Ricardo Salvatore and Jörg Baten[1]

A Most Difficult Case of Estimation: Argentinian Heights, 1770-1840

In John Komlos and Joerg Baten, eds., <u>The Biological Standard of Living In Comparative Perspective</u>, vol.1 (Stuttgart: Steiner Verlag, 1998).

This is the last working paper version before this study was submitted and accepted.

Please cite as Salvatore, Ricardo and Baten, Joerg. "A Most Difficult Case of Estimation: Argentinian Heights, 1770-1840", *The Biological Standard of Living In Comparative Perspective, vol.1*, John Komlos and Joerg Baten, eds. Stuttgart: Steiner Verlag, 1998.

Estimation of the mean heights of Argentinian recruits for the period 1770-1840 poses considerable difficulties, not only because the measurements were truncated, but also because of unusually large heaping at the minimum height requirement (Figure 1).[2] In order to subvert the minimum height requirement, the height of an obviously implausible number of shorter recruits were reported as 60 inches in order to allow them into the military. The aim of this article is to estimate intertemporal, interregional and -occupational differences of Argentinian heights and to provide a modal estimate of height levels. A more detailed description of the recruitment procedure as well as the economic interpretation of the results is given in Salvatore (1998). For the moment, a distinction between the "earlier" (1810-51) and the "later" (1852-65) recruitment period is important: the earlier pertains to birth cohorts from the 1770s to 1810s, and the latter pertains to recruits born from the 1800s to the 1830s, with a modest overlap 1805-1814. Prior to 1852, officers very often rounded up the heights of the recruits 60 inches to get as many soldiers as possible, even if their true height was well below that threshold. This is evident from the excessive peak at 60 inches in all the histograms of the earlier (1810-51) recruitment period (Figure 1).Thus, 60 inches is not an actual height measurement, but includes all individuals who were shorter than 60 inches.

---

[1] We thank Markus Heintel and John Komlos for helpful suggestions.

[2] In some sources, the Argentinian inch is reported as being lower than the English inch, but some recruitment records of the 1850s report both measures. They suggest that the Argentinian foot was 30.5 cm, i.e. equal to the English inch.

Consequently, it is impossible to attribute true heights to this group, and therefore will be discarded in further analysis.[3]

However, above 60 inches the height distribution appears normal (Figure 1, see also figure 4). Among the 1780s birth cohort, a modest heaping at 65 inches is visible, otherwise heaping seems to be absent. In a few cases, the 61 inch group is smaller than the 62 inch group, which might imply a mode of around 62 inches.[4] Heights of the later sample (recruitment period 1852-65), are much less truncated, and the distributions are close to normal (Figure 2).[5] The distribution of the birth cohots of the 1830s (Figure 2, panel c) is good enough so that a OLS regression analysis can be used to analyze the data.[6]

The Komlos & Kim method (K&K) is the most robust trend estimator for severely truncated samples such as the early sample (Komlos & Kim 1990). In contrast, the Reduced Sample Maximum Likelihood Estimator (RSMLE) is inefficient if the

---

[3] As we do not know how many individuals of a given height below the minimum height requirement were accepted or rejected.

[4] It could also imply heaping on even numbers, but heaping of this sort is not evident from other parts of the distribution (Figure 1 and 4).

[5] With the exception of the 1800s birth cohort.

[6] This is done in Salvatore (1998). It is important for the following analysis to determine the age at which growth became insignificantly small. Today growth ceases at age 17.5 in the developed world, but in the 19th century men often continued to grow well into their early 20s. Three OLS-regressions of the nearly normally distributed cohort of the 1830s supports the hypothesis that growth at age 22 and above was insignificant. Regressions of height on age dummy variables, birth cohort 1830s:

Regression No.	(1)	(2)	(3)	Share of the sample
Age 21	-2.6	-2.6	-2.8	9%
	(0.01)	(0.01)	(0.01)	
Age 22		-0.3	-0.4	17%
		(0.72)	(0.61)	
Age 23			-0.6	14%
			(0.46)	
Constant	161.0	161.1	161.2	60%
	(0.00)	(0.00)	(0.00)	
R2	0.01	0.01	0.01	
N	978	978	978	

The constants refers to a recruit aged (1) 22-49, (2) 23-49, (3) 24-49. P-values of t-tests in parentheses. The share of these age groups 21, 22 and 23, compared to the other ages 24-49, was about 40%. Therefore, the insignificant coefficients are not caused by small subsample size. Instead, we tentatively conclude that growth was negligible after age 21 among the 1830s birth cohort. Modal estimates (explained below) of the 1780s and 1790s cohort yielded nearly identical heights for ages 21, 22, 23 and 24-49. We conclude that including ages 22-23 among the adults will in the following analysis not bias the estimates.

truncation point is close to the true mean, and we found this to be the case through visual inspection of the histograms.[7]

The K&K trend results for both the earlier and later sample are graphed in figure 3. Both samples were restricted to that part of the sample above 61 inches to allow unbiased intertemporal comparisons. Interestingly, during the overlap of the 1810-51 and 1852-61 recruitment period (birth cohorts of 1805-09 and 1810-14) the K&K estimated heights were very similar, even if the number of cases was small (see Salvatore 1998, this volume).

Heights during the late 18th century were declining and reached the lowest point in 1800-04. The period until the early 1820s can be characterized by increasing heights, which continued after a short dip in the late 1820s until 1835-39.

This time trend is confirmed by modal estimation techniques based on the kernel density estimate of the height distributions (Heintel/Baten (1998), Heintel 1996). Kernel density estimation bears some similarities with the much more popular calculation of moving averages of time series, insofar as it smoothes the height distribution. Figure 4 compares the histogram to the smoothened height distribution for the 1780s. A visual inspection of the histogram enables us to determine the mode only roughly at between 62 and 62.9 inches. However, the kernel density estimate uses the information of the surrounding height intervals to provide a more accurate estimate of 62.1 inch (Table 1). This method enables us to determine also the average height.

In general, the two estimates are highly correlated with one another (Figure 3).[8] In the case of the 1800s, the modal estimate is surely an upper bound estimate. As the truncation point is closest to the mean during this period, the kernel density does not have enough short recruits to the left of the mode to provide an accurate estimate. In contrast to modal estimates, the K&K estimates do not reflect actual mean heights, but provides a trend and the differences in human stature. A way to transform K&K

---

[7] We also tried a large number of possible specifications using the RSMLE, but the results were implausible, yielding height increases of more than 5 centimeters, for example.

[8] Minor deviations pertain to the 1770s and 1800s. For the 1770s, sample size might be too small for a modal estimate.

estimates into estimates of true means is to regress the modal estimates on the K&K values. This regression yields:

Mode =   0.49 * K&K value

p-value: 0.00

R^2:   0.61

N   13

The regression was standardized for means.

This means that a one unit K&K decrease from the K&K result of 162.7 equals a decline from the average of 158.7 cm down to 158.2 cm. Figure 3 compares modes and standardized K&K estimates. While the standardized K&K estimates are more volatile and indicate a stronger decline down to the early 1800s, the increase in height this time period is evident in both series.[9]

We conclude that there was a declining trend in height during the 18th century, and a recovery and continued increase during the 1810s and 1830s birth cohorts. Both the modal estimates and the K&K estimates yield similar results, while truncated regression methods using RSMLE were impossible to use given the close proximity of minimum height requirement to the mean in the earlier sample.

References

Heintel, Markus (1996). Historical Height Samples with Shortfall. A Computational Approach. History & Computing 8/1: 24-37.

Heintel, Markus/Joerg Baten (1998). Smallpox and Nutritional Status in England, 1770-1873. On the Difficulties of Estimating Historical Heights. Economic History Review 51-2.

---

[9] It is importnat to note that the K&K or modal estimation method does not standardize results with respect to birth region of the recruit. Therefore, we also estimated the height trends using truncated OLS to make sure that the trends were not caused by compostional effects. The results confirmed the K&K and modal estimates.

Komlos, John/Joo Han Kim (1990). On Estimating Trends in Historical Heights. In: HM 23 (1990), S. 116-120.

Salvatore, Ricardo (1998). Heights and Welfare in Post-Independence Argentina. In this volume.

Table 1: Mode estimates of Argentinian heights by birth decade

	Mode	N
1770	159.8	48
1780	157.8	318
1790	157.6	340
1800*	157.6	140
1810	158.7	231
1820	160.0	556
1830	160.5	538

N= Number of cases above the minimum height requirement (60 inches excluded)
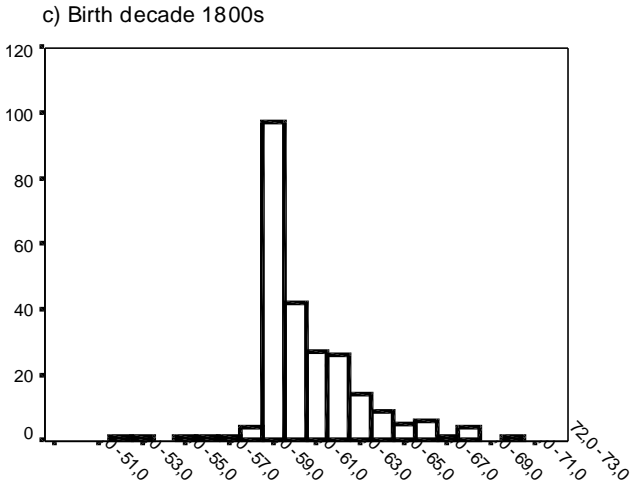* 1800-04 only: 157.5 cm

Figure 1: Height distributions of the earlier sample (recruited 1810-51), Argentinians aged 21-49

a) Birth decade 1780s

Inches

b) Birth decade 1790s
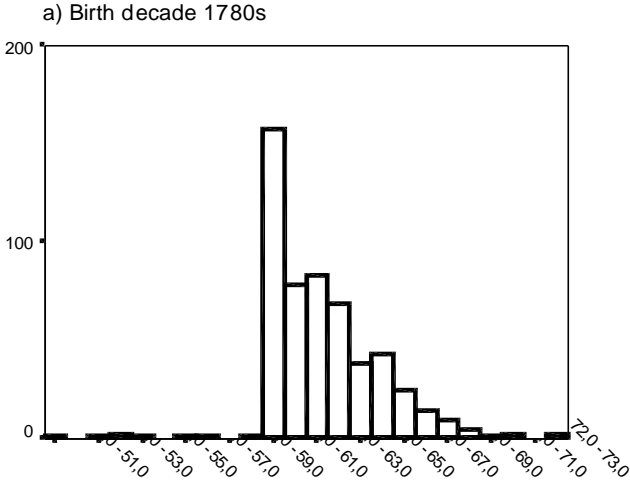
Inches
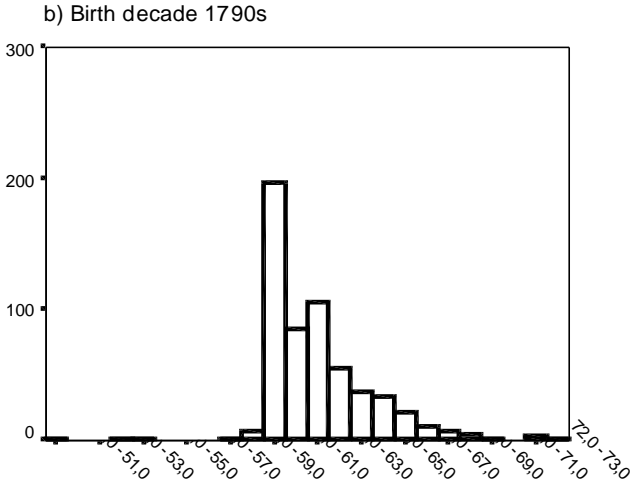
c) Birth decade 1800s

Inches

Figure 2: Height distributions of the later sample (recruited 1852-65), Argentinians aged 21-49

a) Birth decade 1810s

Inches

a) Birth decade 1820s

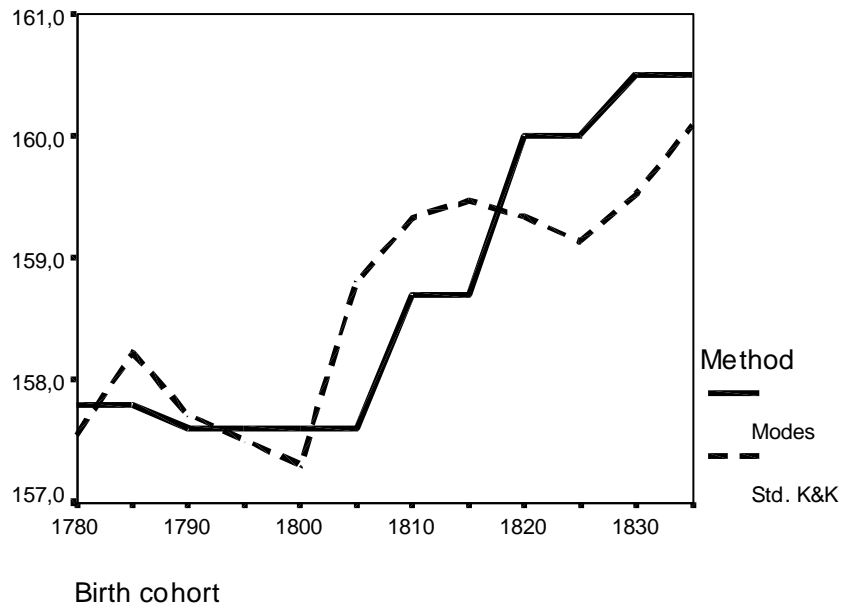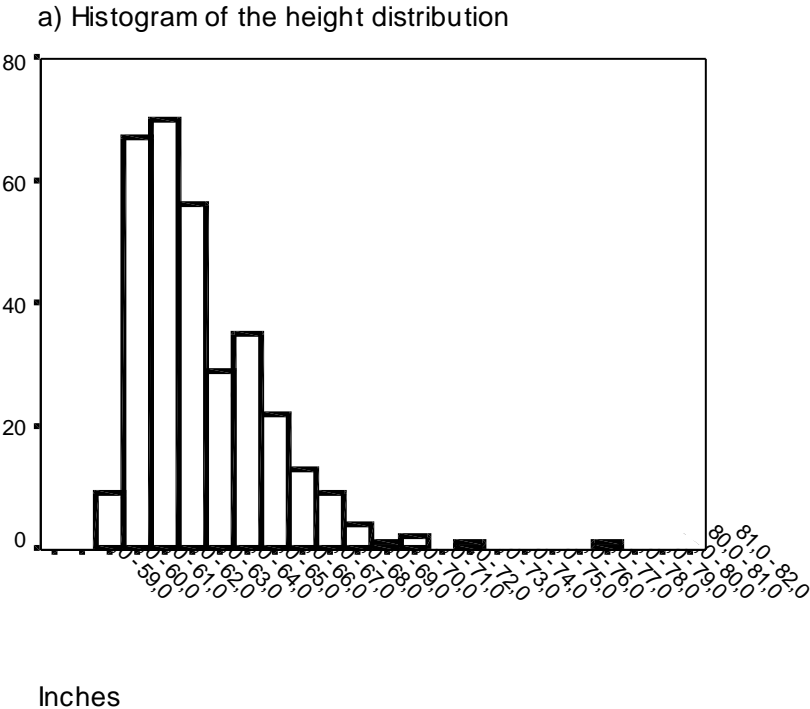

Inches

a) Birth decade 1830s
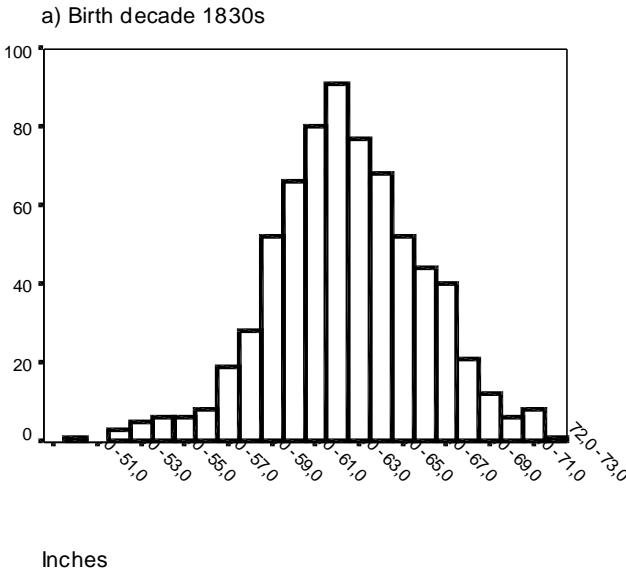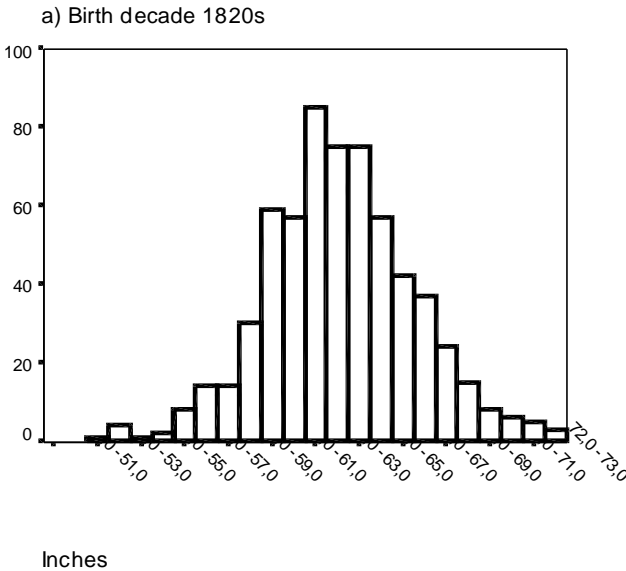
Inches

Figure 3: Height trends in Argentina

Birth cohort

Age 22-49; K&K are standardized, see text; kker5.cht

Figure 4: Height distributions: comparison of histogram and kernel density estimates (Argentinians aged 21-49, birth decade 1780s, truncation point 61 inches)

b) Kernel density estimate

a) Histogram of the height distribution

Inches

in %