ELSEVIER

Available online at www.sciencedirect.com

ScienceDirect

Journal of Macroeconomics 30 (2008) 1173–1187

Journal of MACROECONOMICS

www.elsevier.com/locate/jmacro

Per-capita income gaps across US states and Canadian provinces

Syed A. Basher, Nils-Petter Lagerlöf *

Department of Economics, York University, 4700, Keele Street, Toronto, M3J 1P3 Ontorio, Canada

Received 7 November 2006; accepted 24 July 2007
Available online 19 August 2007

Abstract

We propose that per-capita income gaps across US states and Canadian provinces can be explained by university education. Our ordinary least-squares regressions show university education having a robust positive and significant effect on per-capita incomes, when controlling for, e.g. taxes, unionization, government spending, and the sectoral composition of the economy. To control for endogeneity we instrument education today with different historical variables: population density in 1900, railway density in 1900, the sex ratio (men per woman) in 1900, and the fraction slaves in 1850. Our results support a causal link from education to incomes, and the hypothesis that these historical variables are valid instruments. We also find that the Canada dummy is mostly insignificant, and always positive.
© 2007 Elsevier Inc. All rights reserved.

JEL classification: N11; O51; R12

Keywords: Geography; Population density; Per-capita income gaps; Canada; USA

1. Introduction

We combine PPP adjusted per-capita income data across 50 US states and 10 Canadian provinces with data over the fraction of the adult population having a university degree. We first run a number of ordinary least-squares regressions showing that university education has a robust positive relationship with per-capita incomes. It stays significant when

* Corresponding author.
E-mail addresses: basher@econ.yorku.ca (S.A. Basher), lagerlof@econ.yorku.ca (N.-P. Lagerlöf).

0164-0704/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmacro.2007.07.010

controlling for "political" variables (such as the size of government and unionization), sectoral composition (like fisheries employment), and a Canada dummy.

We then run a number of two-stage least-squares regressions where university education is instrumented with different historical variables: population density in 1900, railway density in 1900, the sex ratio in 1900, and the fraction slaves in the population in 1850. The positive effect of education on per-capita incomes stays significant. Standard statistical tests also suggest that these historical variables are good instruments for university education, in the sense that they affect per-capita incomes through education, rather than directly.

We believe that the fraction of the population with a university degree proxies also for a quality dimension in higher education. Notably, many of those US states which have the highest fraction with a university degree also have the most prestigious and best ranked colleges and universities. Politicians obviously cannot create a top-ranked university over night. (If they could, they would do so more often.) Rather, we believe that the location of top educational institutions (centered around New England) is mostly due to historical coincidence. This has given rise to the type of path dependence which motivates our use of historical instruments.

Another interesting result is that the Canada dummy comes out as mostly insignificant, and always with a positive sign. Canada's per-capita income is lower than that of the US, but this seems to be due to its low educational levels. We do not find any direct negative Canada effect; whatever effect can be found is positive.

This paper seeks to contribute to an empirical literature on per-capita income gaps across countries, using historical variables – such as settler mortality and pre-colonial population density – as instruments for different measures of institutions.[1] We differ by looking at the "neo-European" region of the USA and Canada.

When analyzing per-capita income gaps across US states and Canadian provinces one may argue that we compare "locations" (as opposed to countries), between which migration is costless, and thus we should not observe any income differences (see, e.g. Roback, 1982 for such a model). However, our data display a strong persistence in population density over time, and observed per-capita income gaps today of a factor of about 2. Thus, migration (although it evidently does take place) does not seem to fully equalize incomes. In that sense, we believe our exercise lends itself to thinking also about cross-country income differences.

There is some empirical literature on income gaps and convergence across US states (Barro and Sala-i-Martin, 1992; Ciccone and Hall, 1996; Mitchener and McLean, 2003; Berkowitz and Clay, 2004). However, these typically do not use historical variables as instruments for education, and do not merge data from US states and Canadian provinces. The only income comparisons across US states and Canadian provinces that we know of is an annual report published by the Fraser Institute (see Karabegović et al., 2004).

Finally, there is also a vast literature on the link between geographical distances and skill accumulation, which we discuss in Section 3.1.2.

The rest of this paper is organized as follows. Next Section 2 presents the results when regressing per-capita incomes on university education and a number of control variables using ordinary least-squares. Section 3 then presents the instrumental-variable results. Section 4 ends with a concluding discussion.

[1] See Acemoglu et al. (2001, 2002, 2005), Easterly and Levine (2003), Galor et al. (2006) , Glaeser et al. (2004), Rodrik et al. (2004), Banerjee and Iyer (2005), Lagerlöf (2005), and Tabellini (2005).

2. Ordinary least-squares

Table 3 presents the ordinary least-squares results when regressing log per-capita income (GSP or GPP) on the fraction having a university degree and a number of other variables. University education has a high explanatory power on its own [an R^2 of 52.3% in column (1)]. It also stays very significant when controlling for a range of other control variables.[2]

In all specifications but the first we enter a Canada dummy, which is mostly insignificant. One may note that it is consistently positive: when controlling for levels of university education (and a number of other variables) Canada does not seem to be poorer than the US. If anything, Canada is richer.

Some political variables are significant. Consider first the variables from the Fraser Institute, indicated by FI in Table 3. Among these, the size of government, discriminatory taxation, and union density come out as significant (union density barely). These variables, however, are hard to interpret. They are all 1–10 indices, and it is not even always clear what a high or low score means (see Section A.2.4 in the appendix). Moreover, even though the sources do describe some of the details about how the FI variables are computed, the raw data is not provided and we have not been able to replicate these indices.

There is another problem with the FI variables. We computed an alternative set of political variables, with a clearer interpretation: the ratio of federal, state, and local government expenditure to incomes. As seen in column (10) of Table 3, a high ratio of federal expenditure to income has a significantly negative effect on per-capita income. But the direction of causality is far from obvious. Notably, this variable is the highest and lowest for two Canadian provinces: Prince Edward Island (which is the poorest of all 60 states and provinces), and Alberta (the richest Canadian province). This probably reflects the Canadian federal government's choices in response to existing income gaps, rather than exogenous causes behind the gaps. In other words, Ottawa does not *make* Albertans rich by not giving them money; they *are* rich and thus do not get any money.

The state and local spending ratios, on the other hand, are more plausible causes of per-capita income differences. However, as seen in column (11) and (12) these are insignificant, and of the "wrong" sign, respectively (that is, a bigger local government is associated with higher per-capita incomes).

Moreover, out of the three ratios the federal one is the most strongly correlated with FI's size of government variable (the correlation coefficient is −0.83). This is not strange because the FI variable is based on similar data. However, it does suggest that the size of government as measured by the Fraser Institute shows up as significant in these regressions, not because it causes income gaps, but rather because it is caused by existing gaps through federal expenditure.

Another set of variables measures the sectoral composition of the state's or province's labor force. The employment share in fisheries has a negative, but statistically insignificant, effect on per-capita incomes [column (14)]; the employment share in natural resources industries (mostly oil and gas) has a positive and significant effect [column (15)].

However, data over fishery employment is available for only 33 states and provinces, and natural resource employment for 55 states and provinces. Most of the variation is

[2] All data sources are descried in the Appendix.

Table 1
List of states/provinces and per-capita incomes

Rank	State or province	Code	Capital	Per-capita GSP or GPP (US$), 1999	Rank	State or province	Code	Capital	Per-capita GSP or GPP (US$), 1999
1	Connecticut	CT	Hartford	46,245	31	Tennessee	TN	Nashville	31,017
2	Delaware	DE	Dover	46,008	32	Indiana	IN	Indianapolis	30,659
3	Alaska	AK	Juneau	42,539	33	Kansas	KS	Topeka	30,460
4	Massachusetts	MA	Boston	42,519	34	Arizona	AZ	Phoenix	30,070
5	New York	NY	Albany	41,469	35	Iowa	IA	Des Moines	29,707
6	New Jersey	NJ	Trenton	40,713	36	South Dakota	SD	Pierre	29,505
7	Nevada	NV	Carson City	38,615	37	Louisiana	LA	Baton Rouge	29,496
8	Colorado	CO	Denver	37,900	38	Utah	UT	Salt Lake City	29,411
9	California	CA	Sacramento	37,082	39	New Mexico	NM	Santa Fe	29,328
10	New Hampshire	NH	Concord	36,823	40	Florida	FL	Tallahassee	29,309
11	Illinois	IL	Springfield	36,746	41	Vermont	VT	Montpelier	28,908
12	Wyoming	WY	Cheyenne	36,380	42	Kentucky	KY	Frankfort	28,665
13	Washington	WA	Olympia	36,352	43	South Carolina	SC	Columbia	27,515
14	Minnesota	MN	Saint Paul	36,223	44	Maine	ME	Augusta	27,185
15	Georgia	GA	Atlanta	35,402	45	Idaho	ID	Boise	27,183
16	Virginia	VA	Richmond	35,243	46	North Dakota	ND	Bismarck	26,814
17	**Alberta**	**AB**	**Edmonton**	**34,540**	47	**Quebec**	**QC**	**Quebec City**	**26,432**
18	Hawaii	HI	Honolulu	34,512	48	Alabama	AL	Montgomery	26,333
19	Texas	TX	Austin	34,288	49	**Saskatchewan**	**SK**	**Regina**	**26,094**
20	North Carolina	NC	Raleigh	33,799	50	**British Columbia**	**BC**	**Victoria**	**26,086**
21	Maryland	MD	Annapolis	33,782	51	Oklahoma	OK	Oklahoma City	25,724
22	Oregon	OR	Salem	33,079	52	Arkansas	AR	Little Rock	25,388
23	Rhode Island	RI	Providence	32,848	53	**Manitoba**	**MB**	**Winnipeg**	**25,328**
24	**Ontario**	**ON**	**Toronto**	**32,373**	54	Montana	MT	Helena	23,376
25	Nebraska	NE	Lincoln	32,259	55	Mississippi	MS	Jackson	23,220
26	Ohio	OH	Columbus	32,157	56	West Virginia	WV	Charleston	22,516
27	Pennsylvania	PA	Harrisburg	31,931	57	**Nova Scotia**	**NS**	**Halifax**	**22,336**
28	Wisconsin	WI	Madison	31,708	58	**New Brunswick**	**NB**	**Fredericton**	**22,187**
29	Michigan	MI	Lansing	31,257	59	**Newfoundland**	**NL**	**St. John's**	**21,008**
30	Missouri	MO	Jefferson City	31,174	60	**Prince Edward Island**	**PE**	**Charlottetown**	**20,545**

Notes: Canadian provinces in bold. GSP stands for Gross State Product, GPP for Gross Province Product. Incomes are PPP adjusted.

among a few states and provinces, like Alaska, Alberta, and Atlantic Canada. Endogeneity is also an issue. Canada's Atlantic provinces may have come to rely more on fishery today because when fishery began its decline the labor force did not move into other sectors. Whatever prevented the growth of non-fishery industries should be the ultimate cause of current income gaps; we believe that cause is population density and education.

To sum up, the results shown in Table 3 suggest that university education has a robust positive and significant relationship with per-capita income levels. However, it is not clear whether education causes these income gaps, or if the causality goes the other way around. To address that issue we next turn to instrumental-variable analysis.

3. Instrumental-variable analysis

3.1. Choice of instruments

The instruments we use for education today are: log population density in 1900; the sex ratio (the number of men per woman) in 1900; log railway density (railway miles built by 1900 per square mile of land area); and the fraction slaves in the population in 1850.

Except for slavery, these are all more highly correlated with education than per-capita incomes today (Table 2). This may suggest that the main direct impact has been on education.

Below we discuss the validity of each of the instruments.

3.1.1. Slavery in 1850

One instrument that we use is the fraction slaves in the population in 1850. Across the Americas slavery was used in the Caribbean, Brazil, and the US South. These were relatively rich regions at the time and thus we do not need to worry about reverse causality (e.g., poverty today, and/or in the 19th century, being the cause of slavery).[3] Rather, slavery was used in these regions because geographical fundamentals enabled colonists to grow profitable staple crops, like coffee, sugar and cotton; it is well documented that slave labor was more suitable for such crops.[4]

The reason former slave regions are poor today seems to (at least partly) have to do with the negative effects of slavery on education and the development of educational institutions. School reforms came later in formerly more slave-dependent Caribbean countries and Brazil compared to Canada and the US north of the Chesapeake Bay (Mariscal and Sokoloff, 2000). There are also indications of similar patterns across counties within the US South (Lagerlöf, 2005). Notably, before abolition US slaves were often forbidden (or otherwise prevented) from learning to read or write. Slavery was also associated with high levels of land inequality, which may also have exerted negative effects on public education expenditure (Galor et al., 2006).

However, only a few US states (and no Canadian provinces) practiced slavery. Therefore, slavery can only explain some of the variation in education in the region as a whole. For example, Atlantic Canada is poor, and has relatively low educational levels, without slavery ever being practiced there.

[3] See e.g. Easterlin (1960).

[4] See Sokoloff and Engerman (2000), Mariscal and Sokoloff (2000), Engerman and Sokoloff (2002), Acemoglu et al. (2002), Lagerlöf (2005), and Doepke and Eisfeldt (2007).

Table 2
Correlation matrix

	Log per-capita GSP or GPP	Fraction with university degree	Log population density in 1900	Sex ratio in 1900	Fraction slaves in 1850	Log railway in 1900
Log per-capita GSP or GPP	1					
Fraction with university degree	0.723 (60)	1				
Log population density in 1900	0.176 (59)	0.361 (59)	1			
Sex ratio in 1900	−0.058 (52)	−0.124 (52)	−0.780 (52)	1		
Fraction slaves in 1850	−0.114 (60)	−0.042 (60)	0.214 (59)	−0.292 (52)	1	
Log railway density in 1900	0.311 (57)	0.503 (57)	0.960 (57)	0.701 (52)	0.153 (57)	1

Notes: Correlation coefficients. Number of observations in parentheses.

3.1.2. Population density in 1900

Another instrument for education that we use is (the logarithm of) population density in 1900. (This is the earliest year for which we have population density data for most states and provinces; only data for Newfoundland is missing.)

Our hypothesis is that population density may have spurred the development of educational institutions and norms because shorter geographical distances between people enhances the exchange of ideas and accumulation of skills. This idea goes back at least to Jacobs (1969), and probably much longer (Glaeser (1999) quotes Alfred Marshall on agglomeration effects). Jaffe et al. (1993) show that patent citations are negatively related to distance. Glaeser and Maré (2001) find that wages are higher in cities because cities promote learning rather than the reverse causality by which skilled people choose to live in cities. Theoretical foundations can be found in, e.g. Glaeser (1999).

It is also possible that colleges and universities have come to be located in regions with dense populations due to scale effects in education. Many Ivy League universities lie in the densely populated region around Northeastern USA. Also, vicinity to educational institutions seems to matter for educational choice: Card (1995) finds that men who grew up near a four-year college have higher education and earnings, also when controlling for regional factors and family background. Glaeser and Saiz (2003) find that cities of a given size grow faster if they have more colleges per capita. A skilled labor force can also attract high-technology industries (Henderson et al., 1995).

Coastal regions tend to be more densely populated than inland regions. If waterways exert a direct effect on incomes, aside from the effect working through population density and education, population density would be a poor instrument. For example, Rappaport and Sachs (2003) argue that coastal regions are richer because they have more trading ports.

However, this fails to account for why the Eastern half of North America is richer and more densely populated than the Western half.[5] This probable has to do with the Atlantic being closer to Europe, which is where most early settlers arrived from. Many immigrants stayed in the big cities where their ships landed. Trans-Atlantic trade may also have had an impact on early growth of the Atlantic region of North America, as it did on the European

[5] See e.g. Rappaport and Sachs (2003, Map 2).

side of the Atlantic (Acemoglu et al., 2005), which is less likely to have a direct effect on education and incomes today.

Moreover, population density did not depend only on vicinity to the coast but also on e.g. temperature, which seems to have had a non-monotonic effect on early settlements, especially along the Atlantic coast: the US South and Atlantic Canada were sparsely populated compared to New England. Early settlers in North America, who were mostly farmers, probably avoided colder regions due to their lower agricultural productivity. Hot regions may have been unsuitable for European settlers in particular, since they were not resistent to warm-weather diseases (Acemoglu et al., 2001; Coelho and McGuire, 1997, 1999).[6] It is hard to argue that temperature would have a direct impact on per-capita incomes today, at least not a non-monotonic one.

3.1.3. Railway density in 1900

Obviously, population per unit of land area over a whole state or province may not be the best measure of the mechanism we try to capture. Ideally one would want a measure of how well "connected" people are to the type of social networks which build skills and/or enhance growth of high-skilled industries, or data over how far the average resident of a state or province is from the closest university or college.

Lacking any such data we do look at (the logarithm of) railway density in 1900, i.e. miles of railway per square miles of land area. As seen from Table 2, log population density and log railway density are highly correlated. States and provinces with high population density were more "connected" by railways.

3.1.4. The sex ratio in 1900

We would also like to use early urbanization as an instrument for education. (Indeed, population and railway density are probably indicators of urbanization.) We do not have any historical urbanization data (at least not for both Canada and the US) but a good proxy could be the sex ratio, i.e. the number of men per woman. Edlund (2005) documents that rural areas in the Western world are relatively short on women, compared to urban areas. This seems particularly true in new settlements in colonial times. Guttentag and Secord (1983, Chapter 5) document that in frontier societies of the US men vastly outnumbered women into the 20th century, while the situation was rather the opposite in New England (see also Angrist, 2002). This is confirmed in our data too, where the sex ratio in 1900 is negatively correlated with the income and education variables (Table 2).

3.2. Instrumental-variable regression results

Table 4 shows the results from some two-stage least-squares regressions where we use different sets of the historical variables discussed above as instruments for university education today. The results indicate that per-capita income is affected positively and significantly by the fraction with a university degree. We report a couple of different specifications because data over the sex ratio is missing for eight states and provinces, and over railway density for 3. The educational variable seems significant in all specifications: when using only the slavery

[6] Moreover, as discussed above, much of the migration to the warmer parts of the US were not by free Europeans but by African slaves, without whom population density in the South would have been even lower.

Table 3
Higher education and per-capita income: ordinary least-squares regressions

Dependent variable is log per-capita GSP or GPP

Panel A: specifications (1)–(8)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant	9.811***	9.719***	9.231***	9.206***	9.684***	9.926***	9.566***	9.596***
	(0.068)	(0.108)	(0.107)	(0.155)	(0.119)	(0.150)	(0.238)	(0.139)
Fraction with university degree	**2.297*****	**2.642*****	**1.661*****	**2.311*****	**2.627*****	**2.456*****	**2.716*****	**2.709*****
	(0.287)	**(0.427)**	**(0.348)**	**(0.384)**	**(0.429)**	**(0.428)**	**(0.441)**	**(0.426)**
Canada dummy		0.074	0.128**	0.207***	0.055	0.012	0.164	0.112
		(0.068)	(0.051)	(0.068)	(0.073)	(0.074)	(0.141)	(0.073)
Size of government (FI)			0.100***					
			(0.014)					
Discriminatory taxation (FI)				0.104***				
				(0.025)				
Average unemployment duration (FI)					0.003			
					(0.004)			
Union density (FI)						−0.022*		
						(0.011)		
Felixibility in labor laws (FI)							0.015	
							(0.022)	
Minimum-wage legislation (FI)								0.016
								(0.011)
No. of observations	60	60	60	60	60	60	60	60
R^2	0.523	0.533	0.746	0.643	0.537	0.562	0.537	0.480

Panel B: specifications (9)–(16)

	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Constant	9.184***	10.184***	9.747***	9.540***	10.045***	9.615***	9.595***	9.396***
	(0.194)	(0.129)	(0.112)	(0.141)	(0.165)	(0.161)	(0.107)	(0.120)
Fraction with university degree	**1.347*****	**1.898*****	**2.573*****	**2.698*****	**1.997*****	**3.193*****	**3.015*****	**3.894*****
	(0.327)	**(0.387)**	**(0.435)**	**(0.418)**	**(0.390)**	**(0.606)**	**(0.414)**	**(0.451)**
Canada dummy	0.032	0.032	0.062	0.147*	0.094	0.145	0.124*	0.210***
	(0.101)	(0.057)	(0.069)	(0.076)	(0.069)	(0.092)	(0.063)	(0.066)
Size of government (FI)	0.084***							
	(0.015)							
Discriminatory taxation (FI)	0.069***							
	(0.024)							
Average unemployment duration (FI)	0.006*							
	(0.003)							
Union density (FI)	−0.019**							
	(0.009)							
Felixibility in labor laws (FI)	−0.018							
	(0.021)							
Minimum-wage legislation (FI)	0.008							
	(0.009)							
Ratio of federal expenditure to income		−1.174***			−1.212***			
		(0.234)			(0.252)			
Ratio of state expenditure to income			−0.021		0.026			
			(0.024)		(0.022)			

Table 3 (*continued*)

	Dependent variable is log per-capita GSP or GPP							
Ratio of local expenditure to income				1.146* (0.737)	0.944 (0.663)			
Fraction working in fisheries						−2.479 (1.701)		−2.530 (1.549)
Fraction working in natural resource industries							3.313** (1.458)	5.188*** (1.641)
No. of observations	60	60	60	60	60	33	55	30
R^2	0.825	0.677	0.539	0.560	0.693	0.717	0.595	0.850

Notes: FI indicates data from the Fraser Institute, where a high index is associated with small government and less regulated labor markets. See text for further details. Standard errors are shown in parentheses. Significance at the 90, 95, and 99 percent confidence levels are indicated by *, **, and ***, respectively.

variable and log population density as instruments [column (1)]; when adding railway density [column (2)]; and when also adding the sex ratio [column (3)].

Notably, in Table 4, the Canada dummy comes out as insignificant, and with a positive sign. This confirms the result from the OLS regressions in Table 3. Canada's per-capita

Table 4
Two-stage least-squares regressions

	Dependent variable is log per-capita GSP or GPP						
	Instrumented variable is fraction with university degree						
	Second-stage results						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Constant	9.686*** (0.238)	9.627*** (0.239)	9.535*** (0.230)	9.52*** (0.192)	9.433*** (0.241)	9.533*** (0.676)	9.461*** (0.239)
Fraction with university degree	**2.776*** (0.958)**	**2.992*** (0.962)**	**3.384*** (0.912)**	**3.469*** (0.860)**	**3.532*** (0.794)**	**3.391 (2.573)**	**3.596*** (0.831)**
Canada dummy	0.100 (0.117)	0.129 (0.116)	0.119 (0.113)	0.125 (0.093)	0.133 (0.097)	0.120 (0.328)	0.131 (0.094)
Log population density in 1900				−0.001 (0.019)			
Sex ratio in 1900					0.058 (0.151)		
Fraction slaves in 1850						0.0009 (0.268)	
Log railway density in 1900							−0.008 (0.030)
Hansen *J* statistic	0.393	0.507	1.735	1.673	1.077	1.383	1.211
Degrees of freedom	1	2	3	2	2	2	2
χ^2-test (*p*-value)	**(0.530)**	**(0.776)**	**0.629**	**(0.433)**	**(0.583)**	**(0.500)**	**(0.545)**
F-statistic for joint significance of excl. instr.	6.41	4.78	3.63	4.05	4.69	1.32	4.24
F-test (*p*-value)	(0.003)	(0.005)	(0.011)	(0.012)	(0.006)	(0.278)	(0.010)
No. of obserbations	59	57	52	52	52	52	52

Notes: The first-stage estimates are not shown. Heteroskedasticity-robust standard errors are shown in parantheses, which refer to a *z*-test in the second-stage regression. Fraction with university degree is instrumented. Column (1) uses fraction slaves in 1850 and log population density in 1900 as instruments. In column (2), log railway density in 1900 is added as a third instrument, while column (3) includes sex ratio in 1900 as a fourth instrument. *** indicates significance at the 99 percent confidence level.

income is lower than that of the US because of its low educational levels. There is no (negative) direct Canada effect.

3.2.1. Testing the validity of the instruments

To be valid instruments these historical variables should first of all be highly correlated with the instrumented variable. As reported in Table 4, an *F*-test verifies that the instruments are jointly significant in the first-stage regression.

For the instruments to be valid they should also not influence the dependent variable (log per-capita income) other than through the instrumented variable (the fraction with a university degree). In other words, the instruments should be uncorrelated with the error terms in the second-stage regression. This seems to be the case: Hansen's *J*-test does not reject the null hypothesis that the second-stage error terms are uncorrelated with the instruments.

To verify that the instruments do not exert any direct impact on per-capita incomes we also report the results when letting each instrument enter the second-stage regression [columns (4)–(7)]. As seen, they all come out as insignificant.

All in all, our results suggest that these historical variables are valid instruments and that university education exerts a strong positive effect on per-capita incomes.

4. Conclusions

We have presented regression results from data on US states and Canadian provinces suggesting a causal link from university education to per-capita incomes. The fraction of the population having a university degree shows a robust positive and significant relationship with per-capita incomes, when controlling for all sorts of variables.

The same link from education to per-capita incomes shows up when instrumenting education with historical variables (log population density in 1900, log railway density in 1900, the sex ratio in 1900, and the fraction slaves in 1850). We have argued that these are probably valid instruments, in the sense that they affect economic outcomes today only through education. The data support this hypothesis.

Merging data from Canadian provinces and US states also allows us to examine if there is a direct Canada effect, when controlling for educational levels. There is not: the Canada dummy comes out as mostly insignificant, and always with a positive sign.

We measure education as the fraction of the (adult) population with a university degree. This does not mean that an exogenous rise in that fraction is necessarily good for growth. Our measure is likely to also proxy for a quality dimension which is much harder to capture; recall that the best educated US states according to our measure also have the highest ranked colleges and universities. A top-ranked university cannot be created over night through policy decisions, but is more likely to be the result of many historical coincidences. Indeed, we have argued that such path dependence makes our instruments good.

Some may suggest, or emphasize, explanations which have nothing to do with education. Within the US–Canada region the poorest locations lie in Atlantic Canada (Table 1). It may thus be tempting to attribute income gaps across this region to variations in the dependence on fishery. But this is not an exhaustive explanation, we argue. There are other poor regions which do not rely on fishery (West Virginia is landlocked and almost as poor as Nova Scotia); Alaska is rich and has a relatively large fraction of its employment in fishery. Fishery data is only available for 33 states and provinces, only a

few of these have any significant fishery industries, and our regression results are not too supportive of a fishery explanation (see Section 2).

Moreover, it is not obvious why the decline of one sector of the economy would necessarily mean the decline of a whole region. If fishery has declined, why have people previously employed in the fisheries not moved to other sectors? Consider Massachusetts, a maritime region which once had a fishery and whaling industry, and is rich today and not dependent on fishery. What makes Nova Scotia different from Massachusetts? Why could not Halifax be like Boston? A good explanation should be deeper than simply pointing to the decline of some sector of the economy. It should point to fundamental causes (like geography), rather than proximate.

Acknowledgements

We are grateful for comments from Tasso Adamopoulos, Ahmet Akyol, David Andolfatto, Berta Esteve-Volart, Paul Klein, Shannon Seitz, Andrei Semenov, and two anonymous referees. Lagerlöf thanks SSHRC for financial support. We also thank the following persons for help in locating data sources: Katie Davis, Élise Laplaine and Greg LeBlanc of Statistics Canada; Karen D. Thompson of the US Census Bureau; Brian Dahlin and Serah Hyde of the Bureau of Labor Statistics; David Sutherland of NOAA; and Wendell Cox of Demographia.

Appendix A. Data appendix

Below we list our data sources, some of which are available online only. When it is not self-explanatory we try to describe in as much detail as possible what steps to take to access the online data. All data are also available as a STATA file at: http://www.arts.yorku.ca/econ/lagerloef/HP/DataSummer2007.dta.

A.1. Historical variables

A.1.1. The fraction slaves in 1850

The fraction slaves in the population is calculated as the total number of slaves in 1850 over total population in 1850. This data is made available by the Geospatial and Statistical Data Center at the University of Virginia Library. Their Web site is at: http://fisher.lib.virginia.edu/collections/stats/histcensus/.

Canada did not use slavery.

A.1.2. The sex ratio in 1900

The numbers of males and females in Canada refer to the year 1901 and are from Census of Canada (1902, Table III).

The corresponding data for the US were extracted from the Geospatial and Statistical Data Center, University of Virginia Library (see the previous section; click on 1900 and follow the links).

Note that the Canadian data is for 1901 and the US data is for 1900, but we refer to this variable as the sex ratio in 1900.

A.1.3. Population density in 1900

The Canadian population density data are from Series A54–66: Population density per square mile, Canada and provinces, 1871–1976, Statistics Canada (1983).

US population density is from Table 5, Statistical abstract of the United States 1901, which is available at: http://www.census.gov/statab/www/.

Note that the Canadian data is for 1901 and the US data is for 1900, but we refer to this variable as population density in 1900.

A.1.4. Railway density in 1900

The Canadian railway data are obtained from Bladen (1932, 1934). Canadian land area by province (in square miles) are available at: http://en.wikipedia.org/wiki/List_of_Canadian_provinces_and_territories_by_area#Land_area.

The US railway data are from Stover (1961). US land area by state (in square miles) is available at: http://en.wikipedia.org/wiki/List_of_U.S._states_by_area#Land_area.

A.2. Contemporary variables

A.2.1. GDP per capita

PPP adjusted GDP per capita for both Canadian provinces and US states are from the Web site Demographia run by the Wendell Cox Consultancy, at: www.demographia.com. All figures are for 1999 and in current US dollars. The exact links are: http://www.demographia.com/db-cangdpr99.htm (for Canada); http://www.demographia.com/db-usgdpr99.htm (for the US).

Note that the state and province level equivalents of GDP are called GSP (Gross State Product) and GPP (Gross Province Product), respectively. In the text we also call this variable per-capita income for short.

A.2.2. Fishery and natural resource data

Natural resource employment. By employment in natural resource industries we mean, for Canada, employment in mining and oil and gas extraction. These numbers are from Statistics Canada (2002, Table A.32). The numbers for the US are people employed in natural resource and mining industries, from the Bureau of Labor Statistics, available online at: http://data.bls.gov/PDQ/outside.jsp?survey=sm.

For both Canada and the US the numbers are from 2000.

Fishery employment. The number of people employed in fisheries in Canada are downloaded electronically from the Statistics Canada Web site, at: www.statcan.ca. To access the data, select census; select data; select topic-based tabulations; click on number 11 – "Canada's workforce: paid work." Table 8 provides the number of people employed by industry and province.

The US data are from Pritchard (2003, p. 95), under the category "employment, craft, and plants." The US data refer to total employment in both the fish-processing and wholesale industry.

For both Canada and the US these numbers are from 2001.

Total employment. To get fishery and natural resource employment as fractions of total employment we use the following data. For Canadian provinces, 2000 and 2001 total employment is from Statistics Canada (2003, Table 18), available online at: www.statcan.ca.

For the US total employment in 2000 is from Table 572, Statistical Abstract of the United States 2001, US Census Bureau, available online at http://www.census.gov/statab/www/. Total employment for 2001 is from Table 565, Statistical Abstract of the United States 2002, also available online at http://www.census.gov/statab/www/.

A.2.3. Fraction with a university degree

For Canada, this fraction is given by the number of persons 15 years and over with a university (Bachelor) degree, divided by total population 15 and over. The number of persons with a degree are from Statistics Canada, 2001 Census, downloaded electronically from the Statistics Canada Web site: www.statcan.ca, following these steps: select census; select search by topic; select education in Canada: school attendance and levels of schooling; click on number 1 (under topic-based tabulations) – detailed highest level of schooling. Total population numbers are from the same Web site: select census; select search by topic; select age and sex; click on number 2, "profile of age and sex".

The US data is the fraction of the population with a Bachelor degree or more for persons 25 years old and over, based on the 2000 census, from Table 231, Statistical Abstract of the United States 2003, US Census Bureau, available at: http://www.census.gov/statab/www/.

The Canadian data is from 2001 and the US data from 2000.

A.2.4. Political variables

Fraser Institute indicators. We use six political indicators constructed by the Fraser Institute (FI), a Canadian think tank. These are meant to measure "economic freedom" and/or the flexibility of labor markets and take values on a scale from 1 to 10. Indicators A–D below are from Karabegović et al. (2004); indicators E and F from Clemens et al. (2004).

Indicator A: "Size of the government" is an index measuring general government consumption expenditures relative to GSP or GPP. A higher score means smaller government.
Indicator B: "Discriminatory taxation" is an index measuring how "discriminatory" the tax system is; taxation is considered discriminatory if, for instance, the link between taxes paid and services received is weak, or marginal taxes are high. A higher score means less discriminatory taxation.
Indicator C: "Minimum-wage legislation" is an index measuring the annual income earned by someone working at the minimum-wage relative to per-capita GSP or GPP. A higher score means a higher minimum wage.
Indicator D: "Union density" is an index measuring the fraction of the work force who is unionized. A higher score means that a larger fraction is unionized.
Indicator E: "Average duration of unemployment" is an index measuring just that. A higher score means longer unemployment duration.
Indicator F: "Flexibility in labor-relation laws" is an index measuring the flexibility in different areas of labor law. A higher score means that labor laws are more flexible.

Government expenditure. Aside from the variables from the Fraser Institute, we calculated three other political variables: the ratio of federal, state/provincial, and local government expenditure to income.

Data over government expenditures across Canadian provinces are from Statistics Canada (2003, Tables 7–9). We first divided expenditure by population to get it in per-capita terms; we then divided by per-capita personal income to get the total expenditure as a fraction of income. Both population and per-capita personal income by province were collected from Statistics Canada (2003, Table 18).

The same data for US states are from Sagoo (2005, Tables C18, E15, and F14). To get total expenditure as a fraction of income we divided per-capita expenditures by per-capita personal income (collected from Table A14 in Sagoo (2005)).

These sources are partly the same as those used to calculate FI's index over the size of government (Indicator A above).

For both Canada and the US, the numbers are from 2002.

References

Acemoglu, D., Johnson, S., Robinson, J.A., 2001. The colonial origins of comparative development: An empirical investigation. American Economic Review 91, 1369–1402.

Acemoglu, D., Johnson, S., Robinson, J.A., 2002. Reversal of fortune: Geography and institutions in the making of the modern world income distribution. Quarterly Journal of Economics 117, 1231–1294.

Acemoglu, D., Johnson, S., Robinson, J.A., 2005. The rise of Europe: Atlantic trade, institutional change and growth. American Economic Review 95, 546–579.

Angrist, J., 2002. How do sex ratios affect marriage and labor markets? Evidence from America's second generation. Quarterly Journal of Economics 117, 997–1038.

Banerjee, A., Iyer, L., 2005. History, institutions, and economic performance: The legacy of colonial land tenure systems in India. American Economic Review 95, 1190–1213.

Barro, R.J., Sala-i-Martin, X., 1992. Convergence. Journal of Political Economy 100, 223–251.

Berkowitz, D., Clay, K., 2004. Initial conditions, institutional dynamics and economic performance: Evidence from American states. mimeo. University of Pittsburgh.

Bladen, M.L., 1932. Construction of railways in Canada to the year 1885. Contributions to Canadian Economics 5, 43–60.

Bladen, M.L., 1934. Construction of railways in Canada part II: From 1885 to 1931. Contributions to Canadian Economics 7, 61–107.

Card, D., 1995. Using geographic variation in college proximity to estimate the return to schooling. In: Christofodes, L.N., Grant, E.K., Swidinsky, R. (Eds.), Aspects of Labour Market Behaviour – Essays in Honour of John Vanderkamp. University of Toronto Press, Toronto.

Census of Canada, 1902. Fourth census. In: Population, vol. 1. S.E. Dawson, Ottawa.

Ciccone, A., Hall, R.E., 1996. Productivity and the density of economic activity. American Economic Review 86, 54–70.

Clemens, J., Godin, K., Karabegović, A., Veldhuis, N., 2004. Measuring Labour Markets in Canada and the United States. Fraser Institute.

Coelho, P.R.P., McGuire, R.A., 1997. African and European bound labor in the new world: The biological consequences of economic choices. Journal of Economic History 57, 83–115.

Coelho, P.R.P., McGuire, R.A., 1999. Biology, diseases, and economics: An epidemiological history of slavery in the American south. Journal of Bioeconomics 1, 151–190.

Doepke, M., Eisfeldt, A., 2007. Colonies. Manuscript. UCLA and Northwestern University.

Easterlin, R.A., 1960. Interregional differences in per-capita income, population, and total income, 1840–1950. Trends in the American Economy in the Nineteenth Century. Princeton University Press (for the NBER).

Easterly, W., Levine, R., 2003. Tropics, germs, and crops: The role of endowments in economic development. Journal of Monetary Economics 50, 3–39.

Edlund, L., 2005. Sex and the city. Scandinavian Journal of Economics 107, 25–44.

Engerman, S.L., Sokoloff, K.L., 2002. Factor endowments, inequality, and paths of development among new world economies. Economia 3, 41–109.

Galor, O., Moav, O., Vollrath, D., 2006. Inequality in land ownership, the emergence of human capital promoting institutions, and the great divergence. mimeo. Brown University and Hebrew University.

Glaeser, E.L., 1999. Learning in cities. Journal of Urban Economics 46, 254–277.
Glaeser, E.L., Maré, D.C., 2001. Cities and skills. Journal of Labor Economics 19, 316–342.
Glaeser, E.L., Saiz, A., 2003. The rise and fall of the skilled city. mimeo. Harvard University.
Glaeser, E.L., La Porta, R., Lopez-de-Silanes, F., Shleifer, A., 2004. Do institutions cause growth? Journal of Economic Growth 9, 271–303.
Guttentag, M., Secord, P.F., 1983. Too Many Women? The Sex Ratio Question. Sage Publications, Beverly Hills, California.
Henderson, V., Kuncoro, A., Turner, M., 1995. Industrial development in cities. Journal of Political Economy 103, 1067–1090.
Jacobs, J., 1969. The Economy of Cities. Random House, New York.
Jaffe, A.B., Trajtenberg, M., Henderson, R., 1993. Geographic localization of knowledge spillovers as evidenced by patent citations. Quarterly Journal of Economics 108, 577–598.
Karabegović, A., McMahon, F., Samida, D., Mitchell, G., 2004. Economic Freedom of North America, Annual Report. Fraser Institute.
Lagerlöf, N.P., 2005. Geography, institutions, and growth: The United States as a microcosm. mimeo. York University.
Mariscal, E., Sokoloff, K.L., 2000. Schooling, suffrage, and the persistence of inequality in the Americas, 1800–1945. In: Haber, S. (Ed.), Political Institutions and Economic Growth in Latin America – Essays in Policy, History, and Political Economy. Hoover Institution Press, Stanford, California.
Mitchener, K.J., McLean, I.W., 2003. The productivity of US states since 1880. Journal of Economic Growth 8, 73–114.
Pritchard, E.S. (Ed.), 2003. Fisheries of the United States 2002. National Marine Fisheries Services, NOAA, Silver Spring, Maryland http://www.st.nmfs.gov/st1/fus/current/2002-fus.pdf.
Rappaport, J., Sachs, J.D., 2003. The United States as a coastal nation. Journal of Economic Growth 8, 5–46.
Roback, J., 1982. Wages, rents, and the quality of life. Journal of Political Economy 90, 1257–1278.
Rodrik, D., Subramanian, A., Trebbi, F., 2004. Institutions rule: The primacy of institutions over geography and integration in economic development. Journal of Economic Growth 9, 131–165.
Sagoo, S. (Ed.), 2005. Facts and Figures on Government Finance, thirty-eighth ed. The Tax Foundation, Washington, DC http://www.taxfoundation.org/news/show/147.htm.
Sokoloff, K.L., Engerman, S.L., 2000. History lessons: Institutions, factor endowments, and paths of development in the new world. Journal of Economic Perspectives 14, 217–232.
Statistics Canada, 1983. Electronic edition of historical statistics of Canada, second ed. Catalogue No. 11-516-XIE. Available from: <http://www.statcan.ca/english/freepub/11-516-XIE/sectiona/toc.htm>.
Statistics Canada, 2002. Human activity and the environment. Annual Statistics, Catalogue No. 16-201-XIE. Available from: <http://estat.statcan.ca/content/english/articles/other/other-envi1.pdf>.
Statistics Canada, 2003. Provincial Economic Accounts. Catalogue No. 13-213-PPB. Available from: <http://www.statcan.ca:8096/bsolc/english/bsolc?catno=13-213-P>.
Stover, J.F., 1961. American Railroads. University of Chicago Press.
Tabellini, G., 2005. Culture and institutions: Economic development in the regions of Europe. mimeo. York University.