UNIVERSITY OF OXFORD

Discussion Papers in Economic and Social History

Number 48, September 2002

DOES INDUSTRIALISATION PUSH UP INEQUALITY? NEW EVIDENCE ON THE KUZNETS CURVE FROM NINETEENTH-CENTURY PRUSSIAN TAX STATISTICS

OLIVER WAVELL GRANT

UNIVERSITY OF OXFORD

Discussion Papers in Economic and Social History

1 Hans-Joachim Voth and Tim Leunig, *Did Smallpox Reduce Height? Stature and the Standard of Living in London, 1770–1873* (Nov. 1995)

2 Liam Brunt, *Turning Water into Wine — New Methods of Calculating Farm Output and New Insights into Rising Crop Yields during the Agricultural Revolution* (Dec. 1995)

3 Avner Offer, *Between the Gift and the Market: the Economy of Regard* (Jan. 1996)

4 Philip Grover, *The Stroudwater Canal Company and its Role in the Mechanisation of the Gloucestershire Woollen Industry, 1779–1840* (March 1996)

5 Paul A. David, *Real Income and Economic Welfare Growth in the Early Republic or, Another Try at Getting the American Story Straight* (March 1996)

6 Hans-Joachim Voth, *How Long was the Working Day in London in the 1750s? Evidence from the Courtroom* (April 1996)

7 James Foreman-Peck, *'Technological Lock-in' and the Power Source for the Motor Car* (May 1996)

8 Hans-Joachim Voth, *Labour Supply Decisions and Consumer Durables During the Industrial Revolution* (June 1996)

9 Charles Feinstein, *Conjectures and Contrivances: Economic Growth and the Standard of Living in Britain During the Industrial Revolution* (July 1996)

10 Wayne Graham, *The Randlord's Bubble: South African Gold Mines and Stock Market Manipulation* (August 1996)

11 Avner Offer, *The American Automobile Frenzy of the 1950s* (Dec. 1996)

12 David M. Engstrom, *The Economic Determinants of Ethnic Segregation in Post-War Britain* (Jan. 1997)

13 Norbert Paddags, *The German Railways — The Economic and Political Feasibility of Fiscal Reforms During the Inflation of the Early 1920s* (Feb. 1997)

14 Cristiano A. Ristuccia, *1935 Sanctions against Italy: Would Coal and Crude Oil have made a Difference*? (March 1997)

15 Tom Nicholas, *Businessmen and Land Purchase in Late Nineteenth Century England* (April 1997)

16 Ed Butchart, *Unemployment and Non-Employment in Interwar Britain* (May 1997)

17 Ilana Krausman Ben-Amos, *Human Bonding: Parents and their Offspring in Early Modern England* (June 1997)

18 Dan H. Andersen and Hans-Joachim Voth, *The Grapes of War: Neutrality and Mediterranean Shipping under the Danish Flag, 1750–1802* (Sept. 1997)

19 Liam Brunt, *Nature or Nurture? Explaining English Wheat Yields in the Agricultural Revolution* (Oct. 1997)

20 Paul A. David, *Path Dependence and the Quest for Historical Economics: One More Chorus of the Ballad of QWERTY* (Nov. 1997)

21 Hans-Joachim Voth, *Time and Work in Eighteenth-Century London* (Dec. 1997)

22 Tim Leunig, *New Answers to Old Questions: Transport Costs and The Slow Adoption of Ring Spinning in Lancashire* (Feb. 1998)

23 Paul A. David, *From Keeping 'Nature's Secrets' to the Institutionalization of 'Open Science'* (July 2001)

24 Federico Varese and Meir Yaish, *Altruism: The Importance of Being Asked. The Rescue of Jews in Nazi Europe* (May 1998)

25 Avner Offer, *Epidemics of Abundance: Overeating and Slimming in the USA and Britain since the 1950s* (Nov. 1998)

[Continued inside the back cover]

DOES INDUSTRIALISATION PUSH UP INEQUALITY? NEW EVIDENCE ON THE KUZNETS CURVE FROM NINETEENTH-CENTURY PRUSSIAN TAX STATISTICS[1]

OLIVER WAVELL GRANT

(Nuffield College, Oxford)

[1] This paper contains material from chapter 8 of my thesis, *Internal Migration in Germany 1870–1913* (Oxford D.Phil. 2000). Previous versions have been given at seminars at Oxford and Cambridge Universities. I am grateful to participants at these seminars for helpful comments and suggestions. I would also like to thank Charles Feinstein, Anthony Atkinson, Hans-Joachim Voth and Hartmut Pogge von Strandmann for comments on earlier drafts. I am also grateful to Eric Grimmer-Solem for having introduced me to the work of the German historical economists

Abstract

This paper presents new estimates of income inequality derived from Prussian tax statistics for the years 1822-1914. Confidence intervals are also calculated. The results show a rise in inequality in the nineteenth century, with a peak around 1906, thus supporting the view put forward by Simon Kuznets that industrialisation will initially lead to a rise in inequality. The paper goes on to consider whether this was due to factors which were particular to Germany in the period, or whether the Kuznets curve is the result of forces which affect all industrialising societies. The conclusion reached is that the Kuznets curve is an avoidable trap, not an automatic consequence of industrialisation.

I. Introduction

The idea that, during industrialisation, the relationship between growth and inequality follows the path of an inverted "U", with inequality first rising and then falling, has long been associated with Simon Kuznets, whose 1955 article gave it wide currency. However, the idea that the distributional consequences of industrialisation might not be favourable was certainly not a new one. Apart from its role in the socialist critique of capitalism, it was also a concern of the German historical school, and discussion of income distribution was an important component in the German debate over *Manchestertum*: the pros and cons of laissez-faire and free trade economics. Evidence that the income distribution was worsening fuelled fears that the lower and middle classes were not benefiting from industrialisation to the same extent as the rich.[2]

Kuznets was influenced by the same evidence, drawn from Prussian tax records. The Soviet economist Prokopovitch had used these in a study of income distribution, and, at the time of the 1955 article, this was the only empirical support that Kuznets had for the initial rise in inequality. Later Kuznets made a cross-section study of inequality in developing countries which appeared to support his hypothesis. However, this result has been disputed. Some studies of contemporary evidence from developing countries have not found an inverted U-shaped curve, so that one participant at a conference was led to declare: "the Kuznets curve is a fiction".[3] However, more recent studies have found evidence which provides limited support to the Kuznets hypothesis: a "conditional Kuznets curve", a statistical relationship which emerges from regressions when the influence of other factors is controlled (such as education, cohort size, and "openness").[4]

Although cross-sectional or panel data studies using modern data are one way to test the hypothesis, the results show that the Kuznets relationship (if it exists) explains only a relatively small portion of total variation in inequality between countries: some countries are on a much higher path than others, and show little

---

[2] See, for example, Engel (1875), Schmoller (1895) and (1895a).

[3] Dani Rodrik, the discussion was published in Bruno et. al. (1996), the remark is on page 159. For an example of the studies which led to this remark, see Anand and Kanbur (1993).

[4] Barro (1999) finds a Kuznets curve when the estimated regressions include education variables and dummies for Africa and Latin America; Higgins and Williamson (1999) find one when controlling for cohort size and "openness" (the Sachs-Warner index for classifying economies as closed or open according to the extent of tariff barriers and other obstacles to trade).

sign of converging. The results also indicate that the relationship may not be a particularly stable one: the values of continent dummies are not constant over time, indicating that some other factor may be pushing inequality up or down.

One influence could be shifts in the world terms of trade between industry and agriculture. An unanticipated shift which had a damaging effect on agricultural incomes would have little effect on internal inequality in a rich, predominantly industrial economy (all would gain), and little effect on a poor, predominantly agricultural economy (all would lose), but would tend to push up inequality in a middle income country with substantial employment in industry and agriculture (where the poorer agricultural sector would be the main loser). The Kuznets curve would appear or disappear with shifts in the world terms of trade. Omitted variables like this could explain why modern econometric studies have produced results which are erratic or inconsistent.

An alternative approach is to study the evolution of inequality in individual countries over long periods, trying to identify the factors which may or may not be affecting the movement of inequality. There are a number of studies of this type; this paper is an addition to their number.[5] One problem is that many of these studies make use of data sources which only became available once a relatively high level of economic development had been attained. In other words, they provide evidence on the downward phase of the Kuznets curve, not on the upwards swing in the earlier stages of industrialisation. This paper, by contrast, focuses on the earlier phase.

[5] Other examples include Atkinson (2002), Piketty and Saez (2001) and Fei et al. (1979).

II. The Prussian tax data: Problems and opportunities

There are three main problems with the use of tax data to examine income inequality. Firstly, part of the population is normally exempt from the tax, due to income levels below the tax threshold, or some other reason for exemption. So, coverage is inevitably incomplete. Secondly, as the income assessment is linked to the tax payment, there is an obvious incentive to minimise declared income and conceal sources of income wherever possible. Some downwards bias must be anticipated. Thirdly, the tax system tends to change over time, which has effects on coverage, on incentives to evade, on definitions of income, on procedures for checking returns and so on.

These are serious problems, but not insuperable. There are two ways of dealing with the first problem. One is to divide the income received by different groups of taxpayers by total personal income, obtained from national income accounts. This is an approach used in modern studies by Atkinson and Piketty.[6] It has been applied to the Prussian income tax data by Geisenberger and Müller.[7] There are, however, a number of problems with this approach. The main one is that can only be used for periods for which national income accounts are available, so the Geisenberger and Müller study starts in 1874. The second is that the income side of the national income accounts is generally the least reliable. Data on physical production are normally available from an earlier stage than information on wages, rents and profits. Hoffmann's estimates of income from capital and land in the nineteenth century were derived using methods which have come in for criticism.[8] Moreover, for a series which is derived from Prussian tax statistics, the divisor should be based on income accounts for Prussia, and these are problematic, as there is no information on inter-regional trade or transfers of property income (post-1871 Prussia contained 61–62% of the total German population).

A different method, which is the one used here, is to find other sources of information on the incomes of those excluded from the tax system. In nineteenth-century Prussia, the most important low income group was the landless agricultural labourers. Information on these, and on incomes in low paid non-agricultural occupations, can be used to construct an overall distribution. This in turn can be used to derive a gini-coefficient, or other measure of general ine-

---

[6] Piketty and Saey (2001), Atkinson (2002).

[7] Geisenberger, S. and Müller, J.H. (1972).

[8] Problems with Hoffmann's figures are discussed in Fremdling (1988) and (1995).

quality, whereas the first approach only yields a measure of the share of top taxpayers in total income.

The second problem, the problem of evasion, is inherent in the use of income tax data.[9] It is a serious source of possible error, particularly in cross-section comparison with other countries (where enforcement procedures may be different). It is a problem in a time series for a single country if there are changes in the proportion of income which is not declared, as a result of improved methods of oversight for example. It is clear from contemporary comment on the tax system, that there were changes of this type.[10] The reform of the Prussian tax system in 1891–2 led to a more rigorous examination of tax returns. It is important to make some allowance for these changes.

The tax system was not constant in Prussia in the period, but evolved in stages from a *Klassensteuer,* under which households paid a fixed charge, reviewable every three years, based on an assessment of ability to pay (which took account of property owed, and indebtedness as well as expected income), to a modern income tax system. Comparability between the different systems is not to be expected.

These last two problems were addressed by producing estimates of income inequality for periods when the tax system was not subject to major changes (1822–46, 1853–73, 1874–91 and 1892–1914), and then looking for contemporary estimates or other data sources to chain the series together. This proved to be possible for 1873–4 and 1891–2, but not for 1846–53. During this period the *Klassensteuer* went through a series of changes: urban areas which previously paid a different tax were brought into the system and the top bands were changed to a system more explicitly linked to income. It was not possible to find evidence from other sources which could cover this period.

In addition to selecting an approach which should minimise potential sources of error, an attempt has been made to quantify the effect of those errors which remain. By constructing different variants for each period, based on alternative sets of assumptions, it was possible to examine the sensitivity of the figures to the assumptions used in the central estimates, and find extreme values produced by a range of possible assumptions. These variants were then used to calculate

---

[9] For a more detailed examination of the pros and cons of tax data as a source of information on income distribution, see Atkinson (2002).

[10] Meisel (1911) gives data on the number of returns rejected or challenged under the post-1892 system. On average, a challenge to an assessment led to an increase of 25–30% in the amount of taxable income. See also Wagner (1880) p. 612, on the pre-1892 system.

variance estimates which were, in turn, employed to construct confidence intervals for the central inequality measures. [11]

The approach is one which is appropriate for a data source which has serious deficiencies by modern standards. It generates a range of probabilities as well as a single set of point estimates. Taking as an example the weakest period, 1822–46, the tax system in these years was only indirectly related to income and coverage was not universal. Nevertheless, the data do provide certain pieces of evidence. The numbers in *Stufe 1*, the highest band, rose by 60% between 1822 and 1846, while the total population in the areas subject to the tax rose by 39%. The numbers exempt from the tax on the grounds of having insufficient property or income rose steadily, from 36.6% of the population in 1822 (including dependents) to 41.5% in 1846.[12] These two factors strongly suggest a hollowing out of the distribution, with numbers rising at each extreme.

Moreover, there is evidence from other sources which suggests that real wages for those at the bottom of the distribution were falling in this period. Table 1 gives an index derived from figures for the earnings of day-labourers employed in the Prussian state forests, deflated by local food prices. This should be a good indicator of the level of incomes received by landless rural labourers.

Table 1. Rural wages in Prussia 1820/9–1875/9 (post 1815 boundaries)

Index based on earnings of male day-labourers in the Prussian state forests (1875–9 = 100), deflated by local food price indices.[13]

1820–9	104.6
1830–9	101.1
1840–9	88.7
1850–9	81.9
1860–9	91.4
1870–4	91.3
1875–9	100.0

[11] This follows the approach set out in Feinstein and Thomas (2001).

[12] The figures of those exempt has been adjusted to allow for the effects of an extension of exemptions to those aged 14-16 and over 60 in 1831.

[13] Earnings in Prussian state forests come from Eggert (1883); this gives the results produced by a circular sent to all *Oberförster* in November 1879 asking for the movement of wages since 1800. The survey was not repeated.

Taking these factors into account, the balance of evidence strongly suggests a worsening of the income distribution in the period. But by how much? This is less easy to say. The variance calculations produce a 95% confidence range for the increase in the gini-coefficient of 12.2% to 16.2%, assuming that the central estimates are subject to uncorrelated errors. But examination of the highest and lowest variants (considering the possibility that there are correlated errors produced by incorrect assumptions), and applying variance calculations to these estimates produces a much wider range: an increase of between 4.3% and 21.4%. This demonstrates that the statement that inequality worsened in the period can be made with a high degree of certainty, even though the exact scale of the increase is less easy to determine.

Thus, for all these caveats and notes of caution, the Prussian tax data do present a major opportunity, provided that the data are handled with caution For 91 years, 1822–1913, the Prussian population paid some form of income-related tax. Coverage was, at times, remarkably high (74% in 1891). The period covers a society going through a process of transformation from an essentially rural, pre-industrial economy in the 1820s to a position as one of the world's most advanced industrial economies in 1913. Moreover, and in contrast to the twentieth century, there were fewer unexpected shocks to the income distribution: world wars with concomitant rises in taxation and prices, major world depressions on the scale of the 1930s, unanticipated periods of high peacetime inflation as in the 1970s. It is a unique opportunity to study the evolution of inequality in an industrialising society.

III. Estimates of income inequality from Prussian tax statistics

(a) *Prokopovitch and the origins of the Kuznets curve*

In his original article Kuznets made use of a study of the Prussian tax statistics by Prokopovitch. This showed that the share of the top 5% of the Prussian income distribution rose from 21% of total income in 1854 to 26% in 1875, and 30% in 1913. The share of the top quintile rose rather more slowly: from 48% in 1875 to 50% in 1913. In his discussion of the course of income inequality Kuznets referred to US data which showed falling inequality in 1929–1944/50, and British figures showing falling inequality in 1880–1910/13. Only the Prussian results showed a period of increasing inequality. Kuznets thought that inequality had risen in 1780–1850 in Britain, 1840–1890 in the United States and 1840/9 to 1890/9 in Germany, but he had no data for Britain or the United States in these periods with which to prove this. The case for the upwards phase of the Kuznets curve rested on the German data.

Prokopovitch's analysis has some important advantages over previous work in this area. This had tended to look at the movement of the numbers of individuals in different income classes: thus, a faster increase in the numbers earning high incomes than in the low income classes would be thought to represent an increase in inequality.

But as Prokopovitch pointed out: "an income of 1,000 Marks in 1913 is different from one of the same amount in 1896".[14] It is different both in terms of real purchasing power, and also as a share of total national income. Prokopovitch re-calculated the incomes of the different groups in what he called "budget units, that is in terms of the average income per head of population".[15]

This is essentially the approach used in modern studies. The basis for any discussion of income distribution is a table giving the share of total income going to different groups, these groups being sorted according to their income level. This means that the statistician has to make a further assessment of the size of total national income. The tax statistics will show that a number of individuals, N, in a given income class, had average income X: what is also needed is a figure for total income, Y, so that the proportion (XN/Y) can be calculated.

[14] Prokopovitch (1926) p.71.
[15] Ibid. p.71.

This obviously introduces a further source of uncertainty and potential error. It is, however, the only way to produce a consistent series of estimates of income distribution over time. Prokopovitch was one of the first to recognise this.

There are, however, some serious problems with Prokopovitch's figures: they only cover a few years, and there were major changes in the Prussian tax system between these years. These are large sources of potential error.

(b) *German studies since Prokopovitch*

The main contributions to the study of German income inequality since Prokopovitch have come from statisticians associated with the preparation of Walter Hoffmann's series of national income accounts. Hoffmann included figures for inequality in *Das Wachstum der deutschen Wirtschaft.*[16] Geisenberger and Müller published estimates for the share of the top 5%, 1% and 0.1% of households in total household income for Prussia 1875–1913. Hoffmann's figures were used by Dumke in a study published in 1991. [17]

The figures are based on a comparison of the incomes received by the top tax payers, derived from the tax records, with total income figures taken from Hoffmann's accounts. Because of comparability problems arising from changes in the tax system, only data on incomes at the upper end of the distribution were used. But this means that a lot of information was discarded. In some years (1891 and 1913) over 60% of Prussian households paid some form of income-related tax. There are additional data which could be used if the comparability problems could be overcome.

Ignoring changes in the lower part of the distribution is not satisfactory. The movement of rural labourers into urban employment will have effects on the lower end of the distribution which will be picked up only indirectly (as a general increase in total household income). A further problem with these studies is the use of Pareto's α as a measure of inequality.[18] This is not a measure much

---

[16] Hoffman's co-authors included Franz Grumbach, whose doctoral dissertation was one of the sources used in the section on income distribution.

[17] Geisenberger and Müller (1972) Dumke (1991).

[18] Pareto's α is a simple measure of inequality. If a series of points representing the numbers of tax-payers for a range of minimum income levels are graphed on double-log paper and a line drawn through these points, the slope of this line is Pareto's α. If only two points are available then the calculation is even simpler. If there are X taxpayers with at least Z income and Y with at least W income then $\alpha = (\log Y - \log X)/(\log Z - \log W)$. The index rises with increased equality, and has a range of 0 (perfect inequality) to infinity (perfect equality).

used today and it has the problem of being sensitive to the form in which the date is entered.[19]

In 1970 Jeck published a study of income distribution which mainly made use of data for areas outside Prussia. The Saxon tax data are particularly comprehensive. The income tax threshold was relatively low, there were fewer changes in the system, and there are also figures on sources of income: employment, enterprise profits, savings, and property ownership. These figures have also been used by Hentschel and Kiesewetter.[20] However, Saxony was a small and not very typical part of Germany, which industrialised early and had a relatively equal land distribution. Given the importance of Prussia, with 62% of the population in 1907, it is clear that any study which attempts to draw conclusions for Germany as a whole must make use of the Prussian data.

More recently German studies have tended to look at other aspects of inequality, and there has been little work on income distribution.[21] One reason is that views have been expressed which are sceptical of the likelihood of obtaining reliable estimates of income inequality from tax statistics. Kaelble, in particular has raised two main objections: they require the use of extrapolation and interpolation, which is not reliable, and the data source does not cover the incomes of those below the tax threshold or any undeclared income.

These are undeniable problems, and some way has to be found to deal with them. As far as the use of extrapolation or interpolation is concerned, this is a problem inherent in the use of this source. Tax thresholds produce grouped data with groups of varying sizes: one year a certain number of taxpayers came into a given band, the next year it will be a different number. This, however, is not a problem only encountered with tax data. All comparative studies of income distribution make use of surveys from a range of sources, and, unless the original records are used, these will be available in grouped form. The Deininger and Squire data set, compiled by the World Bank, made use of a number of different studies of inequality, which were mostly grouped data. Gini coefficients and other measures of inequality were then produced using a computer programme,

[19] Dumke argues that it will be particularly sensitive to inequality in the upper end of the distribution. This is especially true when the only available data comes from this end of the distribution. In effect, the shape of the upper end of the distribution is used to predict the whole distribution, Dumke (1991) p.128.

[20] Jeck (1970), Hentschel (1979) and Kiesewetter (1988).

[21] Kaelble (1986) surveys work up to that date.

POVCAL, designed to deal with grouped data.[22] This data set has been used in a number of cross-section studies of inequality. If extrapolation is rejected then much modern work on inequality would have to be jettisoned.

This is not to deny that extrapolation can cause problems. The procedures used should recognise this and seek to reduce its impact.

The problem of non-taxpayer incomes requires the use of some additional source of information to cover this gap. Wage data provide one possible source. There are also contemporary studies which made estimates of the incomes of those who were exempt from taxation, because their income fell below the threshold, or for some other reason. The analyses by Biedermann and Helferrich are the most useful of these. Helferrich also makes an allowance for evasion, and for the non-taxation of company income transferred to reserves. [23]

When allowing for the effect of these problems, it is important to consider the use that is intended for the resulting estimates. If the study is a cross-section one, then problems such as tax evasion can be very serious. This may be much higher in some countries than in others. If there is some bias in its incidence (being more prevalent either in higher or lower income groups) then this will affect the comparison. If the intention is to study the movement of income distribution in one country over time, then the problem only matters if there is a bias and this alters over time: evasion by higher income groups rises more than evasion by lower income groups, for example. While this is not impossible, the scale of the effect is likely to be reduced.

(c) *The* Klassensteuer *1822–1846*

The value of the Prussian tax system is that it was in operation for a long period of peace, so that income distribution can be studied over 92 years, from 1822 to 1914. It did, however, change quite considerably during this period. Between 1822 and 1891 a *Klassensteuer* was levied, which classified the population into groups and levied a fixed rate of tax on each group. The classification was

[22] The gini-coefficient is a measure of the extent to which the income distribution differs from perfect equality, all divergences being given equal weight. It is, in effect, an average of the absolute values of all deviations from the average income level, expressed on a scale of 0 to 1,where 0 is perfect equality (all incomes are the same) and 1 is perfect inequality (one person has all the income).

[23] Helferrich (1915), Biedermann (1918); other data sources used in this study are Engel (1868) and (1875), Kuhnert (1911) and (1916) and Wagner (1904) and (1904a).

according to an estimate of ability to pay, "*Wohlstand und Steuerfähigkeit",* based on the type of occupation and the amount of property owned. [24]

Over time the *Klassensteuer* evolved into something much closer to an income tax, with explicit income bands. This revealed that it was a mildly progressive tax, with a cut-off point at the top (nobody paid more than 432 Thalers). In the 1822–1846 period it was supplemented by a second tax: a *Mahl- und Schlachtsteuer* (a tax on flour milling and animal slaughterings) which was imposed on the larger towns. These were then exempt from the *Klassensteuer.* Over the 1822–46 period the population covered by the *Klassensteuer* was, fairly consistently, around 85% of the total Prussian population. Between 1847 and 1852 the exempt towns were brought into the *Klassensteuer* and this means that it is impossible to produce consistent results for the 1847–52 period. There is a break in the income distribution estimates which re-start in 1852.

The results of the estimation procedure are given in table 2. The basic estimates allowed for a fall in the real earnings of the exempt group, in line with the fall given in table 1. The earnings of other groups were assumed to be static, and to correspond to the income bands introduced in the 1850s. In addition to the basic estimates, five other series were also constructed, which incorporated possible errors in the assumptions made about the movement of incomes in the exempt group (which could have been static rather than falling), household size, and the numbers who were exempt because of age. These were then used to generate variance estimates for the estimation of confidence intervals.[25]

---

[24] "Well-being and ability to pay tax": the term is taken from Wagner (1904) as are most of the details about the various tax systems; see also Witt (1970) for a description of the systems.

[25] Variance estimates were derived by pooling the different variants, and calculating the overall variance of these. These were then used to derive the confidence intervals for the individual years, and also the confidence intervals for the rates of increase. The procedure is given in Feinstein and Thomas (2001) and also in Bowley (1946).

Table 2. Estimates of income inequality in Germany 1822–1846.[26]

(a) *Estimates for different years*

	1822	1826	1831	1836	1841	1846
Gini-coefficient as indices (1846=100)	0.260	0.266	0.274	0.278	0.286	0.297
Gini-coefficient	87.6	89.4	92.3	93.5	96.4	100.0
Upper 95% CI of Gini-coefficient	96.5	98.5	101.7	103.1	106.2	110.2
Lower 95% CI of Gini-coefficient	78.7	80.3	82.9	84.0	86.6	89.8

(b) *Estimates of the increase 1822–46*

Increase in Gini-coefficient 1822–46 (%)	14.2
Upper 95% CI of increase	16.2
Lower 95% CI of increase	12.2
Upper 95% CI of highest variant	21.4
Lower 95% CI of lowest variant	4.3

The first part of the table gives the figures for different years. As can be seen, the confidence intervals are fairly wide, reflecting the uncertainties surrounding these estimates. The second part of the table looks at the increase in inequality over the whole 1822–46 period. Confidence intervals have been produced for the rate of change, by combining the variances of the two years. These show that, if the errors are not correlated, or systematic, then there is a relatively narrow confidence interval around the central estimates for the different measures.

More serious problems arise if there has been an error which has a systematic effect, and which is correlated between years: an error is made which biases the

[26] Calculated from data in Wagner (1904) pp. 93-96. It has been pointed out by Atkinson that there is no ideal instrument for the measurement of inequality. It is only when the Lorenz curve of one distribution lies entirely to the left or right of another that an unequivocal ranking is obtained. The gini-coefficient, which effectively applies an equal weight to all deviations from perfect equality, is one commonly-used measure, but it is quite possible to argue that some deviations, for example those which affect the bottom end of the distribution, should be given greater weight than others. The main advantage of the gini-coefficient is that, as a commonly used measure, it provides a useful basis for comparison.

estimated growth rate up or down and which is then repeated. This has been estimated by taking the variant which shows the lowest growth rate over the whole period, and then calculating confidence intervals for this figure. The results are shown in the final line. The rise in inequality is greatly reduced, and not really that important.

The conclusion from this analysis is that, using the available evidence, the balance of probabilities favours the view that there was a rise in inequality in 1822–46, but the possibility that there is some systematic error in the estimates which makes this conclusion unsound cannot be completely excluded, given the problems inherent in the data.

(d) *Combined system 1853–73*

The revisions to the tax system in 1847–53 made the data, in some respects, easier to use. There was now an income tax imposed on higher incomes, and the *Klassensteuer* had explicit income bands: the classification was intended to correspond to a certain level of income, attainable in normal years, even though the assessment was not varied in response to annual income fluctuations. There were, however, two changes which were not so helpful. The first was a raising of the threshold so that around 45% of the population was now exempt; the second was a change in the reporting which makes it more difficult to obtain a household distribution. The pre-1846 results are divided according to whether or not the taxpayer had dependents; this no longer applied after 1853.

The larger numbers of exempt households makes it important to include accurate figures for their income. In this period agricultural wages were tending to rise. The estimates given in table 1 show a real increase of 11% between the 1850s and the early 1870s. The available wage data also show increases.[27] The view that the position of the lower income groups was improving also finds support from the fact that there is a slow reduction in the size of the exempt population, from 46.7% in 1853 to 44.0% in 1873.

More households had incomes which were above the *Klassensteuer* threshold.

At the same time the numbers of high earners was increasing rapidly. The numbers with incomes of over 12,000 Marks rose from 3349 in 1853 to 11,513 in 1873; an increase of 244%. These were the figures which alarmed contemporary observers, such as Engel. However, there was a problem here, which re-

[27] Hoffmann (1965) pp. 492–3 and pp. 468–9 gives figures for various sectors, including agriculture.

sulted from the hybrid nature of the system. The period was one in which there was some inflation and money incomes increased rapidly. Hoffmann's deflator for private consumption rises by 32% 1853–1873; his estimates of money earnings of workers in agriculture and industry show increases of 69% and 103%.

The upper income groups paid income tax, levied on actual money earnings. The lower income groups paid the *Klassensteuer*, paid on an assessment of *steuerfähigkeit* (ability to pay), which was a fixed sum for each *Stufe*, though the allocation of tax-payers to the different *Stufen* was subject to revision. The question is: how well did this assessment procedure keep up with the increase in money wages? If it did not keep up there would be systematic bias, as the earnings of the upper income groups would be fully recorded, but there would be a increasing tendency to under-record the earnings of those in the lower and middle income groups.

There are two general points which can be made. Firstly, one reason why *Klassensteuer*-type taxes (which are cheap to administer) are rare, while most states have some form of income tax, is that it is inevitable that the *Klassensteuer* machinery will creak a bit, and there will be some under-assessment at times of rising money wages. Secondly, this implies that the Prussian tax system was becoming more progressive (the upper limit had been abolished) and this was not the intention of the Prussian government. So, it seems reasonable to assume that the Prussian administration would have attempted to keep the *Klassensteuer* assessment in line with income, even if this was hard to achieve in practice. This was, presumably, one reason why the *Klassensteuer* evolved into a tax which was much closer to an income tax.

The base estimates allow for an increase of the incomes of the exempt of 16% in 1853–73 , from 186 Marks to 216 Marks, which still leaves them below the threshold of 300 Marks (or rather 100 Thalers as the Mark was not in use). There is no allowance for drift in the incomes of those subject to the *Klassensteuer*. Household size was allowed to increase with income. [28]

[28] This is a point which is shown by contemporary studies, such as Soetbeer and Biedermann, and by the evidence of the pre-1846 *Klassensteuer*. The base estimates assumed that the average household size of the exempt was 1.3 persons, 1.7 for *Klassensteuer Stufe* 1, and 2 for all others. The variant estimates changed this to 1.2, 1.6 and 2.1. It does not affect the movement of inequality over time, but it does affect the reliability of the individual estimates

Table 3. Estimates of income inequality in Germany 1853–73[29]

(a) *Estimates for different years*

	1853	1855	1858	1861	1864	1867	1870	1873
Gini-coefficient	0.345	0.348	0.352	0.353	0.354	0.355	0.357	0.363
as indices (1853=100)								
Gini-coefficient	100.0	100.9	102.1	102.3	102.6	102.8	103.4	105.2
Upper 95% CI of Gini-coefficient	111.7	112.7	114.1	114.3	114.6	114.9	115.5	117.5
Lower 95% CI of Gini-coefficient	88.3	89.0	90.1	90.3	90.5	90.7	91.2	92.8

(b) *Estimates of the increase 1853–73*

Increase in Gini-coefficient 1853–73	5.2
Upper 95% CI of increase	6.0
Lower 95% CI of increase	4.3
Upper 95% CI of highest variant	20.2
Lower 95% CI of lowest variant	3.5

Estimated variants allowed for upwards drift in *Klassensteuer Stufen* 1 and 2. These were then used to generate the confidence intervals used for table 3, and the systematic bias variant used in the bottom line of the table. In general, this has no great effect on the inequality estimates, which is mainly because the *Klassensteuer* payers were already in the middle of the distribution, and raising their incomes does not have much effect on overall inequality.

The results show a gradual increase in inequality, though it is not as dramatic as the one which contemporaries thought was occurring. This is mainly due to the rise in incomes attributed to the exempt. Ignoring this factor produces a much greater increase.

Equally, on these figures it appears that there was some increase in inequality even when the possibility of systematic bias is introduced. There is, however, one possibility which would completely overturn this conclusion. That would be if there had been a much larger increase in the incomes of the exempt,

[29] Calculated from data in Wagner (1904) pp. 97-100. Between 1867 and 1870 there were territorial changes (due to the incorporation of Schleswig-Holstein into the Prussian tax system) which do not seem to have had much effect on overall inequality.

which would have taken them into areas where they should have been paying the *Klassensteuer*. It is quite easy to construct a variant on this basis which shows static or even falling inequality.

This variant has not been used. It implies a massive fall in the efficiency of the tax collection system, so that, by 1873, quite large numbers who should have been paying tax were not doing so. This would be a different point to the one which is implied by a rise in under-payments attributable to inefficiencies inherent in the *Klassensteue*r system. It is not inconceivable, but, on balance, unlikely.

(e) *Soetbeer's estimates*

Between 1873 and 1875 there was another revision of the tax system which would have caused a gap in the series if there had not been contemporary estimates of the income distribution by Soetbeer which make allowance for the effect of the revision. The figures give average incomes, the numbers of tax payers, and the numbers of taxpayers plus dependents for 6 different income classes. While this is nothing like the detail available for other years (the published returns for 1873 give numbers for 13 *Klassensteuer* bands and 30 income tax bands) it is still enough to produce income distribution estimates.[30]

These were used to produce the figures given in table 4. As only three years are covered, the full analysis of likely errors was not applied. The results have been chained to the earlier series to create an index with 1853 as the base year.

Table 4. Estimates of income inequality in Germany 1873–5

	1873	1874	1875
Gini-coefficient	0.315	0.325	0.318
as indices with 1853=100			
Gini-coefficient	105.2	108.5	106.2

[30] The only alteration to the data as published in Soetbeer (1882) was to revise downwards his figure for the average income of households in the lowest band (income below 540 Marks). Soetbeer uses a figure of 400 Marks, which is high relative to the threshold and which produces estimates of inequality for 1873 and 1875 which are out of line with those produced using different materials. The figure used instead was 270.

(f) ***Combined system 1875–92***

The changes made in 1873–5 had the effect of bringing the two tax systems closer together: the upper two bands of the Klassensteuer overlapped with the two lowest bands of the income tax system. There was also a change in the treatment of dependents which meant that they were more likely to be counted as part of the household rather than as exempt individuals. This had an effect on the size of the exempt population and on the average household size (important as a means of estimating a household distribution from tax payer data). The result of this was that there was a fall in the numbers exempt from 44.0% of the population in 1873 to 26.9% in 1875.

This in turn poses problems of comparability between the two systems, which is one reason why Soetbeer's estimates, which create a bridge between the two periods, are so valuable. The main effect of the change in classification was to reduce the numbers at the bottom of the distribution as many of those previously recorded as exempt on grounds of low incomes were now put down as dependents of better off taxpayers.

Comparison of the basic, unadjusted results, shows that Soetbeer's data, which follow the post-1875 system, produced a lower estimates of inequality for 1873 than those given in table 3. The adjustment produced by the classification change made the overall estimates rather lower.

Table 5 provides estimates of inequality in the 1875–91 period. The household distribution was obtained by using Soetbeer's figures for household size, for the different income classes. The calculated variants allowed for some variation in this. Over this period wages first fell and then rose. According to Hoffmann agricultural wages fell by 15% in 1875–9, and rose 5% in 1879–91, remaining below their 1875 peak. Hoffmann's figures for industrial wages show a fall of 17% in 1875–9 and a rise of 40% in 1879–91, finishing above the 1875 level. Desai's figures (which use different sources) show a fall of 21% in 1875–9 followed by a rise of 20%. Hourly earnings of unskilled building workers in Berlin, Rostock and Nuremberg fell by 16% in 1875–9 and subsequently rose by 34%.[31] The tax returns support the view that there was an inverted U-shaped movement. The exempt population rose from 26.9% of the total in 1875 to 30.0% in 1882 and then fell to 25.8% in 1891.

The base estimates allow for a movement in earnings of the exempt population which follows Hoffmann's figures for agriculture. There is no movement in

[31] Figures from Hoffmann (1965) pp. 492–3, Desai (1968), and Bry (1960) pp. 334–5. Annual movements given by Hohls (1995) use the same source as Desai.

the earnings of any of *Klassensteuer* classes. Variants include allowing for no movement in the earnings of the exempt population, and allowing the earnings of the lowest *Klassensteuer* class to follow Hoffmann's index of industrial earnings.

Table 5. Estimates of income inequality in Germany 1875–91[32]

(a) *Estimates for different years*

	1875	1876	1878	1880	1882
Gini-coefficient	0.340	0.341	0.357	0.371	0.373
as indices (1853=100)					
Gini-coefficient	107.1	107.6	112.5	116.8	117.7
Upper 95% CI of Gini-coefficient	112.1	112.7	117.8	122.3	123.3
Lower 95% CI of Gini-coefficient	102.1	102.5	107.2	111.3	112.2

	1883	1885	1887	1889	1891
Gini-coefficient	0.370	0.374	0.374	0.379	0.382
as indices (1853=100)					
Gini-coefficient	116.7	117.9	118.0	119.5	120.5
Upper 95% CI of Gini-coefficient	122.2	123.5	123.6	125.1	126.1
Lower 95% CI of Gini-coefficient	111.2	112.4	112.5	113.8	114.8

(b) *Estimates of the increase 1875–91*

Increase in Gini-coefficient 1875–91	12.5
Upper 95% CI of increase	13.3
Lower 95% CI of increase	11.6
Upper 95% CI of highest variant	13.6
Lower 95% CI of lowest variant	5.7

The resulting calculations show an increase in inequality. The lowest variant includes the static estimate for the earnings of the exempt. Even with this variant there is a rise in inequality.

[32] Calculated from data in Wagner (1904) pp. 101–122.

(g) *Chaining 1891–2: estimates based on Geisenberger and Müller*

Between 1891 and 1892 the Prussian tax system underwent major reforms. The *Klassensteuer* was abolished and replaced with an income tax system with a much higher threshold: 900 Marks compared to the previous figure of 420 Marks. This meant that a much higher percentage of the population was now exempt. Quite apart from the problems this involves for the post-1892 period, this also creates major problems of comparability across the years 1891–2. And this time there are no helpful contemporary estimates.

One possibility is to use the Geisenberger and Müller series to chain the two series. [33] This only gives the shares of the top 5%, 1% and 0.1% of the population, but it might be used for just two years without causing any major problems. Getting from these figures to a gini-coefficient involves a major extrapolation exercise. The method adopted was to use the shape of the Saxon income distribution, and adjust this so that it coincided with the cumulative 95% point given by Geisenberger and Müller. The Saxon distribution used was a combination of returns for 1878–1906. The decile points for the gini-coefficient calculation could then be read from this.

The analysis revealed that, inconveniently, the period around 1891 is one where the estimates derived from Geisenberger and Müller are unusually erratic. The estimated gini-coefficient rises 1.4% in 1890–1 and then falls 1.1% in 1891–2

Are these movements realistic? Or did the problems with the comparability of the tax data around this period affect the Geisenberger and Müller figures as well? This would be a more likely explanation of the fall in 1891–2 than the rise in 1890–1.

The base estimates accept the Geisenberger and Müller fall for 1891–2. It does not have a major effect on the general movement of inequality over the whole period, but it fits in with a pattern found for the 1890s, that inequality was tending to fluctuate at a high level.

[33] Geisenberger and Müller (1972).

(h) *The 1892–1914 period*

The period after the reform of the Prussian tax system poses different problems. The main problem is the fact that, under the new system, 76% of the population was initially exempt from income tax, and there was now no *Klassensteuer* covering the lower income groups. The exempt proportion fell to 37% in 1914 as rising incomes carried more of the population into the tax-paying bands.

There are, however, contemporary estimates of the income distribution made by Biedermann for the years 1895 and 1914 (the tax year relating to income received in 1913). These include a breakdown of the exempt population into income classes, for which Biedermann used unpublished data from the *Reichsversicherungsamt* (from the accident and sickness insurance system), and also information from the Saxon income distribution. This provided two base years.

Biedermann also gave data, for all years 1892–1914, from which the average household size of the exempt population as well as the tax-paying population can be calculated. This is important as this fell as the size of the exempt population shrank. Using this it is possible to make estimates of the household distribution.

The main requirement is, therefore, to make estimates for the other years for which only tax-payer data are available. Estimates of the average incomes of the exempt were prepared using Biedermann's figures for 1895 and 1914 and various wage series (adjusted to coincide for these dates). The wage series used were an average of Hoffmann's industrial and agricultural series (the base estimates), Desai's index and the average hourly earnings of unskilled building workers in Berlin, Rostock and Nuremberg. Biedermann's exempt distributions were then applied to the numbers known to be exempt, and the results adjusted to fit the estimates derived from the wage series.

This produced three variants, and a further set was produced allowing for errors in the allocation of the average household numbers. The results of the calculations are shown in table 6.

Table 6. Estimates of income inequality in Germany 1892–1914[34]

(a) *Estimates for different years*

	1892	**1893**	**1894**	**1895**	**1896**	**1897**	**1898**
Gini-coefficient	0.321	0.324	0.324	0.323	0.315	0.318	0.324
as indices (1853=100)							
Gini-coefficient	119.2	120.1	120.1	119.8	117.0	118.1	120.0
Upper 95% CI of Gini-coefficient	123.9	124.8	124.8	124.6	121.6	122.8	124.8
Lower 95% CI of Gini-coefficient	114.5	115.3	115.3	115.1	112.4	113.4	115.3

	1899	**1900**	**1901**	**1902**	**1903**	**1904**	**1905**	**1906**
Gini-coefficient	0.327	0.331	0.333	0.334	0.331	0.332	0.335	0.340
as indices (1853=100)								
Gini-coefficient	121.4	122.9	123.6	124.1	122.6	123.1	124.3	126.0
Upper 95% CI of Gini-coefficient	126.2	127.8	128.6	129.0	127.5	128.0	129.3	131.1
Lower 95% CI of Gini-coefficient	116.5	118.0	118.7	119.1	117.7	118.2	119.4	121.0

	1907	**1908**	**1909**	**1910**	**1911**	**1912**	**1913**	**1914**
Gini-coefficient	0.338	0.334	0.333	0.319	0.316	0.315	0.318	0.329
as indices (1853=100)								
Gini-coefficient	125.5	124.0	123.5	118.2	117.3	117.0	117.9	122.0
Upper 95% CI of Gini-coefficient	130.4	128.9	128.4	122.9	122.0	121.6	122.6	126.8
Lower 95% CI of Gini-coefficient	120.5	119.1	118.6	113.5	112.6	112.3	113.3	117.1

[34] Calculated from data in Biedermann (1918) p. 65–70.

(b) *Estimates of increase in Gini-coefficient 1892–1906, and decrease 1906–1914*

Increase 1892–1906	5.7	Decrease 1906–1914	–3.2
Upper 95% CI of increase	6.0	Upper 95% CI of decrease	–3.1
Lower 95% CI of increase	5.5	Lower 95% CI of decrease	–3.4
Upper 95% CI of highest variant	8.3	Upper 95% CI of highest variant	–5.4
Lower 95% CI of lowest variant	–1.3	Lower 95% CI of lowest variant	–2.7

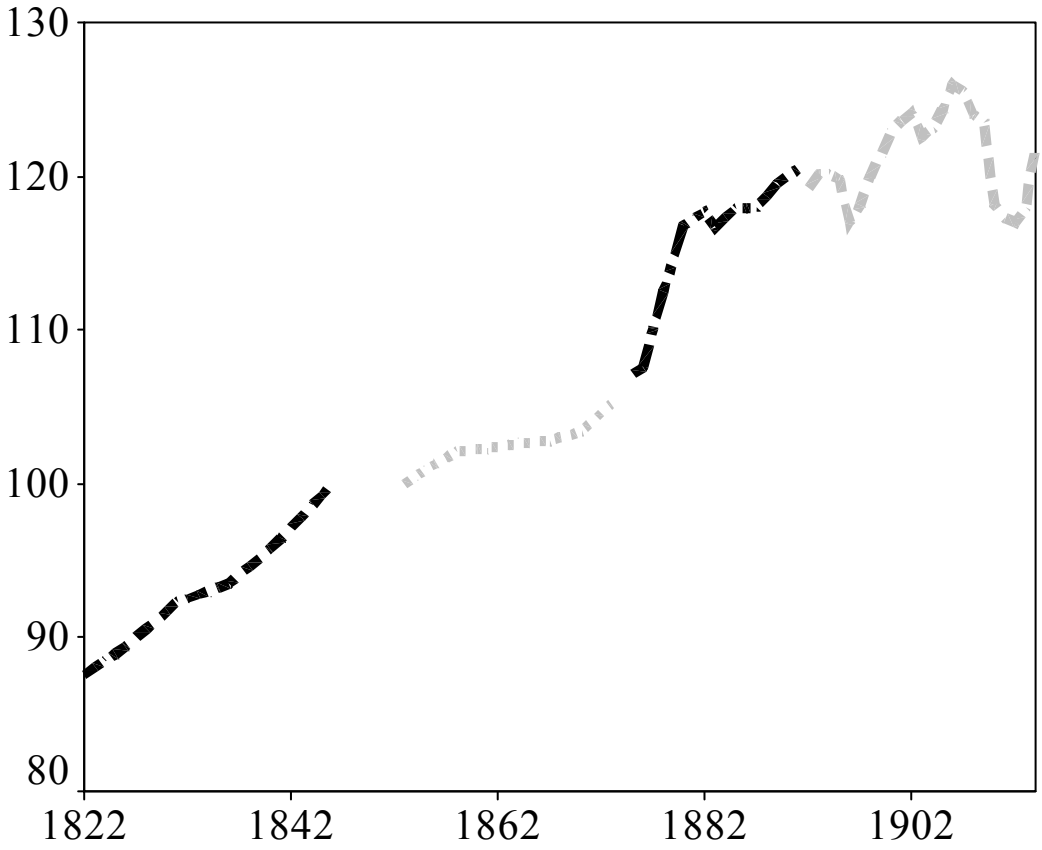
The results show a rise to 1906 followed by a decline. The various variants produce different peaks, some putting it earlier than 1906, which is the main reason why some do not show an increase in 1892–1906. The exact date of the Kuznets curve peak is a matter of some uncertainty. All variants do, however, show a fall starting at some point after 1900, none later than 1906.

IV. The Kuznets curve in nineteenth-century Prussia

(a) *Chained estimates for Prussia*

Figure 1 gives the base estimates from the tables in the previous section chained together to produce an index. Because of the gap between 1846 and 1853 there are two base years, 1846 and 1853. The estimated gini-coefficients are converted to indices with these years set to 100.

Figure 1. Estimates of inequality from Prussian tax data 1822–1914



Indices of estimated gini-coefficients (1846/1853=100)

The shape is approximately that of the first part of Kuznets's "inverted U". There was a downwards movement in inequality after a peak in the early 1900s, though the extent of the fall was not large (a fall of just 5% in the average gini-coefficient estimated for 1910–14 compared to 1903–7). The main downwards phase of the Kuznets curve for Germany, as for other industrialised countries, came in the twentieth century. Modern estimates give a gini-coefficient for

Germany in the 1990s of .25, which is approximately the same level as the chained estimate (using 1892–1914 as the base estimates and making no allowance for changes in 1846–53) for Prussia in the 1820s. [35] So the Prussian or German gini-coefficient appears to have risen 35–40% in the nineteenth century, and fallen by a similar amount in the twentieth. This is in line with Kuznets's hypothesis.

In the first period, 1822–46, industrialisation was just beginning, and the Prussia was still a predominantly rural society. In these years, the main influences on the level of inequality came from developments within the rural sector. These include the effects of the Prussian land reforms, the consequences of a relatively high level of population growth, leading to a period of quite severe demographic pressure, and the effect of rural expansion in the east, bringing into cultivation land in remote areas which were not well-suited to market-orientated peasant agriculture. In addition, agricultural prices fell in the 1820s and 1840s, which would have disadvantaged the sector as a whole, without necessarily leading to a worsening of inequality within agriculture.[36]

The onset of a period of more rapid growth in Germany, and more widespread industrialisation, is generally dated to the 1850s. Contemporary observers were concerned that the fruits of industrialisation were not equally distributed. A wide range of opinion feared that industrialisation would lead to widening inequality unless government action was taken to prevent this. Considerable impetus was given to the pessimistic view by a an article by Ernst Engel, the director of the Prussian Statistical Office, published in the *Zeitschrift des Königlich Preussischen Statistischen Bureaus* in 1875. This contained an analysis of the Prussian tax statistics which appeared to show that the benefits of industrialisation were going, in the main, to the higher income groups. Engel concluded: "Die Konzentration des Reichtums ist hiernach wohl auβer Zweifel".[37]

However, Engel's analysis was probably too pessimistic. The fears expressed by Engel, Schmoller and others, over the widening of the distribution and the disappearance of the *Mittelstand,* were exaggerated by the nature of the tax system. Laissez-faire economics, free trade and *Manchestertum* were unjustly blamed. The tax system in this period was a hybrid. The richer tax payers were subject to an income tax, which was responsive to increases in earned and unearned incomes; the less well-off paid a revised version of the *Klassensteuer.*

---

[35] Atkinson et al. (1995) p. 40 .

[36] On the effect of the land reforms, see Conze (1969) and Saalfeld (1963).

[37] *There is no doubt that the concentration of wealth has increased,* Engel (1875) p. 142.

This was a period in which, according to available evidence, wages were rising in industry and agriculture. This may have led to a lag in the adjustment of *Klassensteuer* rates to improved earnings. More importantly, it means that any estimate of the overall income distribution should allow for some improvement in the incomes of those below the tax threshold.[38]

The base estimates for this period show an increase in the gini-coefficient of between 4.3% and 6.0%, but the systematic error variants show a much wider range, from 20.2% to 3.5%. This uncertainty is largely caused by the question of the extent to which the incomes of the exempt population moved up with other incomes.

In the next period, 1873–1892, there was a change in the terms of trade which was unfavourable to agriculture. The widening gap between agricultural and non-agricultural earnings had the effect of pushing up inequality. As the exempt population contained large numbers of landless agricultural labourers, the assumptions made with regard to this group have a significant effect on the level of overall inequality. The base estimates show an increase in the gini-coefficient of between 11.6% and 13.3%. The systematic error variants show a range from 13.6% to 5.7%. The lower figure comes from a variant which is based on a more optimistic view of the movement of real earnings in agriculture in the period.

The final period, 1892–1914, can be divided into two: a first period in which inequality may have continued to rise, and a second one in which inequality began to fall, albeit slowly and uncertainly. The break came around 1906. Given the nature of the data, there are considerable difficulties with annual movements in the inequality measures. The main problem is the estimation of the movement of the income levels of the exempt. There are a number of possible wage series, but these show different patterns of cyclical movements. [39] Thus, while the base estimates for the period 1892–1906 show an increase in the gini-coefficient of 5.5–6.0%, the range of possible variants runs from +8.3% to –1.3%. In other words, 1906 might not be a peak after all.

The decline in 1906–1914 is rather better founded. The base estimates show a fall of 3.1–3.4%, with a range for the other variants of between

[38] This may not be immediately obvious; if the threshold remains unaltered, why should there be a change in the average incomes of those below the threshold? The important point is that, as a distribution (of a roughly log-normal shape) shifts upwards, the average income of the group below an unchanged threshold will also move up.

[39] The main wage series are those given in Hoffmann (1965), Desai (1968), Bry (1960) and Hohls (1995).

–5.4% and –2.7%. The decline to 1913 was steeper, since there was a cyclical rise in inequality in 1913–14. [40]

Although the timing of some of these movements can be uncertain, the overall trends are rather clearer. Throughout most of the nineteenth century inequality was rising. Sometime in the 1890s this upwards movement came to an end, and, after a period when there was no discernible trend upwards or downwards, inequality began to fall.

(b) *Comparison with other countries*

To get some idea of the scale of inequality in Prussia in the period, it is necessary to make comparisons with other countries. This is the only way to create a standard of measurement, to decide whether a given gini-coefficient is "high" or "low".

The peak gini-coefficient for Prussia in 1906 is estimated to have been .34 (according to the results given in the previous section). The World Bank data set measuring inequality in modern economies shows that only 7 out of the 70 developing economies had lower gini-coefficients than this. So, even at the 1906 peak, Prussian incomes were more equally distributed than in most developing countries today.[41] On this basis the overall level of inequality in Prussia in the period appears to have been relatively low for a country going through the disruptions of industrialisation. However, the increase over levels in the earlier part of the century may well have been a factor creating resentment and pessimism over the future distribution of the rewards of industrial progress.

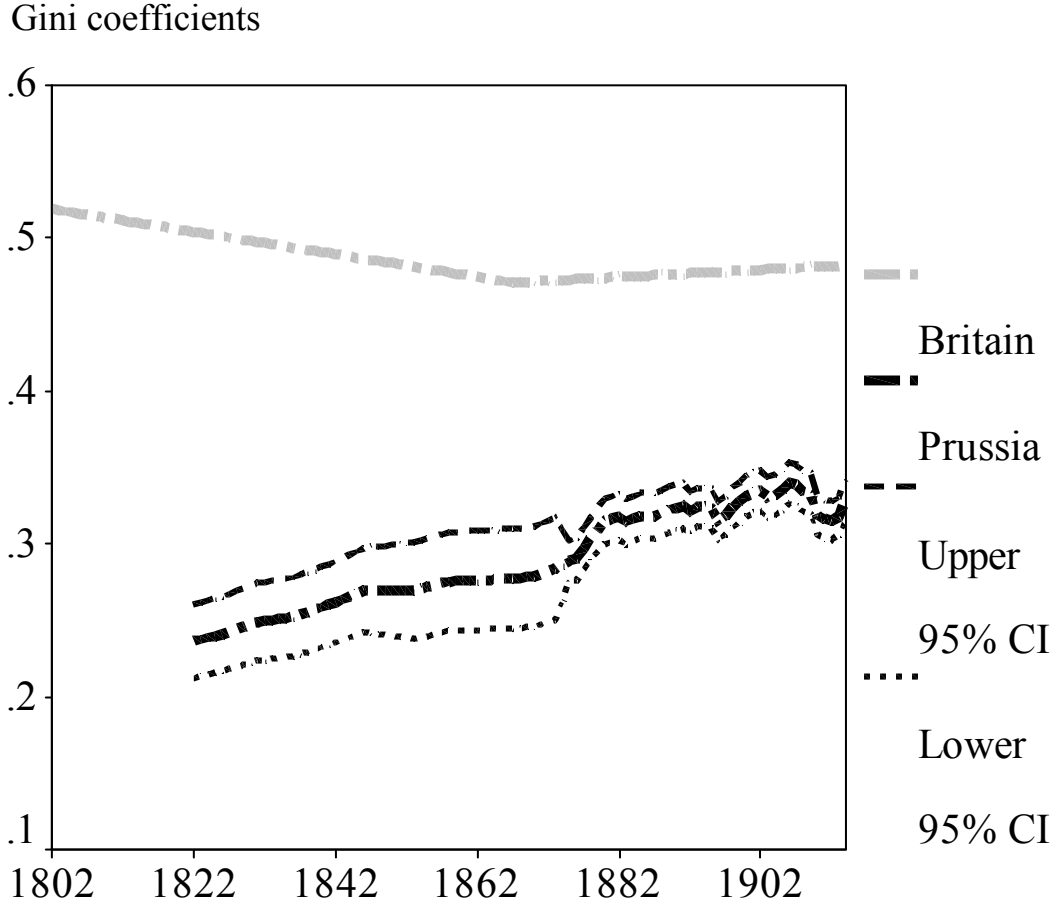
Another point of comparison is the experience of other European countries in the late nineteenth century. Figure 2 compares the gini-coefficients estimated for Prussia with estimates for Britain given by Charles Feinstein. The Prussian estimates are chained backwards from the estimates for 1892–1913 (which are the most reliable). Confidence intervals are also shown. It is not possible to apply the same procedure to the British figures, but, given the nature of the data, it can be assumed that the confidence intervals would be larger than the Prussian ones, even those given for the 1822–46 period.

---

[40] The tax system worked in arrears, so the assessments for 1914 were not affected by the outbreak of war in that year (as would have been the case with a PAYE system).

[41] Deininger and Squire (1996) Table 1; the calculation excludes eastern Europe and the former Soviet economies. There are massive problems of comparability in an exercise of this type. For a discussion of some of these see Atkinson and Brandolini (1999).

Figure 2. Estimates of inequality from Prussian tax data, 1822–1914 (including calculations of upper and lower 95% confidence intervals), compared with figures for Britain.[42]



It appears that the movement of inequality in Britain and Prussia followed different paths. Inequality in Britain was consistently high, and not much affected by industrialisation. Estimates produced by Peter Lindert show a rise in inequality in the late eighteenth century, but Britain was already a highly unequal society before industrialisation.[43] The estimated Gini-coefficient for 1913 (.48) would put Britain in the top half of the World Bank data set, with a higher level of inequality than 45 of the 70 developing economies.

[42] British figures from Feinstein (1988) p.723. The 1846–1853 gap is covered by assuming that the Gini-coefficient for 1846 was the same as that for 1853. An important point to note is that the probability of a movement from the upper limit of the 95% range to the bottom is 1/1600, which means that the confidence intervals for rates of change are tighter than might be assumed from this graph.

[43] Lindert (1994).

This is not surprising given the much less equal distribution of land ownership in Britain. Industrialisation brought about a transfer of population out of an agricultural sector which was itself highly unequal. The fact that inequality did not rise as a result of industrialisation suggests that inequality was just as high outside agriculture.

(c) ***Analysis by sectors***

The Prussian experience was very different. Inequality in the early part of the century was relatively low, reflecting the more equal distribution of land ownership. But industrialisation was accompanied an increase in overall inequality. There are a number of possible reasons for this. There was an increase in the number of landless labourers in the early part of the century, at a time when real wages in agriculture were falling. This probably was the main cause of the increase in inequality in the 1822–46 period. There was a rise in the urban-rural wage gap in the 1870s, which would have increased overall inequality in this period. Both these factors were, however, relatively short-term effects. The longer trend needs a different explanation.

The most likely cause was a shift of population from a relatively equal rural sector to a less equal urban or industrial sector. There are two sources of information on inequality in the urban sector. One is provided by tax data for the period after 1896. A breakdown into two sectors is available for a number of years. These have been used to produce estimates of inequality which are given in table 7.

Table 7. Gini-coefficients estimated for rural and urban sectors from prussian tax data, 1896–1913[44]

	Urban	**Rural**
1896	0.429	0.351
1900	0.438	0.349
1905	0.396	0.321
1910	0.362	0.261
1913	0.355	0.265

Some caveats should be noted. These are based on much cruder breakdown than that used for the estimates given in section III: just six tax-paying classes

[44] Income tax figures from Biedermann (1918); income of the exempt calculated using Hoffmann's figures for wages in industry and agriculture.

plus the exempt. Consequently margins of error are likely to be rather larger than those calculated for the overall figures . In addition, the same figures for household size by income class are used for both sectors which is a possible source of bias.[45]

The results show that urban Gini-coefficients were 28% higher than rural on average. So, a shift of population from one sector to another would have an effect on overall inequality. [46] The table also shows that inequality in both sectors fell substantially in 1900–1913. This, however, would have been partially balanced by the effect of the continuing shift of population into the less equal urban sector.

A second source of evidence comes from budget surveys in urban areas. Results from some surveys of this type are given in table 8. Again there are caveats to note. These are based on household frequency distributions with no information about household size. This has the effect of raising the estimated Gini-coefficient by around 0.08 (many low income households had few dependents). The surveys were undertaken at different times, using methods of recording and selection which could have varied considerably. This may well explain some of the surprising jumps in inequality shown for Hamburg, for example.

Table 8. Gini-coefficients calculated from household budget surveys[47]

Hamburg 1867–8	0.651	Leipzig 1875	0.643
Hamburg 1873–4	0.467	Leipzig 1885	0.613
Hamburg 1881–2	0.720	Leipzig 1900	0.605
Hamburg 1890–1	0.534	Dresden 1880	0.559
Hamburg 1900–1	0.496	Breslau 1880	0.542
		Breslau 1900	0.595

Allowing for these reservations, the results still show a pattern of inequality substantially higher than the estimates given for the rural sector. They indicate that rural–urban migration would have been a powerful force driving up overall inequality.

[45] Household size was definitely larger in agriculture, but this would only be a source of bias if the difference were not the same for all income classes.

[46] Overall inequality is not, however, just a weighted average of the two sectors. The size of the gap between the sectors is also important.

[47] Data from surveys given in Kucyzinski (1908),

V. Migration and the Kuznets curve

(a) *Migration and Kuznets's two sector mechanism*

As a result of Prokopovitch's work, Kuznets was aware of the difference between rural and urban inequality in nineteenth century Prussia. He found a similar pattern in studies of inequality in the United States and India.[48] This led him to a more general statement of the relationship between inequality and economic development.

Migration had a central role in Kuznets's own explanation of his "inverted U" relationship. Kuznets's model had three parts:

1. There are two sectors A (rural) and B (urban); incomes are higher in B; there is migration from A to B;
2. Income inequality is higher in B;
3. Income inequality in B will tend to fall[49].

Kuznets showed, using mathematical examples, that under certain assumptions the first part, on its own, could produce a "Kuznets curve". This effect was intensified by the second part. But Kuznets also thought that income inequality in the three countries he was considering (Britain, Germany and the US) had in fact stabilised at a relatively early stage while the first two components were tending to cause it to rise, and he attributed this to the third component (for which he had no direct evidence).

There are a number of theoretical problems with Kuznets's attempt to explain the curve. Kuznets's assumptions are quite restrictive. In particular he assumed that the income distributions in the two sectors were not affected by migration. This may seem innocuous but the implication is that migrants as a group have average incomes (for each sector) before and after migration, and that these incomes are distributed in line with the distribution for the sector as a whole.

Some evidence on this point can be obtained from the German occupational censuses. There are no comparative figures for migrant earnings, but the 1907 occupational census did record the skill levels of migrant workers. It is important to concentrate attention on those making genuine rural–urban moves. In general there is no tendency for migrants as a whole to have lower skill levels than those born in the region or city, but this reflects the high mobility of skilled urban workers. Many skilled migrants had made city to city moves. Table 9

[48] Kuznets (1955) p. 7; the German figures came from Prokopovitch (1926).

[49] Kuznets (1955) pp. 7–8 for parts 1 and 2, p. 17 for part 3.

provides figures for men residing in urban areas in 1907, distinguishing between those born in cities and those born in rural areas.

The first part shows that, both in industry and in services, migrants supplied proportionately more manual workers than entrepreneurs or managerial staff. The gap is, however, not as large as might be feared. Some migrants did manage to obtain better paid jobs. The implied distribution of migrant earnings is one which overlaps with the distribution of the "urban-born" population, even though the mean is rather lower.[50]

The second part of the table gives some results for different industries. The skill classification is more detailed at the industry level, and skilled workers can be distinguished from unskilled (the German terms "gelernt" and "ungelernt" imply the possession of some formal qualification). This shows that rural migrants did contribute a disproportionate number of unskilled workers, although some managed to obtain skilled jobs.

Table 9. Percentages of the total numbers residing in cities in 1907 who were born in rural areas, by skill category (men only).[51]

	Entrepreneurs & self-employed	Managerial & technical	All manual workers
All industry	34.3	28.1	38.8
All services	39.5	28.8	46.7

	Entrepreneurs & self-employed	Managerial & technical	Skilled manual workers	Unskilled manual workers
Mining & metal production	31.2	32.0	43.4	46.7
Stone & quarrying	30.0	40.4	38.8	54.3
Metal-working	25.6	23.6	30.0	36.8

Results for women (table 10) show a rather different pattern. Relatively few rural migrants obtained white collar managerial jobs, but rather more were self-

[50] There is a question mark over the comparability of figures showing not dissimilar numbers in the "entrepreneurs and self-employed" category. This could include peddling and small-trading activities which might well be a form of disguised unemployment. Migrant earnings in this category may have been well below those of born in urban areas.

[51] Data from the 1907 occupational census, *Statistik des Deutschen Reichs,* n.f. 202–211.

employed than in manual employment. This result, while being of interest, is of less significance when considering the total income of migrant households, mainly because the participation rate of migrant women was low. There were 3,238,900 men, born in rural areas but residing in cites, in non-agricultural employment or self-employment, but only 675,400 women.

Table 10. Percentages of those residing in cities in 1907 that were born in rural areas, by skill category (women only)

	Entrepreneurs & self-employed	Managerial & technical	All manual workers
All industry	29.9	14.7	26.9
All services	39.1	15.7	33.4

These results do not disprove Kuznets's hypothesis. Even if the migrant income distribution has a mean somewhat below that of the "urban-born" population, there could still be a rise in inequality as population shifts from sector A to sector B, provided the gap between average earnings in the two sectors is large enough. However, they make it more difficult to come to any firm conclusions with regard to the likely effect of industrialisation on income distribution. Kuznets's mechanism may work as he described, but it is not a provable proposition once the restrictive assumption about the effect of migration on income distribution in the two sectors is relaxed.

The point about the importance of the size of the gap between the two sectors has been emphasised by development economists concerned with "dualism". Studies have shown that large gaps between rural and urban earnings are associated with high levels of overall inequality.[52] This is not so surprising, the rural–urban gap is a component of overall inequality. Poorer rural workers and peasants account for a disproportionate number of the total numbers on low incomes even in advanced countries, so a worsening of the relative position of agriculture will tend to worsen the overall distribution of income.

There is one possibility that could produce a Kuznets curve from the movement of the earnings gap, and that would be if migration costs themselves followed an "inverted U" course, initially rising and then falling. Then the stretching of the income distribution does not come about as a result of a period of dis-

[52] Bourguignon and Morrison (1998), Bourgignon (1990); Basu (1997) discusses the theory behind this.

equilibrium, or delayed adjustment of expectations, but is itself an equilibrium position.

There are a number of reasons why migration costs might follow such a pattern. Firstly, most migrants in the early stages of industrialisation are drawn from the districts closest to the cities. It is only later, when industrial demand for labour can no longer be satisfied from these sources, that there is more long distance migration. Both pecuniary and non-pecuniary costs (such as social dislocation) of migration are likely to be strongly related to distance, a point confirmed by the negative influence that distance has been shown to exert on the likelihood of migration in the equations estimated in this study.

Secondly, there may well be a dynamic element to urban disamenities. The pace at which a city is expanding will affect the strain placed on urban infrastructure and available housing. In which case the costs of migration will rise with the rate of migration. Costs will only fall when urban sanitation, health provision and residential construction has caught up with the demands of the expanding population. A related point is that some of the benefits of migration may also be affected. Greater social mobility as a result of improved educational and training opportunities is one benefit that migrants (and their children) may anticipate as result of a move to a city. But the development of these institutions may also lag behind the growth of population during the early stages of industrialisation.

There is one aspect of urban disamenities which is easy to measure, and that is the relative death rate. Table 11 gives the ratio between the crude death rate in Berlin and in the whole of Germany.

Table 11. Crude death rate per 1000 inhabitants in Berlin divided by the rate for the whole of Germany (post 1871 frontiers), by decade

1841–1850	0.97	1881–1890	0.96
1851–1860	0.98	1891–1900	0.86
1861–1870	1.12	1901–1910	0.86
1871–1880	1.12		

This does indeed show an “inverted U” curve, peaking in the 1860s and 1870s. It would have had a direct effect on migration costs. The likelihood of a premature death would figure quite heavily in a migrant’s calculation of the potential benefit from migration if migrants were aware of this. But it is also important as a symptom of other problems. A high death rate would be an indicator of poor housing conditions, inadequate sanitation, workplaces which were

dangerous or unhealthy, and other urban problems of which migrants would have had direct experience. The differences shown in this table are indicative of a general rise and fall in urban disamenities.[53]

(b). *Analysis by simulation*

The first year for which separate distributions for both the rural and urban sectors are available is 1896. These can be used to simulate the effect on overall inequality of different scenarios. Between the 1820s and the 1890s there was a total increase in inequality of 35–40%; this is the change that has to be explained.

Table 12. Simulations using 1896 rural and urban income distributions.

I. Using 1896 sector weights: 40% urban, 60% rural

Scenario:		***Effect on overall inequality***
A	All rural incomes fall 10%	Gini-coefficient rises 2.6%
B.	Incomes of poorest 50% in rural sector fall 10%	Gini-coefficient rises 3.8%
C.	10% of total population moves from the rural to the urban sector – no effect on sector distributions (Kuznets process).	Gini-coefficient rises 1.8%
D.	10% of the total population moves from the poorest half of the rural sector to the bottom half of the urban sector.	Gini-coefficient falls 2.1%

II. Using circa 1850 sector weights: 10% urban, 90% rural

E.	10% of rural population moves from the rural to the urban sector – no effect on sector distributions (Kuznets process).	Gini-coefficient rises 7.1%
F.	10% of the total population moves from the poorest half of the rural sector to the bottom half of the urban sector.	Gini-coefficient falls 3.8%

[53] If so, then the rise in income inequality during industrialisation may be, at least in part, an illusion. The shift from low-paid agricultural work to higher-paid industrial jobs, which drives up inequality in Kuznets's model, was not such a large gain when allowances for urban disamenities are included. The gap between urban and rural well-being may have been constant rather than rising and then falling.

III. Overall effect of changing weights

G.	Overall effect of move from 10% urban/90% rural to 40% urban/ 60% rural (Kuznets process).	Gini-coefficient rises 14.5%

The simulations show that, using Kuznets's assumptions, migration will push up inequality (scenarios C and E) but that this effect can be reversed if more realistic assumptions are made, so that migrants come from the lower part of the rural distribution and move to the lower part of the urban distribution (scenarios D and F). So migration, in itself, is not a very convincing explanation of the Kuznets curve. [54]

However, the Kuznets process as a whole, the shift from a predominantly rural society to one in which the urban sector is substantially increased, will increase overall inequality, as scenario G shows. The effect only accounts for around a third of the increase in inequality which actually occurred between the 1820s and the 1890s. Other factors which might have contributed are simulated in scenarios A and B: a shift in the terms of trade which has an adverse on agriculture raises inequality, but the effect is greater if it is strongest on those at the bottom of the rural distribution (scenario B).

In general, the strength of the rise in inequality shown by the analysis of the Prussian income tax data is such that it was almost certainly brought about by a combination of adverse factors. These were:

1. A shift of population from a more equal rural sector to a less equal urban sector (Kuznets process);
2. Increased inequality in the rural sector (due to a rise in the numbers of landless labourers):
3. High inequality in the urban sector ;
4. A shift of the terms of trade against agriculture.

[54] If there is a two-stage, generational process, in which the first generation of migrants remain in the lower part of the urban distribution, but their children move up to the sector average, then migration will not permanently alter the urban distribution. This may be a better justification of Kuznets's assumption.

(c). ***Econometric analysis***

The Prussian tax statistics can also be used to examine some of the causes of inequality in the rural *Kreise.* In 1916 an analysis of tax returns in 1914 at the *Kreis* level was published in the *Zeitschrift* of the Prussian statistical Bureau. The assessments were based on income in the previous year, so this is an analysis of the Prussian income distribution in the final year before the war.

The published sources provide the numbers of tax-payers, and their dependents in four categories: those with annual incomes below 900 Marks, those with between 900 and 3000 Marks, those with between 3000 and 9500 Marks, and those with over 9500 Marks. To get a measure of inequality from this requires extrapolation and an estimate of the likely shape of the overall distribution. The procedure used is discussed and described in appendix A.

Given the nature of the data provided in the published sources, it was considered desirable to use a measure of inequality which minimised the likely errors due to extrapolation. The preferred measure is the U10/L40 ratio: the estimated share of total income in the *Kreis* which went to the top 10% of the population (including dependents), divided by the share which went to the bottom 40%. This measure can be derived using the minimum amount of extrapolation (the upper 10% and lower 40% shares are in the middle of the ranges of the published figures).

Table 13 gives the results of regression analysis using U10/L40H ratios for 583 Prussian *Kreise.* The first column shows a basic analysis which confirms that urban areas had higher inequality (a rise in agricultural employment reduces inequality) and also shows that there is a strong tendency for inequality to be higher in the east: the coefficient on RUSSDISTANCE (the distance in Kilometres from the *Kreis* mid-point to the border with Russian Poland) is significant and negative. The second column adds other geographical variables which explore the effect of remoteness on inequality: CITYDISTANCE(a) measures the distance in Kilometres from the *Kreis* mid-point to the nearest city with over 200,000 inhabitants; CITYDISTANCE (c) is the same measure for towns with at least 50,000 inhabitants.

The results show that different sized cities are associated with different levels of inequality in nearby regions. Larger cities tended to have high levels of inequality, and this appears to spill over into nearby regions (inequality falls as CITYDISTANCE (a) rises). However, the effect of smaller towns appears to be more beneficial: inequality is low in regions which are close to these cities (the dataset includes the *Stadtkreise* themselves for which the variable values are zero).

Table 13. Analysis of inequality in the Prussian *Kreise* 1913/4: dependent variable is the log of the U10/L40 ratio

(t-statistics in brackets)

	a.	**b.**	**c.**	**d.**
Constant	0.736	0.759	0.757	0.818
RUSSDISTANCE	–0.000358	–0.000362	–0.000386	–0.000775
	(8.15)	(7.42)	(5.13)	(3.98)
RUSSDISTSQD				+5.10E–07
				(2.19)
%AGOCCP1907	–0.00805	–0.00900	–0.00899	–0.00908
	(17.0)	(14.7)	(14.6)	(14.8)
CITYDISTANCE (a)		–0.00044	–0.00045	–0.00057
		(1.57)	(1.59)	(1.99)
CITYDISTANCE (c)		+0.00146	+0.00148	+0.00149
		(3.01)	(3.03)	(3.08)
AGVALUE			+0.00011	
			(0.42)	
N	583	583	583	583
Standard error of equation	0.2815	0.2794	0.2796	0.2785
R-sqd	0.348	0.359	0.359	0.364
adj R-sqd	0.345	0.355	0.354	0.359
F-test of regression	154.7	81.1	64.8	66.3
Residual sum of squares	46.0	45.2	45.2	44.8

In remote regions, such as the Prussian east, the effect of distance from both large and small towns or cities was to push up inequality. The positive coefficient on CITYDISTANCE (c) is larger than the negative one on CITYDISTANCE (a). However, this does not appear to explain the east-west division: the coefficient on RUSSDISTANCE is not altered.

The third column introduces a variable to capture the effect of "economic dynamism": the value of *Aktiengesellschaften* founded 1871–1910 by province (per head of population 1895). This measure should pick up the effect of the

relative backwardness of the east in terms of industrial development.[55] It has little discernible effect on inequality.

The final column adds a quadratic term to examine the possibility that the east–west relationship is a non-linear one. This is significant: the relationship is one which is steepest in the east, with a gradual falling-off of the effect in central and western regions

These observations raise an interesting issue: to what extent was higher inequality in the east due to internal factors, unequal land distribution for example, and to what extent to external factors. Contemporary observers like Max Weber were concerned about the effect of immigration from Russian Poland. Weber argued that the use of Polish labourers disadvantaged German workers and led to increased migration out of the rural east into German cities. Eastern landowners, under the influence of capitalist values, rejected the obligations of the pre-capitalist "patriarchal" system of labour relations, and substituted cash payments for harvest shares and other forms of payments in kind. The eastern agricultural worker lost the status of a small sub-contractor and was reduced to that of a wage labourer, who had to compete with cheaper Polish labour.[56]

If this is correct, then the proximity of Russian Poland might well have an effect on overall inequality. In areas close to the border, agricultural wages would be held down, and the profits of "capitalist" landowners would rise.

To examine this, it is necessary to split the sample, dividing the urban areas from those which were predominantly rural. This removes 105 urban *Kreise* (with a population of 15.4 million out of a total of 41.2 million). This makes it possible to introduce terms which capture the effect of agricultural structure. Table 14 gives the results of these regressions.

The new terms give the percentage of the total agricultural area which is in holdings of over 100 hectares, and in holdings of between 20 and 100 hectares (these were the farms which employed significant amounts of outside labour). Interaction terms are also used, as the hypothesis is that an unequal division of land-holding will push up overall inequality in the agricultural sector, so the land distribution terms are multiplied by the relative size of the agricultural population (in a *Kreis* with little agricultural employment the effect of the land distribution on the level of overall inequality would not be expected to be a large one).

---

[55] Regional figures for industrial production would be desirable, but these are not available, so this measure is the best alternative.

[56] Weber (1893); Weber's views are examined in more detail in Grant (2002).

Table 14. Analysis of inequality in the Prussian *Kreise* 1913/14 (urban areas excluded): dependent variable is the log of the U10/L40 ratio
(t-statistics in brackets)

	a.	**b.**	**c.**	**d.**
Constant	0.381	0.315	0.340	–0.072
RUSSDISTANCE	–0.000352	+0.000040	–0.000305	
	(8.27)	(0.69)	(1.84)	
RUSSDISTSQD			+4.50E-07	
			(2.23)	
RUSSDIST<50				+0.00206
				(2.09)
%AGOCCP1907	–0.00190	–0.00702	–0.00707	–0.00115
	(2.75)	(5.32)	(5.38)	(1.76)
CITYDISTANCE (a)	–0.00086	–0.00111	–0.00118	–0.00107
	(3.38)	(4.58)	(4.86)	(4.47)
CITYDISTANCE (c)	+0.00137	+0.00158	+0.00160	+0.00164
	(3.38)	(4.21)	(4.27)	(4.42)
%LAND>100HA		+0.00589	+0.00619	+0.00566
		(8.25)	(8.55)	(11.82)
%LAND>100HA* %AGOCCP1907		+0.000170	+0.000156	+0.0001785
		(6.26)	(5.65)	(6.53)
%LAND20–100HA		+0.00128	+0.00165	+0.00114
		(2.20)	(2.73)	(1.98)
%LAND20–100HA %AGOCCP1907		+0.0000582	+0.0000627	+0.0000628
		(1.56)	(1.66)	(1.70)
N	478	478	478	478
Standard error of equation	0.2182	0.1974	0.1966	0.1966
R-sqd	0.171	0.327	0.334	0.333
adj R-sqd	0.164	0.316	0.321	0.321
F-test of regression	24.32	28.51	26.10	29.23
Residual sum of squares	22.53	18.28	18.08	18.13

The results show that, while the relationship between inequality and RUSSDISTANCE holds for this split sample (column a) using the original variables, it is largely eliminated when the new variables are introduced. Column b shows that the negative relationship between inequality and the distance from Russian Poland is replaced by an insignificant positive one. The introduction of a quadratic term (column c) re-creates a significant relationship, as confirmed by a Wald test on the joint significance of RUSSDISTANCE and RUSSDISTSQD (a p-value to accept that both are zero of 0.0463), but the shape is a perverse one. Inequality falls to the middle of the sample (around 400 km from the border) and then rises. This raises the possibility that the effect from Russian Poland was confined to certain areas close to the border. This in turn was tested using truncated variables (measuring the distance to the border up to a certain value and then limited to this value thereafter). Truncations at 50 km, 100 km and 200 km were tried. The only significant result came from a variable truncated at 50km (column d), and this shows a positive relationship: inequality was lower in the *Kreise* whose mid-points were within 50 km of the border.[57]

Taking these results into account, it appears that the main factors driving up inequality in the Prussian east were the unequal land distribution and the remoteness of the region (CITYDISTANCE (c) remains an important influence). Weber may have been right in his analysis of the effect of changing social attitudes on eastern estates, but the proximity of the border with Russian Poland did not exert a major influence. Inequality in the Prussian east was home-grown not imported.

Turning to the urban areas, the effects of industrial structure can be analysed using the results of the occupational census. In addition, migration figures are available for some cities. These were used to construct an index of mobility, by dividing total recorded moves for 1910–12 by the resident population. These figures are available for a total of 44 cities with 9.4 million inhabitants.

The analysis shows that high mobility is associated with higher levels of inequality. Migrants are likely to be relatively unskilled, and therefore may be disadvantaged in urban labour markets, so that cities with large numbers of recently arrived migrants, and, in consequence, high mobility, will tend to have high inequality.

[57] Comparison of the diagnostic statistics shows that the extra variables introduced in columns c and d add little to the explanatory power of the estimated equations. By contrast, the variables introduced in column a produce a better fit compared to column a.

Table 15. Analysis of inequality in Prussian cities 1913/14: dependent variable is the log of the U10/L40H ratio

(t-statistics in brackets)

	a.	b.	c.	d.	e.
Constant	–0.368	–0.323	+0.482	–0.375	–0.449
	+0.021	+0.020	+0.015		+0.018
MOBILITY	0	8	0	+0.0210	1
	(4.06)	(3.94)	(2.29)	(3.98)	(3.53)
	+0.019	+0.019			
MINING	1	8		+0.0193	**
	(1.99)	(2.01)		(1.91)	**
	+0.009	+0.009			
TEXTILES	6	8		+0.0096	**
	(4.15)	(4.10)		(4.08)	**
LEATHER	–0.0301	–0.0298		–0.0304	**
	(4.29)	(4.18)		(3.79)	**
	+0.041	+0.040			
UTILITIES	3	6		+0.0419	**
	(2.43)	(2.35)		(2.24)	**
	+0.080	+0.079			
PRINTING	4	2		+0.0810	**
	(6.20)	(5.89)		(5.22)	**
		–			
RUSSDISTANCE		0.00021			
		(0.38)			
			–	+0.0001	
FACTORYSIZE			0.00503	1	
			(3.41)	(0.08)	
N	44	44	44	44	44
Standard error of equation	0.233	0.235	0.326	0.236	0.219
R-sqd	0.668	0.669	0.277	0.668	0.777
adj R-sqd	0.614	0.604	0.242	0.603	0.658
F-test of regression	12.38	10.39	7.87	10.33	6.510
Residual sum of squares	2.003	1.995	4.353	2.003	1.343

This in turn suggests one reason for the high levels of inequality found in urban areas in Prussia in the late nineteenth century. This was a period of high migration out of agriculture into industry, a period when Prussia (and Germany) passed through a "labour surplus" phase of economic development, as described in the "Lewis Model" of development economics.[58]

The finding that high mobility is associated with high inequality is robust to the inclusion of variables giving relative employment in different industries. Column a gives the results of regressions run with a few selected industries included in the regressions (selected by a step-by-step procedure). Column e gives the results when the full set of employment variables is included (for 14 industries): the results for the individual industry variables are not reported.[59]

The analysis also shows that urban inequality is not significantly affected by the distance from the border with Russian Poland (column b); this result also holds when the full sample of 105 urban *Kreise* is used (not reported). This confirms the results obtained from the analysis of the rural *Kreise.* The strong explanatory power of the RUSSDISTANCE variable in the regressions given in table 13 is shown to be a spurious effect when the sample is split and additional variables are introduced.

According to column c, inequality is affected by the average enterprise size (FACTORYSIZE – derived by dividing the number of employees from the occupational census returns by the number of surveyed enterprises), but the result is a surprising one: inequality falls as enterprise size increases. Column d shows that this result is not robust to the inclusion of employment variables, which is to be expected as average factory size will be affected by the presence or absence of different industries. However, this does provide a warning that the relationship between industrial structure and inequality is likely to be a complex one, depending on capital intensity, the concentration of capital ownership, whether or not returns to entrepreneurs and share-holders are above average, and the extent of wage dispersion amongst employees. Some of this may feed into the apparent relationship between enterprise size and inequality.

---

[58] The thesis from which this paper is derived was mainly concerned with the application of the Lewis Model to 19th century Germany, Grant (2000).

[59] Collinearity problems mean that the results for individual industries should not be given too much credence. Under the step-by-step procedure quite small differences in the calculated "f-test to reject" can lead to the elimination of variables, and this boosts the significance of the variables left in the estimated regressions.

VI. Is the Kuznets curve a peculiarity of German history?

The results produced so far are in line with the conclusion reached by some modern cross-section studies: there is a conditional Kuznets curve, which can emerge when the historical circumstances are right but which, under different conditions, might be obscured by other factors. The question then is, what were the forces that produced the first stage of the Kuznets curve in nineteenth-century Prussia?

The first point is that, even at the peak, inequality was not particularly high in Prussia by modern standards. Although income inequality in the urban sector was quite high, the overall figure was held down by a relatively egalitarian rural sector. What created the Kuznets curve in Prussia was the shift from a rural society with a low gini-coefficient to an urban or industrial sector which had a level of inequality which was high by comparison with the rural sector though not so unusual when compared to other countries going through the process of industrialisation.

The second point is that, from the 1870s onwards, there was a shift in the terms of trade against the agricultural sector as a result of developments on world markets. This had the effect of widening the gap between rural and urban incomes, and thus worsening the overall income distribution. Although this was not the only cause of the rise in inequality in the period, it contributed to the rise in the gini-coefficient.

A third point is that, in the early part of the century, inequality was rising in the rural sector, due to a range of factors: demographic pressure, expansion into remote eastern areas, the effects of the agrarian reforms, and (according to Max Weber) changing social relations. This was only indirectly connected to industrialisation.

These factors provided a considerable reinforcement to the mechanism specified by Kuznets himself, which as already noted relies on assumptions about migrant characteristics which may not be justified. It would be quite possible for a society which industrialised under different circumstances to escape the Kuznets curve. The Prussian example points to some factors which might help to achieve this: a relatively low level of demographic pressure, a prosperous agricultural sector in which incomes kept pace with those in urban areas, agrarian policies which promoted an equal distribution of income in this sector. These are the "good" ways to escape a Kuznets curve. But the comparison with Britain also reveals that there is a "bad" way to escape it, and that is to start off with

such a high level of inequality in the rural sector that the transition to an industrial society does not push up overall inequality even though incomes in the urban sector are also unequally distributed.

In short, the Kuznets curve, while not just a peculiarity of German history, owes its origins to some features of nineteenth century Prussia which may or may not be replicated in modern developing countries. It is the product of circumstances rather than an automatic consequence of industrialisation.

Appendix A
Estimation of *Kreise* income distributions

The data set used in the regressions given in section IVc is derived from figures given in Kuhnert's 1916 article in the *Zeitschrift* of the Prussian Statistical Bureau.[60] These give percentages both of the total population (including dependents) and of potential taxpayers for four income bands: less than 900 Marks, 900–3000 Marks, 3000–9500 Marks and over 9500 Marks. The problem is how to make the best use possible of this limited amount of data.

The procedure used the overall income distribution derived from Biedermann's 1914 analysis. This can be represented as a cumulative income distribution as shown by the uninterrupted line in figure A.1. This gives a basic shape for the distribution, which can then be manipulated so as to fit the available points on the *Kreis* distribution. The procedure was, firstly, to alter the scale of the X-axis so that the points on the distribution curve corresponded to the available *Kreis* distribution points, and then to read from this the points required to produce a cumulative income distribution. This could then be used to make the U10/L40 calculation. This effectively stretched or compressed the original distribution to make it correspond to the *Kreis* distribution. It meant that the adjustment used all the available data.[61]

The results for one *Kreis*, *Kreis* Labiau in East Prussia, are shown in the diagram. The distribution has been stretched at the lower end. There were large numbers on low incomes: 80.6% of the population were in households with less than 900 Marks income, against a Prussian average of 36.7%. However, there were also some higher earners, so the curves move closer together at the higher income levels. The estimated U10/L40H ratio for this *Kreis* was somewhat above the Prussian average (1.68 against 1.50).

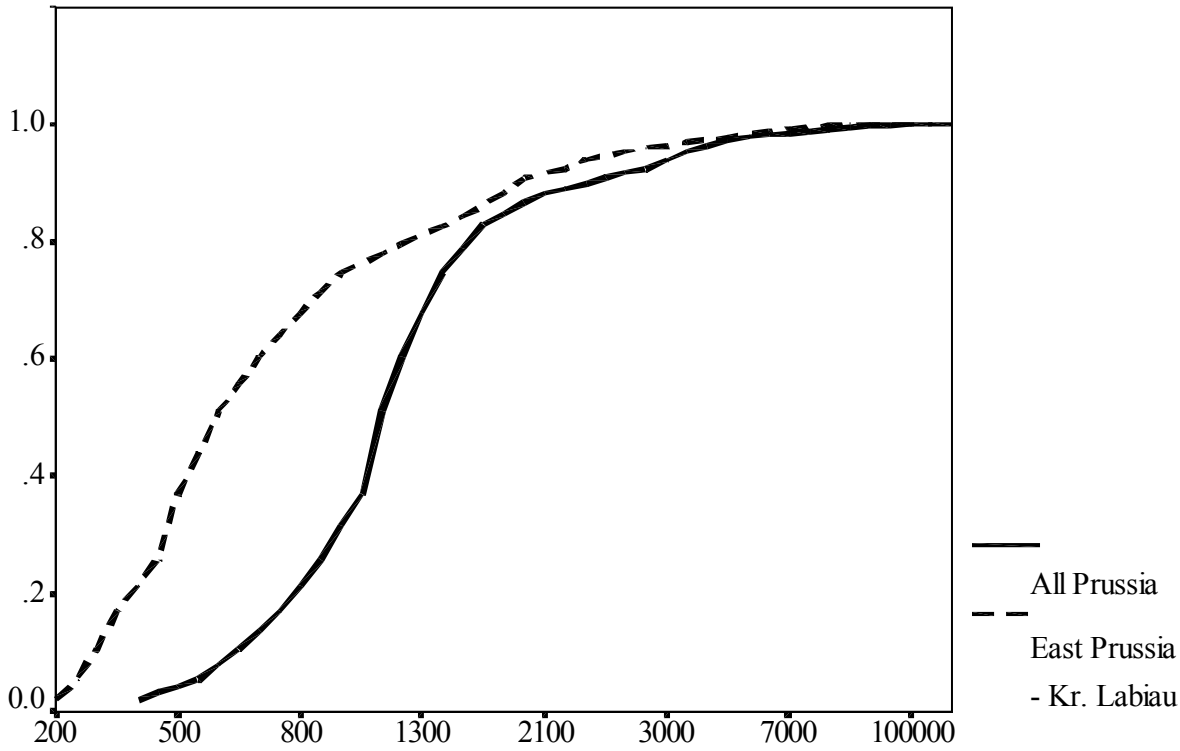
There is obviously scope for error, and this would affect the fit of the estimated equations. But the basic relationships shown by the estimated equations should be reliable, unless the errors are correlated with some of the explanatory variables. There is no evident reason why this should have happened.

---

[60] Kuhnert (1916) pp. 298–303.

[61] An earlier attempt imposed a log-normal curve on the available data points. This effectively restricts the adjustment to two parameters: the mean of the distribution and the standard deviation of logY (Y = income). Although some World Bank economists hold that the log-normal curve is a reliable approximation for the shape of the income distribution, Dollar and Kraay (2000) p.9 is one example, pre-1914 German income distributions do not always fit this pattern. The procedure outlined in this section made better use of the available data.

Figure A.1. Cumulative income distribution curves: Prussia and *Kreis* Labiau in East Prussia, 1914



Appendix B

Description of regression variables and sources

U10/L40	Ratios of the income shares of the top 10% to the bottom 40%, for the taxpayer and the household distributions, estimated as described in appendix B.
%LAND>100HA	The percentage of the total agricultural area in holdings of over 100 hectares, from 1895 Agricultural Census, *Statistik des Deutschen Reichs* n.f. 112, pp. 489–500.
%LAND20–100HA	As above, for holdings of between 20 and 100 hectares.
%AGOCCP1907	The population occupied in agriculture as a percentage of the total occupied population, from the 1907 occupational census, *Statistik des Deutschen Reichs* n.f. 209.
RUSSDISTANCE	Shortest distance in kilometres from the Kreis mid-point to the Russian-Polish border (measured from contemporary maps).
RUSSDISTSQD	Above variable squared.
CITYDISTANCE (a)	Distance in kilometres from the Kreis mid-point to the nearest city with at least 200,000 inhabitants in 1900.
CITYDISTANCE (c)	As above, for towns and cities with at least 50,000 inhabitants.
AGVALUE	The value of *Aktiengesellschaften* founded 1871–1910 by province: from *Vierteljahrsheft zur Statistik des Deutschen Reichs*, 1907 iv, p. 373 and subsequent issues of the VSDR; the figures were then divided by the regional population 1895 (1895 census).
MOBILITY	Data from *Statistisches Jahrbuch deutscher Städte* 1911–1913: total moves for each city in 1910–12 divided by total population. The selected cities were those with complete sets of data for the period

References

Anand. S. and Kanbur, S.M. (1993), "The Kuznets process and the Inequality–Development Relationship", *Journal of Development Economics,* 40(1), pp. 25–52.

Atkinson, A. B. and Brandolini, A. (1999), *Promises and Pitfalls in the Use of "Secondary" Data-Sets: Income Inequality in OECD Countries,* manuscript, Nuffield College, Oxford and Banca D'Italia, Research Department.

Atkinson, A. B., Rainwater, L. and Smeeding, T.M. (1995), *Income Distribution in OECD Countries,* OECD Social Policy Studies no. 18.

Atkinson, A.B. (2002), *Top Incomes in the United Kingdom over the Twentieth Century,* Oxford Discussion Paper in Economic and Social History no. 43.

Barro, R. (1999), *Inequality, growth and investment,* NBER working paper no. 7038.

Basu, K. (1997), *Analytical Development Economics; the less developed economy revisited* [Cambridge, Mass.: MIT Press].

Biedermann, E. (1918), "Die Einkommens- und Vermögensverhaltnisse der preuβischen Bevölkerung im Zeitraum 1895–1914", *Zeitschrift des Königlich Preuβischen Statistischen Bureaus,* 58(1), pp. 60–86.

Bourgignon, F. (1990), "Growth and Inequality in the Dual Model of Development: the role of demand factors", *Review of Economic Studies,* 57, pp. 215–228.

Bourguignon, F. and Morrison, C.(1998), "Inequality and Development", *Journal of Development Economics,* 57, pp. 233–257.

Bowley, A.L. (1946), *Elements of Statistics,* 6th ed. [New York: Staples].

Bruno, M., Ravailllon, M. and Squire, L. (1996), *Equity and Growth in Developing Countries: old and perspectives on the policy issues,* Policy Research Working Paper 1563 [Washington: World Bank].

Bry, G. (1960), *Wages in Germany 1871–1945* [Princeton: Princeton University Press].

Conze, W. (1969), "The effects of nineteenth century liberal agrarian reform on social structure in central Europe" in Crouzet, F. ed., *Essays in European Economic History* [London: E. Arnold].

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

Desai, A.V. (1968), *Real Wages in Germany, 1871–1914* [Oxford: OUP].

Dollar, D. and Kraay, A. (2000), *Growth is Good for the Poor,* manuscript, www.worldbank.org/research.

Dumke, R. (1991), "Income Inequality and Industrialisation: the Kuznets curve re-examined" in Brenner, Y., Kaelbe, H. and Thomas M. eds., *Income Distribution in Historical Perspective* [Cambridge: CUP].

Engel, E. (1868), "Die Ergebnisse der Classensteuer, der classificierten Einkommensteuer und der Mahl- und Sachlachtsteuer im preußischen Staate", *Zeitschrift des Königlich Preußischen Statistischen Bureaus,* 8(1), pp. 25–84.

Engel, E. (1875), "Die Ergebnisse der Classensteuer und der classificierten Einkommensteuer im preußischen Staate", *Zeitschrift des Königlich Preußischen Statistischen Bureaus,* 15(1), pp. 105–148.

Eggert, U. (1883), "Die Bewegung der Holzpreise und Tagelohn-Sätze in den preußischen Staatsforsten von 1800 bis 1879", *Zeitschrift des Königlich Preußischen Statistischen Bureaus,* 23(1), pp. 1–44.

Feinstein, C.H. and Thomas, M. (2001), *A Plea for Errors,* Oxford Discussion Paper in Economic and Social History no. 41.

Feinstein, C.H. (1988), "The Rise and Fall of the Williamson Curve", *Journal of Economic History*, 47(3), pp. 699–729.

Fei, J.E., Ranis, G. and Kuo, S.Y. (1979), *Growth with Equity: the Taiwan Case* [Oxford: World Bank/OUP].

Fremdling, R. (1988), "German National Accounts for the 19th and early 20th Century. A Critical Assessment", *Vierteljahrschrift für Sozial- und Wirtschaftsgeschichte,* 75, pp. 339–354.

Fremdling, R. (1995), "German National Income, Product and Expenditure 1816–1934, a review of the evidence", *Économies et Sociétés,* 21(Dec.), pp. 25–55.

Geisenberger, S. and Müller, J.H. (1972), *Die Einkommenstruktur in verschiedenen deutschen Ländern 1874–1914,* Schriften zu Regional- und Verkehrsproblemen 10 [Berlin].

Grant, O. (2000), *Internal Migration in Germany, 1870–1913,* Oxford D.Phil. thesis.

Grant, O. (2002), "Max Weber and Die Lage der Landarbeiter im ostelbischen Deutschland: a statistical examination", forthcoming in the *Jahrbuch für Wirtschaftsgeschichte.*

Helferrich, K. (1915), *Deutschlands Volkswohlstand*, [Berlin: G.Stilbe].

Hentschel, V. (1979), "Erwerbs- und Einkommensverhältnisse in Sachsen, Baden und Württemberg vor dem Ersten Weltkrieg", *Vierteljahrschrift für Sozial- und Wirtschaftsgeschichte*, 66(1), pp. 26–73.

Higgins, M. and Williamson, J. (1999), *Explaining Inequality the World Round: cohort size, Kuznets curves and openness,* NBER working paper no. 7224.

Hoffmann, W. (1965), *Das Wachstum der Deutschen Wirtschaft seit der Mittel des 19. Jahrhunderts* [Berlin: Springer].

Hohls, R. (1995), "Germany" in Scholliers, P. and Zamagni, V. eds. *Labour's Reward* [Aldershot: E. Elgar].

Jeck, A. (1970), *Wachstum und Verteilung des Volkseinkommens* [Tübingen].

Kaelble, H. (1986), *Industrialisation and Social Inequality in Nineteenth Century Europe* [Leamington Spa: Berg].

Kiesewetter, H. (1988), *Industrialisierung und Landwirtschaft : Sachsens Stellung im regionalen Industrialisierungsprozess Deutschlands im 19. Jahrhundert* [Köln: Böhlau].

Kucyzinski, R. (1908), *Einkommen und Miete* [Berlin: Berlin Statistische Amt].

Kuhnert, F. (1911), "Die Ergebnisse der preußischen Einkommens- und Ergänzungssteuer: Veranlagung nach Städtegruppen für das Jahr 1908", *Zeitschrift des Königlich Preußischen Statistischen Bureaus,* 51(1), pp. 1–54.

Kuhnert, F. (1916), "Einkommensgliederung der preußischen Bevölkerung im Zeitraum 1902–1911", *Zeitschrift des Königlich Preußischen Statistischen Bureaus,* 56(3), pp. 269–309.

Kuznets, S. (1955), "Economic Growth and Income Inequality", *American Economic Review*, 45, pp. 1–28.

Lindert, P. (1994) "Unequal Living Standards" in Floud, R. and McCloskey, D., *The Economic History of Britain; Volume one: 1700–1860* [Cambridge: CUP].

Lindert, P. and Williamson, J.G. (1980) *American Inequality* [New York: Academic Press].

Meisel, F. (1911), "Moral und Technik bei der Veranlagung der preußischen Einkommensteuer", *Jahrbuch für Gesetzgebung, Verwaltung und Volkswirtschaft im Deutschen Reich,* 35(1), pp. 285–374.

Piketty, T. and Saez, E. (2001), *Income Inequality in the United States, 1913–1998,* NBER Working Paper no. 8467.

Prokopovitch, S. (1926), "The Distribution of National Income", *Economic Journal*, 36, pp. 69–82.

Saalfeld, D. (1963), "Zur Frage des bäuerlichen Landverlustes im Zusammenhang mit den preußischen Agrarreformen", *Zeitschrift für Agrargeschichte und Agrarsoziologie,* 11, pp. 163–171.

Schmoller, G. (1895), "Die Einkommensverteilung in alter und neuer Zeit", *Jahrbuch für Gesetzgebung, Verwaltung und Volkswirtschaft im deutschen Reich,* 19, pp. 1067–1094.

Soetbeer, A. (1882), *Umfang und Verteilung des Volkseinkommens im Preußischen Staate, 1872–1878* [Leipzig].

Wagner, A.(1904), "Zur Methodik der Statistik des Volkeinkommens und Volksvermögens", *Zeitschrift des Königlich Preußischen Statistischen Bureaus*, 44(1), pp. 41–122.

Wagner, A. (1904a), "Weitere statistische Untersuchungen über die Verteilung des Volkseinkommen in Preuβen", *Zeitschrift des Königlich Preußischen Statistischen Bureaus*, 44(4), pp. 229–267.

Wagner, A. (1880), *Finanzwissenschaft. Vol.2* [Leipzig].

Witt, P.-C. (1970), *Die Finanzpolitik des Deutschen Reiches von 1903 bis 1913* [Hamburg: Matthiesen].

Witt, P.-C. (1973), "Der preußische Landrat als Steuerbeamter 1891–1918. Bemerkungen zur politischen und sozialen Funktion des deutschen Beamtentums" in Geiss, I. and Wendt, B. *Deutschland in der Weltpolitik des 19. und 20. Jahrhunderts* [Düsseldorf: Bertelsmann]

[Continued from inside front cover]

26 David Stead, *An Arduous and Unprofitable Undertaking: The Enclosure of Stanton Harcourt, Oxfordshire* (November 1998)

27 Oliver Grant, *The Diffusion of the Herringbone Parlour: A Case Study in the History of Agricultural Technology* (December 1998)

28 Antonia Taddei, *London Clubs in the Late Nineteenth Century* (April 1999)

29 Liam Brunt, *Estimating English Wheat Production in the Industrial Revolution* (June 1999)

30 Matthew Braham, *Volunteers for Development: A Test of the Post-Materialism Hypothesis in Britain, c.1965–1987* (June 1999)

31 Paul A. David and Gavin Wright, *General Purpose Technologies and Surges in Productivity: Historical Reflections on the Future of the ICT Revolution* (September 1999)

32 Liam Brunt, *An Arbitrage Model of Crop Rotation* (September 1999)

33 Paul A. David and Gavin Wright, *Early Twentieth Century Productivity Growth Dynamics: An Inquiry into the Economic History of 'Our Ignorance'* (October 1999)

34 Avner Offer, *Economic Welfare Measurements and Human Well-Being* (January 2000). Rev. version, March 2000.

35 Liam Brunt, *'Where there's Muck, There's Brass'. The Market for Manure in the Industrial Revolution* (February 2000).

36 Alasdair Crockett, *Variations in Churchgoing Rates in England in 1851: Supply-Side Deficiency or Demand-Led Decline?* (August 2000).

37 Martin West, *State Intervention in English Education, 1833–1891: A Public Goods and Agency Approach* (October 2000).

38 George Speight, *Who Bought the Inter-War Semi? The Socio-Economic Characteristics of New-House Buyers in the 1930s* (December 2000)

39 Peter Temin, *A Market Economy in the Early Roman Empire* (March 2001)

40 Michael Biggs, *Positive Feedback in Collective Mobilization: The American Strike Wave of 1886* (April 2001)

41 Charles H. Feinstein and Mark Thomas, *A Plea for Errors* (July 2001)

42 Walter Eltis, *Lord Overstone and the Establishment of British Nineteenth-Century Monetary Orthodoxy* (December 2001)

43 A. B. Atkinson, *Top Incomes in the United Kingdom over the Twentieth Century* (February 2002)

44 Avner Offer, *Why has the Public Sector Grown so Large in Market Societies? The Political Economy of Prudence in the UK, c.1870–2000* (March 2002)

45 Natàlia Mora Sitjà, *Labour and Wages in Pre-Industrial Catalonia* (May 2002)

46 Elaine S. Tan, *'The Bull is Half the Herd': Property Rights and Enclosures in England, 1750–1850* (June 2002)

47 Oliver Wavell Grant, *Productivity in German Agriculture: Estimates of Agricultural Productivity from Regional Accounts for 21 German Regions: 1880/4, 1893/7 and 1905/9* (August 2002)

48 Oliver Wavell Grant, *Does industrialization push up inequality? New evidence on the Kuznets cure from nineteenth century Prussian tax statistics* (September 2002)

University of Oxford
Discussion Papers in
Economic and Social History

are edited by:

Robert Allen
Nuffield College, Oxford, OX1 1NF

Liam Brunt
Nuffield College, Oxford, OX1 1NF

Jane Humphries
All Souls College, Oxford, OX1 4AL

Avner Offer
All Souls College, Oxford, OX1 4AL

Papers may be downloaded from
http://www.nuff.ox.ac.uk/Economics/History/