A Reconsideration of the Angrist-Krueger Analysis on Returns to Education

Lennart Hoogerheide & Herman K. van Dijk

July 2006

Econometric Institute report EI 2006-15

Abstract

In this paper we reconsider the analysis of the effect of education on income by Angrist and Krueger (1991). In order to account for possible endogeneity of the education spell, these authors use quarter of birth to form valid instruments. Angrist and Krueger apply a classical method, two-stage least-squares (2SLS), and consider results for data sets on individuals from all states of the US. In this paper the research by Angrist and Krueger is extended both in a methodological and an empirical way. Classical as well as Bayesian methods are used. Bayesian results under the Jeffreys prior are emphasized, as these results are valid in finite samples and because in the instrumental variables (IV) regression model the Jeffreys prior is in a certain sense, truly, non-informative. Further, it is considered how results vary between subsets of the data corresponding to regions of the US. Finally, some assumptions of Angrist and Krueger are investigated and it is examined if one could still obtain usable results if some assumptions are dropped. Our main findings are: (1) The Angrist-Krueger results on returns to education for the USA are almost completely determined by data from a few Southern states; (2) The conclusion of Bound, Jaeger and Baker (1995), that the instruments of Angrist and Krueger give hardly any usable information concerning the causal effect of education on wages, is too strong. A model of Angrist and Krueger (or a slightly modified version) can give usable information on the causal effect of education on income in the Southern region of the US; (3) The instruments for education that are based on quarter of birth are stronger for people with at most 8 or at least 14 years of education than for people with 9-13 years of education. This suggests that quarter of birth does not only affect the number of completed years of schooling for those who leave school as soon as the law allows for it, as these persons usually have completed 9-13 years of education. Therefore, if one intends to increase the understanding of the working of the quarter-of-birth instruments, it is a better idea to focus on differences between states in school entry requirements and/or compulsory schooling laws for children of age 5-7 than to concentrate on the differences in compulsory schooling laws for students of age 16-18.

1 Introduction

Measuring the effect of education on income, the (monetary) returns to education, is a matter of great importance for several decision processes. For example, the results of such analysis are relevant for government agencies responsible for compulsory schooling laws, for school districts considering changes in school entrance policies and also for parents deciding when to enroll their children to school. However, a problem is that intellectual capabilities, which are usually not observed, not only influence education but also directly affect income. Therefore, a simple regression of income on the number of years of education may lead to incorrect conclusions. For example, smarter students find school less difficult and may choose to obtain more schooling to signal their high ability. So, even if extra years of education have no effect on income, people with higher education will on average have higher incomes because of their higher abilities. Therefore, one may expect that an ordinary regression of income on the years of education leads to an upward bias, i.e. an overestimated effect of education on income.[1] Another problem is the measurement error in reported education. First, usually only the completed number of years of education is reported. Second, people may misreport their education spell.[2] If the measurement error would be the only problem, one would expect that a simple regression of income on education would result in a downward bias, i.e. an underestimated effect of education on income, as the part of the variation in education that is merely due to measurement error does not lead to variation in income.

A method for solving these problems is the use of instrumental variables that must be correlated with education but uncorrelated with latent capabilities (and measurement errors). However, it is hard to find variables that are correlated with education but uncorrelated with intellectual capabilities. Angrist and Krueger (1991) use American data and suggest using quarter of birth to form instrumental variables. These instruments exploit that students born in different quarters have different average education. This results since most school districts require students to have turned age six by a certain date, a so-called 'birthday cutoff' which is typically near the end of the year, in the year they enter school, whereas compulsory schooling laws compel students to remain at school until their sixteenth, seventeenth or eighteenth birthday. This asymmetry between school-entry requirements and compulsory schooling laws compels students born in certain months to attend school longer than students born in other months: students born earlier in the year enter school at an older age and reach the legal dropout age after less education. Hence, for students who leave school as soon as the schooling laws allow for it, those born in the

[1]The intellectual capabilities of the persons in the sample may not be the only reason for an overestimated effect of education on income. The (often unobserved) intellectual capabilities, income and education level of their parents may also cause an upward bias, as these characteristics of their parents may also influence their education level and have a direct effect on their income; for example, it may be the case that children of more intelligent and higher educated parents on average learn more at home.

[2]Siegel and Hodge (1968) find that the correlation between individuals' education reported in two surveys is only 0.933.

first quarter have on average attended school for three quarters less than those born in the fourth quarter.

The strength of these instruments clearly depends on the fraction of students that immediately leave school when it is permitted. This is, however, only a small part of the total population of students since most students do not immediately leave school when it is allowed and some leave school before they attain the legal dropout age. Angrist and Krueger (1991) mention several factors that influence the size of the latter group. Compulsory schooling laws allow certain officers to take children into custody and/or punish a child's parents if a child does not attend school; and child labor laws restrict or prohibit children of compulsory school age from participating in the work force, the main alternative to attending school. There are, however, exemptions to compulsory schooling laws: students are exempt from compulsory school attendance if they have a high school degree; and in many states there are exemptions for children suffering from physical or mental disabilities, or if they live far from school.

Alongside that the quarter of birth only affects the years of education for a small fraction of the student population, its influence is also limited since it only implies a maximum difference of one year over the different quarters which is small compared to the overall variation in the education spell. Quarter of birth is therefore expected to be a weak instrument. Bound, Jaeger and Baker (1995), for example, show that randomly generated instruments, designed to match the data of Angrist and Krueger (1991), yield results remarkably similar to those based on the actual instruments. Staiger and Stock (1997) also show that inference on the returns to education is strongly affected by the weakness of the quarter of birth instruments. Hence, although the quarter of birth seems a plausible source for constructing instruments, we should be careful with interpreting the results because of the weakness of the instruments.

Angrist and Krueger (1991) only apply a classical method, two-stage least-squares (2SLS), and consider only results for the whole data set of all states of the US. In this paper the research by Angrist and Krueger (1991) is extended. Results are examined for both classical and Bayesian methods. Bayesian results under the Jeffreys prior are emphasized, as these results are valid in finite samples and because in the instrumental variables (IV) regression model the Jeffreys prior is - unlike the flat prior of Drèze (1976) - truly non-informative; see e.g. Kleibergen and Zivot (2003). Furthermore, it is considered how results vary between subsets of the data corresponding to regions of the US. It is shown that results of Angrist and Krueger (1991) on returns to education for the USA are almost completely determined by data from a few Southern states. Finally, some assumptions made by Angrist and Krueger (1991) are investigated and it is examined if one could still obtain usable results if some assumptions are dropped. It is shown that the instruments for education that are based on quarter of birth are stronger for people with at most 8 or at least 14 years of education than for people with 9-13 years of education. This suggests that quarter of birth does not only affect the number of completed years of schooling for those who leave school as soon as the law allows for it, which is suggested by Angrist and Krueger (1991), as these persons are (mostly) contained in the group with 9-13 years of education. Therefore, if one intends to increase the understanding of the working of the quarter-of-birth instruments, it is a better idea to focus on differences between states in school entry requirements and/or compulsory schooling laws for children of age 5-7 than to concentrate on the differences in compulsory schooling laws for students of age 16-18, as is done by Angrist and Krueger (1991). Furthermore, it is shown that the conclusion of Bound, Jaeger and Baker (1995), that the models of Angrist and Krueger (1991) give hardly any usable information concerning the causal effect of education on wages, is too strong. A model of Angrist and Krueger (1991) (or a slightly modified version) can give usable information on the causal effect of education on income in the Southern region of the US.

The paper is organized as follows. In section 2 the particular model and data that we use are described. In sections 3 and 4 classical and Bayesian results are given, respectively, both for the whole data set and for subsamples corresponding to regions of the US. In sections 2 - 4 it is assumed that the assumptions made Angrist and Krueger (1991) are satisfied by the data. In section 5 some assumptions made by Angrist and Krueger (1991) are investigated, and it is examined if one could still obtain usable results if some assumptions are dropped. Section 6 gives conclusions.

2 Model and data

Angrist and Krueger (1991) use data sets concerning men born in the USA in the years 1920-1929, 1930-1939 or 1940-1949, and consider several model specifications. We use a subset of the data used by Angrist and Krueger (1991): a data set on income, years of education and state/quarter/year of birth consisting of 329,509 men born in the USA in the years 1930-1939.[3] We use the following model:

$$\tilde{y}_i = \tilde{x}_i\beta + \textstyle\sum_{j=1}^{9} D^y_{j,i}\, \delta^y_j + \sum_{t=1}^{S-1} D^s_{t,i}\, \delta^s_t + \pi_1 + \tilde{\varepsilon}_i \qquad i = 1, \ldots, T \tag{1}$$

$$\begin{aligned} \tilde{x}_i = \; & \textstyle\sum_{j=1}^{9} D^y_{j,i}\gamma^y_j + \sum_{t=1}^{S-1} D^s_{t,i}\gamma^s_t + \pi_2 \\ & + \textstyle\sum_{t=1}^{S} \sum_{h=2}^{4} D^s_{t,i} D^q_{h,i}\, \pi^{sq}_{th} + \sum_{j=1}^{9} \sum_{h=2}^{4} D^y_{j,i} D^q_{h,i}\, \pi^{yq}_{jh} + \tilde{v}_i \end{aligned} \tag{2}$$

where $\tilde{y}_i$ is the logarithm of the weekly wage of person i in 1979, $\tilde{x}_i$ is the number of *completed* years of education by person i, and the parameter of interest is the return on education β. The dummy variables $D^s_{t,i}$, $D^y_{j,i}$, $D^q_{h,i}$ are equal to 1 if individual i was born in state t, year $1929+j$, quarter h, and equal to 0 otherwise, respectively. S is the number of states of birth, *i.e.* $S = 51$ (including the District of Columbia) if we use all states. We however also consider four subsamples for which we divide the US into four regions

[3]The source of the data is the 1980 Census, 5 percent Public Use Sample.

Table 1: US Census Bureau Regions

	Census region	number of observations	number of states	states (including D.C.)
1.	Northeast	84484	9	Connecticut, Maine, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, Vermont.
2.	Midwest	102267	12	Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, Wisconsin.
3.	South	114391	17	Alabama, Arkansas, Delaware, D.C., Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Virginia, West Virginia.
4.	West	28367	13	Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, Nevada, New Mexico, Oregon, Utah, Washington, Wyoming.
	USA	329509	51	

that are also used by the US Census Bureau, the source of the data. The states and numbers of observations in each region are given by Table 1. The coefficients π_1 and π_2 are the constant terms; $\tilde{\varepsilon}_i$ and $\tilde{v}_i$ are disturbances that are assumed to be jointly normal distributed and independent across individuals.

The state and year dummies $D^s_{t,i}$ and $D^y_{j,i}$ are included in both equations since state and year of birth both influence the education spell and income. The year dummies in the wage equation (1) incorporate the effect of age (measured in years) on income.

The exogenous variables that are excluded from the wage equation (1) are the interactions of state and quarter of birth dummies, and interactions of year and quarter of birth dummies. The interacted state and quarter of birth dummies reflect that the influence of the quarter of birth on education may differ between states which results since compulsory education laws differ between states. The legal dropout age varies between 16, 17 and 18 years and in some states students have to finish the school term. The rules concerning exemptions from the compulsory school attendance vary as well across states. The average number of years of education that students desire also differs between states (see Tables 2 and 3, which show that on average men born in the Southern region in the period 1930-1939 have less education than men born in the other regions); the more years of education that students on average want to attend, the smaller the fraction of students that leave school as soon as the law allows it, and the smaller the coefficients at the interacted state and quarter of birth dummies. Note that π^{sq}_{th} is interpreted as the effect of the h-th ($h = 2, 3, 4$) quarter on education in state s in 1939, i.e. the difference in the years of education between men born in the h-th quarter and the first quarter in 1939 (on average).

The interacted year and quarter of birth dummies reflect that the influence of the quarter of birth on education may change over time. For example, the average number of years of education that students desire may change over time. In fact the average number of years of education has increased from 1930 to 1939, see Table 4.[4] Note that π_{jh}^{yq} ($j = 1, \ldots, 9$;$h = 2, 3, 4$) is interpreted as the difference in the effect of the h-th ($h = 2, 3, 4$) quarter on education between the year $1929 + j$ and 1939, i.e. the difference between the differences in years of education between men born in the h-th quarter and the first quarter between the year $1929 + j$ and 1939 (on average).

Model (1)-(2) reads in matrix notation:

$$\tilde{y} = W\Pi_1 + \tilde{X}\beta + \tilde{\varepsilon} \tag{3}$$
$$\tilde{X} = W\Pi_2 + \tilde{Z}\Pi + \tilde{V} \tag{4}$$

where $\tilde{y} = (\tilde{y}_1, \ldots, \tilde{y}_T)'$, $\tilde{X} = (\tilde{x}_1, \ldots, \tilde{x}_T)'$, $\tilde{\varepsilon} = (\tilde{\varepsilon}_1, \ldots, \tilde{\varepsilon}_T)'$, $\tilde{V} = (\tilde{v}_1, \ldots, \tilde{v}_T)'$; W is the $T \times (S + 9)$ matrix of year and state of birth dummies and a constant term with rows $W_i = (D_{1i}^y, \ldots, D_{9i}^y, D_{1i}^s, \ldots, D_{S-1,i}^s, 1)$, $\tilde{Z}$ is the $T \times 3(S+9)$ matrix with rows Z_i containing the state-and-quarter of birth and year-and-quarter of birth interactions $D_{ti}^s D_{hi}^q$, $D_{ji}^y D_{hi}^q$ ($t = 1, \ldots, S$; $h = 2, 3, 4$; $j = 1, \ldots, 9$). The parameter vectors are the $(S + 9) \times 1$ vectors $\Pi_1 = (\delta_1^y, \ldots, \delta_9^y, \delta_1^s, \ldots, \delta_{S-1}^s, \pi_1)'$, $\Pi_2 = (\gamma_1^y, \ldots, \gamma_9^y, \gamma_1^s, \ldots, \gamma_{S-1}^s, \pi_2)'$ and the $3(S + 9) \times 1$ vector Π containing the coefficients π_{th}^{sq}, π_{jh}^{yq} ($t = 1, \ldots, S$; $h = 2, 3, 4$; $j = 1, \ldots, 9$).

We respecify (3)-(4) as:

$$y = X\beta + \varepsilon \tag{5}$$
$$X = Z\Pi + V \tag{6}$$

where y, X, Z (and the error terms ε, V) contain the residuals of $\tilde{y}$, $\tilde{X}$, $\tilde{Z}$ (and $\tilde{\varepsilon}$, $\tilde{V}$) after regression on W; that is, the observations are 'corrected' for differences in mean across years and states.[5] The restricted reduced form corresponding to the structural form (5)-(6) is given by:

$$y = Z\Pi\beta + u \tag{7}$$
$$X = Z\Pi + V \tag{8}$$

[4]Angrist and Krueger (1991) conclude that as average income in 1979 is approximately equal across birth years 1930-1939, age has no or little influence on income for men between 40 and 49 years old. However, as the average education has increased over years of birth 1930-1939, age may very well have a positive effect that is (on average) compensated by the lower level of education. Note that this does not immediately imply that the variable age should be included in the model, as the year dummies already incorporate the effect of age (measured in years).

[5]In classical inference the Frisch-Waugh-Lovell theorem implies the equivalence of the results from (3)-(4) and (5)-(6). In Bayesian inference this results since specifying a flat prior for Π_1 and Π_2 and integrating out Π_1 and Π_2 in the model (3)-(4) yields the same posterior as considering the model (5)-(6) (upto some factor that is neglectable if the number of observations T is much larger than $S + 9$, the dimension of W_i, as is the case throughout this chapter).

Table 2: Summary statistics of education and wage per region

	Census region	number of observations	education* average	st.dev.	% ≤ 9	% ≤ 10	log weekly wage average	st.dev.
1.	Northeast	84484	13.27	3.12	9.4%	14.2%	5.96	0.65
2.	Midwest	102267	13.06	2.99	10.0%	14.6%	5.97	0.66
3.	South	114391	11.93	3.52	22.0%	28.0%	5.77	0.71
4.	West	28367	13.63	3.01	6.5%	10.1%	6.00	0.65
	USA	329509	12.77	3.28	13.7%	18.7%	5.90	0.68

*In most states people born in 1930-1939 were obliged to enter school at age 5 or 6 and allowed to leave school at age 16 having completed 9 or 10 years of education.

where $u \equiv V\beta + \varepsilon$. The error terms in the structural form and the restricted reduced form have covariance matrix Σ and Ω, i.e. $(\varepsilon_i, v_i)' \sim N(0, \Sigma)$ and $(u_i, v_i)' \sim N(0, \Omega)$, with

$$\Sigma = \begin{pmatrix} \sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix},$$

$$\Omega = \begin{pmatrix} \omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix} = \begin{pmatrix} 1 & \beta' \\ 0 & I \end{pmatrix} \Sigma \begin{pmatrix} 1 & 0 \\ \beta & I \end{pmatrix} = \begin{pmatrix} \sigma_{11} + 2\Sigma_{21}\beta + \beta'\Sigma_{22}\beta & \beta'\Sigma_{22} + \Sigma_{12} \\ \Sigma_{22}\beta + \Sigma_{21} & \Sigma_{22} \end{pmatrix}.$$

Section 3 shows results for two classical methods, two-stage least squares (2SLS) and limited information maximum likelihood (LIML). In section 4 the results are given for Bayesian methods, using either a flat or Jeffreys prior.

Table 3: Summary statistics of education and wage per state of birth

	number of		education			log weekly wage	
state	observations	average	st.dev.	% ≤ 9	% ≤ 10	average	st.dev.
Alabama	8536	11.71	3.46	23.1	29.5	5.72	0.76
Alaska	78	13.47	3.12	7.7	15.4	6.09	0.83
Arizona	1066	13.11	3.27	11.5	16.0	5.96	0.62
Arkansas	5794	11.85	3.41	21.4	27.8	5.77	0.70
California	11078	13.87	2.93	4.6	7.8	6.04	0.65
Colorado	2818	13.32	3.11	9.1	12.8	5.95	0.65
Connecticut	3844	13.31	3.08	9.7	14.3	5.96	0.63
Delaware	598	12.30	2.94	15.7	21.2	5.85	0.61
D.C.	1237	13.83	3.12	6.4	10.6	6.01	0.65
Florida	3913	12.68	3.35	14.4	19.9	5.78	0.69
Georgia	8411	11.50	3.51	24.8	31.4	5.68	0.72
Hawaii	246	13.23	3.06	10.2	14.2	5.97	0.69
Idaho	1599	13.54	3.02	7.5	11.1	5.95	0.64
Illinois	18375	13.35	3.00	8.0	12.6	6.03	0.65
Indiana	8918	12.77	2.87	10.8	16.0	5.94	0.63
Iowa	6699	13.14	2.96	9.3	12.5	5.91	0.70
Kansas	4807	13.44	2.96	7.7	10.5	5.91	0.67
Kentucky	8933	11.27	3.55	30.7	36.8	5.80	0.70
Louisiana	5975	12.07	3.61	20.0	25.6	5.84	0.71
Maine	2424	12.35	3.10	17.0	22.0	5.75	0.64
Maryland	4139	12.44	3.27	16.7	23.4	5.88	0.67
Massachusetts	9955	13.47	3.16	8.9	13.1	5.95	0.64
Michigan	14077	13.00	2.89	9.4	14.9	6.03	0.62
Minnesota	7170	13.19	3.03	10.0	13.5	5.97	0.67
Mississippi	5864	11.49	3.73	25.9	32.6	5.68	0.75
Missouri	9274	12.69	3.18	14.8	19.6	5.90	0.69
Montana	1407	13.38	3.01	8.0	12.2	5.91	0.70
Nebraska	3488	13.34	2.96	7.5	11.1	5.92	0.70
Nevada	308	13.48	2.95	8.1	11.0	5.99	0.73
New Hampshire	1200	12.59	3.15	16.6	21.0	5.80	0.64
New Jersey	8964	13.43	3.11	8.3	13.1	6.00	0.66
New Mexico	1325	12.59	3.41	15.1	19.9	5.85	0.61
New York	29015	13.70	3.16	7.1	11.4	6.01	0.66
North Carolina	10798	11.70	3.43	24.0	30.5	5.66	0.71
North Dakota	2028	12.94	3.28	15.4	19.5	5.93	0.70

Table 5.3 (continued)

state	number of observations	education average	st.dev.	% ≤ 9	% ≤ 10	log weekly wage average	st.dev.
Ohio	17070	12.95	2.95	10.1	15.4	5.97	0.63
Oklahoma	6950	13.00	3.11	11.4	15.9	5.90	0.66
Oregon	2127	13.65	2.86	5.2	8.9	5.99	0.62
Pennsylvania	26385	12.84	2.97	10.8	16.3	5.93	0.62
Rhode Island	1698	12.91	3.21	15.1	20.8	5.85	0.67
South Carolina	5245	11.30	3.58	27.3	34.6	5.61	0.75
South Dakota	1754	13.18	3.23	12.8	15.3	5.90	0.67
Tennessee	8335	11.54	3.51	27.0	32.8	5.75	0.71
Texas	15932	12.67	3.62	15.5	20.1	5.87	0.70
Utah	1999	13.94	3.04	5.5	9.7	6.01	0.61
Vermont	999	12.40	3.18	18.2	22.4	5.73	0.67
Virginia	7319	11.47	3.62	27.3	34.0	5.73	0.71
Washington	3610	13.66	2.90	5.4	8.7	6.04	0.64
West Virginia	6412	11.81	3.16	22.5	28.9	5.85	0.64
Wisconsin	8607	12.96	2.94	10.2	14.5	5.93	0.66
Wyoming	706	13.60	3.02	6.9	10.9	5.99	0.68

Table 4: Summary statistics of education and wage per year of birth

year	number of observations	education average	st.dev.	% ≤ 9	% ≤ 10	log weekly wage average	st.dev.
1930	33602	12.46	3.44	17.2	22.6	5.90	0.69
1931	30583	12.59	3.38	15.8	20.8	5.91	0.69
1932	32211	12.63	3.40	15.9	21.1	5.90	0.69
1933	30751	12.69	3.35	14.9	20.1	5.90	0.68
1934	31916	12.72	3.32	14.4	19.7	5.90	0.69
1935	32773	12.78	3.26	13.7	18.6	5.89	0.69
1936	32676	12.84	3.20	12.6	17.6	5.90	0.67
1937	33969	12.90	3.17	11.7	16.5	5.90	0.66
1938	35223	12.99	3.15	11.1	15.8	5.90	0.67
1939	35805	13.03	3.13	10.6	15.5	5.90	0.66

3 Classical approaches

In this section we first briefly summarize two well-known classical single equation estimators for β, two-stage least squares (2SLS) and limited information maximum likelihood (LIML); for extensive discussions of classical single equation procedures the reader is referred to Hausman (1983) or Phillips (1983). After that the results are discussed of applying the 2SLS and LIML methods to the Angrist-Krueger IV model for data of the US and the four Census regions.

In the two-stage least squares method, due to Theil (1953) and Basmann (1957), an estimator of β is obtained by the following two steps. First, an estimate of Π in (6) is obtained by OLS: $\hat{\Pi}_{OLS} = (Z'Z)^{-1}Z'X$. Second, an estimate of β is obtained by OLS of y on $Z\Pi$ in (7) with Π replaced by $\hat{\Pi}_{OLS}$: $\hat{\beta}_{2SLS} = (X'Z\,(Z'Z)^{-1}Z'X)^{-1}X'Z\,(Z'Z)^{-1}Z'y$. The 2SLS estimator $\hat{\beta}_{2SLS}$ is a consistent estimator of β that is asymptotically normal distributed with covariance matrix $(1/T)\sigma_{11}(\Pi'\Sigma_Z\Pi)^{-1}$, where $\Sigma_Z = \text{plim}_{T\to\infty}(1/T)Z'Z$, under the conditions that β is identified and instruments are not too weak. Staiger and Stock (1997) explore a case of weak instruments, defined as $\Pi = C/\sqrt{T}$ where C is fixed (so that $\Pi'Z'Z\Pi$ converges to a constant as the sample size T grows), where $\hat{\beta}_{2SLS}$ is asymptotically biased. In finite samples $\hat{\beta}_{2SLS}$ is less biased than β_{OLS}, the OLS estimator of β in (5). The tails (and bias) of the finite sample distribution of $\hat{\beta}_{2SLS}$ depend on the degree of overidentification, the number of instruments excluded from the structural equation minus the dimension of β; the moments of the finite sample distribution exist up to/including this degree of overidentification.

In the method of limited information maximum likelihood, due to Anderson and Rubin (1949) and Hood and Koopmans (1953), the estimator for β is the value of β for which the likelihood function of (5)-(6), concentrated with respect to Π and Σ, takes its maximum. It is computed by computing the smallest root λ of the determinantal equation $|\lambda(Y\ X)'(Y\ X) - (Y\ X)'Z(Z'Z)^{-1}Z'(Y\ X)| = 0$ and the corresponding eigenvector, after which multiplying this eigenvector with minus the inverse of its first element yields $(-1, \hat{\beta}_{LIML})'$. Staiger and Stock (1997) show that in their case of weak instruments $\hat{\beta}_{LIML}$ is an inconsistent estimator of β. However, in finite samples $\hat{\beta}_{LIML}$ is (approximately) median unbiased if instruments are not too weak. Staiger and Stock (1997) show several cases in which the LIML estimator is approximately median unbiased whereas the 2SLS estimator suffers from huge biases, which makes Staiger and Stock (1997) conclude that estimator bias is less of a problem for LIML than for 2SLS, so that they suggest using LIML rather than 2SLS point estimates. The tails of the finite sample distribution of $\hat{\beta}_{LIML}$ are Cauchy-type (no matter the degree of overidentification), so that $\hat{\beta}_{LIML}$ has no finite moments.

Angrist and Krueger (1991) report the 2SLS estimate of β in the model (5)-(6): $\hat{\beta}_{2SLS} = 0.0928$ with an asymptotic standard error of 0.0093 (column (2) of Table VII). Next to that they report the OLS estimate of β in (5): 0.0673 (with standard error of 0.0003; column (1) of Table VII).

Table 5 shows the results of 2SLS and OLS for the Census regions. This suggests that the 2SLS estimate for the US is almost completely determined by the region South: the difference between the 2SLS estimates for the US and the South is small, and the asymptotic standard error for the South is not much larger than that for the US. An explanation for this result is that the average education level for men born in 1930-1939 is lower in the region South than in the other regions, see Table 2. The influence of compulsory schooling laws is therefore larger for the South, as more students desire to leave school as soon as it is allowed. Therefore the influence of quarter of birth is larger in the Southern region, so that the instruments are strongest in the South.

One problem that the 2SLS estimator may suffer from is that it is biased in the case of weak instruments; this is illustrated by the last column of Table 5, which shows the mean of 10,000 2SLS estimators for 10,000 data sets simulated from (5)-(6) with parameter values chosen as $\beta = \hat{\beta}_{2SLS}$, $\Pi = \hat{\Pi}_{OLS}$ and Σ the covariance matrix of the residuals. The five means are all biased in the direction of the corresponding OLS estimator where the relative bias, the difference between the mean of the 2SLS estimates and the true β in the simulations divided by the difference between the OLS estimate and the true β, is smaller for the US and the Southern region than for the other three regions. In fact it is smaller for the South than for the US, which reflects that the addition of superfluous (or very weak) instruments results in a 2SLS estimator with smaller variance but larger bias.

Table 6 shows LIML estimates for the four Census regions, and quantiles of the estimated finite sample distribution, where maximum likelihood estimates substituted for β, Π and Σ in the finite sample density of $\hat{\beta}_{LIML}$ for the case of one explanatory endogenous variable that is given by Kleibergen (2000) and Kleibergen and Zivot (2003). Again the results are dominated by the Southern region: the difference between the LIML estimates for the US and the South is small, and the 95% and 50% density intervals for the South are not much larger than those for the US. Since LIML is known to focus on the strongest available instruments, this confirms that the instruments are much stronger in the South than in the other regions. Also notice that the median of the finite sample distribution of the LIML estimator is approximately equal to the 'true value', the ML estimate, which reflects that the LIML estimator is approximately median unbiased. So, Tables 5 and 6 illustrate that the LIML estimator is a better point estimator than the 2SLS estimator, although in this case the 2SLS estimator is also 'smart enough' to indicate that the strongest instruments stem from the region South.

The results for the four Census regions suggest that a further division of the data set into states (or groups of states) may be interesting. We first take a closer look at the first stage regression, the OLS results in model (5). Table 7 shows the estimated coefficients $\hat{\pi}_{th}^{sq}$ at the interactions of state and quarter of birth, which are interpreted as the effect of the quarter of birth (as compared with the first quarter) in the year 1939. It shows t-values larger than 3 for Arkansas, Kentucky and Tennessee (and t-values larger than 2 for Alabama, Arizona, California, Colorado, Georgia, Illinois, Louisiana, Maryland, Massachusetss, Mississippi. North Carolina, North Dakota, Texas, Virginia). The effect of quarter of birth on education should be on average smaller than 0.75, which is not satisfied by the estimated coefficients of Alaska, Hawaii and Nevada; this is obviously caused by the small numbers of observations for these states. Table 8 shows the estimated coefficients $\hat{\pi}^{yq}_{jh}$ at the interactions of year and quarter of birth. This shows that the influence of quarter of birth is clearly strongest for men born in 1930. One explanation is that men born in 1930 have on average less education than men born in 1931-1939, so that compulsory schooling laws are more important for this group. Another, more specific, reason could be that these men attain age 16 in 1946, right after World War II when there is arguably a lot of work for young men.

Another way to look at the strength of the quarter of birth instruments for each state is to consider the p-value of the multiple F-test in the first stage regression when considering only data of one state. These F-statistics (and corresponding p-values) are given by Table 9, which also shows the estimated concentration parameter $\Pi'Z'Z\Pi/\sigma^2_{22}$ (with $\sigma^2_{22} = \text{var}(v_i)$; see Basmann(1963)), where $\hat{\Pi}_{OLS}$ and the variance of the residuals in the first stage regression are substituted for Π and σ^2_{22}.

Table 9 shows that three states in the Census region South, Arkansas, Kentucky and Tennessee, have the largest concentration parameter, and the smallest p-values in the multiple F-test ($p < 0.001$). For Kansas the p-value in the multiple F-test is smaller than 0.01. We have $p < 0.1$ for Arizona and three Southern states, Georgia, South Carolina and Texas.

The results of the multiple F-test in the first stage regression for data of one state are graphically illustrated by Figure 1. The three states with p-value smaller than 0.001, Arkansas, Kentucky and Tennessee, are neighboring states in the region South.

Kentucky has the highest concentration parameter (and smallest p-value at the F-test); it is no coincidence that for men born in 1930-1939 those born in Kentucky have the lowest education on average, so that the influence of compulsory schooling laws is relatively large in Kentucky. Arkansas and Tennessee also have relatively low average education levels. However, Virginia and Mississippi have lower average education levels; for Tennessee the states of Alabama and West Virginia also have lower average education. This suggests that the average amount of education desired by people is not the only factor influencing the strength of the quarter of birth instruments: there are also other factors playing a role, which may include the power of government agencies enforcing schooling laws and the exemptions from these schooling laws that vary between states.

Tables 5 and 6 show the results of 2SLS and LIML for data of men born in Kentucky, Arkansas or Tennessee. Notice that the uncertainty in the LIML estimator, reflected by the 95% and 50% density intervals of the (estimated) finite sample distribution, increases by a relatively small amount, as compared with the US. The width of the 95% and 50% density intervals are only 1.93 and 1.92 times larger than for the US while the whole data

Table 5: OLS and 2SLS estimates for β in (5)-(6) for data of US and Census regions

		OLS		2SLS		2SLS (10000 simulations)	
	Region	$\hat{\beta}_{OLS}$	(st.error)	$\hat{\beta}_{2SLS}$	asympt. std.error	mean $\hat{\beta}_{2SLS}$	relative bias
	USA	0.0673	(0.0003)	0.0928	(0.0093)	0.0858	0.275
1	Northeast	0.0738	(0.0007)	0.0707	(0.0234)	0.0721	0.452
2	Midwest	0.0621	(0.0007)	0.0796	(0.0224)	0.0724	0.411
3	South	0.0691	(0.0006)	0.0931	(0.0120)	0.0874	0.238
4	West	0.0559	(0.0012)	0.0506	(0.0206)	0.0530	0.453
	Kentucky, Arkansas & Tennessee	0.0653	(0.0013)	0.0970	(0.0168)		

Table 6: LIML estimates for β in (5)-(6) for data of US and Census regions

			Quantile finite sample dist. $\hat{\beta}_{LIML}$				
	Region	$\hat{\beta}_{LIML}$	median	2.5%	97.5%	25%	75%
	USA	0.1064	0.106	0.0877	0.1256	0.0999	0.1129
1	Northeast	0.0650	0.065	0.0163	0.1124	0.0487	0.0810
2	Midwest	0.1298	0.128	0.0836	0.1825	0.1135	0.1468
3	South	0.1071	0.107	0.0828	0.1324	0.0986	0.1156
4	West	0.0449	0.045	0.0014	0.0874	0.0302	0.0593
	Kentucky, Arkansas & Tennessee	0.1046	0.104	0.0694	0.1420	0.0922	0.1170

set of the US has over 14 times as many observations (329509 vs. 23062). Further, these 95% and 50% density intervals are tighter for the data set of Kentucky, Arkansas and Tennessee than for the region Northeast, Midwest or West. This stresses the importance of the observations on men born in the states of Arkansas, Kentucky and Tennessee for the inference on return on education.

Table 7: Estimated coefficients $\hat{\pi}_{th}^{sq}$ at interactions of state and quarter of birth in first stage regression

State	Second quarter		Third quarter		Fourth quarter	
	coefficient	(t-value)	coefficient	(t-value)	coefficient	(t-value)
Alabama	0.0087	(0.0799)	0.2339	(2.1891)	0.2591	(2.4017)
Alaska	2.0370	(1.9631)	1.6165	(1.4696)	-0.0412	(-0.0342)
Arizona	0.6195	(2.2070)	-0.1240	(-0.4569)	-0.2899	(-1.0466)
Arkansas	-0.2182	(-1.6787)	0.0271	(0.2175)	0.4230	(3.3300)
California	0.2272	(2.3296)	0.1102	(1.1609)	0.1029	(1.0606)
Colorado	0.2718	(1.5384)	0.3377	(1.9393)	0.3958	(2.2194)
Connecticut	0.2548	(1.6819)	0.0123	(0.0807)	0.0875	(0.5661)
Delaware	0.4562	(1.2199)	0.5355	(1.4456)	0.0847	(0.2232)
D.C.	-0.4522	(-1.7005)	-0.4764	(-1.8166)	-0.4318	(-1.6783)
Florida	0.2481	(1.6166)	0.1023	(0.6891)	0.2073	(1.4024)
Georgia	-0.2805	(-2.5634)	-0.0438	(-0.4112)	0.0417	(0.3827)
Hawaii	-0.0748	(-0.1191)	1.4881	(2.6138)	0.7564	(1.3884)
Idaho	0.1301	(0.5699)	-0.0874	(-0.3848)	0.1135	(0.4925)
Illinois	0.0294	(0.3604)	-0.1606	(-2.0227)	0.0464	(0.5712)
Indiana	-0.0111	(-0.1038)	-0.0385	(-0.3720)	0.0140	(0.1315)
Iowa	-0.0846	(-0.7066)	-0.1300	(-1.1160)	0.0574	(0.4814)
Kansas	0.2322	(1.6350)	0.2251	(1.6784)	0.1554	(1.1369)
Kentucky	0.0492	(0.4631)	0.2550	(2.4374)	0.5142	(4.8730)
Louisiana	0.0434	(0.3328)	0.1172	(0.9531)	0.2686	(2.1695)
Maine	-0.0241	(-0.1276)	0.1437	(0.7740)	0.0746	(0.3942)
Maryland	0.3283	(2.2095)	0.3100	(2.1400)	0.2962	(2.0239)
Massachusetts	0.0894	(0.8809)	0.0678	(0.6789)	0.2404	(2.3322)
Michigan	0.1278	(1.4397)	0.0133	(0.1521)	0.1005	(1.1230)
Minnesota	-0.2100	(-1.8148)	-0.2733	(-2.3791)	-0.1192	(-1.0244)
Mississippi	0.0230	(0.1813)	0.1199	(0.9743)	0.3059	(2.4262)
Missouri	-0.1274	(-1.2029)	-0.0186	(-0.1827)	-0.0304	(-0.2922)
Montana	-0.0333	(-0.1376)	0.0058	(0.0234)	0.3234	(1.3071)
Nebraska	-0.1431	(-0.8979)	-0.2160	(-1.3662)	-0.1432	(-0.8997)
Nevada	-0.1005	(-0.1757)	-0.0600	(-0.1113)	0.7544	(1.3369)
New Hampshire	-0.1475	(-0.5482)	-0.0027	(-0.0104)	0.1464	(0.5420)
New Jersey	0.0372	(0.3524)	-0.0448	(-0.4308)	0.1848	(1.7262)
New Mexico	0.1925	(0.7601)	0.0274	(0.1052)	0.4056	(1.6175)
New York	0.0571	(0.8143)	-0.0631	(-0.9140)	-0.0776	(-1.1021)
North Carolina	-0.0857	(-0.8653)	0.0385	(0.3969)	0.2136	(2.1662)
North Dakota	-0.4979	(-2.4204)	-0.3019	(-1.4854)	-0.1407	(-0.6779)

Table 5.7 (continued)

State	Second quarter		Third quarter		Fourth quarter	
	coefficient	(t-value)	coefficient	(t-value)	coefficient	(t-value)
Ohio	-0.0847	(-1.0210)	-0.0583	(-0.7150)	0.0232	(0.2794)
Oklahoma	-0.0800	(-0.6610)	0.0400	(0.3501)	0.2066	(1.7793)
Oregon	0.0550	(0.2705)	-0.0238	(-0.1208)	0.0233	(0.1145)
Pennsylvania	-0.0288	(-0.3996)	-0.0959	(-1.3526)	0.0174	(0.2408)
Rhode Island	-0.3987	(-1.8053)	0.0585	(0.2624)	0.1230	(0.5475)
South Carolina	-0.1087	(-0.8025)	-0.0978	(-0.7348)	0.2983	(2.2239)
South Dakota	0.2971	(1.3435)	0.1932	(0.8849)	0.5111	(2.2999)
Tennessee	-0.1076	(-0.9929)	0.0848	(0.7879)	0.4465	(4.0963)
Texas	-0.0935	(-1.0690)	0.1063	(1.2806)	0.2505	(2.9811)
Utah	-0.0264	(-0.1254)	-0.2249	(-1.0941)	-0.2249	(-1.0827)
Vermont	0.1709	(0.5889)	0.3196	(1.0907)	0.2557	(0.8531)
Virginia	0.0300	(0.2608)	0.1905	(1.6867)	0.2881	(2.4945)
Washington	0.1001	(0.6368)	0.0280	(0.1802)	0.0009	(0.0054)
West Virginia	-0.0901	(-0.7396)	-0.0093	(-0.0774)	0.2320	(1.9120)
Wisconsin	0.1516	(1.4148)	-0.1032	(-0.9689)	0.0702	(0.6479)
Wyoming	0.0806	(0.2302)	0.1759	(0.5205)	-0.1905	(-0.5308)

Table 8: Estimated coefficients $\hat{\pi}_{jh}^{yq}$ at interactions of quarter and year of birth in first stage regression (1939 = reference year)

State	Second quarter		Third quarter		Fourth quarter	
	coefficient	(t-value)	coefficient	(t-value)	coefficient	(t-value)
1930	0.1538	(2.2272)	0.1881	(2.7767)	0.2191	(3.1690)
1931	-0.0378	(-0.5337)	0.1470	(2.1176)	-0.0472	(-0.6650)
1932	0.0804	(1.1494)	0.0977	(1.4301)	0.0962	(1.3841)
1933	-0.0677	(-0.9584)	0.0681	(0.9805)	-0.0989	(-1.4070)
1934	0.0792	(1.1212)	0.0574	(0.8356)	0.0340	(0.4865)
1935	0.1162	(1.6620)	0.1856	(2.7248)	0.0170	(0.2433)
1936	0.0274	(0.3931)	0.1067	(1.5671)	0.0209	(0.3009)
1937	0.0120	(0.1729)	0.1372	(2.0335)	0.0472	(0.6843)
1938	0.0396	(0.5765)	0.0291	(0.4354)	-0.0348	(-0.5096)

Table 9: Summary statistics of first-stage regression for data of individual states

State	concentration parameter	R^2	F-stat.	# obs.	p-value F-stat.
Alabama	18.68	0.0022	0.62	8536	0.9463
Arizona	45.64	0.0426	1.52	1066	0.0365
Arkansas	60.06	0.0103	2.00	5794	0.0009
California	35.93	0.0032	1.20	11078	0.2107
Colorado	31.68	0.0113	1.06	2818	0.3836
Connecticut	38.25	0.0100	1.27	3844	0.1448
Delaware	29.74	0.0506	0.99	598	0.4814
D.C.	27.64	0.0226	0.92	1237	0.5892
Florida	36.19	0.0093	1.21	3913	0.2031
Georgia	41.29	0.0049	1.38	8411	0.0827
Hawaii	27.70	0.1185	0.92	246	0.5853
Idaho	22.57	0.0143	0.75	1599	0.8309
Illinois	32.74	0.0018	1.09	18375	0.3341
Indiana	27.12	0.0030	0.90	8918	0.6169
Iowa	25.64	0.0038	0.85	6699	0.6931
Kansas	52.35	0.0109	1.75	4807	0.0072
Kentucky	68.54	0.0076	2.28	8933	0.0001
Louisiana	31.18	0.0052	1.04	5975	0.4071
Maine	24.15	0.0100	0.80	2424	0.7645
Maryland	34.10	0.0083	1.14	4139	0.2777
Massachusetts	36.02	0.0036	1.20	9955	0.2080
Michigan	23.78	0.0017	0.79	14077	0.7818
Minnesota	15.82	0.0022	0.53	7170	0.9841
Mississippi	30.43	0.0052	1.01	5864	0.4442
Missouri	28.46	0.0031	0.95	9274	0.5460
Montana	26.99	0.0194	0.90	1407	0.6233
Nebraska	30.94	0.0089	1.03	3488	0.4189
Nevada	25.42	0.0866	0.85	308	0.6988
New Hampshire	25.18	0.0212	0.84	1200	0.7147
New Jersey	31.73	0.0035	1.06	8964	0.3803
New Mexico	26.10	0.0199	0.87	1325	0.6690
New York	38.23	0.0013	1.27	29015	0.1440
North Carolina	40.10	0.0037	1.34	10798	0.1033
North Dakota	30.95	0.0153	1.03	2028	0.4191

Table 5.9 (continued)

State	concentration parameter	R^2	F-stat.	# obs.	p-value F-stat.
Ohio	32.16	0.0019	1.07	17070	0.3601
Oklahoma	36.85	0.0053	1.23	6950	0.1822
Oregon	20.37	0.0097	0.68	2127	0.9054
Pennsylvania	27.70	0.0011	0.92	26385	0.5862
Rhode Island	33.96	0.0201	1.13	1698	0.2850
South Carolina	50.91	0.0097	1.70	5245	0.0102
South Dakota	38.39	0.0219	1.28	1754	0.1428
Tennessee	64.06	0.0077	2.14	8335	0.0003
Texas	50.01	0.0031	1.67	15932	0.0125
Utah	31.85	0.0160	1.06	1999	0.3761
Vermont	36.54	0.0367	1.22	999	0.1960
Virginia	28.61	0.0039	0.95	7319	0.5384
Washington	18.54	0.0052	0.62	3610	0.9487
West Virginia	34.36	0.0054	1.15	6412	0.2673
Wisconsin	27.27	0.0032	0.91	8607	0.6090
Wyoming	29.27	0.0421	0.98	706	0.5050

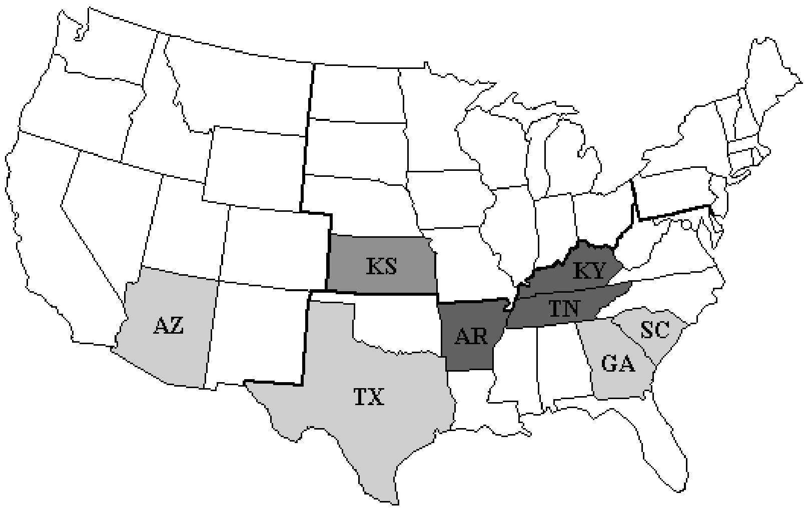

Figure 1: p-value of multiple F-test in first stage regression for data of individual states: p-value < 0.001: dark grey, p-value < 0.01: grey, p-value < 0.1: light grey.
(AR = Arkansas, AZ = Arizona, GA = Georgia, KS = Kansas, KY = Kentucky, SC = South Carolina, TN = Tennessee, TX = Texas)

4 Bayesian approaches

In this section we first briefly discuss the posterior distributions under two commonly used prior density kernels, the flat prior of Drèze (1976) and the Jeffreys prior. For an extensive discussion of these Bayesian approaches (and their relations to the 2SLS and LIML estimators) the reader is referred to Kleibergen and Zivot (2003). After that the results are discussed of applying these Bayesian methods to the Angrist-Krueger IV model for data of the US and the four Census regions.

Drèze (1976) specifies the following flat prior on the parameters of the structural form (5)-(6):

$$p(\beta, \Pi, \Sigma) \propto |\Sigma|^{-1/2\,(k+m+2)} \tag{9}$$

where k is the number of instruments in Z and m is the number of explanatory endogenous variables in X. The primary motivation of this flat prior is that it has an invariance property in the sense that the prior on the structural form implies the same kind of prior on the parameters of the restricted reduced form (which is proportional to $|\Sigma|^{-1/2\,(k+m+2)}$). The marginal posterior of β resulting from the prior (9) is given by:

$$p(\beta|y, X, Z) \propto \left(\frac{(y - X\beta)' M_Z (y - X\beta)}{(y - X\beta)'(y - X\beta)}\right)^{T/2} (y - X\beta)'(y - X\beta)^{-k/2}, \tag{10}$$

where $M_Z \equiv I - Z(Z'Z)^{-1}Z'$. The tails of this posterior of β become thinner when (possibly superfluous) instruments are added to the model, see *e.g.* Maddala (1976) and Kleibergen and Zivot (2003). Further, the location of the posterior mode moves towards the OLS estimator when superfluous instruments are added. Bayesian inference under the flat prior of Drèze (1976) shares these properties with the small sample distribution of the 2SLS estimator which made Kleibergen and Zivot (2003) conclude that this approach has more in common with 2SLS than with LIML.

The Jeffreys prior, the square root of the determinant of the information matrix, is given by:

$$p(\beta, \Pi, \Sigma) \propto |\Sigma|^{-(m+1)} |\Pi' Z' Z \Pi|^{1/2} |\Sigma_{22.1}|^{-1/2(k-m)} \tag{11}$$

with $\Sigma_{22.1} \equiv \Sigma_{22} - \Sigma_{21}\sigma_{11}^{-1}\Sigma_{12}$ for the structural form (5)-(6) or equivalently by:

$$p(\beta, \Pi, \Omega) \propto |\Omega|^{-(m+1)} |\Pi' Z' Z \Pi|^{1/2} |(\beta : I_m)\Omega^{-1}(\beta : I_m)'|^{1/2(k-m)} \tag{12}$$

for the corresponding restricted reduced form (7)-(8); see for example Appendix A of Hoogerheide, Kleibergen and Van Dijk (2006) for a derivation of this Jeffreys prior. In the case of $m = 1$ and for moderate values of T ($T > 20$), an accurate approximation of the marginal posterior of β can be obtained by

$$\begin{aligned} p(\beta|\Omega, y, X, Z) \;\propto\; & [(\beta - \phi)^2 \omega_{11.2}^{-1} + \omega_{22}^{-1}]^{-1} \\ & \times \sum_{j=0}^{\infty} \frac{\Gamma[(k + 2j + 1)/2]}{j!\,\Gamma[(k + 2j)/2]} \left(\frac{(\beta : 1)\Omega^{-1}\hat{\Phi}' Z' Z \hat{\Phi}\Omega^{-1}(\beta : 1)'}{2[(\beta - \phi)^2 \omega_{11.2}^{-1} + \omega_{22}^{-1}]}\right)^j \end{aligned} \tag{13}$$

where $\Omega = (y : X)'(y : X)/T$ is substituted for Ω, and where $\phi \equiv \omega_{21}/\omega_{22}$, $\omega_{11.2} \equiv \omega_{11} - \omega_{21}^2/\omega_{22}$, $\hat{\Phi} = (Z'Z)^{-1}Z'(y : X)$, see Kleibergen and Zivot (2003). The primary motivation of the Jeffreys prior is its universal invariance property with respect to parameter transformations. Kleibergen and Zivot (2003) show that Bayesian analysis using a Jeffreys prior leads to, when there is only $m = 1$ explanatory endogenous variable, a functional expression of the marginal posterior of β that is identical to the finite sample density of the LIML estimator. Just like the finite sample distribution of the LIML estimator, the posterior based on the Jeffreys prior retains Cauchy type tails when (possibly irrelevant) instruments are added, and the location of the mode is insensitive to the addition of superfluous instruments.

Table 10 shows some summary statistics of the posterior distribution of β under the flat or Jeffreys prior. Just like the results for the 2SLS and LIML estimators, the posterior distribution of β for the US under the flat or Jeffreys prior is almost completely determined by the region South; the difference between the means or medians for the US and the South is small, and the posterior standard deviation and 95% and 50% posterior density intervals for the South are not much larger than those for the US, whereas the posterior density intervals are relatively large for the other regions.[6] It is shown in Hoogerheide, Kleibergen and Van Dijk (2006) that Bayesian analysis using the Jeffreys prior, similar to the LIML estimator, focusses on the strongest available instruments. So, the posterior results under the Jeffreys prior once more indicate that the quarter of birth instruments are strongest in the South. Figure 2 shows the graphs of the posterior densities under the flat and Jeffreys prior, respectively.

For all four approaches that have been considered in this chapter, the two classical methods as well as the two Bayesian approaches, inference on the return to education for the US is almost completely determined by the returns to education in the South. If the effect of the return on education is different for the other regions, which can not a priori be ruled out given the large economic differences between these regions, inference using data of the US is not representative for the average returns on education across the US. One should thus be careful when drawing such conclusions.

Notice that the results for the flat prior are remarkably similar to those for the 2SLS estimator: for each region the posterior mean of β is close to the 2SLS estimator $\hat{\beta}_{2SLS}$ and the posterior standard deviation is close to the asymptotic standard error. This agrees with the conclusions of Kleibergen and Zivot (2003) that Bayesian analysis using the flat prior is closer to 2SLS than to LIML. Also note that the results for the Jeffreys prior are similar to those for the LIML estimator: the posterior median of β is close to the LIML estimator $\hat{\beta}_{LIML}$, and for US and the region South the difference in the 95% and

[6]These 95% and 50% posterior density intervals are not equal to 95% and 50% Highest Posterior Density (HPD) regions, although the differences are small in these cases of unimodal, almost symmetric distributions.

Table 10: Posterior results under flat or Jeffreys prior for US and regions

	Region	Posterior β under flat prior		Quantile posterior β under Jeffreys prior				
		mean	st.dev.	median	2.5%	97.5%	25%	75%
	USA	0.092	0.009	0.106	0.083	0.129	0.098	0.114
1	Northeast	0.071	0.023	0.064	-0.024	0.150	0.037	0.092
2	Midwest	0.081	0.023	0.129	0.041	0.246	0.099	0.163
3	South	0.095	0.012	0.107	0.077	0.138	0.096	0.117
4	West	0.051	0.020	0.044	-0.018	0.105	0.024	0.065
	Kentucky, Arkansas & Tennessee	0.095	0.016	0.104	0.068	0.143	0.092	0.117

50% intervals between the methods is close, although for the other three Census regions the 95% and 50% posterior intervals under the Jeffreys prior are somewhat larger than the corresponding intervals for the (estimated) finite sample distribution of the LIML estimator.

We also consider the posterior of β under the flat and Jeffreys prior based on only the observations on men born in the states of Arkansas, Kentucky and Tennessee. Table 10 shows summary statistics of these posteriors. Notice that the uncertainty in the posterior under the Jeffreys prior, reflected by the 95% and 50% density intervals of the (estimated) finite sample distribution, increases by a relatively small amount, as compared with the US. The width of the 95% and 50% posterior density intervals are only 1.63 and 1.56 times larger than for the US while the whole data set of the US has over 14 times as many observations (329509 vs. 23062). Further, these 95% and 50% posterior density intervals are tighter for the data set of Kentucky, Arkansas and Tennessee than for the region Northeast, Midwest or West. Figure 3 illustrates the relative strength of the quarter of birth instruments in the states of Arkansas, Kentucky and Tennessee. If we divide the data set of the US in three subsamples, Arkansas-Kentucky-Tennessee (23062 observations), the other 14 states of region South (91329 observations) and the other three Census regions (215118 observations), then the resulting posteriors of β under the Jeffreys prior are about as tight for these three subsamples. These results again stress the importance of the states of Arkansas, Kentucky and Tennessee for the inference on return on education: to a large extent inference on education for the US is determined by the return on education for men born in these three states.

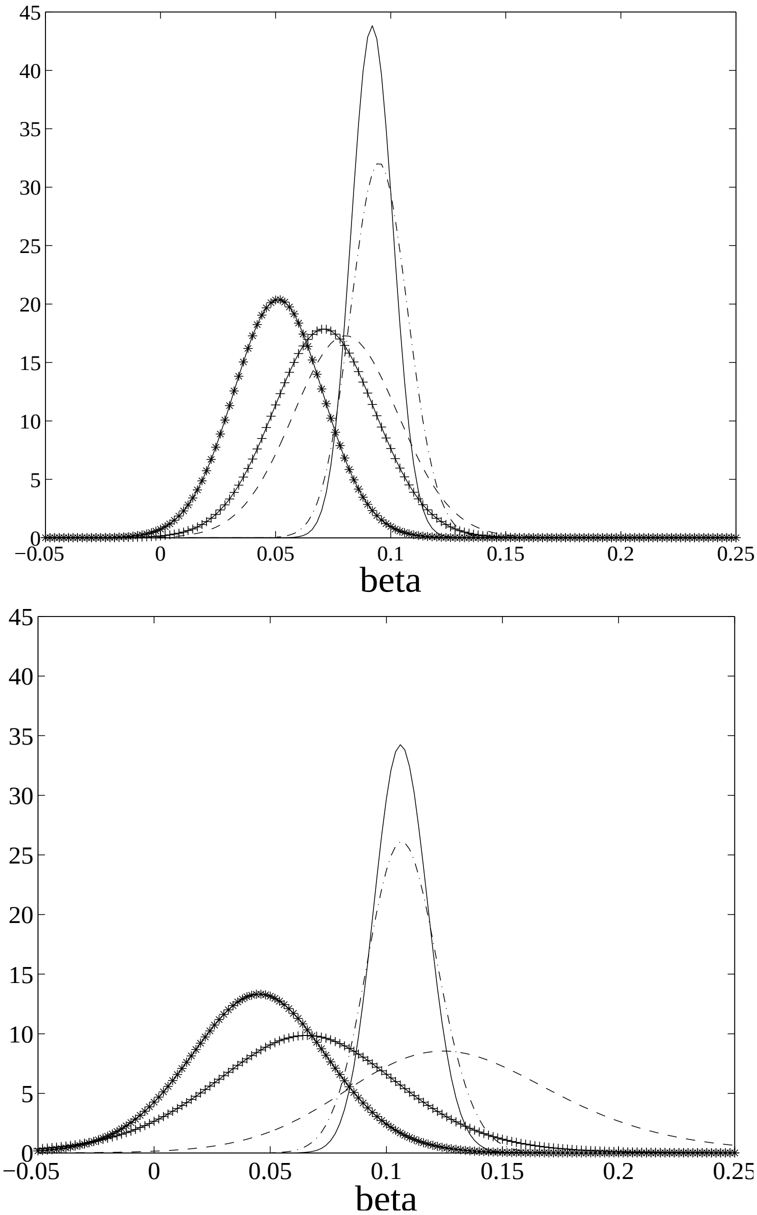

Figure 2: Marginal posterior of return on education β under flat prior (above) or Jeffreys prior (below) for US (solid), Northeast (solid-plusses), Midwest (dashed), South (dash-dot), West (solid with stars).

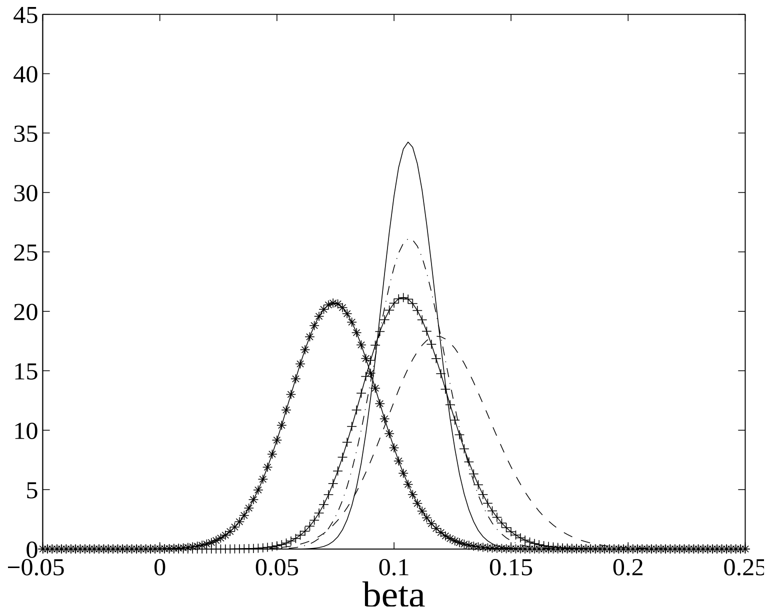

Figure 3: Marginal posterior of return on education β under Jeffreys prior for US (solid), region South (17 states, dash-dot), Kentucky-Tennessee-Arkansas (solid with plusses), rest of region South (14 states, dashed), other three Census regions (34 states, solid with stars).

5 Investigation of some of the assumptions made by Angrist and Krueger (1991)

Angrist and Krueger (1991) make the assumption that the only reason for the influence of quarter of birth on education is the asymmetry between the school-entry requirements and compulsory schooling laws: a child's birthday determines whether the school district allows the child to enter school at age 5 or age 6, whereas compulsory schooling laws generally allow students to immediately leave school when they reach a certain age (mostly 16, sometimes 17 or 18). Reasoning in this way, the quarter of birth should only yield valuable instruments for education for those individuals who have completed 9 - 13 years of education, as all persons who have left school as soon as the law allowed for it should be contained in this group.

We first inspect the empirical cumulative distribution function (CDF) of years of education for the four quarters of birth. If quarter of birth would only affect the education spell for those who leave school as soon as the law allows for it, the CDF of education should only differ for the range $9 \leq$ education ≤ 13. Figure 4 shows the difference between the empirical CDF of education between quarter 2/3/4 and quarter 1. This shows that also for education ≤ 8 and for education ≥ 14 the CDF substantially differs between the quarters of birth. Notice the negative values reflecting that men born in the first quarter have completed less years of education on average (also conditional on education ≤ 8 or

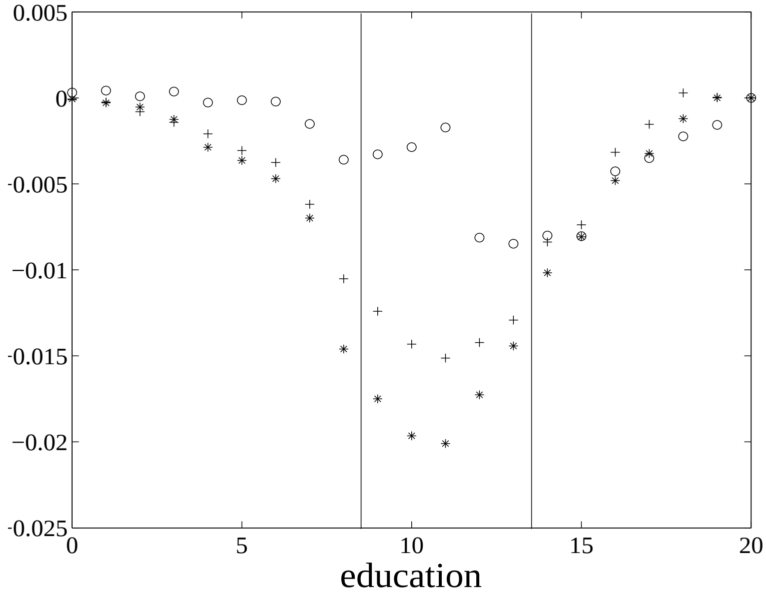

Figure 4: Empirical cumulative distribution function (CDF) of (completed years of) education for men born in different quarters: difference between CDF for men born in second quarter (circle)/ third quarter (plus)/ fourth quarter (star) and CDF for men born in first quarter.

education $\geq$ 14).

In order to investigate the importance of the observations with education $\leq$ 8 or education $\geq$ 14 for the results in the IV model we now divide the data set of the Southern region (that determines the results for the US) into a subsample of men with 9 - 13 years of education and subsamples of men with at most 8 years or at least 14 years of education. Figure 5 shows the number of observations per years of education in the region South. Table 11 shows the R^2 and F-statistic of the multiple F-test in the first stage regression for several subsamples of the region South based on education levels. Notice that the R^2 is even lower for 9 $\leq$ education $\leq$ 13 than for the groups with education $\leq$ 8 and education $\geq$ 14, suggesting quarter of birth instruments are even stronger for people with education outside the interval 9 - 13 than inside this interval. If we look at the group with either education $\leq$ 8 or education $\geq$ 14, which consists of a number of observations comparable to the group with 9 $\leq$ education $\leq$ 13, then we see that the p-value at the multiple F-statistic is also smaller for men with years of education outside the interval 9 - 13. This suggests that the influence of compulsory schooling laws on students who want to leave school as early as possible is certainly not the only factor causing the effect of quarter of birth on (average) education spell.

This is further illustrated by the posterior of β under the Jeffreys prior in Figure 6. The posterior under the Jeffreys prior is tighter for observations with education outside the interval 9-13 than inside this interval. In fact, this posterior is even tighter than for the

Table 11: First stage regressions for subsamples of the region South: R^2, F-statistic of multiple F-test and corresponding p-value

	R^2	F-statistic	obs.	p-value
region South	0.0023	3.3333	114391	0.0000
education ≤ 8	0.0045	1.1034	19214	0.2495
$9 \leq$ education ≤ 13	0.0019	1.5519	63637	0.0013
education ≥ 14	0.0029	1.1790	31540	0.1339
education ≤ 8 or education ≥ 14	0.0043	2.8060	50754	0.0000

region South or the US; intuitively speaking, this is possible since not only one's knowledge on β is updated by the extra observations, but also on Π (which occurs as the product $\Pi\beta$ in the restricted reduced form of the model) and Ω. That the posterior is much tighter for the group with less than 9 or more than 13 years of education and that the posterior for all observations of the South is much closer to this posterior than to the posterior for data on men with $9 \leq$ education ≤ 13, suggests that instruments are truly (much) stronger for men with education ≤ 8 and education ≥ 14. For completeness, Figure 7 shows the posterior under the flat prior, which shows approximately the same shapes. These results suggest that quarter of birth does not only affect years of schooling for those who leave school as soon as it is allowed.

A possible explanation is that the probability that a student leaves school during a quarter depends not only on the number of quarters of schooling that the student has already had, but also (positively) on age (measured in quarters of years): children born in the first quarter enter school at a later age (measured in quarters), so that in each cohort the students born in the first quarter are the oldest. Reasoning in this way, the influence of quarter of birth on age at school entry is enough to cause exogenous variation in years of education, even without requiring laws keeping (a certain percentage of) students at school until they reach a certain age. In other words, quarter of birth influences the age at school entry, so that it causes an exogenous variation in education level as long as students with different ages (with age measured in quarters of years) have a different 'hazard rate' of quitting school after a certain amount of education. So, the results suggesting that the influence of quarter of birth on education is certainly not restricted to men who have completed 9-13 years of education does not imply that the model is useless. It only suggests that the strength of the quarter of birth instruments is not so much caused by the asymmetry between school entry requirements and compulsory schooling laws keeping students at school until they reach a certain age; the value of the quarter of birth instruments seems to stem to a larger extent from the school entry requirements in combination with the dependence of the 'hazard rate' of leaving school on age (measured in quarters).

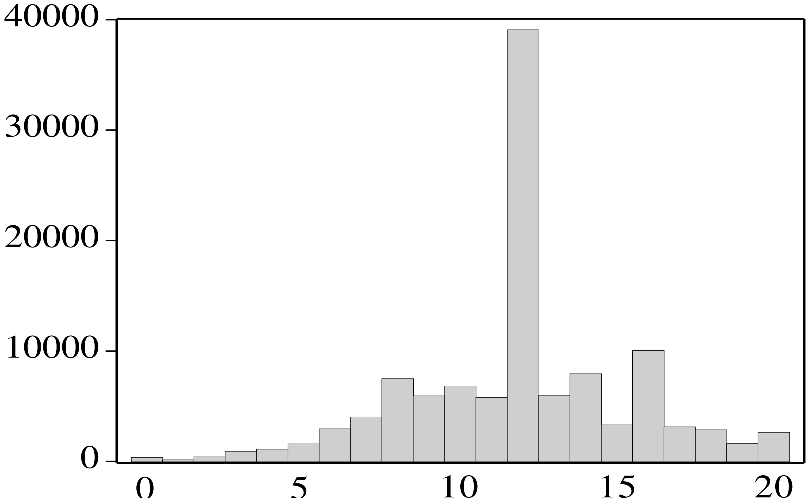

Figure 5: Histogram of the number of completed years of education in US Census region South

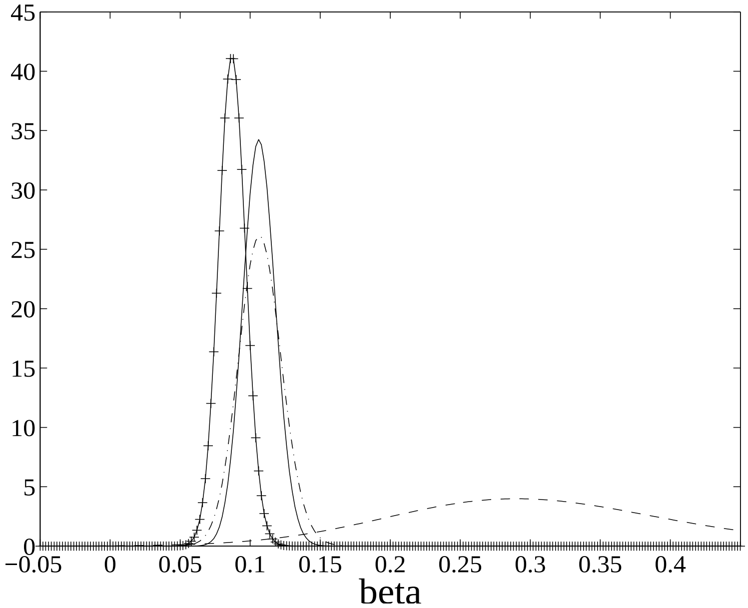

Figure 6: Marginal posterior of return on education β under Jeffreys prior for US (solid), South (17 states,dash-dot), South for education ≤ 8 or education ≥ 14 (solid-plusses), South for $9 \leq$ education ≤ 13 (dashed).

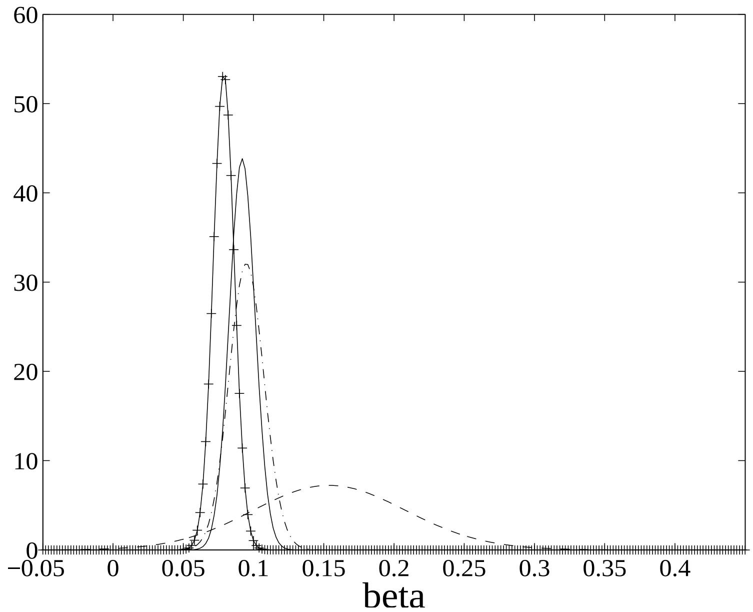

Figure 7: Marginal posterior of return on education β under flat prior for US (solid), South (17 states,dash-dot), South for education ≤ 8 or education ≥ 14 (solid-plusses), South for $9 \leq$ education ≤ 13 (dashed).

Bound, Jaeger and Baker (1995) criticize the assumptions of Angrist and Krueger (1991). They draw attention to two problems associated with the use of the 2SLS estimator in the case of weak instruments. First, the use of weak instruments may lead to large inconsistencies in the 2SLS estimator even if there is only a weak relationship between the instruments and the error in the structural equation.[7] Second, in finite samples, the 2SLS estimator is biased in the same direction as the OLS estimator, where the bias of the 2SLS estimator approaches that of the OLS estimator as the R^2 between instruments and explanatory endogenous variable approaches 0.

We first consider the problem of the finite-sample bias. Bound, Jaeger and Baker (1995) show that the 2SLS estimators of Angrist and Krueger (1991) may suffer from substantial finite-sample bias even with the large sample size, because the correlation between quarter of birth and years of education is only small. The simulations in table 5.5 confirm this; the 2SLS estimators for the US and the Census regions seem to have biases between 0.2 and 0.5 times the bias of the OLS estimator. However, the LIML estimator seems to be approximately median unbiased in these cases, suggesting that the problem of the finite-sample bias caused by the weakness of the instruments could be solved by using LIML instead of 2SLS in order to obtain a point estimate of β.

[7]Bound, Jaeger and Baker (1995) consider the case of weak instruments in the sense of instruments explaining little of the variation in the endogenous explanatory variable(s); this differs from the weak instrument defined as $\Pi = C/\sqrt{T}$ where C is fixed (so that $\Pi' Z' Z \Pi$ converges to a constant as the sample size T grows) considered by Staiger and Stock (1997).

We now consider the problem of large inconsistencies even if there is only a weak relationship between the instruments and the error in the structural equation. Bound, Jaeger and Baker (1995) argue that a weak correlation between quarter of birth and wage (independent of the effect of quarter of birth on education) exists and that this correlation is large enough to have substantial effects on the estimates of Angrist and Krueger (1991). Bound, Jaeger and Baker (1995) mention several publications containing evidence suggesting that quarter of birth directly influences wages for four reasons: there is some evidence that (1) quarter of birth influences a student's performance at school, for example performance in reading, writing and arithmetic; (2) quarter of birth affects the probability that an individual will suffer from certain mental or physical diseases/disabilities such as schizophrenia, multiple sclerosis, manic depression and dyslexia; (3) there are regional patterns in birth seasonality; (4) children born in families with high incomes are are less likely to be born in winter months. Therefore, Bound, Jaeger and Baker (1995) conclude that it is questionable whether the assumption of no direct effect of quarter of birth on income is justified.

At points (2) it should be noted that men with no income in 1979 are excluded from the data set of Angrist and Krueger (1991), so that some of the men suffering from the mentioned diseases/disabilities may be excluded from the data set. At point (3) it should be noticed that part of the regional patterns in birth seasonality are 'filtered' by the state dummies that are included in the wage equation. However, these two factors obviously do not take away the doubt on the assumption of no direct effect of quarter of birth on income.

This doubt and the finite-sample bias of the 2SLS estimator made Bound, Jaeger and Baker (1995) conclude that *'the "natural experiment" afforded by the interaction between compulsory school attendance laws and quarter of birth does not give much usable information concerning the causal effect of education on earnings'*.

Bound, Jaeger and Baker (1995) even report that differences in family income at time of birth (point (4)) would seem to account for virtually all of the association between quarter of birth and wages, which results from the following reasoning. It is argued that the difference in mean log per capita income between those born in the first quarter and the others is at least -0.0238, as this difference of -0.0238 is observed for men born in more recent years and the seasonal variation in fertility has declined since the 1930s (in which the men of our data set are born). Further, Solon (1992) and Zimmerman (1992) both found an intergenerational correlation in long-run income of at least 0.4, so that men born in the first quarter are expected to earn about 0.95% more than those born during the rest of the year. Bound, Jaeger and Baker (1995) report that for men born during the 1930s, those born in the first quarter earn 1.1% lower wages on average, hardly more than the 0.95% resulting from differences in family income at birth.

We now take a closer look at this result that differences in family income at birth between those born in the first quarter and those born during rest of the year would seem to explain all of the effect of quarter of birth on income. First, for men born in the region

South (who determine the results for the US of both the classical and Bayesian methods used in this chapter) the difference between those born in first quarter and the rest is higher: 1.65% (measured as difference in mean log income). For men born in Arkansas, Kentucky, Tennessee (who have a substantial influence on the results within the region South) this difference is even 1.99%. Of course, the difference in family income at birth and/or the intergenerational correlation may also be larger for these regions; this is a topic for research.

Second, the phenomenon that those born in rich families are less likely to be born in winter months can be modelled by including a dummy variable indicating whether a person is born in the first quarter in the wage equation of the model (and dropping one of the interactions of state and quarter dummies from the set of instruments). We consider this model for observations on men born in Arkansas, Kentucky or Tennessee.[8] In the second stage regression of 2SLS the first quarter dummy has an insignificant (and even positive) estimated coefficient of 0.0028 (with standard error 0.0111). The results of 2SLS and LIML are given by Tables 12 and 13, where the quantiles of the finite-sample distribution of the LIML estimator are estimated by substituting the ML estimates into the formula for this finite-sample density in Kleibergen (2000) and Kleibergen and Zivot (2003). Table 14 gives the results of Bayesian inference under the flat and Jeffreys prior. For the Jeffreys prior Figure 8 shows the marginal posterior of β. This shows that the uncertainty in the classical estimators and the posteriors under a flat or Jeffreys prior does not increase much by including a first quarter dummy in the wage equation. In other words, a rather tight posterior for β is obtained using quarter-of-birth information, even if we drop assumption of no influence of first quarter on income.

We can also go somewhat further in the sense of including three quarter-of-birth dummies in the wage equation, so that not only a difference between the first quarter and quarters 2-4 is allowed, but differences are permitted between all four quarters. Tables 12, 13 and 14, and Figure 8 also show the results for this model. The inclusion of two more dummies clearly increases the uncertainty in the estimators and posteriors for β as the instruments are weaker in this case. Intuitively, this can be explained as follows: in this adapted model the strength of the instruments depends on the variation between the effects of quarters of birth across states (and years) instead of the size of these effects. For example, if the effect of quarter 2-4 versus quarter 1 is substantial but (approximately) the same for all states, then the instruments Z (residuals in the regression of $\tilde{Z}$ on W) will be (almost) superfluous in this adapted model, while the instruments may be rather

[8]There are two reasons for confining ourselves to data of Arkansas, Kentucky and Tennessee, the three states for which the quarter-of-birth instruments are strongest, in this case. First, the addition of extra variables in the wage equation substantially increases problems of multicollinearity, which are smaller when considering less states and hence less state-of-birth dummies and interacted state-and-quarter-of-birth dummies. Second, the assumption that if a direct effect of quarter of birth on wages exists, that this effect is constant across states, seems to be more realistic for a region of three neighboring states than for other (sub)samples.

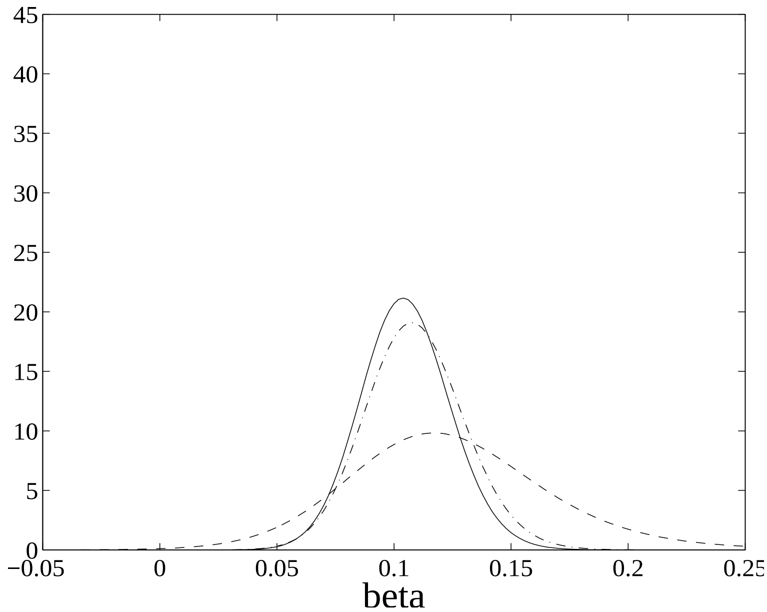

Figure 8: Marginal posterior of return on education β under Jeffreys prior for data of Kentucky, Arkansas and Tennessee: original model (solid), model with first quarter dummy in wage equation (dash-dot), model with three quarter dummies in wage equation (dashed).

strong in the original model.

Including quarter-of-birth dummies in the wage equation may result in (much) wider posterior intervals. Next to that, if a direct effect of quarter of birth on income exists, this may not be constant across states/years; in that case more terms should be added to the wage equation. So, an important question remains if such terms should be included in the wage equation and if so, how these should be specified. We leave this as a topic for further research.

Still, it should at least be noted that if there exists a direct effect of quarter of birth on income, it is not likely that the factors causing this effect differ between states/years in the same way as compulsory schooling laws and the degree to which these are enforced. So, even if there exists a direct effect of quarter of birth on income which varies across states/years, the difference between these effects and the effect of compulsory schooling laws may be exploited, so that the model may still give usable information on the causal effect of education on income.

Notice that since we can use the (approximately median unbiased) LIML estimator instead of the (biased) 2SLS estimator, and since we may still obtain a rather tight posterior for β if we allow for a direct effect of birth during the first quarter on income, it seems that the conclusion of Bound, Jaeger and Baker (1995) that the interaction between compulsory school attendance laws and quarter of birth does not give much usable information concerning the causal effect of education on earnings may have been too strong.

Finally, Bound, Jaeger and Baker (1995) note that random instruments yield results

Table 12: 2SLS estimates for β for data of Kentucky, Arkansas and Tennessee

Model	2SLS $\hat{\beta}_{2SLS}$	asympt. std.error
original model	0.0970	(0.0168)
+ dummy for first quarter of birth in wage equation	0.0986	(0.0182)
+ 3 dummies for quarters of birth in wage equation	0.0928	(0.0274)

Table 13: LIML estimates for β for data of Kentucky, Arkansas and Tennessee

Model	$\hat{\beta}_{LIML}$	Quantile finite sample dist. $\hat{\beta}_{LIML}$ median	2.5%	97.5%	25%	75%
original model	0.105	0.104	0.069	0.142	0.092	0.117
+ 1st quarter dummy	0.109	0.108	0.070	0.149	0.095	0.121
+ 3 quarter dummies	0.121	0.121	0.065	0.187	0.101	0.141

Table 14: Posterior results under flat or Jeffreys prior for data of Kentucky, Arkansas and Tennessee

	Posterior β under flat prior		Quantile posterior β under Jeffreys prior				
Model	mean	st.dev.	median	2.5%	97.5%	25%	75%
original model	0.095	0.016	0.104	0.068	0.143	0.092	0.117
+ 1st quarter dummy	0.097	0.018	0.108	0.068	0.152	0.094	0.122
+ 3 quarter dummies	0.090	0.026	0.121	0.043	0.220	0.094	0.150

similar to those for real data for four model specifications. For each specification the (mean) asymptotic standard error (over 500 simulations) of the 2SLS estimator for β is only somewhat larger for random data than for real data: 2.3, 1.3, 1.7 and 1.4 times larger. For our specification the asymptotic standard error of the 2SLS estimator for β is also only 1.5 times larger for random instruments than for real instruments (for the whole data set of the US). However, the 95% posterior interval under the Jeffreys prior is 3.0 times wider for random instruments than for real instruments, so the use of the Jeffreys prior shows a clear difference between results for random and real data. This reflects the relative insensitivity of Bayesian analysis under the Jeffreys prior to the addition of irrelevant instruments as compared to the flat prior (and the 2SLS estimator), as mentioned by Kleibergen and Zivot (2003).

6 Conclusions

We have shown results of two classical methods, the two-stage least squares (2SLS) and limited information maximum likelihood (LIML) estimators, and two Bayesian approaches, using the flat prior of Drèze (1976) and the Jeffreys prior, for an IV regression model of Angrist and Krueger (1991) for the return on education. It is shown that for these four methods the results for the US crucially depend on the results for the Census region South. A possible explanation for this is that the average education spell for men born in 1930-1939 is lower in the South, implying a larger influence of compulsory schooling laws as these do not concern education above a certain number of years, and hence a stronger effect of quarter of birth on education. A further division shows that results for the South substantially depend on three states: Kentucky, Tennessee and Arkansas. This suggests that the average level of education is not the only factor influencing the strength of the instruments, as men born in Alabama, Mississippi, Virginia and West Virginia have on average completed less years of education than those born in Tennessee: there are also other factors playing a role, which may include the power of government agencies enforcing schooling laws and the exemptions from these schooling laws, which vary across states.

If the effect of the return on education differs between the four Census regions, which may not a priori be ruled out given the large economic differences between these regions, inference using data of the US is not representative for the average returns on education across the US. Therefore one should be careful when drawing such conclusions.

We have further shown that quarter of birth is a stronger instrument for education for people with at most 8 or at least 14 years of education than for people with 9-13 years of education, which suggests that quarter of birth does not only affect the number of completed years of schooling for those who leave school as soon as it is allowed, as these are (mostly) contained in the group with 9-13 years of education. This suggests that the strength of the quarter of birth instruments is not so much caused by the asymmetry between school entry requirements and compulsory schooling laws keeping students at school until they reach a certain age; the value of the quarter of birth instruments seems to stem to a larger extent from the school entry requirements in combination with the dependence of the 'hazard rate' of leaving school on age (measured in quarters). Therefore, if one intends to increase the understanding of the working of the quarter-of-birth instruments, it is probably a better idea to pay more attention to school entry requirements and/or compulsory schooling laws for children of age 5-7 than to concentrate on the differences between compulsory schooling laws for students of age 16-18.

Finally, Bound, Jaeger and Baker (1995) have concluded that the interaction between compulsory school attendance laws and quarter of birth does not give much usable information concerning the causal effect of education on wages for two main reasons. First, the weakness of the instruments may lead to large inconsistencies in the 2SLS estimator even if there is only a weak relationship between the instruments and the error in the structural equation; Bound, Jaeger and Baker (1995) mention evidence casting doubt on the assumption that no such correlation is present. Moreover, Bound, Jaeger and Baker (1995) even report that differences in family income at time of birth would seem to account for virtually all of the association between quarter of birth and wages: they argue that the difference in income between those born in the first quarter and those born during the rest of the year can almost completely be explained by differences in family income at time of birth and an intergenerational correlation. Second, the 2SLS estimates reported by Angrist and Krueger (1991) may suffer from substantial finite sample biases because of the weakness of the instruments (despite the large sample size). However, we can use a Bayesian approach under the Jeffreys prior or the LIML estimator, which is approximately median unbiased in this case, instead of the 2SLS estimator. Furthermore, we may still obtain a rather tight posterior for β if we allow for a direct effect of birth during the first quarter on income. It should be noted that including quarter-of-birth dummies in the wage equation may result in (much) wider posterior intervals, and that if a direct effect of quarter of birth on income exists, this may not be constant across states/years; in that case more terms should be added to the wage equation. So, an important question remains whether the inclusion of such terms in the wage equation is necessary and if so, how these should be specified. This is left as a topic for further research. Still, it should at least be noted that if there exists a direct effect of quarter of birth on income, it is not likely that the factors causing this effect differ between states/years in the same way as compulsory schooling laws and the degree to which these are enforced. So, even if there exists a direct effect of quarter of birth on income which varies across states/years, the difference between these effects and the effect of compulsory schooling laws may be exploited, so that the resulting model may still give usable information on the causal effect of education on income in (regions of) the US.

So, it seems that the conclusion of Bound, Jaeger and Baker (1995), that the interaction between compulsory school attendance laws and quarter of birth does not give much usable information concerning the causal effect of education on earnings, may have been too strong, as the model of Angrist and Krueger (1991) (or a slightly modified version) may give usable information on the causal effect of education on income in (regions of) the US.

We end this chapter mentioning two topics for further research. First, an obvious question is whether the results reported in this chapter can also be found for other model specifications considered by Angrist and Krueger (1991). Second, an interesting idea is to apply the approaches used in this chapter under the assumption that the disturbances obey a different distribution than the normal, and thus investigate whether the results in this chapter are robust with respect to this distributional assumption.

References

Anderson, T.W., Rubin, H., 1949. Estimators of the parameters of a single equation in a complete set of stochastic equations. The Annals of Mathematical Statistics 21, 570–582.

Angrist, J.D., Krueger, A.B., 1991. Does compulsory school attendance affect schooling and earnings? Quarterly Journal of Economics 106, 979–1014.

Angrist, J.D., Krueger, A.B., 1992. The effect of age at school entry on educational attainment: an application of instrumental variables with moments from two samples. Journal of the American Statistical Association 87, 328–336.

Basmann, R.L., 1957. A generalized classical method of linear estimation of coefficients in a structural equation. Econometrica 25, 77–83.

Bound, J., Jaeger, D.A., Baker, R.M., 1995. Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. Journal of the American Statistical Association 90, 443–450.

Drèze, J.H., 1976. Bayesian limited information analysis of the simultaneous equations model. Econometrica 44, 1045–1075.

Hausman, J.A., 1983. Specification and estimation of simultaneous equations systems. In: Griliches, Z., Intrilligator, M.D., editors, Handbook of Econometrics, volume 1. Elsevier Science, Amsterdam.

Hood, W.C., Koopmans, T.C., 1953. Studies in Econometric Method, volume 14 of Cowles Foundation Monograph. Wiley, New York.

Hoogerheide, L.F, Kleibergen, F.R., Van Dijk, H.K., 2006. Natural conjugate priors for the instrumental variables regression model applied to the Angrist-Krueger data. Journal of Econometrics, forthcoming.

Kleibergen, F.R., 2000. Exact test statistics and distributions of maximum likelihood estimators that result from orthogonal parameters with applications to the instrumental variables regression model. Discussion paper TI 2000-039/4, Tinbergen Institute, Rotterdam.

Kleibergen, F.R., Zivot, E., 2003. Bayesian and classical approaches to instrumental variable regression. Journal of Econometrics 114, 29–72.

Maddala, G.S., 1976. Weak priors and sharp posteriors in simultaneous equation models. Econometrica 44, 345–351.

Phillips, P.C.B., 1983. Exact small sample theory in the simultaneous equations model. In Griliches, Z., Intrilligator, M.D., editors, Handbook of Econometrics, Vol.1. North-Holland Publishing Co., Amsterdam.

Siegel, P., Hodge, R., 1968. A causal approach to the study of measurement error. In: Blalock, H. and Blalock, A. (Eds.), Methodology in Social Research, McGraw Hill, New York.

Solon, G., 1992. Intergenerational income mobility in the United States. American Economic Review 82, 393–408.

Staiger, D., Stock, J.H., 1997. Instrumental variable regression with weak instruments. Econometrica 65, 557–586.

Theil, H., 1953. Estimation and Simultaneous Correlation in Complete Equation Systems. Mimeographed Memorandum of the Central Planning Bureau, The Hague.

Zimmerman, D.J., 1992. Regression toward mediocrity in economic stature. American Economic Review 82, 409–429.