish water access that was directed towards the Baltic, Black Sea, or large cities (with one dummy variable "Baltic/Black Sea river system"), as opposed to water access directed toward the Caspian Sea, which might have provided less access to consumers (another variable "Caspian Sea river system"). We will also separately study the effect of the first railway lines. It is clear that it is also possible that the railways were build to the large industrial centers, making it an endogenous variable. This is in the case of Russia not an important problem as the early railway lines were mainly built for strategic purposes and not according to the needs of the industry.
6. Special demand factors are important for some local industries. For example, in the proximity of mining, mining machinery firms will be created, which can react quickly to local demand and provide services to mining firms.
7. Regional human capital can provide an incentive to create a firm in a specific region, if the wages are not too high there. Even more important, the firm creator himself needs substantial knowledge and the ability to set up the firm and survive the first difficult years (Acs, Florida and Lee 2004, Mathur 1999). We measure regional human capital with the age-heaping indicator (see Crayen/Baten 2006, 2007, A'Hearn, Baten,

Crayen 2009), and alternatively with the more traditional literacy rates of those born before 1850. Both variables stem from the Imperial census of 1897.[42] The importance of human capital has been demonstrated in the literature by studying the unequal pay of skilled and unskilled workers (Gregory and Borodkin 2000). Our two indicator variables are modestly correlated, which is visible in the scatter plot of Figure A1 (in the appendix). Only Kurland (today a part of Latvia) and Kowno (today North Lithuania) stood out as being better in literacy terms than in numeracy terms, whereas North and East Russian regions such as Archangelsk, Wologda, Perm, Olonez, and Wjatka were somewhat better with numbers than with words.

8. Alternative specialization in agriculture is another determinant. If agricultural profitability is high, there might not be enough incentives for creating industrial firms (Audretsch and Vivarelli 1996; they show that low wages encourage the creation of new firms). Agricultural productivity is difficult to measure for the pre-1870 period. We use the grain and potato yields in kg per hectare in the 1880s (Materialy 1903[43]), and assume that soil quality, ability of workers, demand and export possibilities, cities etc. remained similar between regions between the pre-1870 and the 1880s period.

[42] The data was friendly provided to us by Andrej Volodin and Leonid Borodkin.
[43] Information was provided by Boris Mironov.

5.4 Results

The results of the estimations are shown in table 1. The dependent variable is the log of the number of firms compared to the number of potential firm creators. All in all our variables capture the most important determinants of firm creation, as we can explain over half of the variation in the creation-ratio with the variables used here.

Human capital was clearly very important in the creation of Russian firms. Regions with a population that could report their age more exactly had a substantially higher rate of firms that exhibited on the St. Petersburg exhibition in 1870. This holds true even after controlling for urban demand and supply, railway and water access, and similar variables which might have stood behind the human capital proxies. Exchanging the numeracy proxy in the second model with log of literacy does not change this result. Our numeracy proxy (the ABCC Index) has almost the same explanatory power than the literacy indicator. Both variables influence significantly the firm creation rate.

We observe that Poland had a significantly higher rate of firms per capita than the other regions. Whether this might be due to trade policy that attracted entrepreneurial migrants or to the Tsarist tax policy, which turned landowners into factory owners, or to other factors, cannot be distinguished. Moreover, urban supply and demand factors mattered, even after controlling for human capital differences – the big city dummy is highly significant and robust. The effect is also substantial and its magnitude larger than the transport variables.

Table 1: Determinants of Firm Creation in the regions of the Russian Empire before 1870

Model	1	2	3
ABCC (numeracy)	0.056*** (0.00)		0.052*** (0.00)
Literacy (Log)		1.19*** (0.00)	
Poland	2.54*** (0.00)	2.39*** (0.00)	
Big City	2.10*** (0.00)	2.20*** (0.00)	2.14*** (0.00)
Near big city	0.76** (0.01)	0.75** (0.02)	0.52* (0.10)
Railway access	0.36 (0.29)	0.34 (0.38)	
Baltic/Black Sea river system	-0.59* (0.05)	-0.96*** (0.00)	-0.29 (0.37)
Caspian Sea river system	-0.19 (0.59)	0.03 (0.92)	
Sugar gouv.	1.40*** (0.00)	1.38*** (0.00)	1.08*** (0.00)
Mining gouv.	-0.52 (0.12)	-0.34 (0.28)	0.03 (0.97)
Agr Land Productivity			-0.00 (0.20)
Constant	5.44*** (0.00)	-0.89 (0.33)	5.53*** (0.00)
N	60	60	50
Adjusted R-squared	0.63	0.65	0.53

Notes: Robust p values in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Agricultural Land Productivity is expressed in grain and potato yield in kg per year.

Interestingly, locations near big cities did benefit from some urban agglomeration effects, as can be seen from the significance of the “near big city” variable. The transport infra-

structure variables, as expected for railways, did not exert a substantial positive effect. The Baltic/Black Sea river water system variable even had a negative effect, but its significance disappears as soon as agricultural productivity is controlled for (see below). The local availability of sugar beets, which provided ample possibilities to set up sugar processing firms (especially in the Ukraine) was very important. The finding that railways did not matter may be due to the timespan under consideration. In 1861, only 2000 km of railways had been built, which grew to 60000 km in 1905 (Thomas 1992), so its impact at the considered point in time might be too small to get captured by the regression.

If we restrict the sample to those gouvernements on which we have data on agricultural land productivity (proxied here with the typical harvest per hectare of grain and potatoes), we have only 50 cases left (Model 3). The opportunity costs of high agricultural productivity did not influence industrial plant creation in a significant way. The human capital, big city, and near-big-city effect are robust to the changes in the specification.

5.5 Conclusion

In general, we can explain the number of firms created before 1870 relatively well with our variables. What might be surprising is the degree to which educational differences decided about industrial firm creation in Russia, and this was not only a rural-urban effect. This could also be a hint that regional conditions actually mattered, and the industrial structure was not government-determined, as some older views reported (on the discussion see, Mironov 2003,

Gerschenkron 1978, Gregory 1982). Agglomeration effects that included also the neighboring gouvernements around the big cities were visible in mid-19th century. In contrast, transport infrastructure did not have a visible impact on firm creation rates, which might be due to the comparably new nature of the railway system.

References

A'Hearn, B., Baten, J. and D. Crayen (2009). Quantifying quantitative literacy: age heaping and the history of human capital. The Journal of Economic History 69 (3), 783-808.

Acs, Z., Florida, R. And S. Y. Lee (2004). Creativity and Entrepreneurship: A Regional Analysis of New Firm Foundation. Regional Studies 38 (8), 879-891.

Arrow, K. (1962), Economic welfare and the allocation of resources for invention, in R. Nelson (ed.), The Rate and Direction of Inventive Activity. Princeton: Princeton University Press.

Audretsch, D. and M. Vivarelli (1996). Determinants of New-Firm Startups in Italy. Empirica 23 (1), 91-105.

Baten, J. (2003). Creating firms for a new century: Determinants of firm creation around 1900. European Review of Economic History 7 (3), 301-329.

Baten, J., Spadavecchia, A., Streb, J. and S. Yin (2007). What made southwest German firms innovative around 1900? Assessing the importance of intra- and inter-industry externalities. Oxford Economic Papers 59 (1), 105-126.

Christaller, W. (1933). Die zentralen Orte in Süddeutschland. Eine ökonomisch-geographische Untersuchung über die Gesetzmäßigkeit der Verbreitung und Entwicklung der Siedlungen mit städtischer Funktion. Reprint Darmstadt, Wissenschaftliche Buchgesellschaft, 1980 (original: Jena 1933).

Gerschenkron, A. (1978). Agrarian Policies and Industrialization in Russia, 1861-1917, in Mathias, P. and Postan, M.M. (eds.), The Cambridge Economic History of Europe, Cambridge, vol. 6, part 2, 706-800.

Gregory, P. (1982). Russian National Income, 1885-1913. Cambridge: Cambridge University Press.

Holl, A. (2004). Transportation infrastructure, agglomeration economies, and firm birth: empirical evidence from Portugal. Journal of Regional Science 44 (4), 693-712.

Jacobs, J. (1970). The Economy of Cities. London: Penguin.

Leonard, C.S. and L. Borodkin (2000). The Rural Urban Wage Gap in the Industrialization of Russia, 1885-1913 2000, University of Oxford, Department of Economics, Economics Series Working Papers: 014.

Marshall. A., (1920). Principles of Economics. London: MacMillan.

Materialy (1903) vysochaishe uchrezhdennoi 16 noiabria 1901 g. Komissii po issledovaniiu voprosa o dvizhenii s 1861 g. po 1900 g. blagosostoiania sel'skogo naselenia srednezemle-del'cheskikh gubernii sravnitel'no s drugimi mestnostiami Rossii. S.-Petersburg: P. P. Soikin, 1903. Vol. 1., 176-177.

Mathur, V. (1999). Human Capital-Based Strategy for Regional Economic Development. Economics Development Quarterly 13 (3), 203-216.

Matthäi, F. (1873). Die Industrie Russlands in ihrer bisherigen Entwickelung und in ihrem gegenwärtigen Zustande mit besonderer Berücksichtigung der allgemeinen russischen Manufaktur-Ausstellung im Jahre 1870. Leipzig: Verlag H. Fries, Vol. 2.

Mironov, B. (1999). The Biological Standard of Living in Russia. Conference Paper, presented at the IEHA pre-conference, Munich.

Mironov, B. (2003). Russia: Modern Period, in Oxford Encyclopedia of Economic History, Oxford University Press, 417-422.

Romer, P. M. (1986). Increasing returns and long-run growth. Journal of Political Economy 94 (5), 1002-37.

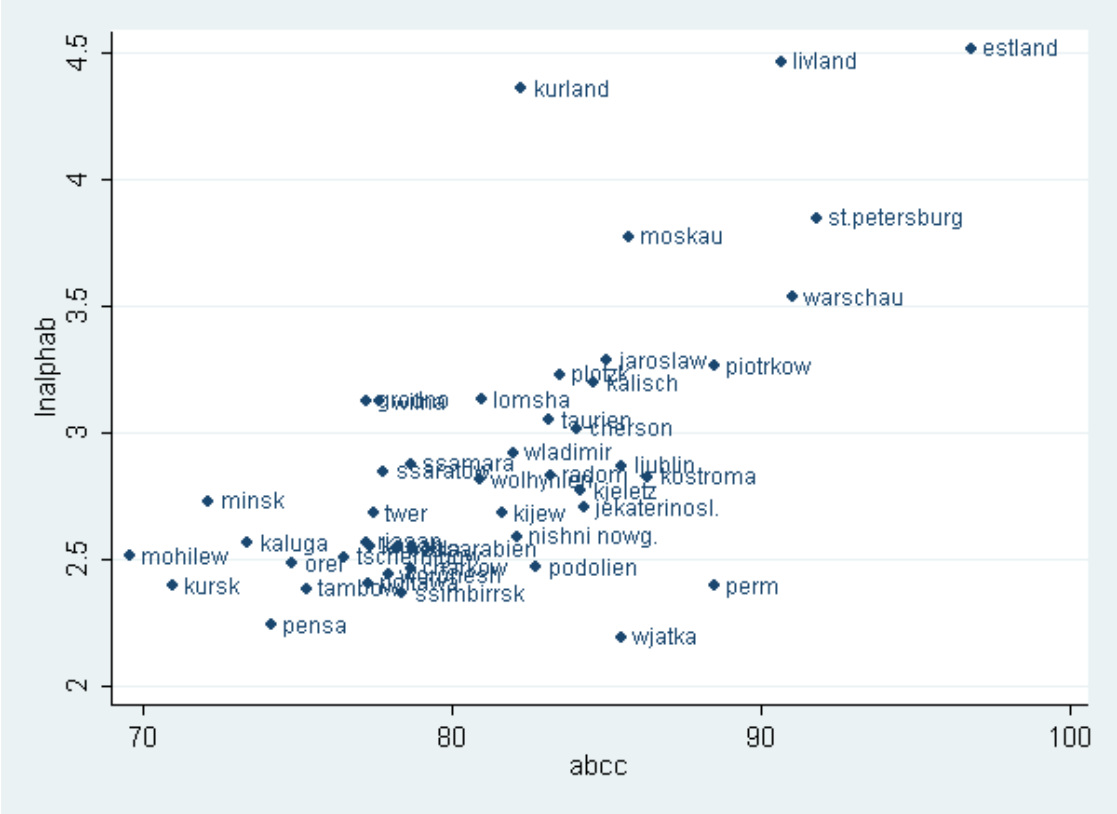
Scherner, J. (2001), Eliten und wirtschaftliche Entwicklung – Kongreßpolen und Spanien im 19. Jahrhundert. Münster: Lit Verlag.

Thomas, L. (1992). Rußland im Jahre 1900. Die Gesellschaft vor der Revolution. Inaugural lecture. 13.10.1992. http://dochost.rz.hu-berlin.de/humboldt-vl/thomas-ludmila/PDF/Thomas.pdf, 6.07.2007.

Wolf, N. (2006). Local Comparative Advantage: Agriculture and Economic Development in Poland 1870 – 1970. Diskussionsbeiträge des Fachbereichs Wirtschaftswissenschaft der Freien Universität Berlin.

Appendix

Figure A1. Comparison of Whipple Index and Literacy



Source: own calculations.

6 On the Determinants of a successful State: Good Governance between the 1850s and the 1980s.

Abstract:

This paper estimates the quality of governments. Our measure of quality is the efficiency with which a country transforms a given number of endowments, such as capital, labor, and land, into general welfare. We use both an anthropometric yardstick, namely adult male height, and a monetary one (per-capita GDP) in order to capture different kinds of human well-being. We use the Data Envelopment Analysis (DEA) in order to calculate efficiency values for 62 countries on a decadal basis between the 1850s and the 1980s.[44]

Our results suggest that generally efficiency increased among industrialized countries. Both Latin American and African economies start on a relatively high level but show a systematic decline in the course of the 20th century. In Africa, former British colonies have done worse compared to their counterparts under French rule. In many countries we find wars, especially occupation, violence and political instability to have a remarkably negative and sometimes permanent influence on efficiency, as in the cases of Spain, Greece and Colombia. Typical characteristics of successful states in this regard seem to be monocultures in the agricultural sector, redistribution and a homogeneous population.

This Chapter is based on a working paper with the same title that was written together with Matthias Blum and Luis Huergo. Matthias Blum and I developed the idea. Luis Huergo implemented the bootstrapping algorithm and provided methodological input. Writing was equally shared between me and Matthias Blum.

[44] We thank Scott Baier, Jörg Baten, Dorothee Crayen, Jerry Dwyer, Jari Eloranta, Kerstin Enflo, Jakob Madsen, Christian Morrisson, Fabrice Murtin and Robert Tamura for providing their data. In addition, we thank Jörg Baten and Bas van Leeuwen for valuable comments on an earlier draft of this paper.

6.1 Introduction

Governments strongly intervene in economic activity and they have a strong influence on the welfare of their populations. During the twentieth century, public spending ratios increased in many countries, reaching levels of about 40% in the European Union (Joumard, Kongsrud, Nam and Price 2003). Among others, this includes social spending, health care systems, and education. The huge amounts that are spent by governments and the strong influence they exert on the whole economy point out that it is a question of central political importance that governments use their resources efficiently to enhance prosperity of the people. One of the reasons for the low welfare level in many developing countries, even in those that have rich natural endowments, is bad governance (Sachs and Warner 2001; Gylfason 2001).

This paper aims at measuring government efficiency over a long time span for many countries. The empirical literature about government efficiency is concerned with several questions. The first one refers to the development and use of indices that are comparable across countries and time periods. Probably the most popular approach are GDP-estimations, but especially economic historians have become quite innovative in finding alternatives, e.g. anthropometric indicators such as height, weight or mortality, in order to go back further in time. Another branch of research tries to categorize countries according to their welfare regimes (Herrmann, Tausch, Heshmati, Bajalan 2008). The idea behind this approach is to distinguish between different kinds of social states, e.g. the Scandinavian Model versus the models of the new EU member states, using a factor analysis or index number approach. Studies that employ this methodology combine different outcome variables into few numbers to distinguish between countries. For example Denmark and Ireland have high expenditure for housing as a direct action against social exclusion, while the Netherlands show low state expenditure on social security systems but high individual expenditure on insurance.

A third field of literature is concerned with the measurement of efficiency of governments or institutions. There are two principal approaches. The first defines good institutions (or an efficient government) according to theoretical considerations or according to factors

that were found to foster economic growth. Examples of such institutions are the protection of property rights, the degree of democracy, the rule of law or a market economy (Kaufmann, Kraay and Mastruzzi 2008; Kaufmann and Kraay 2008). This literature also tries to identify factors that lead to better institutions. In this regard, Kaufmann, Kraay and Mastruzzi (2008) find that ethnic heterogeneity, a more socialist regime or centralized religions, such as Catholicism or the Islam, lead to worse institutions.

The second approach, which we will follow here, measures government efficiency by estimating the relationship between the resources that a country possesses (preconditions or 'inputs') and the actually achieved level of welfare ('output'). The higher the amount of output a country achieves with a constant quantity of inputs the more successful this country is in terms of efficiency. The idea behind this approach is not only to count the money spent on welfare or the achieved average welfare level, but rather to look at the process involved in converting inputs into outputs. In other words: we do not want to measure general welfare or answer the question which country was the wealthiest, but to see which country used its available resources reasonably and which did not.

The question of efficiency is of central importance as a more efficient use of endowments releases resources that can be used elsewhere in order to increase welfare. The European Union is strongly encouraging performance comparisons between its member states in hope that negative results will motivate countries to find ways to increase their performance (see Pochet 2005 for more information on the so called Open Market Competition method). Hence, and especially due to data limitations, many studies cover the OECD countries after 1990; only a few studies look at developing countries (Rayp and Van de Sijpe 2007) or historical periods (Henderson and Zelenyuk (2007) go back until 1965). There are two possible ways to estimate efficiency. One possibility is to regress welfare on the inputs and evaluate the residuals. Positive residuals indicate high governmental efficiency, while strongly negative residuals show low levels of efficiency (WHO 1999). The second approach, which we use in this paper, uses non-parametric frontier models to estimate efficiency with multiple outputs (see Afonso, Schuknecht and Tanzi 2008 for an example). This paper expands the existing literature by going back in time even before the onset of the modern social security sys-

tems. This is one central contribution of this paper to the debate on the development of government performance and its determinants. We calculate efficiency values for 62 countries between the 1850s and the 1980s on a decadal basis in order to get an impression of how governmental performance evolved over time. For more modern times, we are also able to include some countries that are not taken into account by the existing literature, like former African and Asian colonies along with Latin America during the 20th century.

After estimating efficiency with the frontier models (the Data Envelopment Analysis or 'DEA'), we identify determinants of the efficiency values using regression techniques. Simar and Wilson (2007) show that conventional approaches are not valid due to the bounded nature of the efficiency scores (they cannot increase over 100%).

To overcome this problem they propose a bootstrapping-algorithm that leads to consistent results, and we will employ this approach here. Therefore, we are able to provide reliable regression results. This has rarely been the case in the literature so far, as pointed out by Simar and Wilson.

The paper develops as follows. First we will explain the methodology in more detail. We then discuss the choice and sources of our variables we use in the first and second stage. The next section shows long-run efficiency development. Afterwards we discuss the factors affecting government efficiency. The last section concludes.

6.2 Methodology

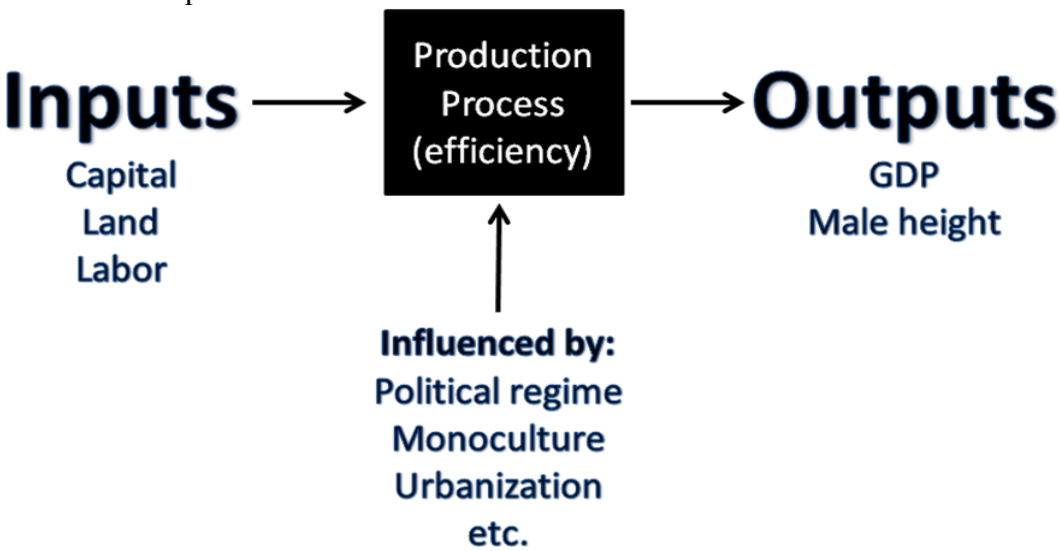
Most authors who aim at measuring government efficiency apply regression techniques. These methods regress output - e.g. life expectancy - on its determinants (inputs, e.g. GDP/c; see Gerring, Thacker, Enikopolov and Arévalo 2008).

The resulting residuals of these regressions can be interpreted as efficiency. However, the most significant disadvantage is that the measurement of governmental success is restric-

ted on one single yardstick. The Data Envelopment Analysis (DEA) overcomes this problem and allows the inclusion of multiple inputs in combination with multiple outputs.

We use it, as it is a flexible method that requires only few assumptions about the production process and can handle inputs and outputs for which no price information is needed. Therefore, we are able to use several outcome variables and reduce the risk of obtaining results that depend severely on the choice of one single outcome variable. We consider more than one measure of welfare, namely a GDP-oriented and an anthropometric proxy for well-being. We first use the DEA to estimate the efficiency of the welfare production process of single countries over time. In the second stage, we will use regressions to assess the determinants of the measured government efficiency (see figure 1). To avoid biases due to the truncated character of the distribution of our efficiency values, we apply truncated regressions suggested by Simar and Wilson (2007).

Figure 1: the model setup.



Source: own figure.

The DEA assumes the existence of a concave production frontier, against which all observations (called 'decision-making units' - DMU) are evaluated. It is important to note that the DEA offers the choice between two different estimation techniques. The first one is based

on the assumption that - ceteris paribus - every additional unit of input increases the output level at the same rate. The second one - which we apply - recognizes the law of diminishing marginal utility and allows for variable returns to scale, so that additional inputs might increase outputs at a diminishing rate. We do not have to specify a functional form a priori, which reduces the risk of getting spurious results which rely on the assumed production function. It is possible to specify the DEA models from an input- and an output-oriented perspective. In an input-oriented model the efficiency score shows the proportion to which all inputs have to be reduced so that the DMU is on the efficiency frontier. In the output-oriented model the score shows by which proportion the outputs have to be increased to reach the frontier.[45] The production frontier is a measure of what would be possible with the technology at hand (which is determined by all DMUs in the sample). Since we investigate internal processes of countries which influence its outputs at a given number of inputs, it is useful to apply the output-oriented model.

One basic DEA-model for *n* DMUs can be represented by the following linear program (Cooper, Seiford and Tone 2006).

$$\max_{\mu,\nu} \theta = \mu_1 y_{1o} + \ldots + \mu_s y_{so}$$

subject to

$$\nu_{1o} x_{1o} + \ldots + \nu_m x_{mo} = 1$$

$$\mu_1 y_{1j} + \ldots + \mu_s y_{sj} \leq \nu_1 x_{1j} + \ldots + \nu_m x_{mj} (j = 1,\ldots,n)$$

$$\nu_1, \nu_2, \ldots, \nu_m \geq 0$$

$$\mu_1, \mu_2, \ldots, \mu_s \geq 0$$

Here θ denotes the efficiency score, μ refers to the weights for the *s* outputs, and *y* the corresponding output; ν describes the weights for the m inputs, while *x* indicates the inputs.

It is possible to assume constant or variable returns to scale amongst others. In the case of a welfare-producing process, the first unit of input (e.g. capital, land) has a larger positive

[45] The two models produce identical results if constant returns to scale are assumed, while for variable returns to scale the results may differ.

impact than the second (see Fernald 1999), so we will employ a variable returns to scale model.

DEA estimates may be biased upwards, as missing DMUs sometimes lower the production frontier and therefore the efficiency scores of the other DMUs are higher than they should be (see Simar and Wilson 2000). Additionally, as each efficiency score depends on the other observations the error terms in any second stage regression are correlated. While the regression is still consistent, standard errors no longer decrease at the normal rate of $\sqrt{n}$, but on a much slower rate instead. Therefore, standard regression approaches with few observations are not reliable.[46] We deal with this problem by using the bootstrapping algorithm proposed by Simar and Wilson (2007).[47] They suggest using a parametric bootstrapping algorithm in the truncated regression model and show that this procedure leads to more reliable results than the Tobit approach that is commonly employed in the literature. Appendix A gives a more detailed explanation of this method.

Ravallion (2005) raises a number of concerns about the existing literature. One point is that the reasons for the choice of variables for the first or second stage are often not clear. We will discuss our variables and the idea behind them below. It is also claimed that with an insufficient number of observations, this method is likely to produce erroneous results. We cannot entirely eliminate this problem here, but using only few inputs and two representative outputs in the first stage (which gives us a comparably large number of observations) and the bootstrapping procedure in the second stage we are able to provide more reliable estimates than previous papers do.

There are two possible alternatives to the DEA model (see Coelli et al 1998 and Rayp and Van de Sijpe 2007). The first is the stochastic frontier (SF) model that allows for noise. In DEA all deviation from the production frontier is considered as inefficiency, so the scores are vulnerable to measurement error. Even though this is an issue when applying the technique to

---

[46] Simar and Wilson (2007) show the rate to be $n^{\frac{-2}{(p+q+1)}}$, p being the number of inputs and q the number of outputs. Even for a small model with 3 inputs and 2 outputs and rate is $n^{-\frac{1}{3}}$, and it decreases with more variables being added.

[47] Estimations were performed using R and the FEAR package by Paul W. Wilson.

historical data, advantages of the DEA outweigh this disadvantage: The SF only outperforms the DEA when the assumed functional form is close to the actual one. Unfortunately we could only guess the functional form of the production function between inputs and outputs making it unlikely to obtain reliable results.

A second alternative would be the free disposal hull model (FDH, see Deprins, Simar and Tuskens 1984 who introduced this technique), which uses only existing DMUs to evaluate the present observation. The DEA uses the whole frontier, thereby also using linear combinations[48] of existing DMUs. We do not see any reasons for the assumption that only some levels of welfare are attainable while others in between are not. Welfare usually increases gradually and does not jump from one level to another. Several scholars who simulated and experimented with the existing techniques proposed the use of the Data Envelopment Analysis (Cooper and Tone 1997; Resti 2000; Banker, Chang and Cooper 2004).

6.3 Data

In this section we discuss the inputs and outputs that are used in the DEA-estimation. The regressors for the second stage will be discussed below.

Macroeconomic Input Measures:

Every single component of an economy can be classified and aggregated into three categories (Feenstra 2010): capital - including physical and human capital (Mankiw, Romer and Weil 1992) – the number of workers, and land. This is the most common level of aggregation, and therefore we follow this scheme. These are the resources a country's economy has to deal with and it is the efficient use of them what makes an economy successful.

[48] If we observe country A with a life expectancy of 76 years and a GDP of 12000 $ and country B with a life expectancy of 68 years and a GDP of 11000 $, a linear combination could be for example country C with a life expectancy of 72 years and a GDP of 11500 $.

Especially in the early period included here, one of the most important factors of production is the area of arable land. Mitchell (1993) provides information on the area of arable land used for the cultivation of crops. Furthermore, he offers information on the number of livestock available by country and year. Both are elementary components of an economy's primary sector and are therefore included into our analysis. We include both the land available for farming and the number of cattle per capita.[49] The latter serves as a proxy for farmland used for the production of animal products (Moradi and Baten 2005, Moradi 2005, Koepke and Baten 2008). This is an essential complement to Mitchell's arable land estimations, since livestock is a valuable supplier of animal proteins. In the course of economic development the consumption of animal proteins increases leading to improved nutrition and taller adult stature (Grigg 1995, Steckel 1995). In the case of land abundant countries, the export of animal proteins may even enable economic growth in terms of GDP/c (Jonsson 1998).

Moreover, we use Baier, Dwyer and Tamura's (2007) estimates of real[50] physical capital stock in order to include another important cornerstone of an economy into our analysis. In their study they apply the perpetual inventory method to use investment rates in order to estimate the stock of physical capital per worker. They provide estimations on capital endowment on a decadal basis for 155 modern countries starting as early as 1830.[51]

In order to include comparable data on human capital, we also avail ourselves of estimations provided by Baier, Dwyer and Tamura (2007). We chose their measure of human capital (per worker) not only because of its broad and comprehensive character, but also because it covers both education in schools, universities and working experience. This is an important source of practical knowledge which is often neglected by the literature. They

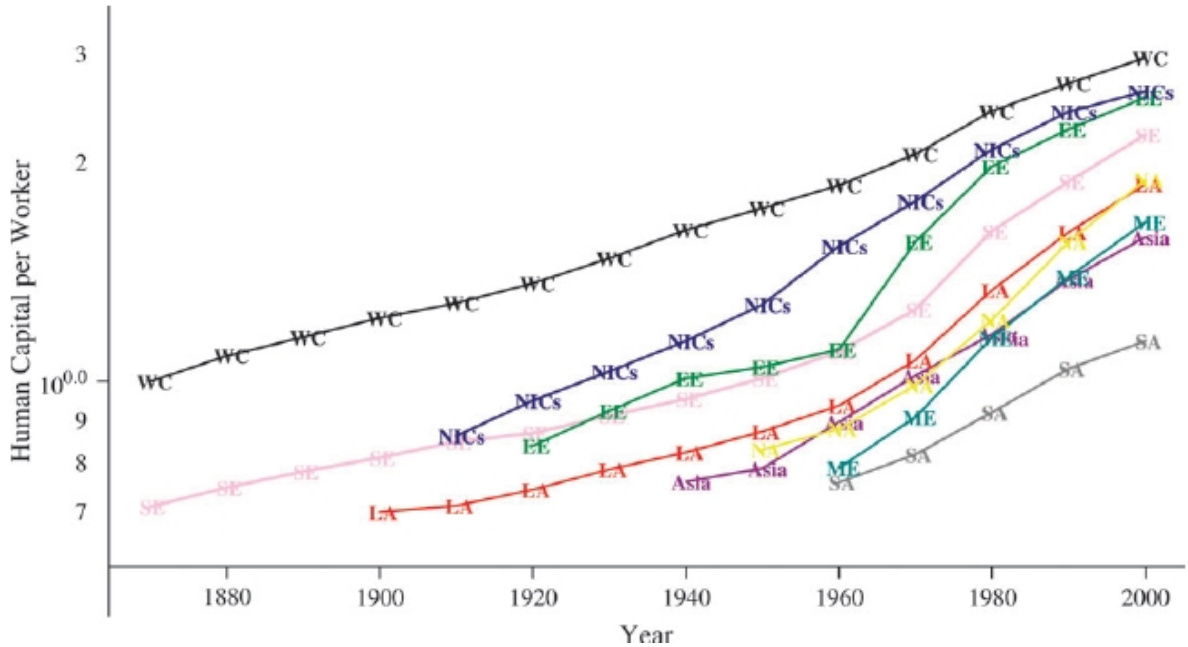
Figure 2: Human Capital Development by regions.

[49] Data come from Mitchell (1993) as well as Baten and Blum (2010) who compiled additional statistics on cattle stocks.

[50] They use Purchasing Power Parity exchange rates from Summers and Heston to convert the estimates into international dollars, see their footnote 3.

[51] There are different estimations for capital and human capital available. Appendix B compares the different estimates.

FIGURE 5
Human Capital



Source: Baier, Dwyer, Tamura (2006), Figure 5.

measure the actual stock of human capital in the population by decade and do not only look at those currently being in education.[52] By including both of these measures Baier, Dwyer and Tamura avoid one problem that sometimes appears when using basic indicators like school enrollment rates. Often the values of these yardsticks are limited to 100 per cent which leads to insufficient increases in human capital estimations during advanced development stages. Figure 2 shows that even during the second half of the 20th century, the human capital stock is still rising without being limited by an upper bound. Our panel consists mainly of rather developed countries, where populations have already reached a certain level on human capital.

As a last step before running our analysis we need to qualify all of the above mentioned endowments by including the current number of inhabitants (World Bank 1999 and Maddison 2001). Only by including information on the **size of the population** we are able to

[52] As we use the stock of human capital and not the investments, which transfer into welfare with a lag, we also do not use lags in the estimations.

compare international data. Hence, livestock, capital, and arable land are always expressed in per capita figures.

A common concern in this discussion is the direction of causation (Holtz-Eakin 1994, p.13). He states that the positive correlation between the prosperity of a state and its investments in public capital might occur just because rich countries are able to invest and poor countries are not. In this study this problem does not play a role, since we analyze the quality of governance at any given instant and do not want to make statements about its inter-temporal role and its influence on future development.

Macroeconomic output measures:

One of the advantages of the DEA methodology is that it is possible to define more than one target variable. In this paper, we want to measure the success of a country's policy and therefore we need one or more proxies that are able to perform this task. Probably the most popular measure of a successful (economic) policy is the national income. As an absolute number it indicates a nation's economic weight and power, as a per-capita-measure it serves as an indicator for the level of economic development. In order to allow a comparison we use Maddison's (2001) GDP/c estimates as one target variable.[53]

Since there are a number of periods in the history of mankind that make clear that measures based on purchasing power do not fully describe a population's welfare and GDP per capita values do not necessarily mean a high standard of living, we need an alternative measure of human well-being. Among others, Sen (1999) and Haq (1995) argue convincingly that only focusing on economic growth and material well-being might be too narrow. In this sense, Inglehart et al. (2008) state that economic development is only one element of happiness. Since the use of the corresponding measures - happiness or the Human Development Index - is rather problematic in economic history, we have to apply another alternative yardstick in order to measure human well-being. Several scholars have highlighted that adult stature

[53] As our model allows for variable returns to scale it is not necessary to use GDP in logs.

serves as an excellent indicator of this kind (Komlos 1985; Steckel 1995, 2009; Baten 1999; Komlos and Baten 2004). Among others, the so called 'biological standard of living' is correlated with high-quality nutrition (positively) and the disease environment (negatively). One of the most convincing examples can be found in 20th century Germany. In order to get prepared for war, Nazi-Germany increased its military expenditures at the expense of public health measures. In addition, food imports were impended and prices were partly under state control. Baten and Wagner (2003) report that during this period, especially in highly developed urban areas and regions in northern Germany near the coast, heights decreased due to their dependence on foreign trade. Height values are adopted from the recent study by Baten and Blum (2010).[54]

Our view is that by including both of these measures as our target variables we are able to evaluate a nation's welfare and in combination with the applied DEA methodology the success or failure of economic policy.

6.4 Results – the Development of Government Performance

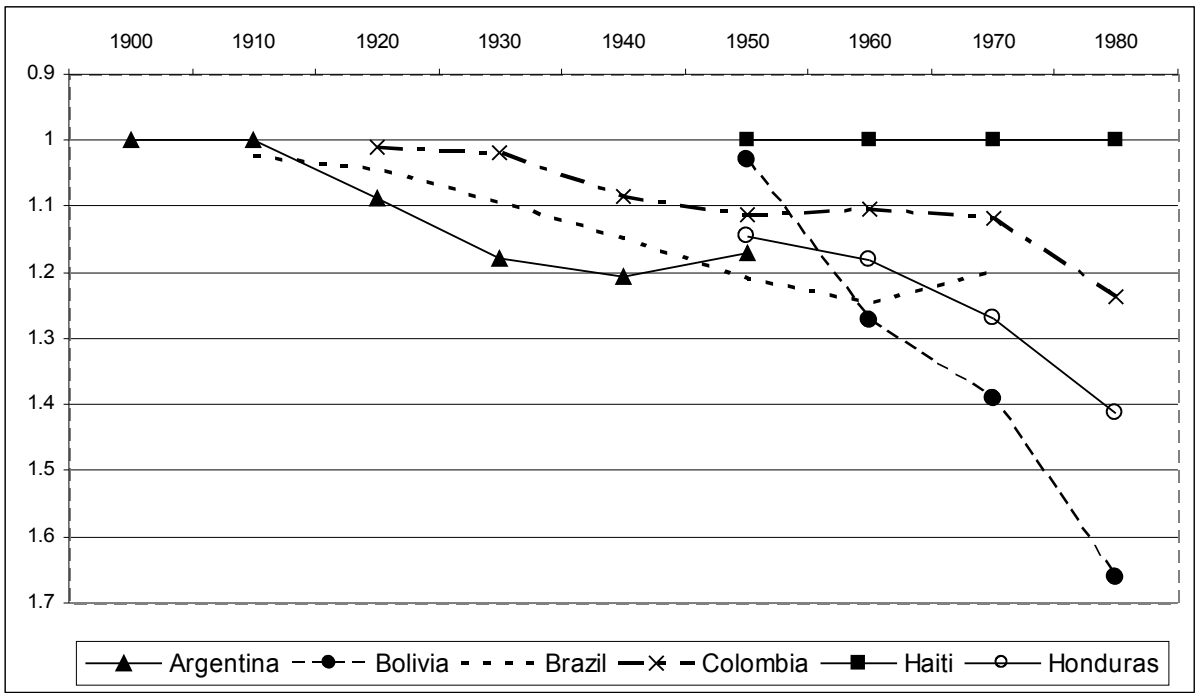
In general, government efficiency was increasing during the period under observation among industrialized countries. Figures 3, 4 and 5 give an overview of efficiency trends among selected countries, the other trends are shown in appendix C. The trends of some countries were interrupted by crises or wars.[55] Latin American as well as African countries experienced a continuous decline during the 20th century. Looking at the trends, Latin America (including the Caribbean) shows a relatively uniform picture. The most successful country is Haiti, reaching 100 percent efficiency between the 1950s and 1980s. Haiti was certainly poorer than most other countries, but given the even lower endowments of physical and human capital, its effi-

[54] We prefer to use male heights, since they are widely available due to conscription, convict lists and anthropological surveys.

[55] The impact of wars often starts before its formal declaration and may continue after the peace treaty. Therefore, we can not simply control for war periods in the regression analysis by including dummies for those periods. We therefore prefer to have a detailed look at the efficiency trends.

ciency was surprisingly good. The other countries in this region were at best stagnating. Mexico's government efficiency decreased after the 1920s. This is true even after the end of the Mexican civil war probably due to the ongoing interior conflicts and the Great Depression. Only after the 1940s did Mexico at least manage to stop this downfall. Peru appears to be very successful until the 1930s. Not even the Colombian-Peruvian War in 1932-33 changed the exemplary figures. Peru's decline of government efficiency started in the 1950s after the

Figure 3: Governmental Performance in Latin America 1900s to 1980s (selection)

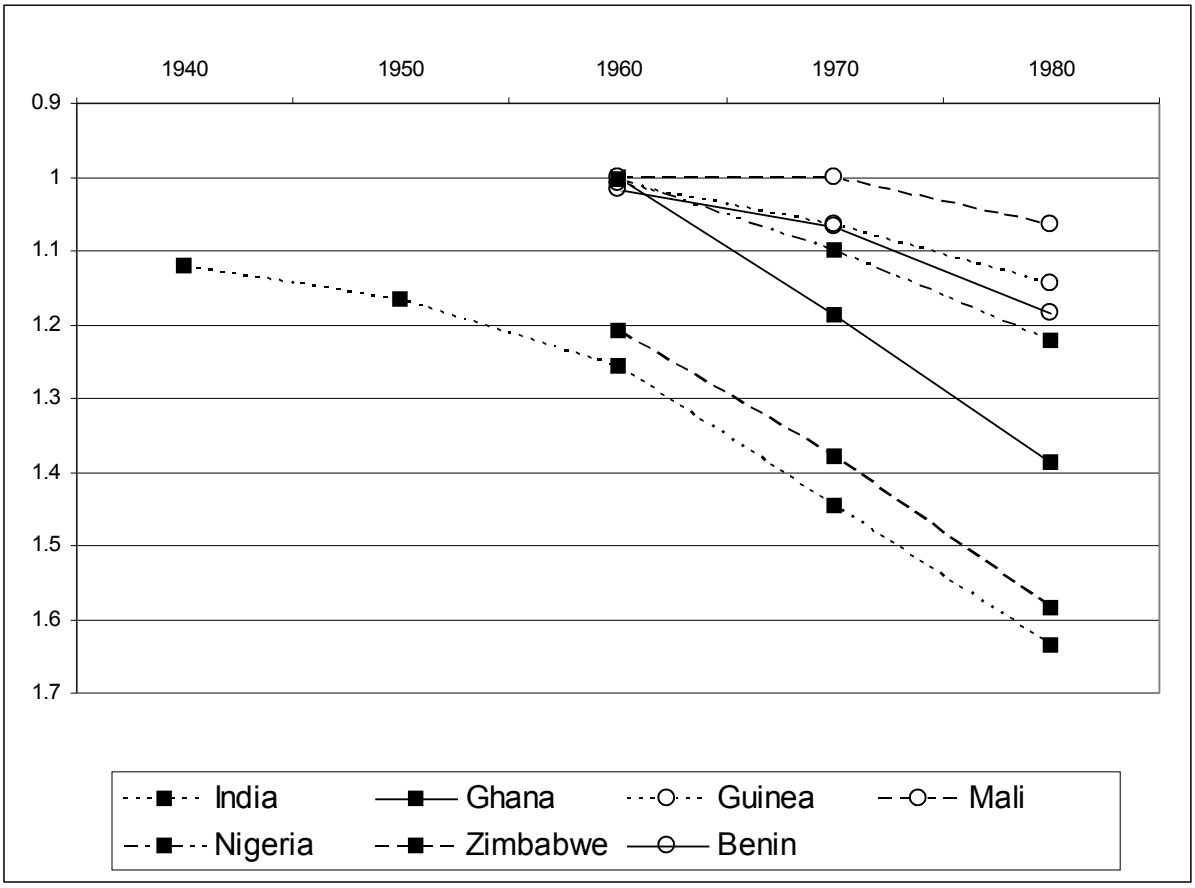


Note: A value of 1 means a fully efficient country. Higher values indicate a worse performance.

military coup in 1947. Until the 1960s, it only stagnated on a lower level but dropped tremendously in the 1970s and 1980s after the junta took over and conducted economic experiments, such as the nationalization of Peru's oil deposits. The efficiency trends of Colombia and Brazil suggest a development similar to Peru. Brazil's governmental performance also started on a remarkably high level in the 1910s and was declining over more than half a century until the 1960s. Beginning in the 1940s with a civil war, the so-called 'La Violencia', Colombia experienced a permanent downturn in governmental efficiency until the 1980s. The

most dramatic decline of government performance can be observed in Bolivia, Honduras and the Dominican Republic. After the nationalization of some national resources and a land reform in the 1950s Bolivia's efficiency fell dramatically. During the whole period between the 1950s and the 1980s Bolivia's performance dropped sharply.

Figure 4: Governmental Performance in former colonies 1850s to 1980s (selection)



Note: A value of 1 means a fully efficient country. Higher values indicate a worse performance.

For Sub-Saharan countries data are only available for the period between the 1950s and the 1980s. Apart from Rwanda, all countries show a permanent and sometimes an even dramatic decline in government performance. Rwanda started at a relatively high efficiency level in the 1960s, but experienced a drop in efficiency in the 1970s. In the 1980s Rwanda's efficiency rose again and reached an optimal level. The rest of the countries in this world re-

gion did not do equally well. Starting from different base levels in the 1950s and 1960s every single of the remaining countries in Sub-Saharan Africa experienced a steady decline.

Among the countries of northern African and the Middle East the results suggest that Egypt has the highest efficiency. Beginning in the 1950s its value is already on an extremely high level and it reached an optimal level in the decades afterwards. Morocco and Turkey did not do as well. In the decades after the Second World War Turkey experienced a continuous downturn of government efficiency despite the political, social and economic reforms by Mustafa Kemal Pasha. In Morocco, the influences of the Algerian war of independence and the Moroccan independence from France left their traces in the efficiency trend. Between the 1940s and the 1980s Moroccan governmental performance decreased constantly after it became a sovereign state. Iran's efficiency values were already on a high level in the 1960s and even increased further in the 1970s. The Islamic Revolution and the Iran-Iraq war, however, caused a sharp decline in the 1980s.

In Asia, we can trace Korean (1940s, the Republic of Korea from then on) government performance from the 1940s to the 1980s. (South) Korea's efficiency values reach a perfect 1 even during the Japanese occupation and the devastating events of the Korean War. Only after the military coup and reign of Park Chung-hee in the 1960s did Korea experience a temporary decline. Another interesting case study in this regard is India. India's efficiency at the time of independence was relatively high, but during the following years India was preoccupied with restructuring its economy according to socialist doctrine, and its involvement in several wars. Therefore, it is no surprise that India's efficiency values declined dramatically in the subsequent decades. In this regard, India can be added to a number of (mainly African) examples when independence caused major turmoil and a declining government performance.

Among the industrialized countries Denmark serves as an exemplary country. Its efficiency value stagnates on a low level until the turn of the twentieth century. From then on, it increases moderately, but is disrupted by the dramatic events of the Second World War. Immediately after the war, its governmental performance improves to a perfect 1.

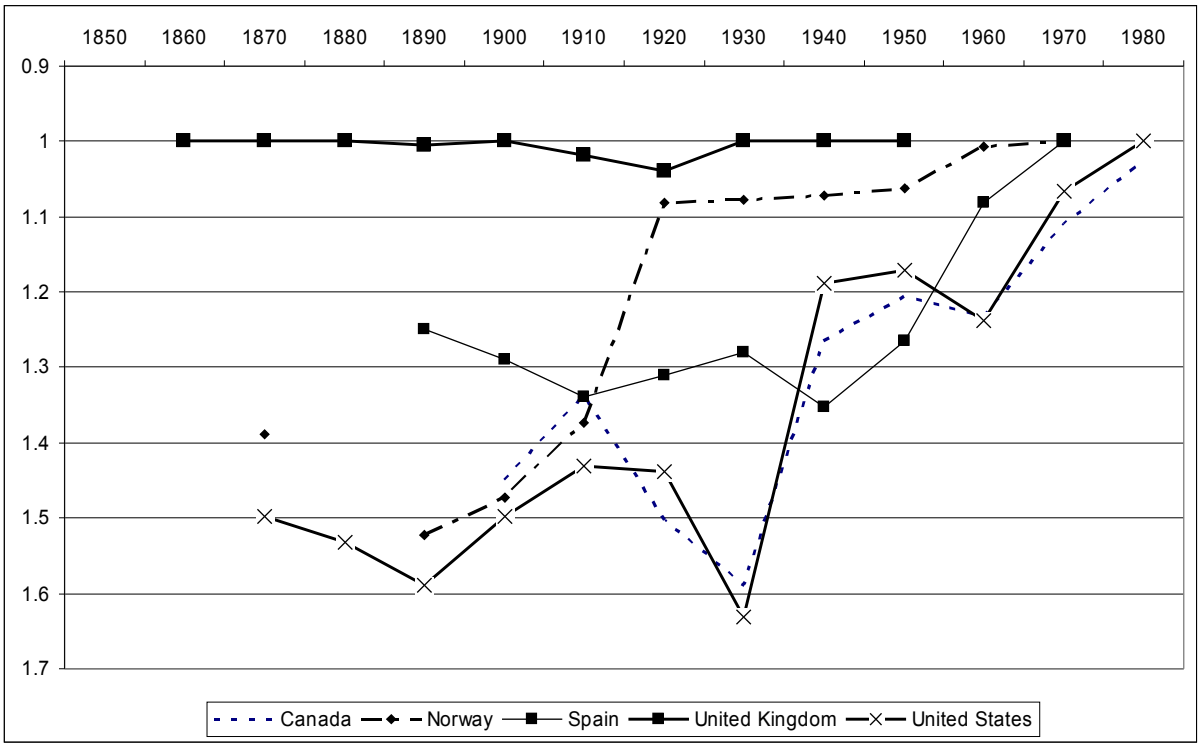
This pattern can be observed in a number of countries. The overall Spanish efficiency trend, for instance, fluctuates moderately on a rather low level and declines remarkably in the

aftermath of the Spanish civil war and during the first years of Franco's authoritarian regime. Although Portuguese governmental performance was higher than the Spanish one, it fluctuated during the time of the Republic, after the military coup and the first years of Salazar's reign. Beginning in the 1950s, both Spain's and Portugal's government efficiency increase. Basically the same is true for Finland, the Netherlands, Italy, Norway, Sweden, Canada and the United States. Their efficiency values move at a low level until the 1930s and the 1940s. This period seems to be a crucial point in the economic history of those countries. From then on, negative trends turn to the better or even make giant leaps forward. Of course, there are differences regarding the exact point in time or the initial level of efficiency, but the general pattern of those countries' development of governmental performance is obvious.

The trends of Japan and the United Kingdom could serve as a model for efficiency. Our estimation reveals perfect values in almost every single decade for the UK, for which we have a continuous trend from the 1860s to the 1950s. The same is true for Japan, although we can not trace Japan’s efficiency through the time between the 1940s and the 1960s.

Furthermore, the consequences of the First World War and the Great Depression cannot be ignored. Especially Canada, the U.S., but also Germany and Greece experience a remarkable decline of efficiency during this period. Surprisingly, the Second World War does not seem to influence the trends among the industrialized countries. Only Denmark and Greece seem to suffer from the German occupation in the 1940s. During these years Denmark's governmental efficiency was remarkably lower compared to the previous years. Greece's governmental performance begins to decline in the aftermath of the Greco-Turkish war and reaches a low during the German occupation.

Figure 5: Governmental Performance in industrialized countries 1850s to 1980s (selection)



Note: A value of 1 means a fully efficient country. Higher values indicate a worse performance.

It is also interesting to note that former French colonies tended to do better both in terms of their initial level of government efficiency and the severity of the decline after independence, as is visible in the graphs.[56]

Possibly the most surprising result is that Canada and the United States belong to the most inefficient economies until the mid-twentieth century. Although they are often considered as flagships of modern and efficient economies we find that their performance in terms of efficiency is rather disappointing. Of course, their overall wealth was remarkable and still is, but it seems that they did not always use their enormous potential efficiently. By looking at the trends one can also see the close links between Canada and the U.S. Both countries show an almost identical path of efficiency. They share low levels of efficiency until the Second World War and a common increase during the 1940s as well as a slow increase during

[56] Due to the small number of cases an econometric test is not possible.

the Cold War. The "New Deal" and the high war-induced demand had a positive impact on those economies.[57]

6.5 How did we distinguish Inputs, Outputs and explanatory Variables?

At first sight, it is not entirely clear which variables should be considered as inputs for an economy and which ones should serve as explanatory variables in order to explain governmental efficiency in the second step.

We use several criteria to draw a line between inputs and explanatory variables. Classical inputs or preconditions are physical and human capital, land and labor force according to economic theory. The explanatory variables can also be divided into two sub-groups. The criterion for this is whether the variable can be influenced by government policy in the short or medium run – such as military burden, the degree of democracy and social spending - or if it has to be considered as a country's characteristic, such as the heterogeneity of the population.

In contrast, the definition of outputs is clear. In general, the success of a country is often connected to its monetary wealth and the welfare its population enjoys. Hence, we use per-capita GDP and mean adult (male) height as our measures of success.

6.6 Data used in the Second Step – Determinants of Efficiency

One of the most important influences for government efficiency is the **heterogeneity of the population**. Different ethnic groups are often separated by different languages and/or religious beliefs and often even have an unsightly common history. In some cases - Switzerland is probably the most successful example - heterogeneity is nothing more than cultural diversity and several minorities complement each other. Unfortunately, this is not always the

[57] This result is robust to all specifications and does not depend on a particular variable.

case. Cyffer (2001) reports that even a taxi ride from one outlying district of an African capital to another may require skills in several languages, just because not everyone speaks a common standard language. Hence, one has to be able to speak the local African language in order to communicate with the local population. In the northern parts of Nigeria, English played almost no role at least until the 1970s (Cyffer 1977). In sum, Cyffer reports that there are almost 400 different languages in Nigeria only. One can imagine how hard it is for the central government to implement working institutional structures in every province of the country. Moreover, linguistic fractionalization may exclude minorities from education just because one part of the population does not speak the official language used in school or university. It may also make it easy to identify a member of an ethnic group and therefore lead to discrimination or nepotism.

Easterly and Levine (1997) find that ethnic diversity helps to explain political instability, underdeveloped financial systems, distorted foreign exchange markets, high government deficits and insufficient infrastructure. Mauro (1995, 1998) even goes one step further: he uses an index of fractionalization as an instrumental variable for corruption and finds that it goes along with lower investments, lower economic growth and lower government spending on education and may lead to adverse budgetary consequences due to tax evasion. We follow their example and use the ELF Index as an indicator for the heterogeneity of a population.

The ELF is probably the most common index of heterogeneity. It is computed as one minus the Herfindhal-Hirschman Index (HHI) of ethno-linguistic group shares and it can be interpreted as the probability that two randomly selected persons in one country belong to different ethno-linguistic groups. It is constructed from data gathered and published by Soviet anthropologists in the 1960s (Atlas Narodov Mira 1964, Taylor and Hudson 1972) and it therefore reflects a country's ethnic composition during that time. The most common critique regarding the ELF is that it is sensitive to the definition of an "ethnic group". There is a continuous transition between different languages and accents of the same language, which makes it difficult to standardize and measure ethnic differences around the globe comparably and to control for the importance of ethnic differences. Moreover, it neither includes religious

nor racial differences. In order to circumvent these problems we also apply other indicators in alternative models, which are in principal, however, all based on the same idea.

The aim of Fearon's (2003) strategy was to construct a list of ethnic groups mentioned in the literature.[58] This list served as a basis to distinguish ethnic and cultural groups and to take the above mentioned methodological problems into account. He only includes an ethnic group in case it makes up at least 1% of a country's population in order to control for the relative importance of ethnic differences. His indexes obtain similar results than the calculations based on the Atlas Narodov Mira (correlation of 0.75). Only in North Africa/ Middle East and Latin America/ Caribbean he obtains systematically higher estimations of ethnic fractionalization.

Posner (2004), however, argues that ethnic, cultural or religion-based indicators are inappropriate to capture economic or political rivalries within one country because they do not take actual political engagement and competition into account. Instead, he suggests using his index instead – the PREG (Politically Relevant Ethnic Groups) – which is based only on politically relevant ethnic groups. Unfortunately, his index is limited to 42 African countries and is therefore unsuitable for our analysis.

Alesina et al. (2003) provide estimations on about 190 countries. They compute separate indexes of religious, ethnic and linguistic fractionalization and find that their estimations of ethnic and cultural diversity are similar to the ELF. Their measure of religious heterogeneity, however, shows almost no correlation (0.372) to the ELF. Therefore, this yardstick serves as a valuable alternative to investigate governmental efficiency.

Urbanization goes along with many different changes and is therefore only the symptom of a complex process in an economy. On the one hand, cities can have negative effects on the welfare of its citizens because transportation and transaction costs separate them from food growing areas. Especially among economic historians transportation plays a critical role regarding the so called 'urban penalty' (Gould 1998, Woods 2003, Martínez-Carrión and Moreno-Lázaro 2007, Baten 2009). In contrast to cities, people in rural areas enjoy food prox-

[58] Fearon's most important sources to identify ethnic groups were the CIA's *World Factbook*, *the Encyclopedia Britannica*, the *Library of Congress Country Study* as well as country-specific sources (Fearon 2003, p. 202).

imity, especially the proximity to animal proteins, which are more difficult to transport than carbon hydrates. However, even when transportation costs are a small portion of the total costs, the question whether there is enough food for a growing urban population is still important. Probably the best documented country in this regard is the US. Komlos (1987), for example, finds that urbanization and the labor force in Antebellum American cities grew faster than its food production. He reports that due to the nutritional gap the stature of 20-year old cadets declined by 3.2 cm between the 1830s and the 1850s. This phenomenon existed at least until the beginning of the 20th century. Especially meat and milk was mainly available in rural places close to its production facilities. Among conscripts who fought in World War II the ones from locations of 2,500 inhabitants or less were 1.2 cm taller compared to those from cities with population more than 500,000 inhabitants (Steckel 1995, p. 1922). Steckel even goes so far to argue that the ongoing urbanization trend in the late 19th century played a strong role in the U.S. height decline during that time.

On the other hand, cities serve as industrial, educational and social hotspots. Cities often face a higher productivity, better education, higher wages and therefore higher purchasing power. Schools, universities, firms and governmental institutions can be operated with higher efficiency in places where economies of scale are possible. Firms in urban agglomerations face better access to information and tend to be more innovative (Bianchi and Bellini, 1991; Porter, 1990; Pouder and St. John, 1996; Bell 2005). The main reasons for this advantage are educational institutions along with low transaction costs for the exchange of information. Especially firms located in geographic proximity to government institutions are able to deal more efficiently with red tape and sometimes even take advantage of their special location (Henderson 2003, Duranton 2008). Furthermore, competitors are able to observe each other and innovations and mutations become more likely. They also access a common pool of specialized workers and suppliers (Burt, 1987; Pascal and McCall, 1980; March, 1994; Rogers, 1995). Glaeser and Maré (2001) state that cities make workers more productive, which results in wage growth. In addition, urban workers who move to smaller cities or rural areas do not experience wage declines but tend to stay at the same income level. We expect the positive influences to prevail, because the period and the majority of the countries under observation

suggest that governments are most likely able to use the positive aspect to compensate the negative influences. We use the data on the rate of urbanization gathered by Banks (1971), which measures the share of the population living in cities with more than 50.000 inhabitants.

We also include Lindert's (1994, 1996, 2004) data on **social spending** into our analysis. There are several studies that indicate that redistribution leads to an intelligent use of a nation's resources. Of course, there are also fears that redistributing resources away from the person who owns it causes inefficiency just because the one who loses money has less incentive to be productive. Moreover, government bureaucracies which are set up in order to redistribute money might also cause inefficiencies. However, Lindert (1994) argues that social spending causes, if at all, only little deadweight.

In terms of the law of diminishing marginal utility, redistribution primarily causes a more sensible and efficient use of resources. Among richer social strata, an additional unit of purchasing power, for instance, results only in a little increase of welfare compared to lower strata. The explanation is simple: The marginal utility of purchasing power among wealthy people is low because all basic needs are already satisfied. In contrast, poor strata have a high marginal utility of purchasing power because these people have many unsatisfied needs. Therefore, the same amount of purchasing power can have different benefits, depending on the preexisting welfare level of the consumer. This is true although redistribution leads to an absolute loss of rich individuals, because the gains of the poor outweigh the losses of the rich (Steckel 1983, Steckel 2009, Carson 2009, Blum 2010).

We use **military expenditure** data gathered by Eloranta (2007) as an explanatory variable, as more military expenditure is often associated with distortions in the economic sector. Since military expenditures are often considered as a special kind of public expenditure it may exert a neutral or even a positive influence on particular industries. However, in this paper one of the yardsticks we measure welfare with is male adult height. Military expenditure may increase the welfare of some parts of an economy but it is certainly incomparable with public social spending. Aizenmann and Glick (2006) have argued that military spending might be

beneficial to economic growth if a country faces an acute threat, because military spending ensures safety. This is a convincing argument in the case of a certain threat, but compared with a situation without threats this is not more that the second best alternative. In addition, the period under observation covers both times of peace and times of war. Apart from that, large military forces tend to have inefficient bureaucracies as well as large military order volumes. These organizational structures might lead to corruption and rent seeking (Gupta, Mello and Sharan 2001). Therefore, we do not expect military spending to increase governmental efficiency, but rather the opposite.

In order to take different institutional settings into account we include the **polity 2** index (Marshall and Jaggers 2009), which ranges from -10 (perfect autocracy) to +10 (perfect democracy). The idea behind this measure is that in autocracies people cannot force its leaders to keep their commitments (Olson 1991). In countries with less developed democratic institutions the political elite often uses its outstanding position to peculate resources at the expenses of lower classes (Przeworski and Limongi 1993). In democracies, lower social strata have more influence and therefore demand a greater share of a country's resources. Secondly, in a democracy there is always the possibility to replace bad governments. Political leaders have incentives to increase welfare of the population in order to stay in power. In the former Soviet Union, for example, the party or the state decided the size of the government and structure of the economy and bureaucracy. The Soviet Union along with many other countries with similar organizational structures had no mechanism comparable with democratic elections to vote an inefficient government out of office.

On the other hand, democracies are often associated with rule of law, property rights, free markets, government consumption and human capital. However, Barro (1996) analyzes the effects of democracies on growth in a panel of 100 countries. He controls for these factors as well as the initial level of GDP per capita and finds that the overall effect of democracy on growth is slightly negative. His findings correspond with the skeptic arguments regarding the negative influence of democratic regimes on growth. Przeworski and Limongi (1993) describe the most common arguments in this regard. Among others, they conclude that democratic re-

gimes are forced to redistribute income from richer to poorer strata in order to stay in power. In addition, in democracies workers organize more likely and are able to drive up wages at the expense of the entrepreneur's profits. Lower social strata, however, are rather consumption oriented compared to their richer counterparts. As a result, resources spent for consumption cannot be used for investments. In this context, Rao (1984) argues that dictatorships are better able to enforce savings and the investment of surpluses. In our case, about 70% of all observations have positive democracy values. In other words: we are mainly dealing with different democratic systems. Therefore, our results do not indicate a difference between the efficiency of democracies and dictatorships but rather the difference between liberal and interventionist democracies.

We also include the type of agricultural production. The two extremes in this respect are perfect **monoculture** and a great diversity of crops. Since Mitchell (1993) provides arable land per crop type it is also possible to compute a Herfindahl index of the shares of the total land used for farming.

The most obvious advantage a large portfolio of crops brings is a reduced risk of fluctuating market prices. Since farmers are forced to make decisions under uncertainty, diversification means a secure income because high and low market prices are expected to cancel out each other. Moreover, monocultural structures lead to a dependency on the demand of one single crop and susceptibility to climatic, political and technological shocks (Moradi 2005; Moradi and Baten 2005; Sylvester 2009, Moschini and Hennessy 2001). Furthermore, the diversification of agricultural production goes hand in hand with the demand for labor force. Sylvester (2009) argues that diversification requires more labor power than homogeneous production. This means that diversified agricultural sectors go along with greater participation of the population and enables broader strata to participate in productive economic activities.

On the other hand, monocultures allow economies of scale because the equipment, workflow and education of the employee can be designed and used more efficiently. In addition, monoculture goes along with a high degree of specialization and a great deal of experi-

ence of the employees. Therefore, it is not entirely clear what kind of influences dominate agricultural efficiency.

6.7 Regression Results

The most robust result in our **regression models** (see Table 1) is the negative influence of fragmented societies. No matter which measure of heterogeneity we apply the result stays the same. As Mauro (1995, 1998) and Easterly and Levine (1997) state, diverse societies tend to be more corrupt and instable. Our results suggest that these characteristics also influence governmental efficiency due to high transaction costs, for example to overcome many different languages or ethnic prejudices (Cyffer 2001). The available endowment are not spent on all groups of the society but are concentrated on those groups that are able to speak the lingua franca and are not discriminated or secluded by their religion or cultural peculiarities.

At first glance, **democratic institutions** seem to lower governmental efficiency. This coefficient is not always significant but it has a negative sign throughout most of our models. Hence, our results do not confirm Przeworski's and Limongi's (1993) as well as Olson's (1991) expectations. But does that result suggest that authoritarian or dictatorially ruled countries are better off? More than 70 per cent of the underlying data comes from countries and regions that have positive democracy values. The Republic of Korea in the 1970s, for example, has a democracy value of 3.8.[59] At the same time, France and the United States show values of 8 and 10, respectively. Table 2 reveals that countries like the U.S. and Canada show inadequacies in terms of efficiency although they have highly developed and stable democracies. In other words, the negative influence of our democracy yardstick does not imply superiority of dictatorships compared to democracies but rather a disadvantage of liberal democracies compared to more restrictive ones.

[59] Please remember: the range of definition is -10 (perfect autocracy) to +10 (perfect democracy).

One could assume that focusing only on the cultivation of a small number of cash-crops increases the risk of bad harvests or fluctuating market prices. On the other hand, **monocultures** allow specialization and economies of scale. Since there are two opposing effects at work, the clear and unambiguous results are somewhat surprising. Despite all difficulties correlated with monoculture its advantages seem to outweigh its disadvantages.

We expected a nation's **military burden** to decrease governmental efficiency because it is lost for the welfare producing process in an economy and it may lead to corruption and rent seeking (Gupta, Mello and Sharan 2001). In fact, in both models the corresponding coefficient has a negative sign. Unfortunately the number of observations drops to 50 once we include military spending, which might be the reason why we do not obtain statistically significant results.

The inclusion of **social spending** into our analysis indicates the expected benefits. We consider this variable to be important to explain governmental performance since redistribution helps to ease the negative influences of inequality by narrowing the gap between unequal marginal utilities of rich and poor strata. As Steckel (1983, 2009), Carson (2009) and Blum (2010) show, less inequality can – ceteris paribus – increase welfare even when the richer strata lose wealth because their losses are outweighed by the gains of the poor. Therefore, redistribution offers a tool for the increase of government performance. In all of the estimated models this coefficient suggests a positive – often a statistically significant – influence of social spending.[60]

[60] Social spending is the bottleneck-variable in our regression. If we include this variable we lose half of our observations.

Table 1: Determinants of efficiency 1850s – 1980s in a panel of 62 countries

	I	***II***	***III***	***II***	***III***	***IV***	***V***	***VI***	***VII***	***VIII***
Monoculture	0.28*	0.29**	0.38**	0.31**	0.45**	0.38**	0.46***	0.41***	0.31**	0.59**
Degree of democracy	-0.21	-0.78**	-0.76*							
Urbanization	-0.03	0.005	-0.009	-0.0021	-0.015	0.04	0.004	-0.02	-0.002	0.03
Military Burden										1.26
ELF (Fearon)	*-0.022*		*-0.17**		*-0.26***					
Religion (Alesina et. al)						-0.35***				
ELF (Taylor and Hudson)							-0.34***			
Culture (Fearon)								-0.51***		
Language (Alesina et. al)		-0.33***		-0.38***					-0.38***	-0.46***
Social Spending		0.03*	0.04**	0.016	0.03*	0.01	0.027*	0.04**	0.016	0.05
Time Dummies included?	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Constant	-1.07***	-1.25***	-1.38***	-1.17***	-1.35***	-1.19***	-1.35***	-1.33***	-1.17***	-1.56***
N	134	77	78	77	78	77	78	78	77	50
pseudo-R2	0.23	0.34	0.28	0.32	0.25	0.28	0.31	0.32	0.32	0.31

Note: Efficiency values were multiplied by (-1) in order to obtain intuitive results. Therefore, positive coefficients express a positive influence of the independent variables on efficiency. *, ** and *** denotes significant at 10%/5%/1% levels. Time dummies are included in order to control for unobserved trends. As described above we used a bootstrapping approach with truncated regression models to link the efficiency scores to our explanatory variables. This procedure results in valid standard errors. The pseudo-R^2 can be used to compare the models, but they do not have the standard interpretation.

Urbanization does not seem to have a robust influence. The coefficients are neither significant nor do they have a consistent sign. This result could be explained by the miscellaneous influences of urban agglomerations. Positive influences like low transaction costs, high productivity and higher education might be outweighed by the phenomenon economic historians call the 'urban penalty', primarily the lack of fresh foodstuffs and bad hygienic conditions. In this paper we chose both income per capita and biological welfare as criteria for economic success in order to capture two important characteristics of the standard of living. Urbanization may increase the former but it does not necessarily increase the latter.

6.8 Conclusion

This study estimates government efficiency in terms of welfare production for 62 countries between the 1850s and the 1980s. Therefore, we are able to supplement the existing literature which is normally limited to the post-1965 period. We give a first view on the long-run evolution of government efficiency and show increasing government efficiency in many countries over the course of the 19th and 20th century. During the 20th century Latin American countries and from the middle of the 20th century on their African counterparts experienced a significant decline. In addition, our results indicate a negative influence of wars, especially occupation, and political instability. Japan and the United Kingdom had very high efficiency scores throughout the period under observation while the United States, Canada, Norway, the Netherlands and Sweden used their resources in a remarkably inefficient manner, especially in the pre-WW II period.

We offer several methodological improvements compared to existing studies. Our DEA-methodology provides efficiency estimates which give, when used as the dependent variable in a second stage regression, unreliable standard errors. We correct for this bias by the bootstrapping algorithm proposed by Simar and Wilson (2007) which provides consistent inference. Second, we are able to draw a more realistic picture of the economies under observation, since we use several inputs and outputs rather than only rely on GDP as many other studies do.

The regression results show that monocultures and redistribution increase efficiency while heterogeneous populations have the opposite effect. This is because monocultural structures in agriculture allow specialization and economies of scale. Redistribution, however, leads to a convergence of marginal utilities in a society. It leads to an absolute loss of rich social strata, which is, however, outweighed by the gains of the poor. Heterogeneous societies go along with high transaction costs and may sometimes even lead to corruption. A careful interpretation is required in order to interpret the negative influence of liberal democracies in comparison with rather restrictive ones. Liberal democracies, like the U.S. or Canada, seem to produce lower efficiency values than restrictive ones.

The results presented in this paper suggest several framework conditions through which governments can increase the efficiency of the economy and thereby increase welfare. The first point refers to the reduction of the negative side-effects of ethnic differences, for example conflicts over distribution of government resources or obstacles in education for ethnic or religious minorities. Our results regarding redistribution can be interpreted as an appeal to incorporate large sections of the population into the economy by compulsory education and, if needed, social benefits. Our results show the importance of peace and stability, as the trends indicate that government efficiency suffers from wars and crises in the period under observation.

References

Afonso, A., Schuknecht, L. and V. Tanzi (2008). Income Distribution Determinants and Public Spending Efficiency. Working Paper 861, European Central Bank.

Aizenmann, J. and R. Glick (2006). Military Expenditure, Threats, and Growth. Journal of International Trade and Economic Development 15 (2), 129-155.

Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S. and R. Wacziarg (2003). Fractionalization. Journal of Economic Growth 8 (2), 155-194.

Atlas Narodov Mira (1964). Moscow: Miklukho-Maklai Ethnological Institute at the Department of Geodesy and Cartography of the State Geological Committee of the Soviet Union.

Baier, S. L., Dwyer, G. P. jr. and R. Tamura (2006). How Important are Capital and Total Factor Productivity for Economic Growth? Economic Inquiry 44 (1), 23-49. Data is available at http://www.jerrydwyer.com/; Date of last access: 2010/07/15.

Banker, R. D., Chang, H. and W. W. Cooper (2004). A simulation study of DEA and parametric frontier models in the presence of heteroscedasticity. European Journal of Operational Research 153 (3), 624–40.

Banks, A. S. (1971). Cross-polity time-series data, State Univ. of New York. - Cambridge (Mass.): MIT Press.

Barro, R. J. (1996). Democracy and Growth; Journal of Economic Growth, 1 (1):1-27.

Baten, J. (1999). Ernährung und wirtschaftliche Entwicklung in Bayern, 1730-1880. Stuttgart: Steiner.

Baten, J. and A. Wagner (2003). Autarchy, market disintegration, and health: the mortality and nutritional crisis in Nazi Germany, 1933–1937; Economics and Human Biology 1 (1), 1-28.

Baten, J. and K. Enflo (2007). Estimates of capital stocks for 77 countries. Working Paper.

Baten, J. (2009). Protein supply and nutritional status in nineteenth century Bavaria, Prussia and France. Economics and Human Biology 7 (2), 165-180. doi:10.1016/j.ehb.2009.02.001

Bell, G. G. (2005). Clusters, Networks, and Firm Innovativeness. Strategic Management Journal 26 (3), 287–295.

Bergstrand J. H. and P. Egger (2007). A knowledge-and-physical-capital model of international trade flows, foreign direct investment, and multinational enterprises. Journal of International Economics 73 (2), 278-308.

Bianchi, P. and N. Bellini (1991). Public policies for local networks of innovators. Research Policy 20 (5): 487–497.

Blum, M. (2010). The Influence of Inequality on the Standard of Living: Worldwide Evidence from 1810 to 1980. Working Paper, University of Tuebingen.

Burt, R. S. (1987). Social contagion and innovation: cohesion versus structural equivalence. American Journal of Sociology 92 (6), 1287–1335.

Carson, S. A. (2009). Health, Wealth, and Inequality. A Contribution to the Debate about the Relationship between Inequality and Health. Historical Methods 42 (4), 43-56.

Coelli, T., Prasado Rao, D. S. and G. E. Battese (1998). An Introduction to Efficiency and Productivity Analysis. Boston (u. A.): Kluwer.

Cooper, W. W., Seiford, L. M. and K. Tone (2006). Introduction to data envelopment analysis and its uses. With DEA-solver software and references. New York: Springer.

Cooper, W. W. and K. Tone (1997). Measures of inefficiency in data envelopment analysis and stochastic frontier estimation. European Journal of Operational Research 99 (1), 72–88.

Crayen, D. and Baten, J. (2010). Global trends in numeracy 1820–1949 and its implications for long-term growth; Explorations in Economic History; Volume 47 (1), 82-99.

Cyffer, N. (1977). Sprachplanung in Nigeria. Afrika Spectrum 12, 239–262.

Cyffer, N. (2001). Babylon in Afrika. Der Kontinent der zweitausend Sprachen. In: Christoph Plate and Theo Sommer (Eds.). Der bunte Kontinent. Ein neuer Blick auf Afrika, 177-183.

Deprins, D., Simar, L., and H. Tulkens (1984). Measuring labor-efficiency in post offices. In: M. Marchand, P. Pestieau, and H. Tulkens (Eds.). The performance of public enterprises: concepts and measurement. Amsterdam: North-Holland.

Duranton, G. (2008). Viewpoint: From cities to productivity and growth in developing countries. Canadian Journal of Economics 41 (3), 689-736.

Easterly, W. and R. Levine (1997). Africa's Growth Tragedy: Policies and Ethnic Divisions. The Quarterly Journal of Economics 112 (4), 1203-1250.

Eloranta, J. (2007): From the great illusion to the Great War: Military spending behaviour of the Great Powers, 1870–1913. European Review of Economic History 11 (2), 255–283.

Fearon, J. D. (2003). Ethnic and Cultural Diversity by Country. Journal of Economic Growth 8 (2), 195-222.

Feenstra, R. C. (2010). International Trade. W. H. Freeman & Co.

Fernald, J. G. (1999) Roads to prosperity? Assessing the link between public capital and productivity. American Economic Review 89 (3), 619-638.

Glaeser, E. L. and D. C. Maré (2001). Cities and Skills. Journal of Labor Economics 19 (2), 316-342.

Gould, W. T. (1998). African mortality and the new 'urban penalty'. Health and Place 4 (2), 171-181.

Grigg, D. (1995). The Pattern of World Protein Consumption, Geoforum 26 (1), 1-17.

Gupta, S., de Mello, L. and R. Sharan (2001). Corruption and military spending. European Journal of Political Economy 17 (4), 749-777.

Gylfason, T. (2001). Natural resources, education, and economic development. European Economic Review Volume 45 (4-6), 847-859.

Haq, M. ul. (1995). Reflections on Human Development. New York: Oxford University Press.

Henderson, V. (2002). Urban primacy, external costs, and quality of life. Resource and Energy Economics 24 (1-2), 95-106.

Herrmann, P., Tausch, A., Heshmati, A. and C. S. J. Bajalan (2008). Efficiency and Effectiveness of Social Spending. Institute for the Study of Labor Discussion Paper No. 2382.

Holtz-Eakin, D. (1994). Public-Sector Capital and the Productivity Puzzle. The Review of Economics and Statistics 76 (1), 12-21.

Huberman, M. and C. Minns (2007). The times they are not changin': Days and hours of work in Old and New Worlds, 1870–2000. Explorations in Economic History 44 (4), 538-567.

Inglehart, R., Foa, R., Peterson, C. and C. Welzel (2008). Development, Freedom, and Rising Happiness: A Global Perspective (1981–2007). Perspectives on psychological science 3 (4), 264-285.

Jonsson, G. (1998). Changes in Food Consumption in Iceland, 1770-1940. Scandinavian Economic History Review XLVI (1), 24-41.

Joumard, I., Kongsrud, P. M., Nam, Y.-S. and R. Price (2003). Enhancing the cost effectiveness of public spending: Experience in OECD Countries. OECD Economic Studies, No. 37, 2003/2.

Kaufmann, D., Kraay, A. and M. Mastruzzi (2008). Governance Matters VII: Aggregate and Individual Governance Indicators 1996-2007. Policy Research Working Paper 4654, World Bank.

Kaufmann, D. and A. Kraay (2008). Governance Indicators: Where Are We, Where Should We Be Going? Oxford University Press, DOI: 10.1093/wbro/lkm012.

Koepke, N. and J. Baten (2008). Agricultural specialization and height in ancient and medieval Europe. Explorations in Economic History 45 (2), 127-146.

Komlos, J. (1985). Stature and Nutrition in the Habsburg Monarchy: the Standard of Living and Economic Development in the Eighteenth Century. The American Historical Review90 (5), 1149-1161.

Komlos, J. (1987). The Height and Weight of West Point Cadets: Dietary Change in Antebellum America. The Journal of Economic History 47 (4), 897-927.

Komlos, J. and J. Baten (2004). Looking Backward and Looking Forward Anthropometric Research and the Development of Social Science History. Social Science History 28 (2), 191–210.

Larson, D. F., Butzer R., Mundlak Y. and A. Crego (2000). A Cross-Country Database for Sector Investment and Capital. The World Bank Economic Review 14 (2), 371-91.

Lindert, P. (1994). The Rise of Social Spending, 1880 – 1930. Explorations in Economic History 31 (1), 1-37.

Lindert, P. (1996). What Limits Social Spending? Explorations in Exonomic History 33 (1), 1–34.

Lindert, P. (2004). Growing Public. Cambridge: Cambridge Univ. Press.

Maddala, G. S. and S. Wu (1999). A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test. Oxford Bulletin of Economics and Statistics 61 (1), 631-652.

Maddison, A. (1994). Standardised estimates of fixed capital stock: A six country comparison. Research Memorandum 570 (GD-9). Groningen University.
Maddison, A. (2001). The World Economy: A Millennial Perspective. OECD Development Centre.

Madsen, J. B. and E. P. Davis (2006). Equity Prices, Productivity Growth and the New Economy. The Economic Journal 116 (513), 791-811.

Madsen, J. B. (2007). Technology Spillover through trade and TFP convergence: 135 years of evidence for the OECD countries. Journal of International Economics 72 (2), 464-480.

Madsen, J. B. (2010). Growth and capital deepening since 1870: Is it all technological progress? Journal of Macroeconomics 32 (2), 641–656.

Mankiw, N. Gregory, Romer, D. and D. N. Weil (1992). A Contribution to the Empirics of Economic Growth. The Quarterly Journal of Economics 107 (2), 407-337.

March, J. G. (1994). A Primer on Decision Making: How Decisions Happen. Free Press: New York.

Marshall, M. and K. Jaggers (2009). Polity IV Project: Dataset User's Manual. Center for Global Policy, School of Public Policy, George Mason University and Center for Systemic Peace.

Martínez-Carrión, J.-M. and J. Moreno-Lázaro (2007). Was there an urban height penalty in Spain, 1840–1913?; Economics and Human Biology 5 (1), 144-164.

Mauro, P. (1995). Corruption and growth. The Quarterly Journal of Economics 110 (3), 681-712.

Mauro, P. (1998): Corruption and the composition of government expenditure. Journal of Public Economics 69 (2), 263-279.

Mitchell, B. R. (1993). International Historical Statistics. Basingstoke [u.a.] : Palgrave Macmillan.

Moradi, A. (2005). Ernährung, wirtschaftliche Entwicklung und Bürgerkriege in Afrika südlich der Sahara (1950-2000). Inaugural-Dissertation University of Tübingen. http://tobias-lib.uni-tuebingen.de/volltexte/2005/1879/pdf/Dissertation_Moradi.pdf.

Moradi, A. and J. Baten (2005). Inequality in Sub-Saharan Africa: New Data and New Insights from Anthropometric Estimates. World Development 33 (8), 1233-1265.

Morrison, C. and F. Murtin (2009). The Century of Education; Journal of Human Capital, vol.3(1) pp.1-42.

Morrison, C. and F. Murtin (2010). The Kuznets Curve of Education: A Global Perspective on Education Inequalities 1870-2010. PSE Working Paper. Retrieved from www.-fabricemurtin.com. Accessed 01.09.2010.

Moschini, G., and D. A. Hennessy (2001). Uncertainty, Risk Aversion, and Risk Management for Agricultural Producers. In: Garnder, B. L. and G. C. Rausser (Eds.). Handbook of Agricultural Economics, edition 1, volume 1, 88–153. Amsterdam: Elsevier.

Nehru, V. and A. Dhareshwar (1993). A New Database on Physical Capital Stock: Sources, Methodology and Results. Revista de Analisis Economica 8(1), 37-59.

Olson, M. Jr., (1991). Autocracy, Democracy and Prosperity. In Zeckhauser, R. J. (ed.). Strategy and Choice, 131-157. Cambridge: MIT Press.

O'Rourke, K. H. and A. M. Taylor (2006). Democracy and Protectionism. NBER Working Paper 12250. May 2006.

Pascal A. H. and J. J. McCall (1980). Agglomeration economies, search costs, and industrial location. Journal of Urban Economics 8 (3), 383–388.

Pochet, P. (2005). The Open Method of Co-ordination and the Construction of Social Europe. A Historical Perspective, in Zeidlin, J. and P. Pochet (eds). The Open Method of Coordination in Action. The European Employment and Social Inclusion Strategies. Brussels: PIE-Peter Lang.

Porter M. E. (1990). The Competitive Advantage of Nations. Free Press: New York.

Posner, D. N. (2004). Measuring Ethnic Fractionalization in Africa. American Journal of Political Science 48 (4), 849-863.

Pouder, R. and C. H. St. John (1996). Hot spots and blind spots: geographical clusters of firms and innovation. Academy of Management Review 21(4), 1192–1225.

Przeworski, A. and F. Limongi (1993). Political Regimes and Economic Growth. The Journal of Economic Perspectives 7 (3), 51-69.

Rao, V. (1984). Democracy and Economic Development. Studies in Comparative International Development 19 (4), 67-81.

Ravallion, M. (2005). On Measuring Aggregate "Social Efficiency". Economic Development and Cultural Change 53 (2), 273–292.

Rayp, G. and N. Van De Sijpe (2007). Measuring and explaining government efficiency in developing countries. Journal of Development Studies 43 (2), 360 — 381. DOI: 10.1080/00220380601125230.

Reddaway, W. B. (1959). Some limitations of the capital-output ratio. Royal Society: London.

Resti, A. (2000). Efficiency measurement for multi-product industries: a comparison of classic and recent techniques based on simulated data. European Journal of Operational Research 121 (3), 559–578.

Rogers E. M. (1995). Diffusion of Innovations. 4th Edition. Free Press: New York.

Sachs, J. D. and A. M. Warner (2001). The curse of natural resources. European Economic Review 45 (4-6), 827-838.

Sen, A. (1999). Development as freedom. Oxford: Oxford University Press.

Simar, L. and P. W. Wilson (2007). Estimation and inference in two-stage, semi-parametric models of production processes. Journal of Econometrics 136 (1), 31-64.

Simar, L. and P. W. Wilson (2000). A general methodology for bootstrapping in non-parametric frontier models. Journal of Applied Statistics 27 (6), 779-802.

Steckel, R. H. (1983). Height and per capita income. Historical Methods 16 (1), 1-7.

Steckel, R. H. (1995). Stature and the Standard of Living. Journal of Economic Literature 33 (4), 1903-1940.

Steckel, R. H. (2009). Heights and Human Welfare: Recent Developments and New Directions. Explorations in Economic History 46 (1), 1-23.

Summers, R. and A. Heston (1994). The Penn world tables, Mark 5.6. Retrieved from http://pwt.econ.upenn.edu/php_site/pwt_index.php. Accessed 01.09.2010.

Sylvester, K. M. (2009). Ecological Frontiers on the Grasslands of Kansas: Changes in Farm Scale and Crop Diversity. The Journal of Economic History 69 (4), 1041-1062.

Taylor, C. L. and M. C. Hudson (1972). World Handbook of Political and Social Indicators. Ann Arbor, MI: ICSPR.

Wilson, P. W. (2007). FEAR: A package for frontier efficiency analysis with R. Socio-Economic Planning Sciences.

Woods, R. (2003). Urban-Rural Mortality Differentials: An Unresolved Debate. Population and Development Review 29 (1), 29-46.

WHO (1999). The World Health Report 1999 - Making a Difference. World Health Organization.

Appendix A: The Bootstrap

In this paper a two-stage approach is used, in which efficiency is estimated on a first stage and then the estimated efficiency values are regressed on environmental variables.

Simar and Wilson (2007) provide a framework in which this two-stage estimation procedure makes sense. In their setting, the relationship between the efficiency measures δ and the environmental variables z is defined by the following model:

$$\delta_i = \psi(Z_i, \beta) + \epsilon_i \geq 1, \tag{1}$$

where ψ is a smooth function, β is a vector of parameters and ϵ_i is distributed N(0, σ_ϵ^2) with left truncation at $1-\psi(Z_i, \beta)$ for each i.

Like most of the two-stage studies, we specify $\psi(Z_i, \beta) = z_i'\beta$, so that equation (1) results in

$$\delta_i = z_i'\beta + \epsilon_i \geq 1. \tag{2}$$

Simar and Wilson point out, however, that this approach carries a problem arising from the fact that the efficiency estimates on the first stage are serially correlated, with this correlation disappearing very slowly as the sample sizes grow large. Additionally, they show the rate of convergence $\%r$ to be of order $O(\%n^{(\frac{-2}{p+q+1})})$, p being the number of inputs and q the number of outputs. Even for a small model with 3 inputs and 2 outputs the rate is $\%r = \%n^{(-1/3)}$, and it decreases with more variables being added. As a consequence, standard approaches to inference are invalid.

To deal with this problem, Simar and Wilson suggest a special form of the parametric bootstrap that allows for a consistent inference in this two-stage approach. The steps of their method are the following:

1. Using the original data compute $\hat{\delta}_i(x_i, y_i)$ for $i=1,\ldots,n$.

2. Use the method of maximum likelihood to obtain an estimate $\hat{\beta}$ of β as well as an estimate $\hat{\sigma}_\epsilon$ of σ_ϵ in the truncated regression of $\hat{\delta}_i$ on z_i in (2) using the $m<n$ observations where $\hat{\delta}_i>1$.

3. Loop over the next three steps (3-1 to 3-3) L times to obtain a set of bootstrap estimates $A=((\hat{\beta}^*, \hat{\sigma}^*_\epsilon)_b)_{b=1}^{L}$:

   3.1 For each $i=1,\ldots,m$ draw ϵ_i from the N(0, $\hat{\sigma}_\epsilon$) distribution with left-truncation at $(1-z_i'\hat{\beta})$

   3.2 Again for each $i=1,\ldots,m$ compute $\delta_i^*=z_i'\hat{\beta}+\epsilon_i$

   3.3 Estimate the truncated regression of δ_i^* on z_i to get the estimates $(\hat{\beta}^*, \hat{\sigma}^*_\epsilon)$

4. Use the bootstrap values in A and the original estimates $\hat{\beta}$, $\hat{\sigma}_\epsilon$ to construct estimated confidence intervals for each element of β and for σ_ϵ.

Appendix B: The Data Set

Appendix B1: Reliability of the Capital Stock Estimates

In this appendix we compare different capital stock datasets. The most important difference between them is that they rely on different strategies to estimate the capital stock and are based on different assumptions on depreciation rates.

The largest comparable dataset on physical capital stock in existence was compiled by Baier, Dwyer and Tamura (henceforth BDT; 2006). They use the perpetual inventory method, which sums up net investments.[61] Applying this method, they are able to provide physical capital stock estimates for up to 143 countries on a decadal basis between the 1840s and 2000 (see table B1). This data set has already been used by economic historians and international

Table B1: Number of countries per decade for which capital data are available in the BDT-data.

Year	Number of Countries
1840	1
1850	6
1860	9
1870	13
1880	17
1890	20
1900	23
1910	28
1920	33
1930	37
1940	46
1950	66
1960	114

[61] Scholars using this method have to make an assumption about the depreciation rate, on which they disagree. Therefore, scholars arrive at differing estimates based on this method. Some Scholars assume a uniform rate for the whole capital stock (Maddison as well as Nehru and Dhareshwar use 4%), while others differentiate between machinery and equipment on the one hand and non-residential building (Madsen uses 17.6% and 3%, respectively).

1970	122
1980	127
1990	143
2000	143

Source: own calculations.

economists (see for instance Huberman and Minns 2007, Bergstrand and Egger 2007, O'Rourke and Taylor 2006). The same method is also used by other scholars, but with a narrower focus on time and space. Maddison (1994) generates capital stock estimates for 6 large countries. His series go back until 1820 in the case of the UK, but others start later, as in the case of the Netherlands in the 1950s. Madsen (2007, 2010) and Madsen and Davis (2006; henceforth 'Madsen') used data for 22 countries from 1870 to the 2000s for TFP estimations. For the period after 1950, more comprehensive datasets are available. The most prominent data can be found in the Penn World Tables (Summer and Heston 1994). They cover the period from 1965 to 1992 for a panel of 62 countries. Alternative estimates for the period after 1967 were compiled by Larson, Butzer, Mundlak and Crego (2000), covering 62 countries.

Baten and Enflo (2007) base their estimates on energy consumption, which is highly correlated with physical capital. By estimating this relationship for 1965 (using the PENN World Tables) and using this estimate to project backwards in time with data on energy consumption, they can estimate capital stock data for 4 data points between 1925 and 1965 for up to 77 countries.

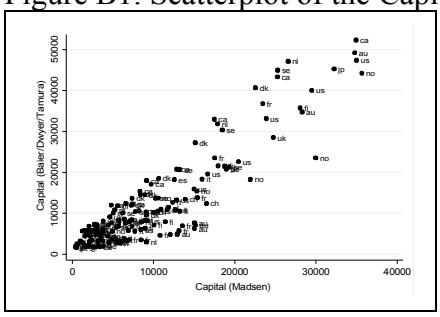
In this Appendix we assemble various alternative capital stock series and compare the data (see notes for sources). Table B2 shows that the different estimations are highly correlated, suggesting that the different estimation strategies employed lead to similar results. In addition, Figure B1 shows the correlation between the capital estimates of BDT and Madsen in the form of a scatter plot.

Table B2: Correlation between various estimation of physical capital

	BDT	BE	PWT	Maddison	LBMC	Madsen
BDT	1					
BE	0.9094	1				
PWT	0.9089	0.9948	1			
Maddison	0.9133	0.8570	0.8485	1		
LBMC	0.7502	0.7486	0.7474	0.8886	1	
Madsen	0.9396	0.9163	0.8851	0.9588	0.8961	1

Notes: The table contains Pearson correlation coefficients. 'BDT' are the data from Baier, Dwyer and Tamura (2006). 'BE' refers to Baten/Enflo (2007), which are supplemented with the Penn World Table data for the periods after 1965. 'PWT' are the Penn World Table data (Summers & Heston 1994). 'Maddison' refers to his 1994 estimates. 'LBMC' are data from Larson, Butzer, Mundlak and Crego (2000). 'Madsen' refers to the dataset used in Madsen (2007, 2010) and Madsen and Davis (2006).

Figure B1: Scatterplot of the Capital estimations from BDT and Madsen



Notes: We included only observations which we used to calculate the efficiency values.

We now look at capital-output ratios, as with this measure in many countries and periods less trend correlation is visible (also according to Fisher tests, see Maddala & Wu 1999 on the test) than in the capital series (Nehru and Dhareshwar 1993). This yardstick is a measure of productivity of capital. It has several weaknesses, such as the fact that productivity of capital is not only a function of investment, but also of e.g. technical progress and the interaction of physical capital with other characteristics of an economy, like the growth of the labor force (Reddaway 1959). Another problem is its susceptibility to variation over time. We calculate both the mean and the coefficient of variation of alternative capital-output measures of the above described measures of physical capital in order to get a first idea of the quality of the

data sets. The descriptive statistics in Table B3 show that, although the mean of the BDT series is somewhat higher, the coefficients of variation (CV) have comparable magnitudes.

Table B3: Descriptive statistics of capital-output ratios

	Mean	CV
BDT	2.56	0.38
BE+PWT	1.76	0.38
Maddison	1.70	0.48
LBMC	2.43	0.31
Madsen	2.08	0.51

Notes: Abbreviations see Table B2. GDP-Data come from Maddison (2001). The sample here is restricted to the countries and time periods for which Madsen provides data to ensure comparability.

Table B4 shows that the capital-output values are less correlated than the capital series. The coefficient between the BDT data and the alternative estimations fit into the general picture. The correlations are of medium strength, but neither significantly lower nor higher compared to other correlation coefficients.

Table B4: Correlation between several measures of capital-output ratios

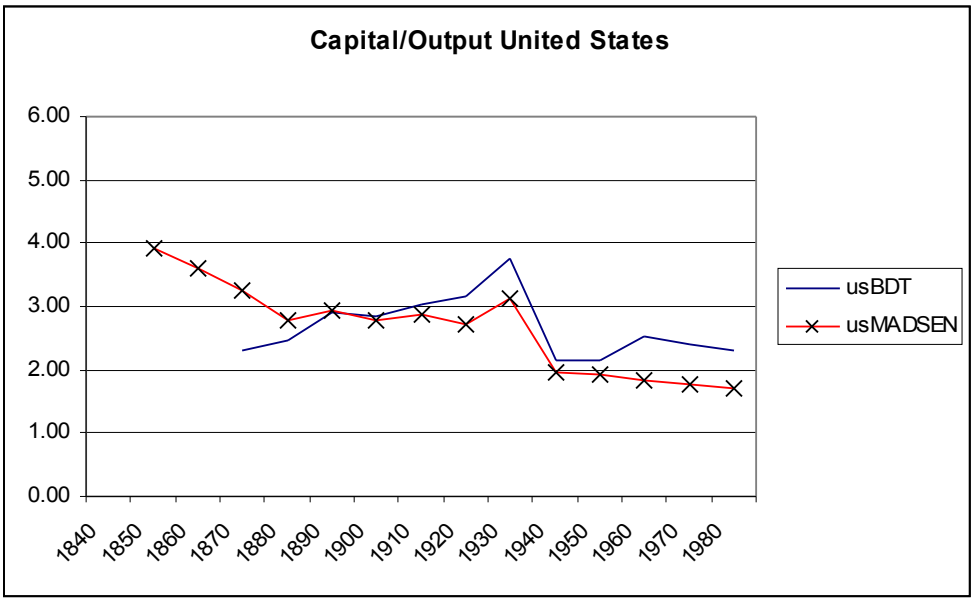
	BDT	BE	PWT	Maddison	LBMC	Madsen
BDT	1					
BE	0.2878	1				
PWT	0.3357	0.8374	1			
Maddison	0.5852	-0.1018	0.4847	1		
LBMC	0.5093	0.1366	0.1451	0.7763	1	
Madsen	0.3616	0.2907	0.6501	0.8958	0.6685	1

Notes: Abbreviations see Table B2. GDP-Data come from Maddison (2001).

Figures B2 to B7 show the development of capital/output ratios for different countries, each graph comparing the series based on the BDT data and the data from Madsen, as these have the highest coverage for the earlier periods. In most series the same pattern is found, although for some countries deviations for some decades are observed. Nevertheless, we cannot

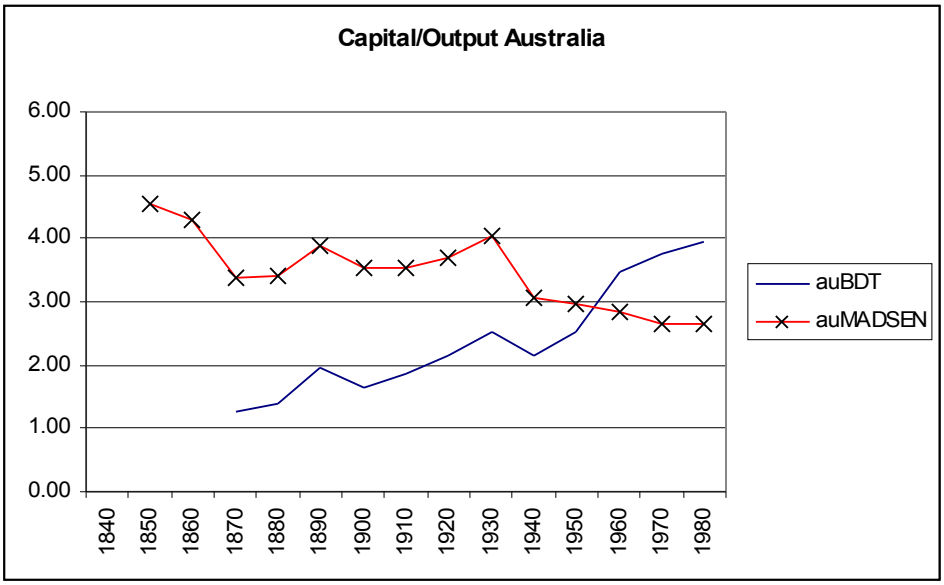
conclude in favor of one of the datasets since the graphs do not help to identify which has a superior quality.

Figure B2:



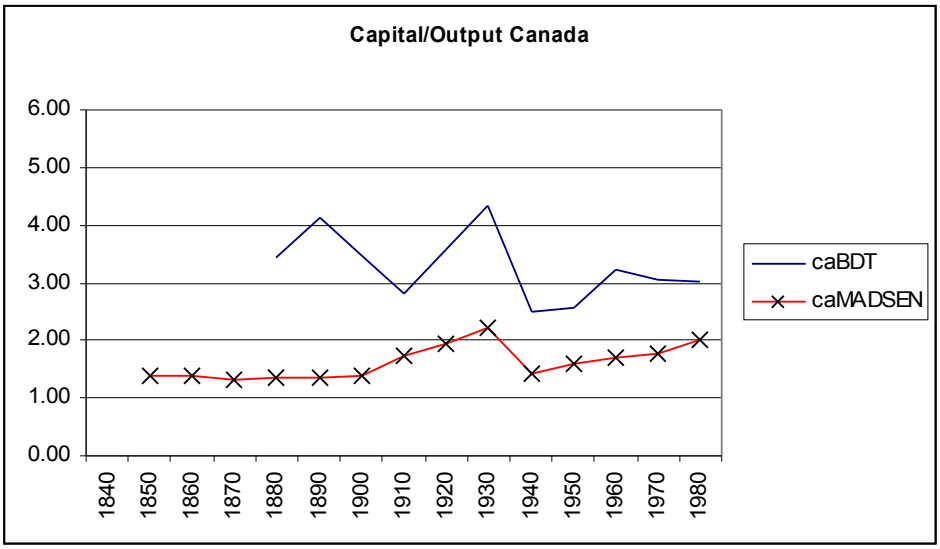
Source: own calculations.

Figure B3:



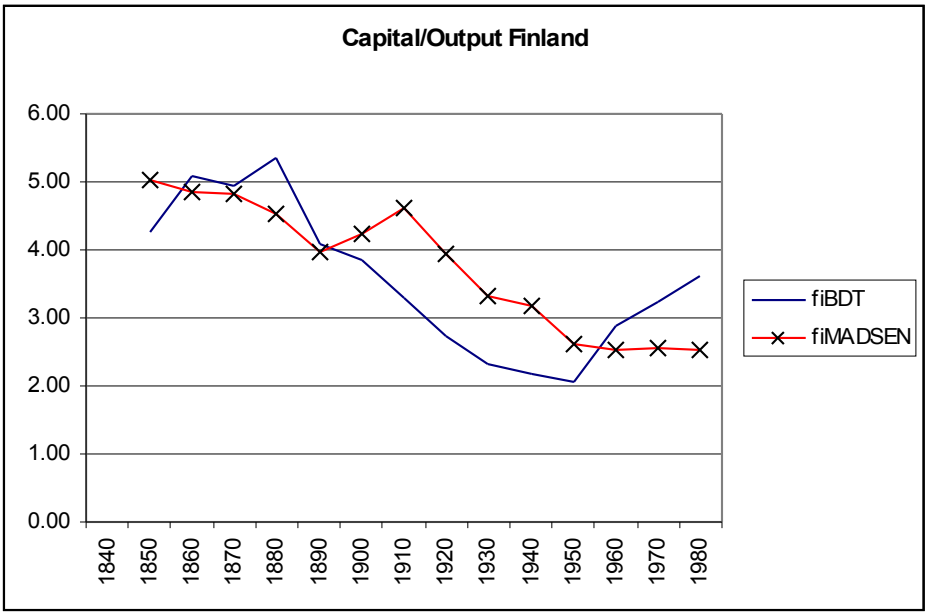
Source: own calculations.

Figure B4:



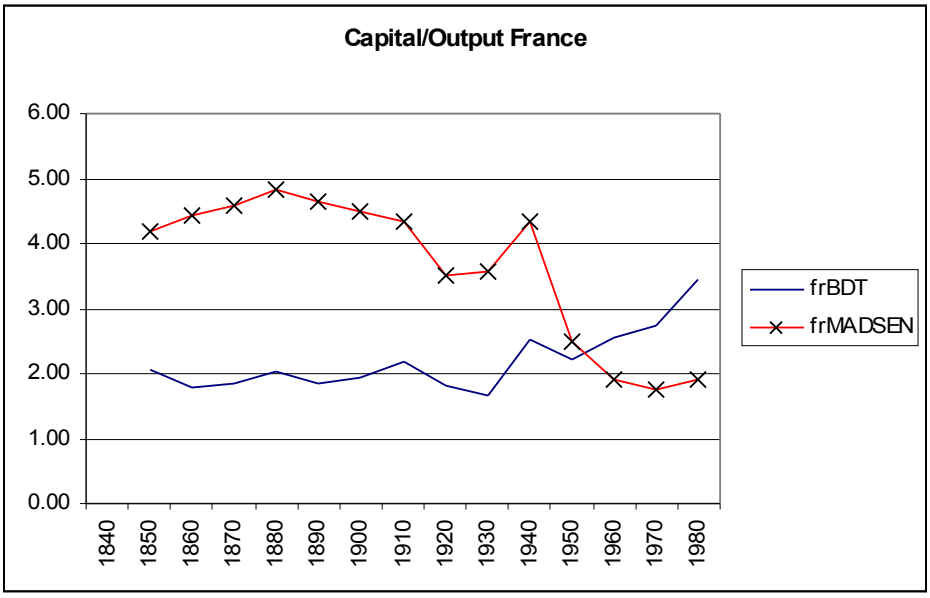
Source: own calculations.

Figure B5:



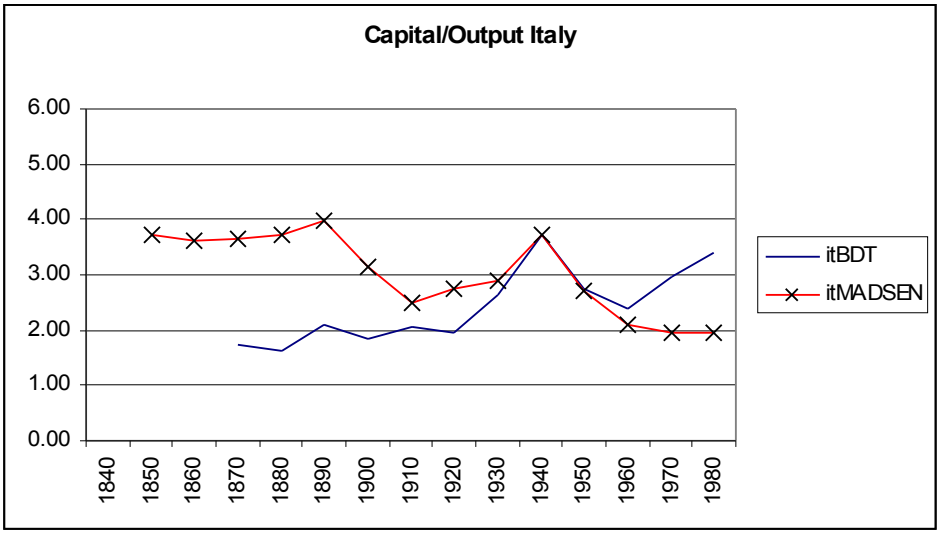
Source: own calculations.

Figure B6:



Source: own calculations.

Figure B7:



Source: own calculations.

We conclude that the dataset compiled by Baier, Dwyer and Tamura (2006) is as reliable as alternative estimations. In addition, it is the only data set that allows us to follow government efficiency for up to 130 years in a panel of 60 countries. BDT is the most suitable

dataset in the prevailing context, since the other ones would not offer improvements in terms of quality or quantity.

Appendix B2: Reliability of the Efficiency Estimates

Table B5 shows the correlation coefficients between alternative estimations of governmental efficiency. BDT, for instance, stands for the efficiency estimation based on the measure of physical capital – the other in- and outputs are the same in all estimations – provided by Baier, Dwyer and Tamura (2006) as it is described above. The correlation matrix suggests that altering the measure of physical capital does not change the estimates of government performance significantly. The correlation coefficients between the DEA values based on our preferred measure (BDT) and its alternatives lie in the range of 0.68 and 0.95. None of our conclusions depend on the choice of the capital measure.

Table B5: Correlation between different efficiency estimations, each based on alternative estimation of physical capital.

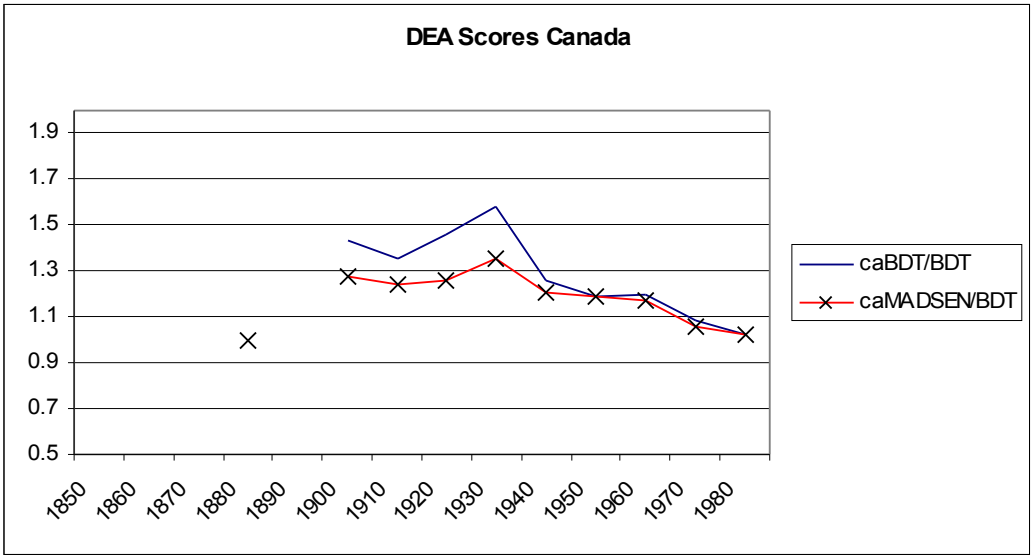
	BDT	BE	PWT	Maddison	LBMC	Madsen
BDT	1					
BE	0.81	1				
PWT	0.74	0.95	1			
Maddison	0.96	0.85	0.89	1		
LBMC	0.68	0.77	0.82	1	1	
Madsen	0.75	0.78	0.94	0.99	0.91	1

Notes: Abbreviations see Table B1.

Figures B8 to B11 show the efficiency trends for selected countries based on the capital data from BDT and Madsen, respectively. Though the detailed forms differ, the general

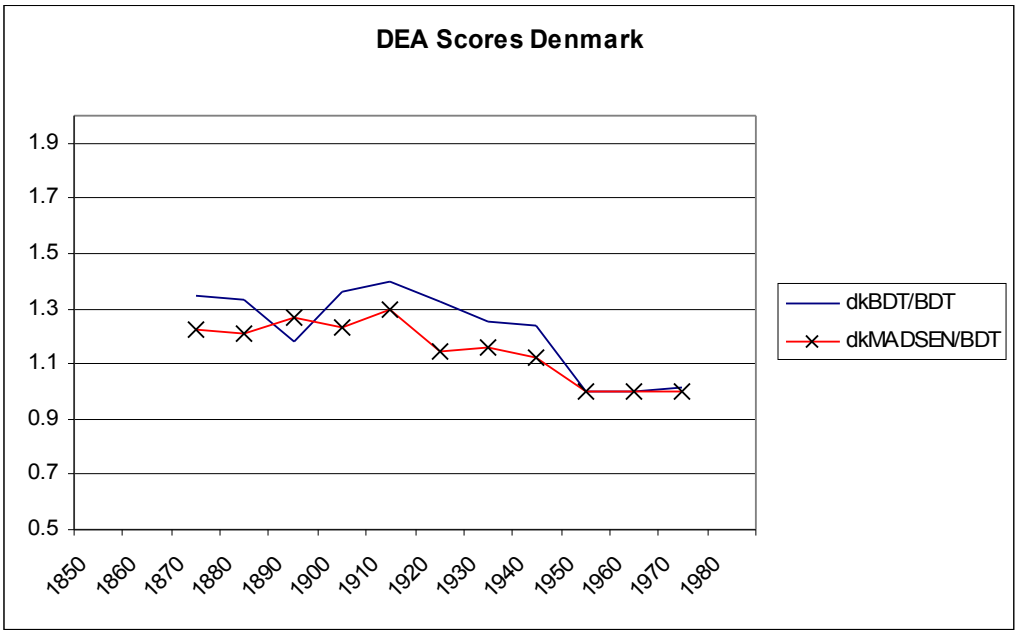
conclusions to be drawn from the series are similar. Our results do not depend on the choice of the capital measure.

Figure B8:



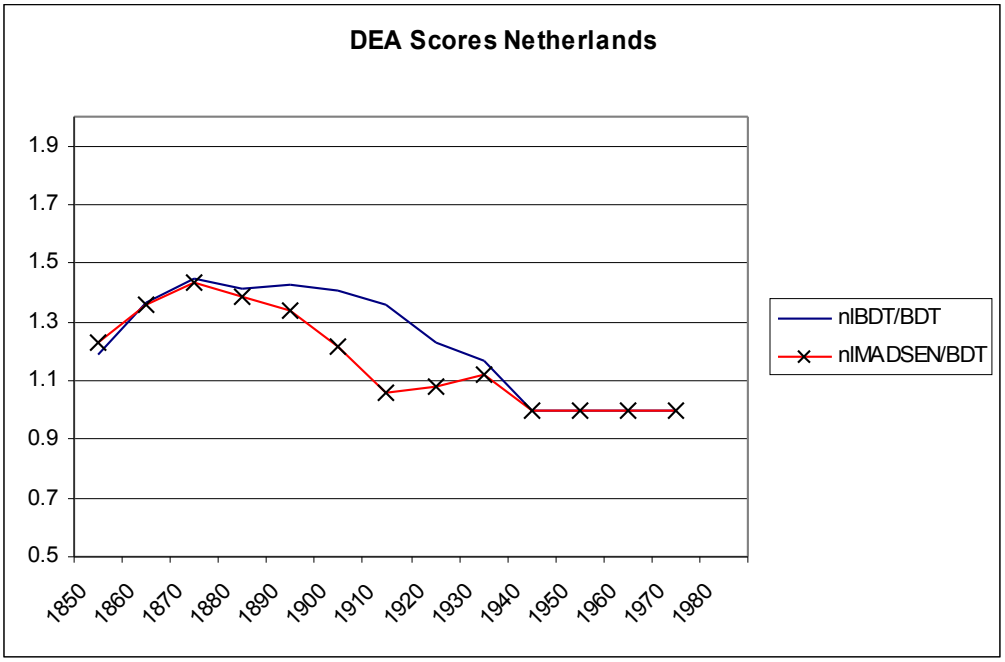
Source: own calculations.

Figure B9:



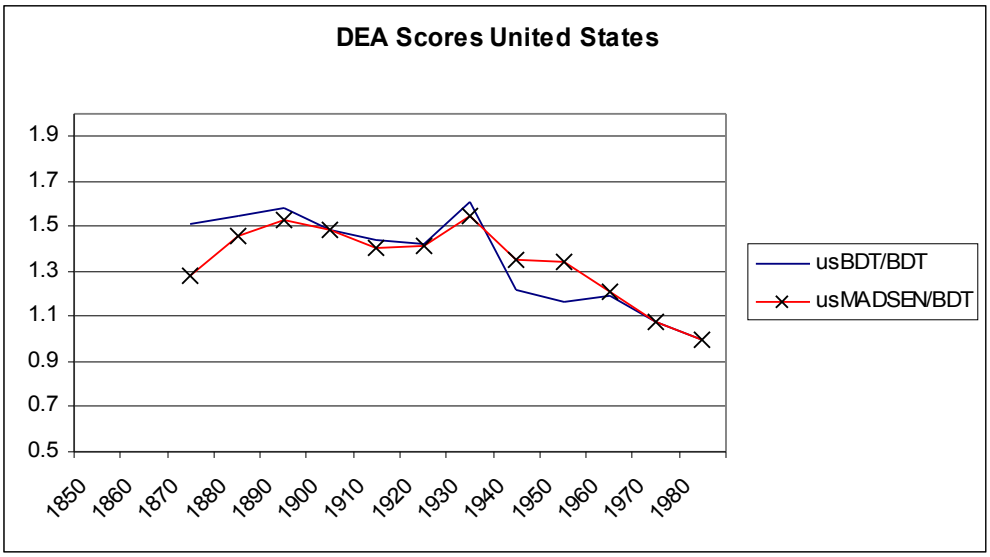
Source: own calculations.

Figure B10:



Source: own calculations.

Figure B11:



Source: own calculations.

Appendix B3: Reliability of the Human Capital Stock Estimates

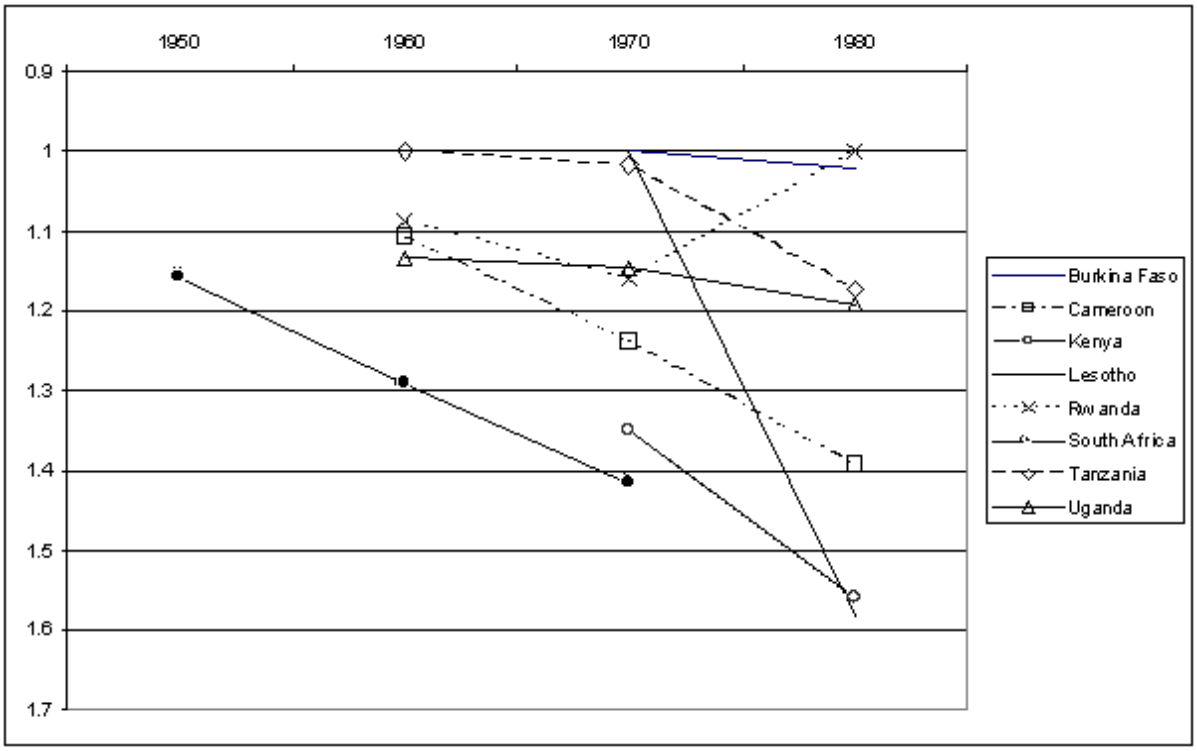
There are several global long-run datasets on human capital available. In general, there are two different types of data. First, there are data such as school enrollment, signature ability, literacy or numeracy, which have an upper bound, as they reach values of 100 percent during the course of development. Second, there are measures such as average years of schooling which do not face such a boundary. For our analysis it might not be reasonable to use datasets which have an upper bound. Both Baier, Dwyer and Tamura (2006) and Morrisson and Murtin (2009) provide estimates of human capital on a global scale. These authors offer a good compromise, as they deliver a series that is not reaching an upper bound during the twentieth century and their data incorporate not only basic skills.

Moreover, in our context we cannot rely on data that is arranged in birth decades of the people holding the human capital. The use of human capital as an input in an economic process ought to include all people in the labor force. Data arranged by birth decades rather indicates the quality of schooling at the time one particular cohort is schooled. This is not exactly what is in demand in this paper. In contrast, BDT estimate the stock of human capital in the way we prefer, since they look at the population in general in order to capture the human capital stock available for the economy on a decadal basis. Their values on human capital are calculated by combining average years of schooling, which were calculated by including enrollment ratios (including primary, secondary and college education), the age distribution of the population and the average working experience. By doing that, they are able to show that human capital was strongly increasing during the 19th and 20th century.

We compare the BDT dataset with the one of Morrison and Murtin (2009, 2010). The Morrison and Murtin dataset offers estimations of a comparable quantity and quality. The estimates are closely correlated, the correlation coefficient is 0.9068. This shows that both sets are useful sources to represent human capital development. However, we use the BDT data as they incorporate working experience as well as education and therefore provide more information than the Morrisson and Murtin data.

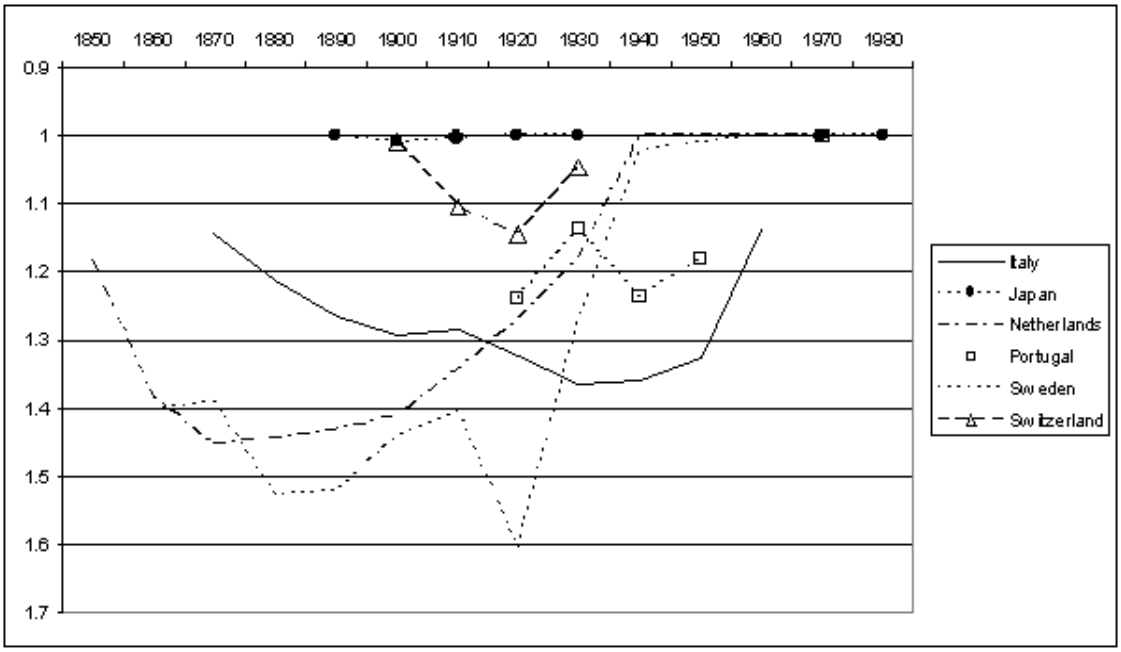
Appendix C: Figures

Figure 6: Efficiency in Africa (continued).



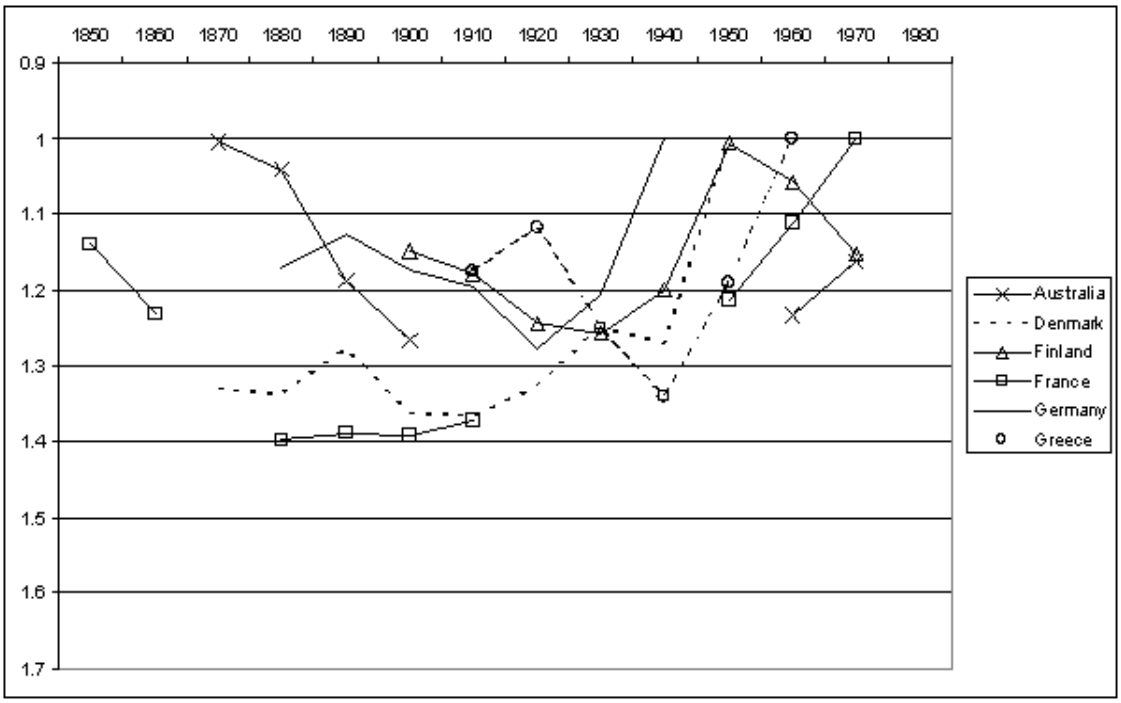
Source: own calculations.

Figure 7: Efficiency in the Industrial Countries (continued).



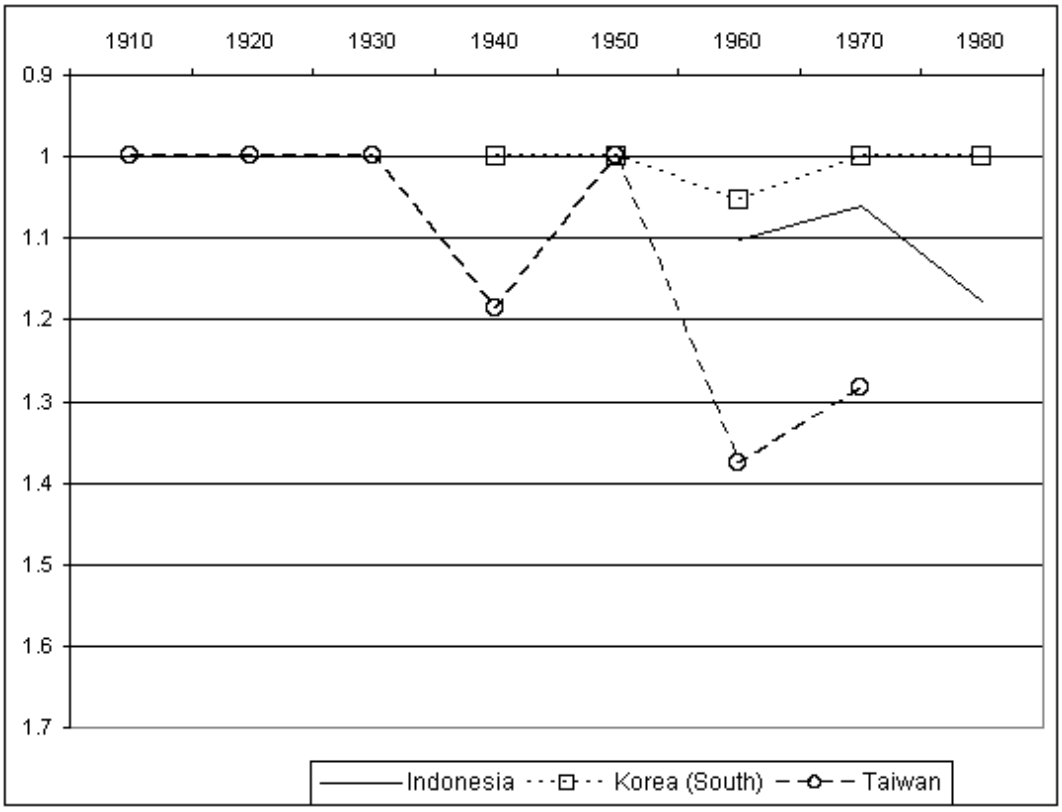
Source: own calculations.

Figure 8: Efficiency in the Industrial Countries (continued).



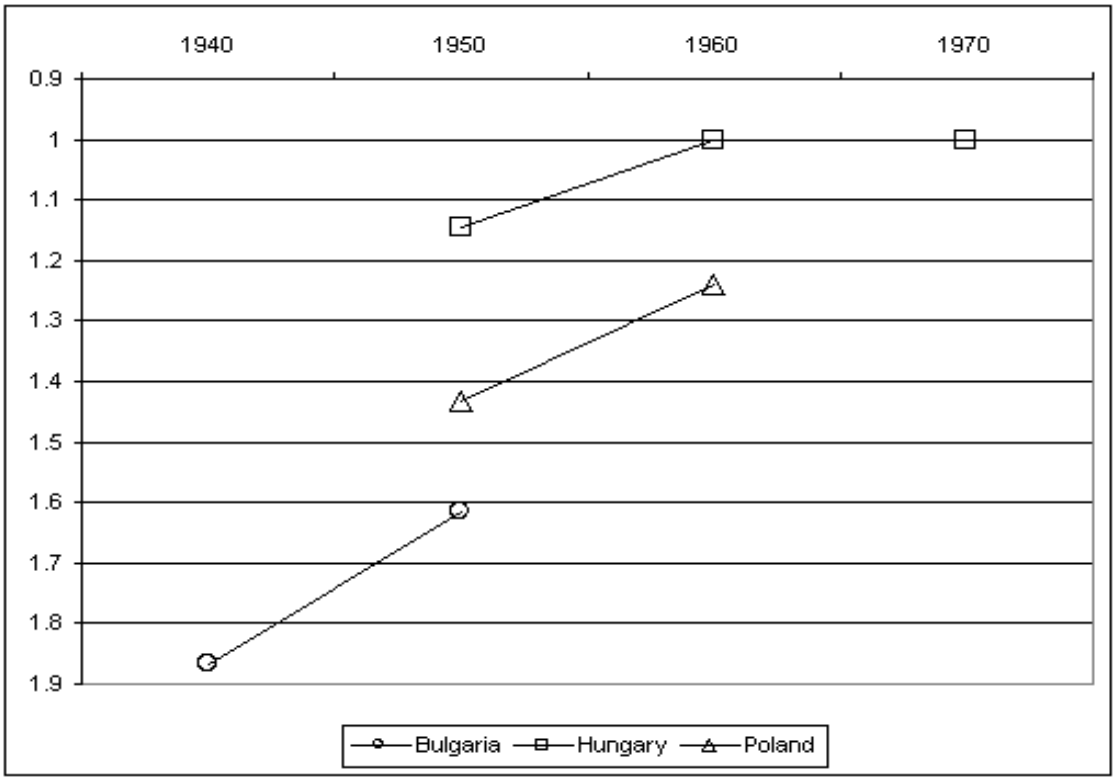
Source: own calculations.

Figure 9: Efficiency in Asia.



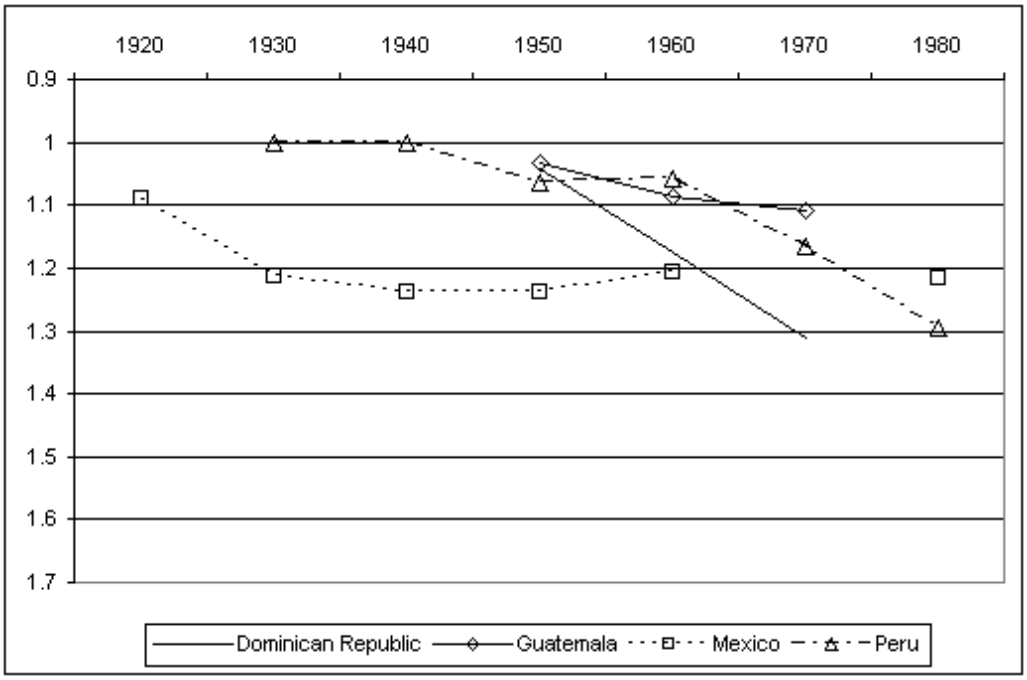
Source: own calculations.

Figure 10: Efficiency in Eastern Europe.



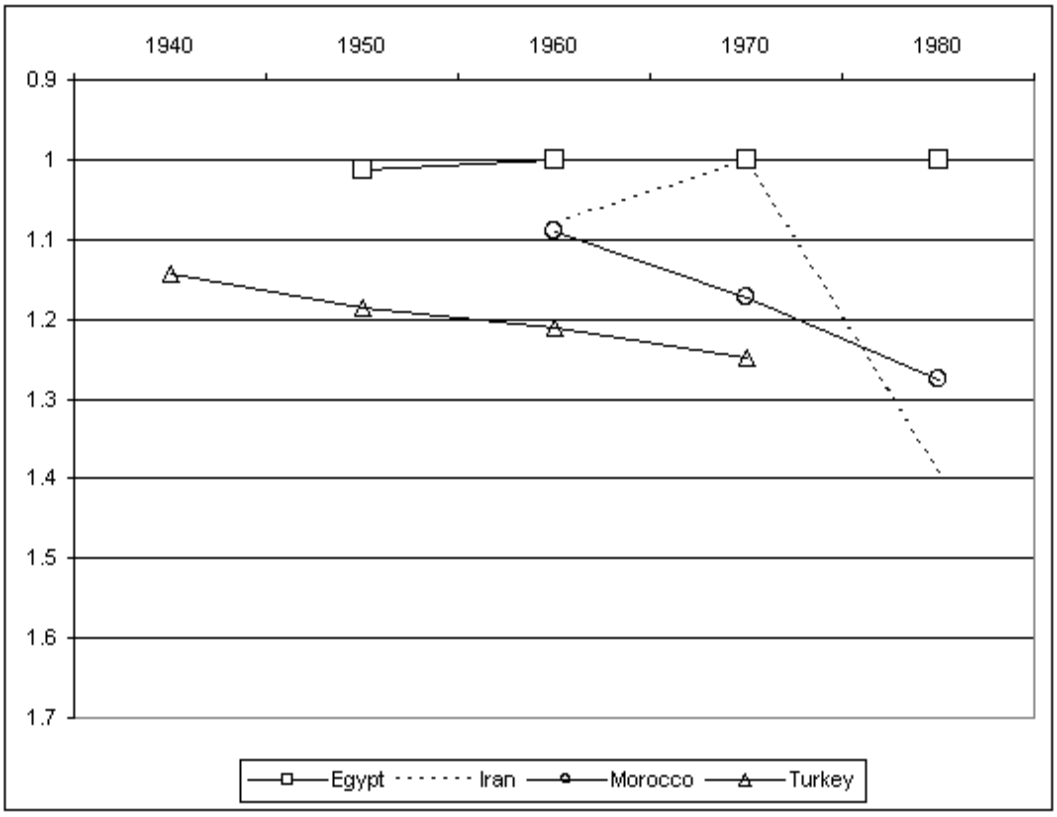
Source: own calculations.

Figure 11: Efficiency in Latin America.



Source: own calculations.

Figure 12: Efficiency in the Middle East.



Source: own calculations.

7 Summary and Research Outlook.

This thesis presented studies concerned with productivity and efficiency in historical settings. The number of firms, the revenue of an industry or the number of workers employed can often be found with relative ease and regularity from census figures or industry reports. These figures are available for the nineteenth century, sometimes further back in time. The US, with regular manufacturing censuses from the beginning of the nineteenth century onward, is one of the best documented countries with respect to this, but also other countries leave quantitative descriptions, sometimes even from colonial bureaucracies.

The historical literature asks questions like “did specialization give larger firms efficiency-advantages” or “was the management in one region better than in the other” which are difficult to answer from census enumerations. It is possible to estimate production functions in an attempt to answer these questions or to seek reports and letters or even newspapers. However, those have several significant disadvantages: the firms that leave details might not do so at random, statements in the press often exaggerate or contain only claims and not the real state of the industry. Production functions are difficult to specify when there are different regions of returns to scale in the industry and the functional form is not known.

The Data Envelopment Analysis, which is a common tool in industrial economics and management science to evaluate the efficiency of firms or branches, is able to circumvent some of the problems mentioned above and give further evidence on the questions posed above. As this method has not yet seen large use in economic history, this dissertation aimed

at paving the road to apply the DEA to some questions not yet answered in the literature and to show its potential for historic inquiry. I showed that the method can be fruitfully employed in the historic literature and can shed light on questions on which only speculations or indirect evidence exist until now .

Therefore, chapter two of this dissertation is concerned with an old question in the history of the United States. Why did the North industrialize while the South remained a more agricultural economy during the late nineteenth century? A number of explanations have been brought forward for this phenomenon and quantitative evidence has been assembled to test these. I add to this literature by directly assessing two hypotheses about efficiency-differences, that have hitherto only been assessed by estimating profit rates. The results show that although the South was among the more inefficient regions, the differences, especially with the Northeast, are too small to explain the difference in the resulting economic structure. It might be the case that inefficient management was a factor that slowed Southern industrial development, but it was surely not the decisive factor.

Chapter three uses the DEA to study the American Civil War. Causes and macroeconomic effects from civil wars have received increasing attention by economists in the last decade and several econometric studies have emerged that are concerned with those effects (Collier 2003). My contribution is to look at efficiency. How large are the disturbances brought in the existing industry through war? My efficiency calculations reveal only small effects from the American Civil War. It seems that the war had even been used to modernize existing machinery, which is confirmed by reports about large firms that were quickly rebuilt and sometimes used imported Northern or even English machinery.

The next chapters assess different, but related topics in economic history, firm creation and labor productivity. They do not use the DEA but more conventional techniques. The results demonstrated the importance of human capital for both processes. Thus, the educational expansion of the nineteenth century was very important for the fast-spreading industrialization and for increasing productivity. Height and health are also shown to be very important for the industry of the nineteenth century.

The sixth chapter, which was concerned with the efficiency of governments, argues that countries differ in their ability to convert endowments into welfare. The estimated efficiency trends cover nearly a century and give the first glimpse that is available on the efficiency of welfare development in the long-run. Wars and occupations were found to be major detrimental events that decrease the efficiency. One result was that fractionalized societies in general have lower efficiency values than other ones.

The different research questions analyzed in this thesis show that there is a wide possibility to use the DEA in economic history. Old claims about the superiority of the industry or economic system of one country may be evaluated by this method. Likewise, the implications of changes in the organizational forms of an economy may be assessed. This has been done in regard to the efficiency of slave agriculture in the Southern United States, but it is surely a promising way for future research to look at other instances where one system of manufacturing or agriculture replaced another. Note that the comparisons with systems from different points in time require a careful handling. But viable approaches for this task are documented in the literature (see for example Brockett and Golany 1996 or Cooper, Seiford and Zhu 2004) and were successfully used in this thesis.

In working with historical data the researcher must always remember that data are subject to errors. Nevertheless, using statistical methods with historical data sets is a necessary and promising way to study history. Statistics can shed light on questions that are not accessible with qualitative methods and random errors do not invalidate the results. Some questions were answered in this thesis, but many more are still unanswered. I believe that the study of efficiency is very promising and its results will provide more important hints for decision makers.

References

Brockett, P. L. and B. Golany (1996). Using Rank Statistics for Determining Programmatic Efficiency Differences in Data Envelopment Analysis. Management Science 42 (3), 466-472.

Collier, P. (2003). Breaking the conflict trap. Civil war and development Policy. World Bank policy research report.

Cooper, W. W., Seiford, L. M. and J. Zhu (2004). Data Envelopment Analysis: History, Models and Interpretations. In Handbook on Data Envelopment Analysis. Boston: Kluwer Academic Publishers.