5-1-2013

Problems of Sample-selection Bias in the Historical Heights Literature: A Theoretical and Econometric Analysis

Howard Bodenhorn

Timothy Guinnane

Thomas A. Mroz

ECONOMIC GROWTH CENTER
YALE UNIVERSITY

P.O. Box 208269
New Haven, CT 06520-8269

http://www.econ.yale.edu/~egcenter

Economic Growth Center Discussion Paper No. 1023

Economics Department Working Paper No. 114

Problems of Sample-selection Bias in the Historical Heights Literature: A Theoretical and Econometric Analysis

Howard Bodenhorn
Clemson University

Timothy W. Guinnane
Yale University

Thomas A. Mroz
Clemson University

May 5, 2013

*This paper was first circulated under a different title. For comments and suggestions we thank Shameel Ahmed, Cihan Artunç, Gerard van den Berg, Claire Brennecke, Jeremy Edwards, James Fenske, Amanda Gregg, Farly Grubb, Sukjin Han, Brian A'Hearn, Philip Hoffmann, Sriya Iyer, John Komlos, John Murray, Sheilagh Ogilvie, Jonathan Pritchett, Paul Rhode, Mark Rosenzweig, Gabrielle Santangelo, Richard Steckel, Jochen Streb, William Sundstrom, Werner Troesken, James Trussell, Christopher Udry, Marianne Wannamaker, David R. Weir, and participants in seminars at the University of Michigan, the University of Nuremberg, the Queen's University (Ontario), the Rhein-Westfälisches Wirtschaftsintitut, Tulane University, and the 2012 Cliometrics meetings. We acknowledge financial support from the Yale University Economic Growth Center. Meng Liu, Yiming Ma, and Adèle Rossouw provided excellent research assistance. Direct correspondence to Guinnane: timothy.guinnane@yale.edu.

Abstract

An extensive literature uses anthropometric measures, typically heights, to draw inferences about living standards in the past. This literature's influence reaches beyond economic history; the results of historical heights research appear as crucial components in development economics and related fields. The historical heights literature often relies on micro-samples drawn from sub-populations that are themselves selected: examples include volunteer soldiers, prisoners, and runaway slaves, among others. Contributors to the heights literature sometimes acknowledge that their samples might not be random draws from the population cohorts in question, but rely on normality alone to correct for potential selection into the sample. We use a simple Roy model to show that selection cannot be resolved simply by augmenting truncated samples for left-tail shortfall. Statistical tests for departures from normality cannot detect selection in Monte Carlo exercises for small to moderate levels of self-selection, obviating a standard test for selection in the heights literature. We show strong evidence of selection using micro-data on the heights of British soldiers in the late eighteen and nineteenth centuries. Consequently, widely accepted results in the literature may not reflect variations in living standards during a soldier's formative years; observed heights could be predominantly determined by the process determining selection into the sample. A survey of the current historical heights literature illustrates the problem for the three most common sources: military personnel, slaves, and prisoners.

JEL Codes: I00, N3, O15, O47, C46, C52, C81

Keywords: self-selection, selection bias, heights, anthropometrics, standards of living, industrialization puzzle, long-run economic growth

> ...the socio-economic composition of the institution studied might have varied over time, even in the absence of explicit changes in the admission criteria. This might be due both to supply and demand considerations. The willingness of individuals to enter the military, for instance, might have varied over time... This problem is quite intractable (Komlos (2004, note 44)).

1 Introduction

Anthropometric history came into its own in the 1980s, and now enjoys a prominent place in quantitative economic history. The central idea of the heights literature is that, at the population level, heights after maturity reflect the net nutrition available to individuals during the growing years. Thus a cohort might be unusually short because its members had less food, or gross nutrition; because hard work during youth made greater caloric demands on gross nutrition, leaving less for growth; or because disease made demands on gross nutrition.[1] The heights literature often presents the anthropometric approach as capable of capturing dimensions of economic well-being that the real wage or GDP per capita cannot. This approach's appeal is obvious. Many historical sources record heights. By studying direct measurements of the human organism researchers can hope to achieve a broader picture of how economic growth and development affected human well-being. Unfortunately, as we show in this paper, the sources used in this literature often suffer from selection bias that cast doubt on its substantive conclusions.

Prominent researchers use height measures to gauge changes over time in the well-being of populations, often referring to the historical literature or making use of the data generated by it. Fogel's (1994) Nobel Lecture, for example, presented a theory of long-run development based in important ways on the historical heights literature. Weil's (2007) detailed analysis of the effect of health on economic growth assumes that compilations of historical heights accurately reflect the true distribution of heights in those populations. One of his estimates of the effect on height on labor productivity depends crucially on an accurate measure of the average height in Great Britain from the late 1700s. Crimmins and Finch (2006) link military samples generated by economic

[1] Most research on heights relies on sources that pertain to men only, such as soldiers. Some historical research uses anthropometric measures other than heights. Our argument applies with equal force to any anthropometric measure used to study human well-being; for brevity here we use the term "heights." Some early papers in this literature viewed height as a proxy for conventional measures of economic output such as GDP per capita. Brinkman, Drukker and Slot (1988), for example, regress the heights of Dutch males on per-capita income in Holland in the period 1900-1940, and then use the resulting equation and measured heights to infer incomes in the nineteenth century.

historians with age-specific mortality to estimate associations between long-run changes in height and mortality. Deaton (2007), for modern Africa, and Bozzoli, Deaton and Quintana-Domeque (2009), for modern Europe and the United States, attribute part of the lack of a substantively meaningful height-income relationship to conclusions drawn from comparisons of the heights of British and Irish volunteers to the East India Company in the 1800s. These and other important papers on economic growth and development treat height measures from possibly non-random, selected samples as if they accurately reflect true measures of economic conditions. Our analysis suggests one should be much more skeptical about the ability of such samples to provide useful information.

Using historical heights forced scholars to contend with several peculiarities of these sources. Militiary forces often recorded individual-level data on soldiers that include height, but many militaries imposed a minimum height requirement. Military height samples are thus often left-truncated at or near that minimum height. Truncation led to a number of approaches to estimating mean height from truncated samples, most notably the Reduced-Sample Maximum Likelihood Estimator (RSMLE) and the Quantile Bend Estimator. While there remain some debates about how this is best done, most recent studies account for truncation using well-tried approaches.[2]

Unfortunately the historical literature has not satisfactorily confronted the consequences of a potentially more serious problem; namely, that many if not most of the samples used are not random draws from the population in question. Non-randomness is distinct from left truncation. For most sources, an individual risks inclusion in the sample only if he or someone else made a decision that led to their inclusion. Examples include soldiers and sailors in volunteer forces; prisoners; runaway or manumitted slaves; slaves that entered into the international or interregional slave trade; students at elite educational institutions; and passport holders.[3] If the probability of inclusion in the sample differs across individuals in ways that are correlated with height, then the sample height distribution may not be an unbiased measure of the mean population height, even after adjustment

[2]The seminal reference is Wachter and Trussell (1982a). The Reduced-Sample Maximum Likelihood Estimator (RSMLE), requires *ex ante* specification of a (maximum) fixed truncation point. The Quantile Bend Estimator relies on fitting a linear relationship between the expected and actual normality probability plots. Some studies discuss another problem, which is "heaping" of heights. One would expect the distribution of heights to be continuous, but measured heights often show both rounding (say, to 5'4" when the true height is 5' 4.1" inches) and a tendency to cluster at certain heights, such as 5 feet. Our argument is distinct from both truncation and heaping.

[3]Section 7 discusses the heights literature in detail. Research relying on volunteer armies includes Floud, Wachter,and Gregory (1990) and Komlos (1989); on prisoners, Nicholas and Oxley (1996); on the slave trade, Steckel (1986, 1987); on students, Murray (1992); on passports, Sunder (2011).

for possible left-truncation. More importantly, *changes in observed heights over time may be driven not only by changes in the actual heights of the population, but also by changes in the probability that individuals of different heights were differentially selected into the sample.* Similarly, cross-sectional differences in height within a selected sample may reflect a particular variable's effect on the probability of inclusion in the sample, rather than population differences in height correlated with a given characteristic. Thus many heights studies could suffer from sample-selection bias (SSB).[4]

We construct a simple Roy-style model of the decision to join the Army and use it to generate two important conclusions. Even modest military-civilian differences in the returns to height can generate (selected) military samples whose mean height is a poor estimate of the population mean height. Simulations based on our Roy model show that statistical tests of normality have almost zero power; they cannot detect departures from normality in selected samples. Thus the height literature's standard approach to detecting selection does not perform well. In addition, both the RSMLE and the QBE estimators rely crucially on the normality assumption. While we are not able, using the data available, to estimate an econometric model of heights that corrects for the selection problems, we report several exercises that when considered together provide strong evidence for selection.

There is compelling evidence of differential selection into the modern U.S. military, which since 1973 has relied entirely on volunteers. Simon and Warner (2008), for example, find that the elasticity of the supply of high quality recruits (those scoring 50 or higher on the AFQT) over the years 1996-2005 with respect to state unemployment rates is 0.42; the elasticity for the overall supply of military recruits is only 0.22. Warner and Asch (2001), also find the supply of high quality recruits to be quite responsive to increases in the unemployment rate over the first 20 years since the end of the U.S. military draft. They attribute much of the decline in high quality recruits during the 1990's to declines in relative military pay and civilian unemployment rate over that decade. Warner, Simon, and Payne (2000) estimated a 15,000 person annual decline in high quality recruits over that decade due to the declines in unemployment rates. Kilburn and Klerman (1999) find that factors associated

[4]Extensive discussion of SSB in the economics literature dates to efforts to estimate models of married women's wages and hours of work. See Heckman (1979). Some sources used in the heights literature are themselves random samples drawn from a larger source, for example, a random sample of men who joined a particular volunteer army. We are not criticizing the method of drawing the sample; our argument is that the Army itself represents a selected sample of men from that time.

with increases in the returns to education reduce military enlistment. Asch, Hosek, and Warner (2007) present a detailed review of how unemployment rates affect high-quality military enlistments; all of the elasticities they report fall in the range of 0.11 to 0.77. If selection is important for modern military recruitment, it was likely important for volunteer forces in the past as well.

Several papers in the historical heights literature recognize that their samples might not be representative, but few have acknowledged that some of the most-discussed results in this literature may be entirely an artifact of SSB.[5] Our message is a critique, not a condemnation. We offer a warning against too-facile connections drawn between height and human welfare. The anthropometric approach retains its promise of offering much-needed insight into human welfare, but existing interpretations need to be re-examined and tempered, given the implications of selection-bias in many heights samples.

2 Declining heights in the presence of rising real wages

The sample-selection problems we discuss affect any study based on a choice-based sample of heights in the past. To fix our ideas, however, we focus on perhaps the most extensive debate that relies on height data, the changes in the standard of living that came with industrialization and after. (In Section 7 we discuss the sample-selection issue as it arises in other heights research more broadly.) Economic historians and others have studied how economic growth, and more specifically industrialization, affected human welfare in the past. The most famous version of this "standard of living debate" focuses on the working classes during the British Industrial Revolution. A long debate that goes back at least to Engels (1845/1897) divides "optimists," who think the working classes benefited from early industrialization, from "pessimists," who think they did not. The earliest literature focused on conditions of work and life – sanitation, diet, disease and mortality – but in its modern incarnation the literature has focused on conventional economic measures such as real wages, or incomes or consumption per capita.[6]

[5] Pritchett and Chamberlain (1993) and Grubb (1999) discuss selection in specific contexts. Mokyr and Ó Gráda (1996) is the only anthropometric study we know of that acknowledges that their result may be a consequences of SSB. We read their conclusions as a precursor of what we say here. See also Ó Gráda (1996). Weir's (1997, p.175-7) discussion alludes to SSB in the British Army data. We probably first developed the ideas presented in this paper in conversation with Weir. Lamm (1988) addresses the question of differential record survival, which is potentially a type of selection bias but not the topic of our paper.

[6] Feinstein's (1998) broadly-accepted estimates of real wages imply little increase until least the first decade of the 19th century. Important references in the large literature on the standard of living during this period include

Floud, Wachter, and Gregory's (1990) finding that British soldiers grew shorter during the British Industrial Revolution, and after, seemed a strong point for the "pessimist" case. Nicholas and Steckel (1991) report similar evidence. The estimated mean height of English males declined by about 1.5 inches during a period in which wages were at worst stagnant or, more likely, growing slowly. If anything, this evidence seems stronger than conventional "pessimist" arguments: Feinstein (1998) reports wage *stagnation*, not wage decline, and Allen (2007) reports slow real-wage growth throughout the nineteenth century. But British soldiers apparently became significantly shorter. Something similar also occurred in the United States. During a period when the U.S. economy experienced significant real GDP growth, estimated mean height declined by more than one inch (Weiss 1992). Growth in estimated heights did not recover until the end of the nineteenth century. The heights literature and much of the profession has taken the finding of rising incomes and shorter people at face value. The US literature calls this negative correlation between heights and income the "antebellum puzzle." Komlos (1998) presents a list of possible explanations, including increasing income inequality, increased income variability, changes in relative prices between nutritious (meat and grains) and non-nutritious (coffee and sugar) foods, transportation innovations and the integration of the disease environment and intensification of labor, among others.

The heights literature focuses on three of these effects. First, changes in the income distribution might reduce net nutrition for the working classes, even if GDP per capita increased (Steckel (1983)). This explanation is potentially related to the SSB problem. Changes in income distribution could be a driving force behind different types of selection into the military; if workers with particular sets of skills saw wage declines, they might more readily enlist.[7] A second explanation for the puzzle focuses on an unhealthy environment. Cities grew with industrialization, and we know that cities were comparatively unhealthy places to live until at least the late nineteenth century. Diarrheal and other endemic infectious diseases common in cities drive a wedge between gross and net nutrition (Fogel et al (1982)). A third explanation emphasizes the price and availability of important food items. The relative price of nutritious, height-enhancing foods might have changed in ways that lead to less consumption even with rising real wages. Komlos (1987), Bodenhorn (1999) and others have

Brown (1990), Crafts (1997), Feinstein (1998), Hobsbawm (1957), and Williamson (1984).Voth (2003) summarizes the debate.

[7]Similarly, Steckel and Prince (2001, p. 291) appeal to the egalitarian practices of the Native Americans of the Great Plains to explain their relatively *tall* stature.

investigated this possibility, but translating food production into consumption estimates is difficult. Floud et al (1990, pp.233-243) doubted whether such studies could contribute much to the debate.

The puzzle provoked a host of studies that provide valuable insights into the standard of living, but we are skeptical of the puzzle itself as a real phenomenon. To date, the negative correlation of income and height only emerges in choice-based (selected) samples. Both the UK and the US had volunteer Armies for most of the period in question. France, on the other hand, drafted men according to lotteries so that all young men had an approximately equal chance of being called for service. Consider Figure 1, from Weir (1997), which presents the heights of British and French males over the same period. There is no French equivalent of the industrialization puzzle; French men grow without interruption. A simple regression of the French figures on a linear time trend yields an R-square of about 0.97. What accounts for the difference between the US and British case on the one hand, and the French, on the other? The heights literature focuses on France's slower urbanization, which reduced the disease burden and thus the impediments to converting calories into height. There may be something to these differences, but we think selection bias in the British and U.S. heights samples offers a more satisfactory explanation.

The enormous variations in estimated mean heights of British soldiers reflected in Figure 1 cannot reflect just short-term variations in the standard of living. Consider the cohorts that reached age 20 around 1810. Between 1804 and 1812, French heights fluctuate between 163.5 cm and 163.9 cm, less than 0.5 cm, without obvious trend. British heights, after having already experiencing a 2 cm (=0.79 inches) decline 20 years earlier, first decline by about 2 cm and then as quickly recover. The heights of men potentially eligible for the French military do not oscillate as wildly as British heights even though both militaries imposed minimum (but not common) height standards. We return to this issue in Section 6 below.[8]

[8]Militia conscription records for Sweden for the period 1820–1965 cover nearly the entire male population, and, like France, show no deterioration in heights (save for a .5 centimeter dip from 1835–40, which was made good by 1845 (Sandberg and Steckel 1997, Table 4.1). Using the settled Army, which was apparently a volunteer force, Sandberg and Steckel show a sharp decline in height for the cohort born in the 1840s (Sandberg and Steckel 1997, p.135; Sandberg and Steckel 1988, Figure 1).

3 Modeling Military Heights

We begin with a simple model in which an individual of given characteristics decides whether to join the Army or remain a civilian. This model could be adapted to fit several other contexts that generate height data. The model delivers expressions describing the degree of selection into the Army, by height. In Sections 4 and 5 below we use the model to simulate height distributions for specific parameters. Our model draws on Roy (1951), and resembles Heckman and Sedlacek's (1985) two-sector occupational choice model. Each person's utility depends on his "wage" or material compensation in the Army or in the civilian world. Each individual also has parameters that describes his taste for military life and for civilian life. Thus two individuals with identical "wages" can make different decisions.

We treat both Army and civilian wages as (possibly) a function of the individual's height. This assumption has two interpretations that are equivalent for our purposes. The first is that the Army or some civilian occupations might reward height itself. Promotion might come faster to taller soldiers, and in some military and civilian occupations a tall person might be more productive. A second interpretation seems more natural, and more consistent with the basic tenets of the heights literature. Height is correlated with a person's "biological standard of living" and with other aspects of "health human capital" (Schultz 2002). The Army might reward a tall person not for being tall per se, but because taller individuals are generally healthier and more productive than shorter people. Similarly, the civilian world might reward a tall person for his health in addition to any skills correlated with height.[9]

The military pays soldiers as a function of their height, h, and their observable set of productive military traits:

$$\ln (w_M) = \alpha_M + \beta_M h + \gamma_M \varepsilon_M. \tag{1}$$

where w_M is military compensation, α_M is a constant, β_M is the return to height in the military, and γ_M is the return to individual-specific productive traits. Those traits (ε_M) can entail any trait other than height, such as literacy or some specific skill such as the ability to shoot straight. In our

[9]The military might reward human capital and cognitive ability, though not necessarily at the same rate as civilian markets. Komlos (1989, p.237) reports enlistment bonuses into the Hapsburg Empire's army, circa 1809. These bonuses increase in heights, from 3 fl. for soldiers just 5'-0" to 45 fl. for soldiers 5'-5" and above. Persico et al (2005) report nontrivial returns to height in civilian employment. Case and Paxson (2008) attribute part of that return to cognitive abilities associated with health and height.

model this trait can affect returns in either the military or the civilian world. Civilian wages are given by:

$$\ln(w_C) = \alpha_C + \beta_C h + \gamma_C \varepsilon_M + \delta_C \varepsilon_C. \tag{2}$$

Where w_C, α_C, β_C are defined for the civilian world by analogy to their military equivalents. γ_C is the civilian sector's return to ε_M, δ_C is the civilian return to the individual characteristic ε_C that is relevant only to the civilian sector. Note that both β_M and β_C might be zero. Table 1 sets out all notation used here and in the simulations reported in Sections 4 and 5. Individuals differ in their tolerance for other features of military and civilian life, and we summarize those preferences under τ_M and τ_C. An individual's utility is additive in the log-wage and his tastes for the military or civilian life: $U(M) = \ln(w_M) + \tau_M$ and $U(C) = ln(w_c) + \tau_{C.}$ For the moment, we make no assumptions about the relationship between height and other productive military traits (ε_M). Height and military productivity could be positively correlated (tall people might be better with a rifle) or negatively correlated (tall people might find it harder to find cover in combat). Height and other military characteristics (ε_M) might be valued by both the military and the civilian economy, but another set of characteristics, ε_C, are valuable only in the civilian world. All productive traits can be correlated. For simplicity we assume that tastes (the τ terms) are independent of the productive traits (h and the ε terms). Relaxing this assumption only requires the introduction of additional covariance terms in the following derivations.

We view this primarily as a static model of occupational choice. But one could interpret the functions $U(M)$ and $U(C)$ in a dynamic context. $U(C)$ could represent the expected utility from choosing the civilian sector today when there is a possibility (option) that one might choose to enter the military at some later date. However, throughout this analysis we assume that an individual has just one opportunity to enlist in the military. Similarly, the military pay function could represent "expected" lifetime payments that could include, in part, the expected increase in pay if the volunteer were to be promoted at some later date from soldier to sergeant.

An individual enters the military if $U(C) \leq U(M)$, or (equivalently) if $\eta \leq (\alpha_M - \alpha_C)$,where

$$\eta = \beta h + \gamma \varepsilon_M + \delta \varepsilon_C + \tau \tag{3}$$

and

$$\beta = (\beta_C - \beta_M), \gamma = (\gamma_C - \gamma_M), \delta = \delta_C, \text{ and } \tau = \tau_C - \tau_M. \qquad (4)$$

The decision to join the military in the model reflects *differences* between features of the military and civilian worlds. For example, if height is more highly rewarded in the civilian sector than in the military sector, i.e., $\beta = (\beta_C - \beta_M) > 0$, then taller individuals prefer the civilian sector over the military sector, all else equal. Similarly, if the civilian sector's reward to the military trait ε_M increases (i.e., γ or γ_C increases), then *ceteris paribus* those with higher values of ε_M will be less likely to join the military.

Suppose that all productive traits follow standard normal distributions, not necessarily independent of each other or of heights, and that the taste parameters follow mean zero normal distributions independent of heights and productive traits.[10] Then height h and "selection error" η follow a bivariate normal distribution with

$$E\begin{pmatrix} h \\ \eta \end{pmatrix} = \begin{pmatrix} \mu_h \\ 0 \end{pmatrix}; Var(h) = \sigma_h^2; \qquad (5)$$

$$\begin{aligned} Var(\eta) = \sigma_\eta^2 &= \beta^2 Var(h) + \gamma^2 Var(\varepsilon_M) + \delta^2 Var(\varepsilon_C) + Var(\tau) \\ &\quad + 2\beta\gamma Cov(h, \varepsilon_M) + 2\beta\delta Cov(h, \varepsilon_C) + 2\gamma\delta Cov(\varepsilon_M, \varepsilon_C) \end{aligned} \qquad (6)$$

and

$$Cov(h, \eta) = \sigma_{h,\eta} = \beta Var(h) + \gamma Cov(h, \varepsilon_M) + \delta Cov(h, \varepsilon_C). \qquad (7)$$

Under these assumptions, we can derive the distribution of heights for those who prefer military service. Integrating the bivariate normal distribution of (h, η) described above over the range of $\eta \leq (\alpha_M - \alpha_C)$, and normalizing by the $\text{Prob}(\eta \leq (\alpha_M - \alpha_C))$ yields[11]

[10] The unit variance assumption is innocuous, and the mean zero assumption for all of these factors only requires simple redefinitions of the intercepts in the two wage offer equations. One could easily relax the independence assumption for tastes.

[11] See Kotz et al (2000, p.316).

$$f_{h|mil}(h) = f_{h|\eta\leq(\alpha_M-\alpha_C)}(h) = f_h(h)\,Z(h) \tag{8}$$

where

$$Z(h) = \frac{\Phi\left\{\left[(\alpha_M-\alpha_C-\beta\mu_h)/\sigma_\eta-\rho_{\eta,h}(h-\mu_h)/\sigma_h\right]/\sqrt{1-\rho_{\eta,h}^2}\right\}}{\Phi\left[(\alpha_M-\alpha_C-\beta\mu_h)/\sigma_\eta\right]}. \tag{9}$$

$\Phi(\cdot)$ is the standard normal cumulative distribution function, $f_h(h)$ is the unconditional (population) distribution of heights, $\frac{1}{\sigma_h\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{h-\mu_h}{\sigma_h}\right)^2\right)$, and $\rho_{\eta,h} = \frac{Cov(h,\eta)}{\sigma_h\sigma_\eta}$ is the correlation of height and the population selection error η defined above. $Z(h)$ summarizes the selection process that generates the observed distribution of military heights from the underlying distribution of population heights.[12] Most historical studies of military heights infer the population height distribution $f_h(h)$ from the observed military height distribution ($f_{h|mil}(h)$). Accurate inferences about the population distribution, however, depend on the properties of $Z(h)$. If $Z(h)$ does not depend on height, there is no selection problem. The dependence of $Z(h)$ on height, however, could vary over time (or in the cross-section) in a way that might be unrelated to variations in the population height distribution. Observed variations in $f_{h|mil}(h)$ in this case would reflect only variations in Z(h), perhaps due to changes in the parameters determining w_M and w_C rather than variation in the population height distribution $f_h(h)$.

Our model thus far ignores the minimum height many militaries imposed. The existence of an exact and binding minimum height requirement implies that instead of the conditional height distribution described above, we observe only its truncated analogue. Suppose the military enforces a minimum height requirement that is completely non-binding above some height h^*. All observed heights above h^* can then be considered as random draws from the upper tail of this conditional height distribution. This assumption underlies Wachter and Trussell's (1982a) reduced sample maximum likelihood (RSMLE) and quantile bend estimators (QBE) (i.e., h^* is above the extent of the shortfall in their terminology).

[12] The model here focuses on the supply of soldiers to the military, which we think is the important issue for selection bias. Appendix A extends the model to account for the military's optimization decisions. We show that in constructing the least-cost military force, the Army might impose a minimum height restriction, as occured in practice.

This simple choice model highlights an important assumption in Watcher and Trussel's description of the process giving rise to observed military heights. They assume there is no selection process operating for heights *above* h^*; individuals taller than h^* choose to enter the Army based on a coin flip. Our model suggests an alternative process. Suppose the civilian sector rewards heights and productive traits more than the military, and that the covariances of heights and productive traits are positive, zero, or at most only slightly negative. Then the covariance (or correlation) of heights and the selection term η is positive, and $Z(h)$ approaches zero for taller individuals. This conclusion is the central point of our critique, and has two, related implications. First, Wachter and Trussel's approach implicitly rules out such a possibility. If there is selection, their estimation techniques (whether RSMLE or QBE) generate biased estimators of the unconditional height distribution even if the observed upper tail closely resembles that for a normal distribution. Second, our model provides a clear understanding of why the right-hand tail may have fewer taller individuals than one would observe in a true normal distribution.

Wachter and Trussell's assumptions hold only under fairly extreme conditions. h^* must lie above the unconditional mean height. This implies that the military's height standard rejects at least half the population. But their assumption has an additional implausible implication: for the assumption to be compatible with the selection model requires that the correlation of height and the selection error ($\rho_{\eta,h}$) must approach -1, which seems implausible.[13]

4 Simulating the Model

We use the Roy model to assess the ability of selected samples to yield accurate information about the unconditional population height distributions. The simulations we discuss in this section use a simplified version of the model presented in Section 3. We abstract from the taste parameters (τ),

[13]Specifically, two sets of parameter values can satisfy the requirement that ($\rho_{\eta,h}$) must approach -1. (1) Assume the civilian sector more strongly rewards productive traits than the military sector ($\beta > 0$). In this case, a large negative correlation ($\rho_{\eta,h}$) arises only when the covariances of height with the productive traits ε_M and ε_C are large and negative. Such negative correlations seem unlikely; one expects *positive* correlations of heights and other skills. (2) Assume (implausibly) that the military rewards height and the relevant skills more strongly than does the civilian sector. If heights and skills are *positively* correlated, and if the reward to the civilian sector specific skill, δ_C, is small, then we obtain that the correlation of heights and the selection error is large and negative. Under these assumptions $Z(h)$ would not approach zero for taller people, and the selection operating on the upper tail would be weak or non-existent. This second possibility implies that the military rewarded common skills more than the civilian sector while the civilian sector pays at most a small premium for its sector-specific skill. If so, then the variance of wages in the military would be larger than the variance of wages in the civilian sector, which seems contrary to fact.

and assume that each individual has a given shock to his earnings in the civilian or military sector.[14] We vary α_M, α_C, β_M, and β_C, as described below. As in any simulation exercise, results hinge on an appropriate choice of parameter values. Apppendix B describes our approach to parameter calibration and reports robustness checks.[15] The simulations all assume that the population has height normally distributed with mean 66 inches and standard deviation 2.5 inches. Our primary interest is in the four behavioral parameters, α_C, α_M, β_C and β_M; in this simplified model only the sectoral differences in the intercepts and in the slopes matter for the decision to select into the military.[16]

Figure 2 reports the proportion of the population in the military and the mean heights of soldiers for a range of returns to height in the civilian sector. We consider two values of α_C. The vertical line at $\beta_C = 0.02$ marks the point at which the return to heights in the Army and civilian sector are identical, and there is no selection on height. Consider first the results for the proportion in the military. The constant terms shift the proportion of men who join the Army, but the shapes of the two curves remain similar. As the difference in returns for the two sectors increases, the proportion of the population in the military declines. Now consider the mean height of soldiers. At the point $\beta_C = 0.02$, at which there is no selection on height, the mean height in the military equals the population mean value of 66 inches for both values of α. Above $\beta_C = 0.02$, the region in which the reward to the civilian height exceeds the reward to height in the military, the mean heights of individuals willing to join the Army drops rapidly. When the return to height in the civilian sector exceeds that in the Army by 2 percentage points ($\beta_C = 0.04$), the average soldier is nearly one half inch shorter than the population mean of 66 inches. When β_C equals 0.06, so the return to civilian height exceeds the return to military height by 4 percentage points, the volunteers are more than one inch shorter than the population mean.

Figure 3 contrasts the density of heights for those in the "Army" and the population height distribution for two different returns to height in the civilian sector, namely 0.04 (left-hand graph) and 0.06 (right-hand graph). These values correspond to the civilian labor market rewarding each

[14] That is, our simulation is based on equations (1) and (3), but we set γ_C=0, $\tau_C = 0$,$\tau_M = 0$,$\gamma_M = 1$, and $\delta_C = 1$.

[15] An earlier version of this paper reported results from numeric simulations; here we drop the numeric simulations and report evaluations of analytic expressions. The simulation programs used here (written in Stata) can be obtained from the corresponding author.

[16] We re-parameterize the log-wage specifications as $\ln(w) = \alpha + \beta * (h - 56) + \varepsilon$, which just makes the intercept terms simpler to interpret.

inch of height by two and four percentage points more than the military does. Figure 3 clearly illustrates one of our principal contentions: the military height distribution unambiguously shifts to the left, meaning the Army has relatively few of the tallest men and relatively more of the shortest men. The measured standard deviation of the selected heights for those in the military is also smaller than that in the population. Yet the distribution of heights in the Army "looks normal." Moreover, we know that the military densities reported there reflect deliberate, non-random sampling schemes, and the analytic density formula clearly rules out that military heights follow a normal distribution.

Figure 3 also reports the "adjustment terms" $Z(h)$ (from equation (9)) for our simplified Roy model. The adjustment term for the 2 percentage point differential in the left-hand panel implies, relatively, that five-foot tall individuals on average would be about twice as likely to prefer the military sector as six foot tall individuals. The adjustment term for the 4 percentage point differential is even more striking. Five foot tall individuals are some 5 times more likely to prefer the military sector to the civilian sector. Those differences in selection, of course, are what cause the 0.45 and 1.1 inch mean height differentials of the selected military samples from the true population mean height. These are substantively meaningful differences.

5 The power to reject normality in selected samples

Almost all researchers in the historical height literature accept the view that population height distributions closely follow the normal distribution. Equation (8) shows that the analytic distributions of selected heights derived from the simple application of the Roy model do not follow normal distributions when height differentially affects the rewards in the military and civilian sectors. This implication suggests that a test of the null hypothesis of normality for an observed height distribution, against the alternative hypothesis of observed heights not following a normal distribution, constitutes a useful device for detecting the presence of sample selection bias. Scholars working in the heights literature have used such tests to argue that their samples do not suffer from selection problems. Nicholas and Steckel (1991, pp. 941-2), for example, have used the failure to reject such a null hypothesis as a reason to dismiss concerns about selection biases.[17] In this section we demonstrate that such tests may have little power to uncover the existence of selection biases even

[17] This approach goes back to at least Sokloff and Vilaflor (1982, p.469). See also Margo and Steckel (1982, p.533) and Johnson and Nicholas (1997, p.206).

when the magnitudes of the selection biases are substantively important.

If our model is a reasonable approximation to why individuals choose to enter the military, then we have a puzzle: why does the upper tail of nearly every one of the height distributions examined by Wachter and Trussel appear to follow a normal distribution? The answer follows from the simulations in Section 4: for a wide variety of "reasonable" variances, covariances, and reward structures, the distributions of the self-selected heights in the military follow distributions that closely resemble the shapes of normal distributions *even when there is selection built it into the simulation model.* This result highlights a serious statistical problem. Even with moderate-to-large sized samples by cliometric standards, standard skewness-kurtosis tests have limited power to reject normality for the forms of selection modeled here.[18]

It is vital, therefore, to evaluate the ability of standard tests for normality to detect deviations from normality in this case. A failure to reject normality would provide some reassurance about the representativeness of the observed data only if the power of the normality tests were large enough to reject interesting alternative distributions. To evaluate this we use a test that examines deviations of estimated measures of skewness and kurtosis in an observed sample from their theoretical counterparts derived from a normal distribution with the same mean and variances as estimated in the observed sample.[19] We use our model to generate samples of between 100 and 50,000 soldiers, assuming a 4 percentage-point differential in the return to heights between the military and civilian sectors. These sample sizes correspond to the range of sample sizes found in the literature. This specification of the data generating process implies a 11.34% unconditional probability of preferring the military sector to the civilian sector. For each population size, we draw 1000 samples, and for each of these samples we apply the selection model to obtain the "military" subsample. We then calculate whether the test would reject the normal distribution assumption for the selected "military" subsample. Figure 4 reports power functions (as a function of sample size) for these tests. The results imply that this approach cannot detect selection. For tests at the 5% level, even

[18] Self-selected samples of heights do typically differ from normal height distributions in the *variance* of the observed height. The selected height distributions, even though they appear to be almost normally distributed, typically have standard deviations for heights well below those in the underlying population. Many of the estimated standard deviations in Floud et al's (1990) Table 4.8, especially for the 1800s, seem quite small, suggesting a strong degree of selection into the military consistent with this simple extension of the Roy model.

[19] For most sample sizes considered here, the Shapiro-Wilk (for $N \leq 2000$) and Shapiro-Francia (for $N \leq 5000$) tests performed similarly. Discussion in the text is restricted to the skewness-kurtosis test implemented in Stata ("sktest"). There are many different tests for normality; Appendix C reports similar exercises for a wider range of tests.

for military subsamples as large as 50,000 persons, one would correctly reject the null hypothesis only about 5% of the time. At the 10% level, one would reject only about 10% of the time. Even for the largest sample sizes in these two figures (samples that are large relative to those found in the historical literature), the tests have virtually *zero* power (above the tests' size/level) to reject the assumption that the selected samples come from a normal distribution.[20] This happens even though we *know* that they are not normally distributed.

We can see the reason for the poor performance of these tests by examining $Z(h)$ in the distribution of the selected heights sample. We compare $Z(h)$ to "rejection sampling," in which one samples from an "instrumental distribution," and randomly accepts each draw from that instrumental distribution as if it were from the target distribution. If this rejection sampling rate is proportional to the ratio (density of the target distribution)/(density of the instrumental distribution), then the random draws selected in this way follow the distribution of random samples from the target distribution.[21] The right kind of rejection sampling from a true normal distribution can yield another normal distribution. Suppose the target distribution is a normal distribution with mean and variance equal to the mean and variance for the non-normal selected height distribution. Denote this distribution $f_{nor@mil}(h)$. Let the population height distribution, $f(h)$, be the instrumental distribution. When drawing from the population height distribution acceptance probabilities proportional to $f_{nor@mil}(h)/f(h)$ will yield samples following those that would be drawn from the selected height distribution because

$$\int f(h)\left(\frac{f_{nor@mil}(h)}{f(h)}\right)dh = \int f_{nor@mil}(h)\,dh. \tag{10}$$

A comparison of this relative acceptance probability, $q(h) = f_{nor@mil}(h)/f(h)$, to the $Z(h)$ function (in equation (8)) reveals why the non-normal selected height distribution so closely resembles a normal distribution. Figure 5 provides such a comparison for the case when the reward to height differs by four percentage points in the two sectors and the selected mean height differs from

[20] Some studies do recognize the low power of these tests. See, for example, Sokoloff and Villaflor (1982, p.457).

[21] To make this analogy to rejection/acceptance sampling precise, we need to assume that the height distributions only have positive support over finite subset of the real line, say for adult heights in the range 46 to 86 inches (e.g., the mean plus or minus eight standard deviations). The ratio of the target to the instrumental distribution needs to be bounded at all relevant values such that $f_{target}(h) < Mf_{inst}(h)$ for some fixed, finite value of M; this need not be satisfied for normal distributions with unbounded support. To simplify the notation, throughout this discussion we assume that $M = 1$, satisfies this criteria; allowing M to be larger than 1 only complicates the notation.

the population mean height by over one inch. The two different "adjustments" to the normal population height distribution are nearly identical, even though one adjustment, $q(h)$, yields a normal distribution exactly and the other, $Z(h)$, yields a selected distribution that clearly does not follow a normal probability model. Moderate amounts of selection can generate selected populations that differ substantively from the unconditional population distribution in terms of the first two moments, but tests for normality would be unlikely to provide any evidence of selection biases because of the similarity of the selected distribution to a normal distribution.[22]

5.1 Which selection mechanism?

So far we have shown that it is difficult to distinguish selected height distributions from truly normal distributions. Now we show that it is also difficult to distinguish among a wide variety of selection mechanisms that can give rise to an observed selected height distribution. Consider as a baseline military sample the one generated by the parameters appearing in the first column of Table 2. While the "population" has a mean height of 66 inches, the "observable" military height distribution has a mean height of 64.9 inches, more than one inch below the true population mean height. Columns (2) though (6) in Table 2 report alternative normal population height distributions and selection mechanisms that yield approximately the same distribution of observable military heights. These scenarios vary greatly in mean population height (from 64 inches to 68 inches) and in the form of the selection, from positive selection into the military based on height to negative selection and no height-related selection.[23] Figure 6 displays the population height distributions for the first five models as well as the distribution of observable military heights generated by these population distributions and their associated selection mechanisms. Substantively different population height distributions combined with substantively different selection mechanisms give rise to approximately identical selected height distributions.

To demonstrate how nearly identical these differently selected distributions appear to be, we generated a sample of military heights from a (Monte Carlo) population of 20 million observations

[22] Earlier we noted that the RSMLE and QBE estimators require an assumption of normality, which is not satisfied. Here we show that the selected distributions are so close to normal distributions that standard tests for normality cannot detect the difference. Note that this latter result does not rehabilitate the estimators; in a selected sample, they still estimate the wrong mean and variance.

[23] Model 6 is just a normal distribution with mean and variance given by the selected height distribution generated by the baseline model.

using the baseline model reported as column (1) of Table 2. Applying the selection model, we obtained nearly 1.6 million observed military height outcomes from this baseline specification. For each of these observed heights, we constructed the value of the conditional height probability distribution function for each of the six models described in Table 2 (columns 2-7). For a variety of different sample sizes, we drew random samples from these observed heights (without replacement within each sample size) and determined which of the six models had the highest log-likelihood value for each sample when evaluated at each model's true parameter values. Table 3 contains summary information from these experiments. Most surprisingly, this criterion seldom selects the true data-generating model (model 1) as the best-fitting model.[24] If one were to rely upon a likelihood criterion to select a model as having generated the observed sample, one would often select as "best fitting" a model that implies a population height mean well above or below the true mean. And in a large fraction of cases (19%-26%) one would select a model with no selection on height as the best fitting model.

The simulations indicate that researchers cannot rely on the selected height data alone to make informative statements about the distribution of heights in the population. Identification issues arise mostly from the fact that in this exercise we are attempting to uncover the process determining selection into the military based solely on a sample of heights for those in the military. A key piece of missing information is the fraction of the population that enlists in the military. If researchers had that information and understood the determinants of how selection into the military and civilian sectors operated over time, it might be possible to better discriminate among these models. Without such knowledge, we must rely upon arbitrary assumptions about how individuals end up in the military to tell a story of how birth cohorts heights varied over time. The above analysis indicates that it will be nearly impossible to discriminate among alternative sets of assumptions that can yield substantively different stories about how mean height in the population varied over time when we use just the information typically employed by heights researchers, the heights of those in the military. However, as we demonstrate in the next section, just adding a single piece of information, the fraction of the population that enlists in the military, would make it much easier to discriminate

[24] When we used all 1,573,488 observations as one sample, model 6 (incorporating no selection on height) had the highest value of the log-likelihood function. The true model's log likelihood, however, was only smaller by 0.6 points which is likely within computer round-off error. The p-value for the null hypothesis of normality for this full sample using the Doornik-Hansen test was 0.068 (Stata's sktest also had a p-value of 0.068.)

among thse models.

5.2 Identification when additional information is available

The preceding analysis demonstrates the severity of the identification problem for heights researchers: a single observed military height distribution, to a close approximation, can be generated by a wide variety of self-selection mechanisms, including a mechanism with no self-selection at all. But the situation is not without hope. Appendix D extends our discussion to the case where the researcher has available covariates that affect the relative desirability of the military sector. Even with this additional information, if the dataset contains only observed military heights then one can expect severe problems when trying to uncover the population distribution of heights from a selected subsample. However, if we also observe the fractions of the population entering the military, then it is possible to overcome this severe identification problem.

Here we briefly describe this extension and our results, referring the reader to Appendix D for details. For this exercise we generated a Monte Carlo sample of 20 million observations from a population quite similar to the one we just examined. The new specification assumes that exogenous shifters affect population mean height and the relative attractiveness of the civilian sector arising from observable explanatory variables. While these variables do not appear in many of the main sources used by the heights literature, they are often, in principle, available. For example, one can proxy the relative attractiveness of the civilian sector using data on macroeconomic conditions at the time the individual enlisted (see Section 6). Otherwise the model we employ here is similar to that introduced in Section 3. We assume that observed shocks to height and to the desirability of Army life are symmetric about zero and independent of each other. Under this data-generating mechanism, approximately 8.15 percent of individuals prefer the military to the civilian sector. After imposing minimum height standards of 63 and 65 inches (equally probable and independent of all other covariates), our Monte Carlo dataset includes 1,004,410 observations (5.02 percent of the total) who prefer the Army and are acceptable to the Army. In all estimations we use only information on the integer portion of heights; we adjust each likelihood function to integrate over the continuous heights that would imply the integer portion of height. This is nearly identical to a Heckman selection model, except we assume only integer heights within the selected sample, and impose a binding upper bound on the military height restriction.

Our results (summarized in Appendix Table D.1) imply that many estimates currently found in the literature are biased (as we argue in Section 3 and Section 4) and not identified. Estimating the RSMLE model using our Monte Carlo dataset, we obtain a mean height estimate one full inch less than the true value. We then consider a more general model that conditions on selection into the military, but that makes no use of information drawn from individuals who do not join. This model does not converge easily, suggesting it is not well-identified, and it yields a variety of biased estimates of mean height that are sensitive to starting values. Finally, we consider an unconditional selection model that uses information on the entire population. This model converges well and its parameter estimates adhere closely to the moments built into the simulation. Again, this last model requires additional information not typically used by researchers in the heights literature. But information analogous to what we introduce in the final model may be available, namely, the fraction of each birth cohort joining the military at each point in time. This offers the chance to obtain estimates of mean height that are relatively free from selection bias.[25]

6 Econometric analysis of selection in the British military data

Our theoretical and simulation analysis demonstrates that studies based on volunteer armies and similar sources may suffer from serious selection problems. We now ask whether a closer look at the sources might have identified the problem. Note the central difficulty: we are asking whether we can identify selection bias from the selected alone. That said, we show in this section that there is discernible evidence of selection bias in at least one well-known study. Scholars rightly consider Floud et al (1990) a central contribution to the historical heights literature. Their discussion of living standards in Britain in the period 1750–1980 relies heavily on three distinct samples that are all "military." The Royal Army and the Royal Marines recruited adult men to serve in the forces. Floud and his co-authors also use samples of younger males to study age-patterns of growth. We use two of their "adult" data sources to show that the observed heights of soldiers in these sources are consistent with the importance of selection.[26]

In all models we have two types of regressors, which we call "cohort" and "current conditions".

[25] This information is quite useful in this static model of sector choice. In more realistic situations, when individuals have multiple opportunities to enlist, it should be also provide key information to help control for self-selection.

[26] The military data used in this section underlie the estimates reported by Floud et al (1990) and used by Weil (2007) and others.

The first is a series of dummy variables for the years in which the individuals were born. These variables should, under the "standard of living" interpretation, affect height, because they characterize the individual's experience as a child and young adult. Everyone born into similar socioeconomic circumstances in 1820 faced a similar biological standard of living.[27] The "current conditions" variables reflect not the standard of living for a person's birth-cohort, but conditions obtaining at the time *he joined the Army.* Consider two men, both born in 1820. One joins the Army in 1842 and other in 1843. The heights literature gives no reason to think the former should be taller or shorter, assuming both have stopped growing by the time they joined. We might think, however, that if economic conditions in 1843 were relatively bad, then men from relatively more privileged backgrounds will be more likely to join than were similar men in 1842. *If there is no selection on height, then no current-conditions variable should explain the heights of current Army recruits who have reached their final height.* The "current conditions" variables we employ here enter with signs that make sense in the way just outlined. But we should stress that *any* current-conditions variable that affects conditional mean height, whether positively or negatively, suggest a problem of selection. If the heights of current soldiers yield an unbiased estimate of the heights of the entire male population, then no current conditions variables should enter the regression with meaningful magnitude or statistical significance. On the other hand, the estimation discussed in this section does not reflect an adequate effort to model the selection problem. What follows are diagnostics, not models of the selection process.

We use two datasets pertaining to the Royal Army, which we call "Army" and "AMD" (Army Medical Department, whose reports underlie that source).[28] We experimented with a number of different measures of current economic conditions, and also with different cut-offs for both height and age. The age cut-off is important because younger men might still be growing when they enlist, and the height cut-off is important to be sure we are above h*, the minimum height standard for enlistment. (Floud et al (1990) do not always report the lower truncation point they use, but in one example (the birth cohort of 1806-1809), they use a truncation of 65 inches (Table 3.13)). We

[27]More precisely: in studies such as Floud et al (1990), the main variable affecting height is birthyear. In this case the birth cohort (b) is the only determinant of heights. For those enlisting at date t, in the absence of selection biases $E[h|a<h<c,b,z(t)]=f(a,c,b)$ and $P[(a<h<c|b,z(t)]=g(a,c,b)$. In other studies more information about the person's background is known and used. Appendix G provides a more formal justification for the diagnostics reported in this section.

[28]Appendix F discusses the sources, provides descriptive statistics, and also reports the full models presented in this section.

have direct reports on the minimum height from Spiers (1980) and the AMD reports, and, with the exception of 1861, the cut-offs were always 65 inches or less for the periods we use. Table 4 pertains to the Army data. Here we report the tests for the null hypothesis that different sets of variables have no effect on height. Appendix F reports the full models. All models have single-year birth-year dummies for all years of birth. Column (1) shows that we reject the null that ages do not affect height; they should not, in the absence of selection. This simple diagnostic performs the same way in all specifications in Table 4. In Column (2) we add a set of "macroeconomic" variables as well as the age dummies, and obtain an even stronger result. The macro variables are somewhat sensitive to cut-offs for age and height, at least for the Army data. The most flexible and robust approach is reported in Column (3): we add to the birth-year dummies a full set of dummies for the year of enlistment. This relaxes the implicit restriction imposed by the age dummies (that a given age has the same effect for all periods) and captures the same information as the macro variables, but more flexibly. Exclusion of the recruitment-year variables is rejected for all of the subsamples we consider. Table 5 reports a similar exercise for the AMD data using OLS models and the RSMLE model. Here the results are uniformly stronger, in part because the AMD data create much larger sample sizes. Appendix F discusses results for this approach using the RSMLE estimator.

7 Treatment of sample-selection bias in the historical heights literature

The scientific study of the human physical growth – auxology – emerged in the 1830s, with studies by European scientists, including Louis-René Villermé, Adolphe Quetelet and Eduoard Mallet, who gathered information on the heights of army recruits in France, Belgium and Switzerland respectively (Staub et al. 2011). Villermé drew a connection between height and health; Quetelet introduced the normal distribution to the study of practical scientific questions, including human growth; Mallet noticed a modest urban height advantage and attributed it to Geneva's relative prosperity. A half-century later Danson (1881) published his statistical study of English prisoners, which showed that males did not reach their terminal heights until after age 22. He concluded that armies should eschew 18-year old recruits because slightly older men who had already reached their terminal height would prove to be hardier soldiers. The thread that connects these studies is their belief that height reflected well being. The thread that connects them to the modern literature, besides their concerns with human height and well-being, is that they relied on readily available convenience samples subject to unknown selection. Selection was an issue at the dawn of statistical anthropometrics and, as we argue below, remains an underappreciated issue in the literature. In this section, we review how selection issues have shaped the discussion of three principal sources of height data: military recruits, slaves, and prisoners. With the exception of a fruitful, but underappreciated debate about height-based selection into the slave trade, concerns with selection and how it might affect the interpretation of results have not received sufficient attention.

7.1 Soldiers, military recruiting, and military school students

Fogel et al (1982, pp.29-30) were the first to call attention to the industrialization puzzle, namely that the positive correlation between height and per capita income observed in modern cross sectional studies did not hold in the time series for adult white males in the late antebellum United States. They postulated that the decline in heights was concentrated in the urban-born population. Rapid urbanization, poor public sanitation and more pronounced urban income inequality meant the conditions of city life deteriorated during industrialization. Urbanites were shorter as a consequence.

Several subsequent studies report remarkably similar patterns. Komlos (1989) reports a late eighteenth-century decline in the heights of adult male Hapsburg recruits. The mid-eighteenth century peak was not attained again for nearly 150 years. Floud et al (1990) find a general secular increase in English heights during the Industrial Revolution, with a sharp reversal for cohorts born in the 1840s and 1850s. Komlos (1993) reworks the English figures and argues that that English heights declined between 1770 and 1830. Subsequently Komlos and Küchenhoff (2012) push the onset of English height decline back to the 1750s, a secular decline that persists to 1850. A'Hearn (2003) reports a 3cm decline in Italian heights between 1740 and 1800, and attributes it to Malthusian pressures. The era of early industrialization was an era of shrinking men.

Studies of U.S. soldiers generate similar results. Komlos (1987) and Coclanis and Komlos (1995, fig. 3, p.101), use military school students to demonstrate long height cycles. The mean height of 19-year old West Point cadets falls by a half-inch in the late antebellum era, but recovers by the end of the Civil War. The mean height of 19-year old Citadel cadets is stable up to 1900 and increases by about 2.5 inches up to the 1930s.[29] Steckel (1995) links several independently constructed series of US soldier heights, from the French and Indian War (1754-1763) through the Second World War (1941-1945). His widely-reproduced diagram demonstrates a decline in terminal adult heights for cohorts born between the 1830s and the 1880s.

Once the industrialization puzzle was found in the time series, it was quickly uncovered in the cross-section as well. Komlos (1989) finds that Hapsburg Empire army recruits from the most economically developed regions within the empire were the shortest, while recruits from the least developed regions were among the tallest. Similar patterns emerged elsewhere in Europe. Mokyr and Ó Gráda (1996) report that poor Irish recruits into the English East Indian Company (EIC) army were taller than less poor English recruits. Moreover, Irish EIC recruits from relatively wealthy Ulster were shorter than recruits from elsewhere in Ireland. Sandberg and Steckel (1988) find that mid-nineteenth century Swedish soldiers from the less developed north and east were taller than recruits from the more developed west. Urban Italians paid a height "penalty" relative to their rural-born peers (A'Hearn 2003). In the United States, Union Army troops from less developed Kentucky and Tennessee were taller than troops from the Old Northwest who were taller than troops from industrializing New England (Johnson and Nicholas 1997, p. 208). Among Pennsylvanian recruits, men from less developed and less commercially-oriented regions were taller than men from the industrializing southeast region of the state (Cuff 2005, p.207). Margo and Steckel (1982) also found that ex-slave recruits into the Union Army from the less commercialized inland areas were taller than those from the more commercialized coastal regions.

Explanations of both the time-series and the cross-section puzzles build on those offered by Fogel et al (1982). The resolution of the industrialization puzzle offered in the literature focuses on the decline in net nutrition that occurred in the early stages of industrialization. According to this view, the underlying sources of decline were: (1) increasing income inequality; (2) increasing income variability; (3) increases in the price of food relative to manufactured goods; (4) increasing distance between the production and the consumption of food, with the consequent spoilage waste and loss of nutrients; (5) increased work effort; and (6) increased infection rates and disease incidence (Komlos and Coclanis 1997, p.455). The difficulty lies in accounting for many or most of these effects,

[29] Several studies attempt to estimate per capita food consumption (or disease incidence) and link it to birth cohort heights, but the exercise of recreating diets is fraught with assumptions, judgment calls and error (Floud et al. 1990).

which, in the end, are more commonly asserted than shown to be the cause of height trends.[30] Coclanis and Komlos (1995, p. 92), in fact, insist that any residual controversy concerning the industrialization puzzle centers on the "nature and causal connections of height cycles" because the existence of heights cycles, "is no longer questioned." We beg to differ. While we do not reject the possibility that heights cycled prior to the long secular increase in heights observed in the twentieth century in the developed world, we remain skeptical about the existence and existing interpretations of eighteenth- and nineteenth-century cycles because the literature has not taken the selection problem (as opposed to the truncation problem) seriously.

Much as the literature follows Fogel et al (1982, p.42-45) in documenting the puzzle; it follows them, too, in its discussion of selection issues. They write:

> Much of our work during the past four years has been devoted to assessing the quality of the data .... Volunteer armies, especially in peacetime, are selective in their admission criteria and often have minimum height requirements. Consequently, even if information on rejectees exists, there is the question of the extent to which applicants are self-screened... There is clearly evidence of self-selection bias in volunteer armies. Persons of foreign birth and from cities are overrepresented. Native-born individuals living in rural areas are underrepresented...

Their concerns with the nativity and urban-rural composition of their samples do not address height-based selection into the sample; rather it reflects selection on other observable characteristics within the sample. They then turn their attention to left-tail shortfall and their approach to correct for the nonnormality of the data that followed from the military's minimum height standard. More than two decades later, the discussion is hardly changed. After a brief discussion of existing methods for dealing with left-tail shortfall, A'Hearn (2003, p.359) turns his attention to selection:

> A second complication concerns selection effects. Whatever method is used [to correct for left-hand tail shortfall], the legitimacy of comparisons across groups depends on constant selectivity above the truncation point... [T]here is reason to worry about this

[30]Several studies attempt to estimate per capita food consumption (or disease incidence) and link it to birth cohort heights, but the exercise of recreating diets is fraught with assumptions, judgment calls and error (Floud et al. 1990) See Haines, Craig and Weiss (2003), Komlos (1987), and Bodenhorn (1999) for examples. Gallman (1996) and Haines, Craig and Weiss (2003) express skepticism concerning antebellum nutritional decline.

> for the sample at hand. In the eighteenth century, especially before 1770, it seems likely that recruiters did not get a random sample of individuals exceeding 63" in height...[But] it seems likely that selection disproportionately emphasizes heights in the immediate neighborhood of the statutory minimum.

From there, the discussion returns to the truncation issue, and A'Hearn lets pass an opportunity to explore the implications of selection on heights throughout the distribution of heights; instead, he focuses on selection in the immediate neighborhood of the minimum height standard.

Our survey of the literature uncovered only a few instances in which the selection problem discussed above is recognized and given careful consideration and explored in any meaningful way. Floud (1984, p.12) acknowledges that recruits into volunteer armies represented a self-selected sample that may not be representative of the population. He continues, however, that any selection is "unlikely" to be "large enough to vitiate comparisons over time and between... countries." Weir (1997, p.174-5) disagrees. He recognizes if recruiter selection manifests as a strict exclusion only of those below the truncation point, the methods, such as RSMLE and QBE may correct for it, but if selection is continuous across the entire distribution of heights, these estimators fail to generate accurate estimates of mean height. Meier(1982, p.297), in his comment on the original Wachter and Trussell (1982a) paper introducing the RSMLE and QBE estimators, pointed out that "recruitment effort" and "disparagement of shorter individuals" might vary continuously over a range of heights and could yield a selected military height distribution with a nearly normal shape. Wachter and Trussell (1982b, p.302), in their rejoinder, agreed that some scenarios for recruitment and volunteering could yield "spuriously normal observed distributions whose failure to represent the underlying population would be undetectable from internal evidence."

In his critique of Komlos' (1987) study of West Point cadets, Gallman (1996, p.194) not only refuted Komlos' claim of nutritional decline but raised serious concerns about selection. If Komlos' contention that average cadet height actually declined were true, it would only be interesting "if cadets can be taken to be a random sample of some larger, more interesting group – say all young white men in the United States. That, of course, cannot be." Not everyone was eligible for West Point and, of those who were, young men who attended were interested in either a military or engineering career. Despite the unrepresentative nature of the sample, Gallman (1996, p.194) notes that it might "still have wide meaning if the pool of candidates retained an unchanging character across the full period of this study – that is, if the same segments of the society continued to supply candidates, in roughly the same relative numbers. But is that likely to have been so?" Gallman observes that cohorts born after the early 1840s and interested in a military career would have faced the prospect of serving in an army with a large permanent class of lieutenants, captains and majors who had gained battlefield experience in the U.S. Civil War. This surely pushed many otherwise promising cadets into non-military pursuits.

A stylized fact of the European industrialization puzzle literature is Mokyr and Ó Gráda's (1994; 1996) finding that poor Irish recruits into the English East India Company (EIC) army were taller than similarly situated English recruits. A number of possible explanations have been offered for this counterintuitive insight, including the relatively nutritious Irish diet of milk and potatoes and "epidemiological isolation" (Nicholas and Steckel 1997). Mokyr and Ó Gráda (1994, p.42) offer an alternative explanation: because incomes were lower in Ireland than England, the relative quality of Irish recruits was higher. The poor Irish appear to be taller in the record than they really were because the EIC drew a larger proportion of recruits from better-off families. The taller Irish were not really taller and, therefore, biologically better off than the English. Faced with less attractive civilian employment opportunities for a given height, taller Irish men were disproportionately more willing to enlist than were English men. The tall Irish recruits, they note, were a supply-driven rather than a demand-driven selection phenomenon.

The remarkable feature is not that Irish men presenting themselves for service in the EIC were taller than their English counterparts; the remarkable feature is that the "tall-but-poor" Irish result is widely repeated in the literature absent Mokyr and Ó Gráda's repeated concern that the result is chimerical.[31] Two of the more outspoken critics of the military heights literature have been largely ignored. Ó Gráda (1996) argues that the English East Company's Irish military recruit data is subject to changing selection on height across the business cycle and with military events. When civilian labor demand was weak or when the EIC army was likely to see action, recruit heights fell

[31] Nicholas and Steckel (1997), Johnson and Nicholas (1997), Komlos (1998, p. 780), Deaton (2007), and Bozzoli, Deaton, and Quintano-Domeque (2009), among others, discuss the tall-but-poor Irish without referring to selection concerns. Mokyr and Ó Gráda (1994, p.50) also doubt the time series evidence. The 1810-1814 (birth) cohort of recruits was taller than the 1802-1809 cohort not because the biological standard of living had changed noticeably, but because labor market conditions changed enough for adults that the EIC had greater success attracting taller Irish recruits.

dramatically. Grubb (1999, p. 140) rightly notes that "heroic assumptions about the randomness in height" of unmeasured potential recruits and about the nature of self-selection into the military are necessary to draw inferences about the population from samples of volunteer armies.

Our reading of the historical heights literature that makes use of military samples, which represents a plurality, if not an outright majority of the studies, reveals that the potential for sample selection bias is underappreciated. When the issue does arise, it is given passing notice and attention turns to left-tail shortfall or truncation bias (see Komlos (2003, pp. 166-167) for a recent example). Given our theoretical and empirical discussion above, the failure to take sample selection bias seriously raises substantive concerns about the existence of and explanations for the industrialization puzzle.

7.2 Slavery and the slave trade

Given that one of the principal contributions of the modern heights literature is the demonstration that income and wealth inequality manifests itself in human height-at-age, it is not surprising that historians were intrigued by the anthropometric consequences of slavery. Relatively deprived children are consistently shorter at age than relatively well-off children, and it is hard to imagine a more potentially deprived population than slaves. Economic historians have produced several notable studies of slave heights, which have led to two general results. First, slave children were "extraordinarily small" (Steckel 1995, 1923). The mean height of slave children generally fell below the first percentile of modern stature until about their tenth year. These kinds of heights are sometimes observed in developing countries today, but are practically unheard of in the developed world. As Steckel (1987) observes, a modern American pediatrician would be alarmed if presented with such short children. Second, the growth of slaves in adolescence was remarkably vigorous; adult slaves attained nearly the 20th percentile of modern stature. Such recovery growth generated relatively tall adult slaves. Although they were shorter, by about one inch, than contemporary white Americans, measured adult slaves were taller than many contemporary European populations, especially low-income Europeans. But it is not clear whether the conclusions – recovery growth and relative well-being – are real or the consequence of selection into the sample.

Economic historians have constructed plausible explanations for these two features of slave heights. Poor medical knowledge and even poorer practice led to high infant mortality rates, which is probably indicative of high morbidity rates from both acute and endemic (gastrointestinal and diarrheal) infections.[32] Because diarrheal infection interferes with the nutrient-growth nexus, persistent endemic infection will lead to stunting. If nutrition is simultaneously low, the negative consequences on growth are magnified. Slave children, according to this literature, survived on a poor diet of hominy and pork fat. Growth recovery began around age 10 because the typical slave child entered the plantation labor force around that time. Normally, the demands of heavy work expected of slaves would have further interfered with growth, but once children entered the labor force they received shoes, which reduced the extent of fecal-based gastrointestinal infection, as well as more and better food, perhaps as much as one-half pound of pork per day. The increased meat ration for working slaves was further supplemented by vegetables and legumes, which contributed to the apparent catch-up growth of North American slaves.

Although information on US slave heights comes from a host of sources – "contraband" or runaway slaves that joined the Union Army in the 1860s (Margo and Steckel 1982), notarized certificates of good behavior filed with New Orleans' courts (Freudenberger and Pritchett 1991), manumission and freedom papers (Komlos 1992; Bodenhorn 2011), and runaway advertisements (Komlos 1994), among others – the principal source of information are the data recorded in the coastwise manifests filed by slave traders (Steckel 1979; Steckel 1987). To enforce the prohibition on international slave trading, slave traders moving slaves in the seaborne, interregional, domestic trade were required to provide shipping manifests that included the name, age, sex, color and height (in feet and inches) of the slave, as well as the name and residence of the slave trader. Tens of thousands of slaves were recorded on manifests.

The issue of whether selection bias affects the observed pattern of height-at-age and growth velocity is not discussed in either Steckel (1979) or Steckel (1987), but is given two paragraphs in Trussell and Steckel (1978, 550-551), who use the data to confirm their estimates (drawn from other

[32] Infant survival may create its own self-selection bias in that the ability to resist certain infections, which may manifest itself as vigorous growth capacity in adolescence, may be partly responsible for the observed pattern of slave growth. Instead of recognizing this possibility, vigorous adolescent growth is attributed to difficult to document changes in the slave child's work regimen, disease environment and food allotments. Rees et al (2003) develop a dynamic optimization slave owner model linking planter profit maximization with the pattern of slave growth. A related literature arose around the question of whether smallpox reduced the height of survivors. Voth and Leunig (1996) claim that the near eradication of small pox in nineteenth century England is responsible for about one-third of the reported increase in average heights between 1770 and 1873. Voth and Leunig's finding brings attention to potential survivor bias and the effects it may have on recorded heights over time. See Razzell (1998) and Oxley (2003) for critiques of the Voth-Leunig hypothesis.

sources) of age at menarche and age at first birth among slave women. They discuss two possible selection effects: a downward bias due to a lemon's market in slaves, in which only below-average height slaves entered the interregional trade; or, an Alchian-Allen (1964) "shipping the good apples out" market in which a fixed shipping cost represented a smaller fraction of the higher price received for taller slaves making taller slaves more likely to enter the interregional trade. They contend that age at peak growth velocity (onset of adolescent growth spurt) would not be much affected by either effect and do not further pursue the issue of potential sample-selection bias.

Higman's (1979; 1984) studies of Caribbean slaves provide some empirical evidence that selection may be driving Steckel's adolescent recovery finding. Unlike the heights of US slaves, which was recorded only if slaves entered into the coastwise trade, the British government required universal registration of slaves in anticipation of general emancipation. Registers of Trinidadian slaves were open for public inspection and government officials visited plantations to confirm their validity, making corrections when necessary. Caribbean slaves demonstrate less adolescent recovery. Height differences between US and Caribbean slaves may be due, of course, to differences in genetic potential, disease incidence, work load or diet (Steckel (1995, p. 1925), in fact, attributes them to fetal-alcohol syndrome due to substantial and growth-retarding rum allowances on Caribbean plantations). But we need not turn to difficult-to-prove conjectures when the differences were quite possibly driven by height-based selection into the US coastwise-trade sample and near-universal coverage in the Caribbean sample.

Freudenberger and Pritchett (1991), Pritchett and Freudenberger (1992), and Pritchett and Chamberlain (1993) systematically explore the sample-selection bias problem of the coastwise manifest sample. They argue that the manifest sample is subject to substantial selection on height of the "shipping the good apples out" type. That is, when a fixed transportation cost is applied to similar goods, the high-quality, high-priced good (a taller slave in this instance) becomes relatively less expensive in the destination market. Four features are needed for the good-apples effect to hold: (1) transport costs must be non-negligible; (2) transportation costs are not proportional to price at the source; (3) the goods are close, but not perfect substitutes; and (4), the elasticity of substitution between each of the two goods in question and a composite third good (say, free or indentured labor) not be substantially different (Borcherding and Silberberg 1978). Pritchett (1997) and his coauthors provide evidence on transportation costs that supports conditions (1) and (2). Conditions (3) and

(4), they contend, are defensible: tall and short slaves are (imperfect) substitutes in production; and free or indentured labor, as the historical record shows, were substitutable for slaves, whether tall or short.[33]

Moreover, Pritchett and his coauthors show that the effect will be more pronounced for younger than adult slaves. One prediction of the good-apples slave model is that prices will be a positive, but declining function of age (Pritchett 1997, p.73 note). Because buyers were willing to pay more for taller slaves, traders faced incentives to select taller slaves for the coastwise market. The incentives to select taller slavers for the coastwise market were also stronger, the longer the slave's expected lifetime – that is, the younger the slave. Calomiris and Pritchett (2009) find that slave children shipped with their mothers were shorter than children of the same age shipped alone. This result is puzzling absent selection on height. The available evidence is consistent with the model's prediction; the height differential between slaves traded with their mother and those traded without their mother were most pronounced in childhood, declining in adolescence and largely disappearing among adults.

The conclusion to draw from the exchange between Pritchett and his coauthors and Komlos and Alecke (1996) is that "slaves shipped coastwise were not representative of the general population" of North American slaves (Pritchett 1997, p. 83). In this instance, the consequence of sample-selection bias works in favor of one of existing interpretations. Pritchett and his co-authors argued that because selection on height was stronger for younger slaves than for adult slaves, Steckel's coastwise sample includes proportionately more tall children than tall adults than would be observed in a true random draw of the general slave population. This means that the much-discussed recovery growth of adolescent slaves is underestimated. It is likely that young slaves were even more pathologically short than they appear in the manifest sample.

Unfortunately, we cannot provide as sanguine a conclusion about the second result of the slave height literature. Pritchett's statement of the good-apple model predicts that selection will be less pronounced at older ages and lowest among adults, but it does not predict the absence of selection. Higman's (1979) study of Caribbean slaves, in fact, provides some evidence of selection on height for inter-island trade. The mean height of native-born Trinidadian adult males (25-40 years) measured

[33] See Komlos and Alecke's (1996) response to Pritchett's evidence. The substitutability of slave and indentured or free labor is discussed in Galenson (1981) and Grubb (1994; 2001).

in 1813 was 165.6cm. The mean height of Creole male slaves, or those born in the New World, and imported into Trinidad from sugar islands was significantly taller at 167.3cm (t=2.15). Creole slaves imported from non-sugar islands were taller still at 170.6cm (t=4.42). Imported females, too, were significantly taller than native-born Trinidadian female slaves.

The available evidence, although not definitive, points toward good-apples selection in the interregional slave trade. This type of selection need not be revealed by (non)normality of height distributions, or heaping on height or age, or left-tail shortfall. That the selection process is not readily revealed does not, of course, imply that it is unimportant. Unrepresentative selected samples will yield incorrect inferences when selection is correlated with the variable of interest. Pritchett's concerns with slave samples are well-founded, and it is unfortunate that his contributions have not had more influence on the literature.

7.3 Criminals and prisoners

The anthropometrician's concern with the condition of the working classes during industrialization, in combination with the wide availability of data, led scholars to study the heights of incarcerated or transported criminals. Scholars making use of prisoner data readily acknowledge that prisoner samples are not representative of the underlying population; the professional and middle classes and farmers are underrepresented, and unskilled laborers and other low-wage groups are overrepresented (Nicholas and Steckel 1991; Riggs 1994). In one of a series of studies of nineteenth-century US prisoners, Carson (2008, p.591) attempts to recast the selection vice as a virtue:

> The prison data probably selected many of the materially poorest individuals, although there are skilled and agricultural workers in the sample. While prison records are not random, the selectivity they represent has its own advantages in stature studies, such as being drawn from lower socioeconomic groups, who were more vulnerable to economic change. For the study of height as an indicator of biological variation, this kind of selection is preferable to that which marks many military records – minimum height requirements.

In other words, prisoner heights are preferable to soldier heights because prisoners are not subject to left-tail shortfall.

Although Carson labors to recast a vice as a virtue, it remains a vice if the purpose of the historical anthropometric literature is to speak to the average standard of living. Several studies (which are not without their own selection biases) have shown that the heights of the middle and upper classes in the nineteenth-century United States did not mirror the height cycles of the lower classes (Sunder 2004; Sunder 2011; Lang and Sunder 2003). Thus, if lower-class heights are more responsive to economic changes than middle- and upper-class heights, samples in which the lower classes are overrepresented are likely to overstate any real changes in population heights. Moreover, inferences about changes in the average biological standard of living may reflect changes in the sample composition, which may be pronounced over the business cycle or during long-run secular changes in rates of economic growth. Using a small subset of the available samples, Sunder and Woitek (2005), in fact, show that co-movements between economic and height cycles are most pronounced for American blacks (predominantly unskilled labor), relatively pronounced for Austrian soldiers and Ohio National Guardsmen (relative overrepresentation of laboring classes), less pronounced for late nineteenth-century registered voters, and virtually nonexistent for middle- and upper-class women.

The potential for compositional changes in the heights of individuals selected into prisons to drive estimated temporal height changes has not gone unnoticed. Nicholas and Oxley (1996), Johnson and Nicholas (1997) and others have considered how the relative under- or over-representation of certain groups, which may change over time, may influence estimated versus the true changes in population heights and biological well-being. But their discussions often fail to appreciate the subtleties of selection on height. Nicholas and Steckel (1991, p.949), for example, offer the following:

> Although there is no single powerful test of selectivity bias, the various tests reported hereafter put at risk the hypothesis that the decline in heights was an artifact of our data. Almost one-third of our convicts were tried in a county other than that in which they were born, but it is not possible to know whether the move occurred during childhood, adolescence, or after maturity. Assuming that all convicts tried in the same county as their place of birth were nonmovers, five-year moving averages of height for rural- and urban-born nonmovers showed the same profile as that for the entire sample.

This argument is irrelevant to the issue of selection into their sample; it merely shows that there is no apparent selection into mover and non-mover groups. The issue is whether cohorts, in response to changing economic conditions, differentially selected into criminal activity and were apprehended, tried, convicted, and imprisoned conditional on characteristics correlated with height. Evidence from nineteenth-century US prisons shows that prisoners were short compared even to soldiers (Bodenhorn, Moehling and Price 2012). The issue is whether selection into prison, conditional on selecting into crime, followed from height, which seems likely. Persico, Postlewaite and Silverman (2005) and Case and Paxson (2008) find evidence that height is positively correlated with labor market outcomes (employment and wages), mediated through cognitive abilities and the accumulation of more human capital by taller youth and adolescents. Because taller individuals face relatively better legitimate labor market opportunities than shorter individuals, criminal activities are less attractive to taller people. Prisons are thus populated by short, relatively low-income people.

Researchers using prison samples take solace in the fact that their samples are Gaussian but, as we have demonstrated above, apparent normality alone need not reveal selection even in the presence of selection. The issue that concerns us is whether selection on height was likely to change in a way that would generate the counterintuitive results reported in the historical heights literature, whether those changes are evident in the research, and whether the explanations are consistent. The Roy-type model developed above predicts that better economic conditions are consistent with shorter prisoners in both the cross section and the time series. If legitimate labor market opportunities are conditioned, at least in part, on height, economic expansions will create more and better legitimate opportunities and short individuals who may have selected into crime in bad times will select into legitimate activities in good times. While criminals will exhibit a distribution of heights, they will be disproportionately drawn from the left-hand tail of the population distribution and even more so as economic conditions improve. If we use the heights of prisoners as indicators of the biological standard of living, it will appear that biological times are tough when economic times are good and vice-versa. In fact, what we are observing in this instance is differential selection on height into the subset of the criminal class that gets caught and convicted.

Consider Riggs's (1994) study of Scottish prisoners. He claims the sample is representative of Scottish working-class heights because "in a society of heavy drinkers ... many workers were at risk of being arrested and thus having their physical stature preserved in the historical record" (Riggs

1994, p.64). While his results are in general agreement with the industrialization puzzle, he finds the "curious" result that the heights of those arrested in the 1840s actually increased markedly. Riggs (1994, p.70) considers this curious because the 1840s, known as the "hungry forties," were "notorious for hardship and hunger." When considered in light of our Roy model, the result is not so curious. If the effect of food shortages and unemployment was that men moved from legitimate to criminal activities, the deterioration in legitimate opportunities would draw differentially more men into crime from the right-hand than the left-hand tail of the height distribution because men in the left-hand tail were already disproportionately in the criminal market prior to the downturn. Moreover, if right-tail entrants into the crime market have relatively little criminal human capital, they were probably more likely to have been apprehended and imprisoned. Thus, heights apparently increase during a sharp economic downturn when the "intuitive" connection between the biological and economic standard of living would suggest otherwise.

In a series of studies, Nicholas and co-authors study the heights of late-eighteenth and early nineteenth-century English and Irish prisoners. They acknowledge that their sample is unrepresentative in that the lower classes and the young are overrepresented, but one feature recurs: the Irish are taller than the English, a feature discussed earlier in reference to military heights. Nicholas and Steckel (1991) find that male Irish convicts are taller than male English convicts. Nicholas and Oxley (1996) find that the heights of rural-born English convicts decline while the heights of Irish-born women rise. Johnson and Nicholas (1997) use the Irish as a "control" group against which to compare English women because no Industrial Revolution occurs in Ireland. Yet, Irish women are tall relative to the English. They are "confident that there are no obvious selection biases..." that would generate their results. Their contention is puzzling in that they cite and discuss Mokyr and Ó Gráda's (1994; 1996) result without acknowledging the Mokyr and Ó Gráda conjecture that the Irish height advantage results from the army being more attractive to taller people in a poor economy. Is it not also likely that criminal activity is relatively more attractive to taller people in a poor economy?

Finally, studies of American prisoners yield results similarly at odds with the industrialization puzzle. Komlos and Coclanis (1997) find that the heights of black men born into slavery did not decline in the 1830s and 1840s; Sunder (2004) reports that heights of Tennessee prisoners were stable in the late-antebellum era; Carson (2008) finds no evidence of the antebellum puzzle among

Missouri's prisoners. These studies then attempt to reconcile stable or rising stature with the antebellum puzzle. Komlos and Coclanis (1997, p.452) contend that slaves were insulated from the market and "ate all they were allotted;" Sunder (2004) argues that Tennesseans raised hogs and did not sell meat in the interregional market so that they had plentiful access to protein; and Carson's (2008, p.603) explanation turns on the relatively tall stature of the low-income Ozark Missourians due, in large part, to their reliance on dairy (a variation of the tall-but-poor Irish explanation), whereas wealthier northern Missourians grew less protein-rich grains.[34] Alternatively, our Roy model predicts a south to north Missouri prisoner height gradient that is consistent with a north to south wealth gradient. The relatively less attractive legitimate market opportunities in the Ozarks drew relatively taller men into the criminal market; better opportunities in northern Missouri drew men of comparable height, conditional on other characteristics, into the legitimate labor market.

We acknowledge that we have offered selection-based criticisms of several prominent studies without returning to and using our diagnostic test to analyze the data underlying those studies to determine, to the extent we could, whether selection drives the results. In future work, we will explore the consequences of potential selection bias in some commonly used data sets. The purpose of the review is to show how a Roy model provides a set of internally consistent predictions rather than ad hoc reconciliations of evidence inconsistent with the industrialization puzzle. Military and prisoner heights rise in bad economic times because alternative employments – military service and criminal activity – become relatively more attractive to the legitimate civilian labor market. Some tall men who would find remunerative legitimate civilian employment in a good economy turn to the military or crime in a bad one. Alternatively, some short men who would not have attractive legitimate market opportunities in bad times will be drawn into the civilian market in good times. The result is an observed, but not actual countercyclical pattern of biological well being.

8 Conclusions

The heights literature is now a large and important component of economic history and development economics. We agree that the central issues of interest in this literature, long-term changes in the

[34] Rees et al (2003) provide a dynamic optimization model of slave owner behavior consistent with the "remarkable" catch-up growth uncovered by Steckel (1986; 1987), which offers some insight into care and feeding of slaves in response to changes in the market price of slaves. The Rees et al (2003) model presumes substantial market-oriented responses rather than "insulation" from the market.

standard of living, should occupy a prominent place in the agenda of economists. Unfortunately, much of the current literature that uses heights as an indicator of living standards probably suffers from selection bias. Most of these sources are selected in some way, and the biases that arise because of the selection issues are pernicious and quite possibly could account for many of the interesting "facts" the heights literature claims to have identified.

Our message may seem entirely negative, but that is not how we intend it. First, we do not advocate rejection of the heights literature, its themes or basic approaches. We do think that taking selection into account will require tempering of some of the broader claims made. We also hope that attention to selection will guide the literature away from complicated, ad-hoc explanations of surprising findings; as we note, some of these might reflect little more than selection. Second, and more importantly, we hope that careful attention to selection issues will allow these sources to speak in new and interesting ways. The nature of selection into the British Army (for example) is a statistical problem for those who want to use soldiers' heights to infer changes in the standard of living standards during the Industrial Revolution. But the selection process itself can yield rich insights into the nature of labor markets at the time: what led some young men into the Army, and what does this tell us about working-class life and the opportunities offered by an industrializing economy? Taking selection seriously, instead of assuming it away, promises yet more insights from an already mature literature.

Appendix A: The military's decision to accept a recruit

The occupational choice model developed in Section 3 implicitly assumes that the Army accepts everyone who wants to volunteer. We know this assumption to be false; the most obvious restriction was the minimum height requirement that generates the need for estimators such as the RSMLE and the QBE. Here we extend our model to allow the military, too, to be an optimizing agent, one that can reject for military service some individuals who volunteer.

We assume the Army produces a service called "security" using a constant-returns-to-scale production function that depends on the sum of security services provided by all current soldiers.[35] Each individual soldier produces a security service that is a function of his height h and a set of military abilities ε_M. Again, we are agnostic on whether height *per se* makes a better soldier, or whether height is simply correlated with characteristics that make men better soldiers. The Army pays each soldier in the manner described earlier. The Army must also pay additional fixed costs for each serving soldier over and above what that man receives as a "wage." These costs include uniforms and training. The net security value (security value minus costs of hiring) of a soldier with height h and military abilities ε_M is $S(h, \varepsilon_M)$. One would expect $S(h, \varepsilon_M)$ to be positive, but one can imagine soldiers who cost more than they are worth.

Now abstract from our log-linear formulation of the utility of being a civilian and the utility of being a solider. The civilian economy values these same two traits along with a third trait that is specific to the civilian sector, ε_C. Civilian wages are given by

$$w_C(h, \varepsilon_C, \varepsilon_M). \tag{11}$$

Suppose the military can pay soldiers a wage that is a function of all three characteristics. Assuming the military can observe h, ε_C, and ε_M, let

$$w_{M,h,C,M}(h, \varepsilon_C, \varepsilon_M) \tag{12}$$

be the wage that the military pays a soldier. We continue to assume that each individual has

[35] In our model, different types of soldiers can produce different amounts of security, but there are no externalities or spillovers; a good (bad) soldier does not increase (decrease) the effectiveness of other soldiers.

preference parameters for the civilian and military life τ_C and τ_M.

This implies that a man joins the Army if

$$w_{M,h,C,M}\left(h,\varepsilon_C,\varepsilon_M\right)+\tau_M \geq w_C\left(h,\varepsilon_C,\varepsilon_M\right)+\tau_C \tag{13}$$

or

$$\tau \leq w_{M,h,C,M}\left(h,\varepsilon_C,\varepsilon_M\right)-w_C\left(h,\varepsilon_C,\varepsilon_M\right) \tag{14}$$

where $\tau = \tau_C - \tau_M$.

Let $f_{h,M,C,\tau}\left(h,\varepsilon_M,\varepsilon_C,\tau\right)$ be the joint distribution of the three characteristics that can be rewarded, as well as relative tastes, and assume that the military minimizes the cost of producing any particular level of total security. The total population eligible for military service is N. The military's objective is to choose a wage payment function $w_{M,h,C,M}\left(h,\varepsilon_C,\varepsilon_M\right)$ to minimize the cost of producing a given amount of security S, or

$$\begin{aligned}&\min_{w_{M,h,C,M}(h,\varepsilon_C,\varepsilon_M)} Cost = \\ &N\int_0^\infty\int_{-\infty}^\infty\int_{-\infty}^\infty\int_{-\infty}^{w_{M,h,C,M}(h,\varepsilon_C,\varepsilon_M)-w_C(h,\varepsilon_C,\varepsilon_M)} w_{M,h,C,M}\left(h,\varepsilon_C,\varepsilon_M\right) f_{h,M,C,\tau}\left(h,\varepsilon_M,\varepsilon_C,\tau\right) d\tau d\varepsilon_C d\varepsilon_M dh\end{aligned} \tag{15}$$

subject to providing some specific level of security

$$S = N\int_0^\infty\int_{-\infty}^\infty S\left(h,\varepsilon_M\right)\int_{-\infty}^\infty\int_{-\infty}^{w_{M,h,C,M}(h,\varepsilon_C,\varepsilon_M)-w_C(h,\varepsilon_C,\varepsilon_M)} f_{h,M,C,\tau}\left(h,\varepsilon_M,\varepsilon_C,\tau\right) d\tau d\varepsilon_C d\varepsilon_M dh \tag{16}$$

Let λ be the Lagrange multiplier corresponding to the security constraint. Assuming interior solutions, for each combination of $(h,\varepsilon_C,\varepsilon_M)$ the military will set the wage, $w_{M,h,C,M}\left(h,\varepsilon_C,\varepsilon_M\right)$, such that

$$\begin{aligned} & N\int_{-\infty}^{w_{M,h,C,M}(h,\varepsilon_C,\varepsilon_M)-w_C(h,\varepsilon_C,\varepsilon_M)} f_{h,M,C,\tau}\left(h,\varepsilon_M,\varepsilon_C,\tau\right) d\tau \\ + \quad & N w_{M,h,C,M}\left(h,\varepsilon_C,\varepsilon_M\right) f_{h,M,C,\tau}\left(h,\varepsilon_M,\varepsilon_C,w_{M,h,C,M}\left(h,\varepsilon_C,\varepsilon_M\right)-w_C\left(h,\varepsilon_C,\varepsilon_M\right)\right) \\ = \quad & \lambda N S\left(h,\varepsilon_M\right) f_{h,M,C,\tau}\left(h,\varepsilon_M,\varepsilon_C,w_{M,h,C,M}\left(h,\varepsilon_C,\varepsilon_M\right)-w_C\left(h,\varepsilon_C,\varepsilon_M\right)\right).\end{aligned} \tag{17}$$

The left hand side of this expression is the marginal cost of paying soldiers with productive traits equal to $(h, \varepsilon_C, \varepsilon_M)$ each one more penny; it is the sum the Army must pay each soldier already in the military (conditional on having this particular set of traits) plus the cost of expanding the Army, that is, paying for additional soldiers with those same traits. The right hand side of the expression measures the marginal benefit of the increased security provided by the additional soldiers (with these traits) induced into the military by the higher wage. For interior solutions, the ratio of the marginal costs of soldiers with different sets of characteristics should equal the ratio of their marginal benefits.

This formulation implies that the military can extract some surplus from those enlisting in the military. That is, the optimal military function may depend on observed characteristics that have no impact on a potential recruit's ability to provide military services. To see this, consider two sets of individuals whose military-relevant characteristics are identical, but who differ in characteristics valued by civilian jobs, ε_C. Assume the military enlists some positive fraction from both groups. Because they differ in their ε_C values, these groups have different potential earnings in the civilian sector. If they are equally represented in the population, then the first order conditions will be satisfied as long as the military keeps the military-civilian pay differential constant for the two groups. The military will pay the group with the lower valued civilian sector trait, ε_C, less than it would pay enlistees from the group with the civilian sector trait that is more highly valued, even though this trait has no impact on their performance as a soldier.

For a variety of reasons, the military may not be able to use such fine details to make pay offers. Suppose, for the moment, that the military can only observe height, h, and so it can only base its pay policy on this one characteristic. In this case $w_{M,h,C,M}(h, \varepsilon_C, \varepsilon_M) = w_{M,h}(h)$, and the first order conditions become

$$\begin{aligned}
& N \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{w_{M,h}(h) - w_C(h, \varepsilon_C, \varepsilon_M)} f_{h,M,C,\tau}(h, \varepsilon_M, \varepsilon_C, \tau)\, d\tau d\varepsilon_C d\varepsilon_M \\
+ \quad & N w_{M,h}(h) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{h,M,C,\tau}(h, \varepsilon_M, \varepsilon_C, w_{M,h}(h) - w_C(h, \varepsilon_C, \varepsilon_M))\, d\varepsilon_C d\varepsilon_M \\
= \quad & \lambda N \int_{-\infty}^{\infty} S(h, \varepsilon_M) \int_{-\infty}^{\infty} f_{h,M,C,\tau}(h, \varepsilon_M, \varepsilon_C, w_{M,h}(h) - w_C(h, \varepsilon_C, \varepsilon_M))\, d\varepsilon_C d\varepsilon_M. \qquad (18)
\end{aligned}$$

The left hand side of this first order condition is the expected cost of increasing the military pay by one penny for all people with height h. It consists of two components. The first is the additional penny paid to all individuals of height h who would have joined the military even without the higher pay. The second term is the expected full cost of the marginal enlistees with height h. The right hand side again measures the expected marginal benefit from the new enlistees with height h who enter the military because of this increased level of pay.

Note that both of these military pay formulations place no *a priori* restrictions on the sign or magnitude of the military pay for any observed set of characteristics. For some characteristics, $(h, \varepsilon_C, \varepsilon_M)$, $S(h, \varepsilon_M)$, could be negative. This would imply that the recruit has to pay the Army for the privilege of being a soldier. We know of no such practice and think it is a principal reason for minimum height requirements.

More realistically, military compensation schedules during the eighteenth and nineteenth centuries usually did not include explicit height gradients. But they usually had a minimum height requirement in addition to a fixed level of compensation. To incorporate these features into the cost minimization model, we allow the military to choose a non-varying level of pay and to restrict military service to those above some optimally-chosen minimum height threshold D. Now the military's cost minimization problem can be written as

$$\min_{w_M, D} Cost = w_M N \int_D^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{w_M - w_C(h, \varepsilon_C, \varepsilon_M)} f_{h,M,C,\tau}(h, \varepsilon_M, \varepsilon_C, \tau)\, d\tau d\varepsilon_C d\varepsilon_M dh \qquad (19)$$

subject to the specified security level

$$S = N \int_D^{\infty} \int_{-\infty}^{\infty} S(h, \varepsilon_M) \int_{-\infty}^{\infty} \int_{-\infty}^{w_M - w_C(h, \varepsilon_C, \varepsilon_M)} f_{h,M,C,\tau}(h, \varepsilon_M, \varepsilon_C, \tau)\, d\tau d\varepsilon_C d\varepsilon_M dh. \qquad (20)$$

The two first order conditions are, in this instance

$$\begin{aligned} & N \int_D^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{w_M - w_C(h, \varepsilon_C, \varepsilon_M)} f_{h,M,C,\tau}(h, \varepsilon_M, \varepsilon_C, \tau)\, d\tau d\varepsilon_C d\varepsilon_M dh \\ +N \quad & w_M \int_D^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{h,M,C,\tau}(h, \varepsilon_M, \varepsilon_C, w_M - w_C(h, \varepsilon_C, \varepsilon_M))\, d\varepsilon_C d\varepsilon_M dh \\ = \quad & \lambda N \int_D^{\infty} \int_{-\infty}^{\infty} S(h, \varepsilon_M) \int_{-\infty}^{\infty} f_{h,M,C,\tau}(h, \varepsilon_M, \varepsilon_C, w_M - w_C(h, \varepsilon_C, \varepsilon_M))\, d\varepsilon_C d\varepsilon_M dh \end{aligned} \qquad (21)$$

and

$$\begin{aligned} & -w_M N \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{w_M - w_C(D,\varepsilon_C,\varepsilon_M)} f_{h,M,C,\tau}\left(D,\varepsilon_M,\varepsilon_C,\tau\right) d\tau d\varepsilon_C d\varepsilon_M \\ = \; & -\lambda N \int_{-\infty}^{\infty} S\left(D,\varepsilon_M\right) \int_{-\infty}^{\infty} \int_{-\infty}^{w_M - w_C(D,\varepsilon_C,\varepsilon_M)} f_{h,M,C,\tau}\left(D,\varepsilon_M,\varepsilon_C,\tau\right) d\tau d\varepsilon_C d\varepsilon_M . \qquad (22) \end{aligned}$$

Consider the first of the two first-order conditions (21). The first term on the left hand side is the marginal cost of paying each soldier not at the margin of joining the military an extra penny, and the second term in that left hand side is the full cost of paying the marginal soldier who enters the military at the wage w_M. Their sum is the cost of inducing an additional "expected" soldier into the military at this wage. The right-hand side of that first first-order condition is the expected marginal value of the security services provided by the marginal soldier induced into the military at this wage.

The second first-order condition (22)describes how the minimum height threshold is set when the optimal military wage is given. The left hand side is the expected cost savings gained by requiring the marginal soldier to be one inch taller, and the right hand side is the loss in security from excluding this marginal soldier from the military. Rearranging the second first order condition yields:

$$\int_{-\infty}^{\infty} \left[\lambda S\left(D,\varepsilon_M\right) - w_M\right] \int_{-\infty}^{\infty} \int_{-\infty}^{w_M - w_C(D,\varepsilon_C,\varepsilon_M)} f_{h,M,C,\tau}\left(D,\varepsilon_M,\varepsilon_C,\tau\right) d\tau d\varepsilon_C d\varepsilon_M = 0. \qquad (23)$$

The military will set the minimum height requirement to the level where the expected value of the military service for a soldier of height D, given the military wage and the occupational choice decision, just equals the expected wage payments to the individuals of minimum height D choosing to enter the military. Note that this truncation argument reflects the Army's limited information on potential enlistees. If the Army could observe $S(\cdot)$ perfectly it would not need to have a the minimum height standard, at least for this reason.

We observe changes over time in the mean height of soldiers as the Army expands and contracts. This fluctuation in heights is evidence of selection and can be derived from our choice model. Changes in the demand for military services due to the outbreak of war increase the Army's target level of security, S. Since the military is a cost minimizer, one would expect an increase in S to result in a higher marginal cost of military service, so the multiplier λ would increase. This would result in the right hand side of the first order conditions to increase, *ceteris paribus*, and this could be offset by the military increasing its pay offer and relaxing the minimum height requirement, thereby increasing the size of the military and the level of military services provided in these more demanding times.

Appendix B: Calibration of the simulations

The simulation exercises reported in Section 4 show that moderate values of differences in returns to height ($\beta_c - \beta_{m)}$ imply the selection problems that are the subject of our paper. Like any simulation exercise, ours requires care that we are not generating implausible results by inappropriate choice of parameter values. This appendix discusses our calibration strategy, showing that the message of our simulation comes through for most reasonable sets of values. We approach this problem in three ways. First, where possible, we try to match empirical moments. The most important is the fraction of men in a cohort who join Army; Floud et al (1990, Table 2.8) report values that range from 6.4 to 17 percent of a cohort of 18 year-olds for the period 1862-1902. We keep the simulation parameters to values that imply this range. We assume $\alpha_C \geq \alpha_M$; this restriction implies that *ceteris paribus*, men prefer the civilian sector or are indifferent between the two sectors at short heights and/or absent any height-reward differential. Second, economic logic implies some restrictions on parameters. For example, the way militaries determined pay and the operation of civilian labor markets suggests sd(ε_C) $\geq$ sd(ε_M). Similarly, we assume Corr($\varepsilon_M, \varepsilon_C$) ≥ 0, as the military and civilian worlds would value some traits in common.We also limit our attention to cases where the military return to height does not exceed the civilian return ($\beta_C \geq \beta_M$).

The Army's relative size and these parameter restrictions provide considerable guidance, but not enough. We further focus our attention by considering the economic implications of the relationship between the deterministic and stochastic parameters. The decision to join the Army reflects the *difference* between the military and civilian log-wage. Define D as the difference between the military and civilian valuations for an individual of a given height and realization of the ε terms. We can write:

$$D = w_m - w_c = \alpha_M + \beta_M(h - h_L) + \varepsilon_M - \alpha_C - \beta_C(h - h_L) - \varepsilon_C \tag{24}$$

where h_L=56, the value used to normalize heights in Section 3. We can write the of variance of D as:

$$Var(D) = Var(\alpha_M - \alpha_C + (\beta_M - \beta_C) * (h - h_L) + \varepsilon_M - \varepsilon_C) \tag{25}$$

Since the α terms are constants and we assume height is independent of the ε terms, (25) can be rewritten

$$Var(D) = (\beta_M - \beta_C)^2 Var(h) + Var(\varepsilon_M) + Var(\varepsilon_C) - 2 * Corr(\varepsilon_M, \varepsilon_C) * sd(\varepsilon_M) sd(\varepsilon_C) \quad (26)$$

Analogous to a linear regression equation, we can define the R^2 for this equation as the proportion of the variance in D explained by height. This R^2 provides an intuitive check on our parameter-value selections; very high values of R^2 are implausible, as one woud not expect differences in height to explain most of the difference between military and civilian wages. On the other hand, if R^2 is close to zero, then we are assuming parameter values that cannot imply much selection, because with those parameter values, height has little role in the decisions to join the Army. Since the variance of the explained portion of Var(D) is $(\beta_M - \beta_C)^2 Var(h)$,

$$R^2 = \left[\frac{(\beta_M - \beta_C)^2 Var(h)}{(\beta_M - \beta_C)^2 Var(h) + Var(\varepsilon_M) + Var(\varepsilon_C) - 2 * Corr(\varepsilon_M, \varepsilon_C) * sd(\varepsilon_M) sd(\varepsilon_C)} \right] \quad (27)$$

The important quantities for our simulation exercises are the variances and standard deviations of ε_M and ε_C as well as their correlation, as well as their differences in returns to height (β). (Our value for the variance of the *population* height comes from the estimates reported in the literature; because Floud et al (1990) report standard deviations in the range of 2.5 inches, we assume a value of 6.25 for the variance.)

Table B.1 reports R^2 values computed from (28) for a number of scenarios that fit the criteria outlined above. Example 6 corresponds to the "baseline" used in Sections 3 and 4 above. Examples 1-5 and 7-10 were constructed to illustrate the implications of varying the relevant parameters. Figures B.1-B.3 graph five of these examples, illustrating graphically the implications of differing parameter values. Figure B.1 shows that doubling Corr($\varepsilon_M, \varepsilon_C$) does not dramatically alter our results; the slightly higher R^2 in Example 6 implies only slightly stronger selection in the form of a steeper Z(h). Figure B.2 illustrates the role of changing β. Values that imply $\beta_C - \beta_M = .06$ (as in Example 10) reach the upper limit of plausibility; given the other parameter values, that example implies that height differences can explain more than one-quarter of the variation in returns to the military and civilian sectors. But Example 3, which assumes the very modest $\beta_M - \beta_C$=.02, still implies an Army almost a full inch shorter than the population.

Note the shape of Z(h) in Example 3: its slope is very shallow, yet even that degree of selection produces the differences in mean heights. Example 3 implies an R^2 of .038. We do not need to assume that heights explain much of the decision to join the Army for the selection process to matter. Figure B.3 shows the effect of doubling $sd(\varepsilon_C)$. The implied R^2 in Example 5 is less than half that of Example 6, and the mean military height in the first case is 1.3 inches taller than in the second. This is a significant change when the population standard deviation is 2.5 inches. On the other hand, even the parameter values used in Example 5 imply considerable selection; Z(h) declines from about 5 percent at a height of 56 inches to nearly zero by 72 inches, and the implied military height is a full inch shorter than in the population.

Appendix C: Alternative tests of normality

Section 5 uses two common tests of normality to show that such tests cannot effectively detect sample-selection. This appendix extends the results reported there to consider a wider range of tests for normality. We report tests for two situations. We first construct an extreme example of selection, assuming $\beta_C - \beta_M = 0.1$ and $\alpha_C - \alpha_M$ =-0.3. Thus 7 percent of men join the military, yielding a mean military height of 63.47 inches when the population mean is 66 inches. These parameter values imply that nearly all shorter men will prefer the military, while few taller men will. Our second example has more modest selection, as in the text and Appendix B. We assume $\beta_C - \beta_M = 0.04$, and $\alpha_C - \alpha_M = 0.1$. The military has 11 percent of the population and the mean military height is 65 inches.[36] Figures C.1 and C.2 report results. For each example, we use expected sample sizes ranging up to 50,000 persons, and report results for two different test sizes. In Figure C.1, which assumes extreme selection, we see that virtually all of the tests have rejection probability no larger than the test size for samples of fewer than about 2000 individuals, and even for the larger samples, the rejection probability remains modest. A sample of 2000 is not unknown in the heights literature, but such a large number is not common. Figure C.1 simply reinforces the text Section 5: researchers working in this literature should not think that standard normality tests will help them detect selection bias in historical height samples. Even when that bias is extreme, the resulting height distributions can be apparently normal. Figure C.2, which reports results for the case of less extreme selection bias, not surprisingly implies the same conclusion. Even the best test rarely does better than the assumed test size in this case.[37]

[36]For both the "extreme" and the "modest" selection examples, the population has a mean height of 66 inches with a standard deviation of 2.5 inches, as in all our simulations. We measure the return to height by the difference between height and 56 inches. The standard deviation of $\varepsilon_M = 0.1$,the standard deviation of $\varepsilon_C = 0.4$,and Corr(ε_M,ε_C)=0.1.

[37]The "sktest" is a test for skewness and kurtosis with an adjustment due to Royston (1991). The "sktest, noadjust" is the same test without Royston's adjustment, and is described in D'Agostino, Belanger, and D'Agostino (1990). "swilk" and "sfrancia" are Stata's implementation of two forms of the Shapiro-Wilk W test. swilk (the W test) is appropriate for 4 through 2000 observations, while sfrancia (the W' test) is appropriate for 5 through 5000 observations. The Doornick-Hansen omnibus test is a skewness-kurtosis test with a transformation to assure independence. Doornick and Hansen (2008) also propose an asymptotic form of this test, but view it as "unsuitable, except in very large samples." (p.928). The Henze-Zirkler (1990) test uses transforms of skewness and kurtosis to better approximate a χ^2distribution. The computation time for this test increases with the square of the number of observations. We perform the Doornick-Hansen tests using the Stata module "omninorm" (C.F. Baum and N.J. Cox, 2007. "Omninorm: Stata module to calculate omnibus test for univariate/multivariate normality."). The other tests are part of Stata.

Appendix D: Identification of selection with additional information

Section 5.2 describes an exercise that explores the conditions under which one can estimate a model of heights taking account of possible selection into the Army. This appendix decribes the model and results in detail. The empirical models follow exactly from the specification of the selection model in the data generating process as outlined in Section 5.2. Let h be observed height, x_i be a height mean shifter, and z_i be the variable measuring the relative attractiveness of the civilian sector, holding height rewards constant. Define I^* as a latent variable measuring preferences for the civilian sector over the military sector. The statistical model for heights and selection into the military is given by:

$$h_i = \mu + \gamma x_i + u_i \tag{28}$$

$$I_i^* = \alpha^* + \beta^* h_i + \delta^* z_i + v_i \tag{29}$$

We assume all error terms follow mean-zero, homoscedastic normal distributions. An individual enters the military if $I_i^* < 0$. The relative attractiveness of the civilian sector can vary by the individual's height, as in the selection model discussed above. For simplicity we assume that the error terms u_i and v_i are independent normal random variables. We begin with the most common situation, that we only observe heights for soldiers. For civilians we substitute the "expected height" (28) into the latent variable equation that descibes preferences for the Army (29). Normalizing the resulting error in the latent outcome equation to have variance one yields

$$h_i = \mu + \gamma x_i + u_i \tag{30}$$

$$I_i = \alpha + \beta[\mu + \gamma x_i] + \delta z_i + \varepsilon_i \tag{31}$$

where $\alpha = (\alpha^*/\sigma^*)$; $\beta = (\beta^*/\sigma^*)$; $\delta = (\delta^*/\sigma^*)$;$\varepsilon_i = (v_i + \beta^* u_i)/\sigma^*$; and $\sigma^* = \sqrt{(Var(v) + \beta^2 Var(u))}$. We observe h_i if and only if $I_i < 0$. These assumptions imply a correlation of the height error and the re-specified selection error of $\rho = (\beta^* \sigma_h)/\sigma^*$, where σ_h is the standard deviation of u. In all following estimations we allow the Army to impose a minimum height standard. Let Min_i be the known military height standard in effect for observation i. We observe an individual in the military if and only if $h_i \geq Min_i$ and $I_i < 0$. Using the above joint distribution of h and I, let $P[Mil_i] = Pr[h_i \geq Min_i, I_i < 0]$ be the probability that we observe a given individual as a soldier. $P[Mil_i]$ is a function of x, z, and the parameters defining the bivariate normal distribution of u and ε.

We consider three estimation models. The first model is Wachter and Trussell's (1982a) reduced sample maximum likelihood estimator (RSMLE). This model uses only the heights of those in the military and relies upon the assumption of no correlation between the height and selection error terms. Thus an individual's height does not affect the choice of sector, or $\rho = 0$.

The second model relaxes the restriction of zero error covariance, but retains the assumption that we only observe height for soldiers. The likelihood in this case is the distribution of an observed soldier's integer height conditional on entering the military. We call this the "Conditional on Military" (COM) selection model, and its likelihood is $f(\llcorner h_i \lrcorner | Mil_i = 1])$ or:

$$\frac{\int_{-\infty}^{\bar{l}_i} \int_{-\infty}^{\underline{l}_i} f(\varepsilon, u; \rho) du d\varepsilon - \int_{-\infty}^{\bar{l}_i} \int_{-\infty}^{\underline{l}_i - 1} f(\varepsilon, u; \rho) du d\varepsilon}{\int_{-\infty}^{\bar{l}_i} \int_{M_i - (\mu + \gamma x_i)}^{\infty} f(\varepsilon, u; \rho) du d\varepsilon} \tag{32}$$

Where $\bar{l}_i = -(\alpha + \beta[\mu + \gamma x_i])$ and $\underline{l}_i = \llcorner h_i \lrcorner + 1 - (\mu + \gamma x_i)$. This expression is just the conditional distribution of heights for those who prefer the military, taking into account the restriction $h_i \geq \text{Min}_i$. For $\rho = 0$ equation (32) simplifies to the likelihood function for the RSMLE with integer heights. Note that if $\rho = 0$, the joint distribution factors into the product of the two marginal distributions; the probability of preferring military service cancels out in the numerator and the denominator. In this case, we cannot possibly identify any of the parameters describing one's preferences for the military because we only observe individuals in the military and their observed heights provide no information about how strongly they prefer military service.

The third model we consider ("REP") assumes we have a representative sample of the entire population. We assume that we have the x_i and z_i variables for all observations, but only observe heights for those who join the military. This is not necessarily an unrealistic assumption. The x_i, variables can be functions of birth cohort, which most sources include. The z_i, variables could measure macroeconomic factors such as the employment opportunities in the civilian sector at the time of the enlistment decision. Assume that we have the sample of observed military heights as well as the proportion of each birth cohort joining the Army, in addition to year of enlistment macro economic data. The likelihood function for this REP specification is $f(\llcorner h_i \lrcorner, Mil_i|x_i, z_i)=$ $\left[P_i^1\right]^{Mil_i}\left[1-P_i^2\right]^{1-Mil_i}$ where $\mathrm{P}_i^1$is equal to the numerator in equation (32) and $\mathrm{P}_i^2$is the denominator in (32).

We estimate the three models using the 20 million-observation Monte Carlo dataset derived from the data-generating progress described above. Appendix Table D.1 reports descriptive statistics and summarizes our attempts to estimate the three models. Column (1) reports parameter information and summary statistics when the data are evaluated at the true parameter values. About eight percent of individuals prefer military service to the civilian sector, but because of the minimum height restrictions only five percent actually enlist in the military. The parameter values used to generate these data imply a 0.24 correlation between the height error and the selection error. Column (1) also reports the value of the log-likelihood function at the true parameters for the two models that are consistent when $\rho \neq 0$.

Column (2) provides information about the estimation of the RSMLE when only the variable directly influencing height (the x variable) is used to model the observed heights. Not surprisingly, its log-likelihood function value is much lower (by almost 220 points) than that for the correct model evaluated at the true parameter values. The estimated parameters here imply a mean population height more than one inch shorter than the true mean height in the population. The height mean shifter's impact is slightly underestimated. Column (3) reports estimates from the RSMLE model where we allow the selection variable z to also enter (linearly) the function determining mean height. Controlling for this "selection" variable barely changes the incorrect estimate of the mean height in the RSMLE in Column (2). The log-likelihood function improves by 220 points with the estimation of just one additional parameter. Surprisingly, this model "fits" the observed data better than the true data generating process (conditional on only observing those in the military) evaluated at the true parameter values.

Columns (4) through (8) summarize results obtained from the Conditional on Military Selection (COM) model using a variety of different starting values and parameter constraints. The results echo our findings for the RSMLE model: there are severe identification issues in a model of the population height distribution that relies on data drawn only from soldiers. The model can barely distinguish among a wide variety of different candidate population height distributions. The log-likelihood function values are nearly identical for the five models reported in columns (4) through (8), even though the estimated mean height varies between 62 and 68 inches. In some cases, the log-likelihood value is slightly less than the one from the RLMS model including the selection covariate (3), and other cases they are slightly larger. The estimated impact of height on selection varies considerably, from large positive to large negative values, and in none of these five models is it close to the true value of 0.094 reported in column (1). The estimated effect of the height shifter x also exhibits considerable variability. The fact that we could not bring the models to convergence without constraining the value of the error covariance also indicates a potentially severe identification problem. These estimates suggest that it might be nearly impossible to obtain accurate inferences about the population height distribution if one only has access to data from subsamples that might be selected on the basis of height.

The estimates in column (9), imply something more positive: incorporating information on the proportion of individuals actually joining the military can resolve nearly all of these identification issues. The only additional information we used for these estimates is the fraction of individuals joining the military for each combination of the $x_i$and z_i, a total of 25 combinations here. The only height information comes from the military subsample. This stunning improvement in estimation model performance reflects a simple fact. With information only on military enrollees, it is impossible to assess whether an increase in a particular "taste" variable (z_i) makes military service more desirable, less desirable, or neither more nor less desirable. Once one can observe how military enrollments fluctuate with variations in the exogenous variables x_i and z_i, it is possible to pin down the parameters of the selection model. Given this identification of the selection process, it is relatively simple to distinguish between shifts in observed military heights due to changes in selection variables z_i from changes in the mean population height due to x_i.

Appendix E: Sources for econometric estimates

In Section 6 we report OLS models of mean height using two distinct sources that pertain to the British Army in the 18th and 19th centuries. Here we provide further information on those samples. We augmented the information on soldiers with annual series that provide an indication on the civilian economy and recruiting conditions. We also summarize those series here.

The "Army" data consists of soldiers included in the public-use sample Floud et al (1990) created in connection with the study discussed in some detail in Section 2.[38] This sample includes enlistees in both the Royal Army and the Royal Marines, but we use only the Army (the soldiers). Royal Marines were recruited on a different basis, and Komlos (1993) and others argue that combining them for analysis is inappropriate. Because some younger enlistees might still be growing, we use only those aged 22-27 at the time of enlistment; thus our estimation sub-samples are much smaller than the total sample. To contend with minimum height restrictions, we exclude individuals who were less than 65 inches tall (in some specifications) and less than 67 inches tall (in other specifications). We also exclude individuals who enlisted in 1861, because in that year the minimum height was shifted to 68 inches. These restrictions produced a sub-sample of the men underlying the OLS regressions in Table 4. Appendix Table E.1 provides descriptive statistics for this sample, and Appendix Table E.2 provides full regression results for the models summarized in Table 4.

The Army Medical Department (AMD) published annual reports summarizing the heights of men inspected for potential for Army service starting in the 1860s. (Floud et al (1990) also used this information, but did not include it in the public use sample, so we tabulated them separately on our own). The AMD reports tabulate the heights of recruits, by age, in cells one-inch wide. Men 72 inches and taller are "top-coded," so we only know they are at least 72 inches tall. We use the reports starting in 1879. We again limit the regression sub-sample to men aged 22-27 who were at least 67 inches tall when they enlisted. Appendix Table E.3 provides descriptive statistics for the AMD data, while Appendix Table E.4 reports the full regressions summarized in Table 5.

The regressions also include information on current conditions when men enlisted in the Army. One group of such variables is simply the men's ages, which come directly from the military sources. In addition, we use two variables that pertain to current military conditions. WAR is a dummy

[38] The files are available from the UK Data Archive. The full record can be found here: http://discover.ukdataservice.ac.uk/catalogue/?sn=2131&type=Data%20catalogue

equal to one if the U.K. was at war when a man enlisted. We code WAR from Floud et al (Table 2.7); it takes a value of one from 1775-1783, then from 1793-1813, then from 1855-6 and finally from 1899-1902. DEFENSE is total nominal UK expenditure on domestic defense, in billions of pounds sterling. GDP is nominal GDP per capita in current pounds sterling. (This is a period of price stability, so there is little difference between nominal and real series; in any case, the identification here is off year-to-year variation, so using real series makes little difference.) We obtained DEFENSE from the website www.ukpublicspending.co.uk, but the series was originally reported in Mitchell (1988). We obtained GDP from www.measuringworth.com. Indoor relief is the estimated proportion receiving indoor relief, using Williams (1981, Table 4.5). Unemployment is Feinstein's (1972, Table 57) series. The Feinstein index is based on the Board of Trade estimates of unemployment among trade-union members. Boyer and Hatton (2002) provide an alternative unemployment series that better-represents the British working classes; trade unionists were a relative elite, and worked disproportionately in industries facing high variance in unemployment rates. We prefer the Feinstein series to Boyer and Hatton for our specific purpose because we also employ the indoor-relief series, which MacKinnon (1986) argues is the best single indicator of unemployment for unskilled workers. (In other circumstances the Boyer-Hatton series is probably preferable). Thus we hope our unemployment series and our indoor-relief series capture the labor-market conditions for two different types of workers. Neither the relief nor the unemployment variable are available for the entire period covered by the Army source, so for this period we use Clark's estimate of the real wages for building laborers.[39]

[39]We obtained the wage series from http://gpih.ucdavis.edu/Datafilelist.htm#Europe.

Appendix F: Selection diagnostics with the RSMLE model

The regressions discussed in Section 6 above are only an approximation to the model most heights studies have in mind. To get closer, we estimate the RSMLE model Floud et al (1990) use for the AMD data only. Note the complication here; the RSMLE is only correct under the assumption that the distribution above the truncation point is free of any selection. We have shown this not to be true. So while our RSMLE estimates replicate and expand upon those reported by Floud et al (1990), they are not "correct," because the statistical assumptions upon which they rest are violated. But the same observation applies to the Floud et al (1990) estimates.

Following Wachter and Trussell (1982a), we assume heights are normally distributed conditional on a set of covariates x, where the covariates can influence either or both the mean and the standard deviation of heights. The AMD source reports heights in integers. For each observation i, define the mean conditional on x and standard deviation conditional on x as $\mu_i = x_i'\pi_M$ and $\sigma_i = \exp(x_i'\pi_\sigma)$. The likelihood function for the integer values of heights, conditional on height being at or above a possibly person specific lower limit L_i, is given by

$$\prod_{i=1}^{N} \left[\frac{\Phi\left(\frac{\lfloor h_i \rfloor + 1 - \mu_i}{\sigma_i}\right) - \Phi\left(\frac{\lfloor h_i \rfloor - \mu_i}{\sigma_i}\right)}{1 - \Phi\left(\frac{L_i - \mu_i}{\sigma_i}\right)} \right]^{\mathbf{1}(h_i \geq L_i)} \tag{33}$$

where $\mathbf{1}(\cdot)$ is the indicator function and $\lfloor \cdot \rfloor$ is the floor function that extracts the integer portion of the height. Columns (7) and (8) of Table 5 in the text reports the results of exclusion restriction tests similar to those discussed above for the OLS estimates. (Full model results are found in Appendix Table F.1.). Again we find that the "current condition" variables have an important impact on the heights of soldiers, which would not be the case if soldier heights were a random sample of the population height distribution. In this case, we only consider the constraint on the mean; if we also consider the analogous restriction on the standard deviation, the results are even stronger.

Appendix G: The logic of the selection diagnostic

This appendix provides a more formal justification for the selection diagnostics reported in Section 6. Suppose we draw a random sample from the distribution of heights at time t only for those observations taller than an arbitrary minimum height threshold H_m. We hold this threshold constant across birth cohorts by choosing a fixed level of H_m to be above the maximum observed threshold for any year t. (Thus $H_m$corresponds to Wachter and Trussell's h*). Suppose this distribution can vary by birth cohort b, so the distribution function for observed heights is f(h;b). The expected value of height conditional on birth cohort, time period t, and height being H_m or greater is given by:

$$E_f(h_{b,t}|_h \geq H_m, b, t) = \frac{\int_{H_m}^{\infty} hf(h;b)dh}{\int_{H_m}^{\infty} f(h;b)dh} = a(b) \tag{34}$$

Thus a(b) is implicitly a function of the minimum height threshold, but does not depend on time period t as this expression (34) is assumed to be a random sample from the distribution of heights H_m and greater in the birth cohort. In general, a(b) is a function of both the birth-period-specific unconditional mean height and the unconditional standard deviation of the birth-period-specific height distribution. By the definition of the conditional expected value, if we estimate a regression function for heights of people H_m and taller on a set of birth cohort dummies, the coefficients on the dummy variables would measure the mean height of each birth cohort b for those at least as tall as H_m. This approach corresponds to the implicit assumptions used by researchers who rely upon height data from samples with minimum height restrictions. It allows us to use information on the functions a(b) to make inferences about variations in macroeconomic conditions for birth cohorts provided they have reached full adult height.

Suppose there are year t variables that could affect the relative attractiveness of the military to persons of different heights. A labor market with a high unemployment rate, for example, might make the military relatively more attractive to a tall person than a short person in such a year when compared to a year with a low unemployment rate. In this instance, the distribution of heights to the military would no longer be independent of variables measuring economic conditions at the time of enlistment. Let g(h; b,t) be the height distribution of those willing to enter the military. The expected value of height conditional on birth cohort, time period t and height being H_m or greater is given by:

$$E_g(h_{b,t}|_h \geq H_m, b, t) = \frac{\int_{H_m}^{\infty} hg(h;b)dh}{\int_{H_m}^{\infty} g(h;b)dh} = c(b,t) \tag{35}$$

Without explicit assumptions about how variables contemporaneous to military enlistment affect selection into the military by height, it would be nearly impossible to separate how variations in "childhood growth" periods (as measured by b) could be separated from the effects of factors influencing the desirability of military enlistment at time t by examining a regression function of observed military heights on birth cohort dummies and time of enlistment explanatory variables. A *comparison* of these two expected values, however, can provide a test of whether there is no non-height related selection into the military. Consider the following regression model for those individuals at least as tall as height H_m:

$$h_{i|h \geq H_m} = a(b) + r(b,t) + \eta_i \tag{36}$$

Under the null hypothesis of no differential selection by height into the military due to economic factors at time of enlistment, the function r(b,t) should equal zero and should not depend on date t variables. As a simple test of this hypothesis, we approximate the above regression model by

$$h_{i|h \geq H_m} = h(b) + q(t) + \xi_i \tag{37}$$

A rejection of the null hypothesis H_0: q(t) = 0 versus the alternative that it is not zero can be interpreted as a rejection of the null hypothesis of no differential selection into the military by height in response to changing economic conditions. Note, however, that estimates of h(b) from this regression *cannot* be considered to be estimates of the true mean hight by birth cohort for those H_m or taller. None of the above derivations would be substantively altered if we assumed heights were observed only falling in a set of ranges. In addition, the logic carries through even if there is an exact censoring of heights above some upper cutoff point.

References

A'Hearn, Brian. 2003. "Anthropometric Evidence on Living Standards in Northern Italy, 1730-1860." *Journal of Economic History* 63(2): 351-381.

A'Hearn, Brian, Franco Peracchi, and Giovanni Vecchi, 2009. "Height and the Normal Distribution: Evidence from Italian Military Data." *Demography* 46(1): 1-25.

Alchian, Armen A, and William R. Allen. 1964. *University Economics*. Belmont, Cal.: Wadsworth Publishing Company, 1964.

Allen, Robert C. 2007. "Pessimism Preserved: Real Wages in the British Industrial Revolution." Oxford University Department of Economics Working Paper 314.

Angrist, Joshua D. and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton and Oxford: Princeton University Press.

Asch, Beth J., James R. Hosek, and John T. Warner, 2007, "New Economics of Manpower in the Post-Cold War Era." *Handbook of Defense Economics*, Vol. 2. Todd Sandler and Kweith Hartle (eds.): 1077-1140.

Baum, C. F. and N. J. Cox. 2007. "Omninorm: STATA Module to Calculate Omnibus Test for Univariate/Multivariate Normality."

Bodenhorn, Howard. 1999. "A Troublesome Caste: Height and Nutrition of Antebellum Virginia's Rural Free Blacks." *Journal of Economic History* 59(4): 972-996.

Bodenhorn, Howard. 2011. "Manumission in Nineteenth-Century Virginia." *Cliometrica* 5(2): 145-164.

Bodenhorn, Howard, Carolyn Moehling and Gregory N. Price. 2012 (forthcoming). "Short Criminals: Stature and Crime in Early America." *Journal of Law & Economics* 55(2).

Borcherding, Thomas E. and Eugene Silberberg. 1978. "Shipping the Good Apples Out: The Alchian and Allen Theorem Reconsidered." *Journal of Political Economy* 86(11): 131-138.

Boyer, George R. and Timothy J. Hatton. 2002. "New Estimates of British Unemployment, 1870-1913." *Journal of Economic History* 62(3): 643-675.

Bozzoli, Carlos, Angus Deaton, Climent Quintana-Domeque. 2009. "Adult Height and Childhood Disease." *Demography* 46(4): 647-669.

Brinkman, Henk Jan, J.W. Drukker, and Brigitte Slot, 1988. "Height and Income: A New Method for the Estimation of Historical National Income Series." *Explorations in Economic History* 25: 227-264

Brown, John C., 1990. "The Condition of England and the Standard of Living: Cotton Textiles in the Northwest, 1806-1850." *Journal of Economic History* 50(3): 591-614.

Calomiris, Charles and Jonathan Pritchett. 2009. "Preserving Slave Families for Profit: Traders' Incentives and Pricing in the New Orleans Slave Market." *Journal of Economic History* 69(4): 986-1011.

Carson, Scott Alan. 2008. "Inequality in the American South: Evidence from the Nineteenth Century Missouri State Prison." *Journal of Biosocial Science* 40(4): 587-604.

Carson, Scott Alan. 2009. "Health, Wealth and Inequality: A Contribution to the Debate about the Relationship between Inequality and Health." *Historical Methods* 42(2): 43-56.

Case, Anne and Christina Paxson. 2008. "Stature and Status: Height, Ability, and Labor Market Outcomes." *Journal of Political Economy* 116, 499-532.

Coclanis, Peter A. and John Komlos. 1995. "Nutrition and Economic Development in Post-Reconstruction South Carolina." *Social Science History* 19(1): 91-115.

Crafts, N. F. R. 1997. "Dimensions of the 'Quality of Life' during the British Industrial Revolution." *Economic History Review* 50(4): 617-639.

Crimmins, Eileen and C. E. Finch. 2006. "Infection, Inflammation, Height, and Longevity." *Proceedings of the National Academy of Sciences* 103(2): 498-503.

Cuff, Timothy. 2005. *The Hidden Cost of Economic Development: The Biological Standard of Living in Antebellum Pennsylvania*. Burlington, Vt.: Ashgate Publishing Company.

D'Agostino, Ralph B., Albert Belanger, and Ralph B. D'Agostino, Jr. 1990. "A Suggestion for Using Powerful and Informative Tests of Normality." *American Statistician* 44(4): 316-321.

Danson, J. T. 1881. "Statistical Observations on the Growth of the Human Body (Males) in Height and Weight from Eighteen to Thirty Years of Age, as Illustrated by the Records of the Borough Gaol of Liverpool." *Journal of the Statistical Society* 44(4): 660-674.

Deaton, Angus. 2007. "Height, Health and Development." *Proceedings of the National Academy of Sciences* 104.33, 13232-13237.

Doornik, Jurgen A. and Henrik Hansen. 2008. "An Omnibus Test for Univariate and Multivariate Normality." *Oxford Bulletin of Economics and Statistics* 70(S1): 927-939.

Engels, Friedrich. 1945/1987. *Condition of the Working Class in England*. New York: Viking Penguin.

Feinstein, Charles. 1972. *National Income, Expenditure, and Output of the United Kingdom, 1855-1965*. Cambridge: Cambridge University Press.

Feinstein, Charles. 1998. "Pessimism Perpetuated: Real Wages and the Standard of Living in Britain during and after the Industrial Revolution." *Journal of Economic History* 58(3): 625-658.

Fogel, Robert W. 1994. "Economic Growth, Population Theory, and Physiology: The Bearing of Long-Term Processes on the Making of Economic Policy." *American Economic Review* 84(3): 369-395.

Fogel, Robert W., Stanley L. Engerman, Roderick Floud, Richard H. Steckel, T. James Trussell, Kenneth W. Wachter, Kenneth Sokoloff, Georgia Villaflor, Robert A. Margo, and Gerald Friedman. 1982. "Changes in American and British Stature since the Mid-Eighteenth Century: A Preliminary Report on the Usefulness of Data on Height for the Analysis of Secular Trends in Nutrition, Labor Productivity, and Labor Welfare." NBER Working Paper No. 890.

Floud, Roderick. 1984. "The Heights of Europeans since 1750: A New Source for European Economic History." NBER working paper 1318.

Floud, Roderick, Kenneth Wachter and Annabel Gregory, 1990. *Height, health and history: Nutritional Status in the United Kindgom, 1750-1980*. Cambridge: Cambridge University Press.

Freudenberger, Herman and Jonathan B. Pritchett. 1991. "The Domestic United States Slave Trade: New Evidence." *Journal of Interdisciplinary History* 21(3): 447-477.

Galenson, David W. 1981. *White Servitude in Colonial America: An Economic Analysis*. Cambridge: Cambridge University Press.

Gallman, Robert E. 1996. "Dietary Change in Antebellum America." *Journal of Economic History* 56(1): 193-201.

Grubb, Farley. 1994. "The End of European Immigrant Servitude in the United States: An Economic Analysis of Market Collapse, 1772-1835. *Journal of Economic History* 54(4): 794-824.

Grubb, Farley. 1999. "Lilliputians and Brobdingnagians, Stature in British Colonial America: Evidence from Servants, Convicts, and Apprentices." *Research in Economic History* 19: 139-203.

Grubb, Farley. 2001. “The Market Evaluation of Criminality: Evidence from the Auction of British Convict Labor in America, 1767-1775.” *American Economic Review* 91(1): 295-304.

Haines, Michael R., Lee A. Craig, and Thomas Weiss. 2003. “The Short and the Dead: Nutrition, Mortality, and the ‘Antebellum Puzzle’ in the United States.” *Journal of Economic History* 63(2): 382-413.

Heckman, James J., 1979. “Sample Selection Bias as a Specification Error.” *Econometrica* 47(1): pp.

Heckman, James J. and Guilherme Sedlacek. 1995. “Heterogeneity, Aggregation, and Market Wage Functions: An Empirical Model of Self-Selection in the Labor Market.” *Journal of Political Economy* 93(6): 1077-1125.

Henze, N. and B. Zirkler. 1990. “A Class of Invariant Consistent Tests for Multivariate Normality.” *Communications in Statistics: Theory and Methods* 19(10): 3595-3617.

Higman, Barry W. 1979. “Growth in Afro-Caribbean Slave Populations.” *American Journal of Physical Anthropology* 50(3): pp. 373-385.

Higman, Barry W. 1984. *Slave Populations of the British Caribbean, 1807-1834*. Baltimore: Johns Hopkins University Press.

Hobsbaum, E. J. 1957. “The British Standard of Living, 1790-1850.” *Economic History Review* 10(1): 46-68.

Johnson, Paul and Stephen Nicholas. 1997. “Health and Welfare of Women in the United Kingdom, 1785-1920.” In *Health and Welfare during Industrialization,* 201-250. Richard H. Steckel and Roderick Floud (eds.). Chicago: University of Chicago Press.

Kilburn, Rebecca M., and Jacob Klerman. 1999. *Enlistment Decisions in the 1990s: Evidence from Individual Level Data.* MR -944-OSD/A RAND, Santa Monica, CA.

Komlos, John. 1987. "The Height and Weight of West Point Cadets: Dietary Change in Antebellum America." *Journal of Economic History* 47(4): 987-927.

Komlos, John. 1989. *Nutrition and Economic Development in the Eighteenth-Century Hapsburg Monarchy: An Anthropometric History*. Princeton, N.J.: Princeton University Press.

Komlos, John. 1992. “Toward an Anthropometric History of African-Americans: The Case of the Manumitted Slaves of Maryland.” In *Strategic Factors in Nineteenth-Century American Economic History: A Volume to Honor Robert W. Fogel*, 297-329. Edited by Claudia Goldin and Hugh Rockoff. Chicago: University of Chicago Press.

Komlos, John, 1993. “The secular trend in the biological standard of living in the United Kingdom, 1730-1860.” *Economic History Review* 46(1): 115-144.

Komlos, John. 1994. “The Height of Runaway Slaves in Colonial America, 1720-1770.” In Stature, Living Standards, and Economic Development: Essays in Anthropometric History, 93-116. Edited by John Komlos. Chicago and London: University of Chicago Press.

Komlos, John, 1998. “Shrinking in a Growing Economy? The Mystery of Physical Stature during the Industrial Revolution.” *The Journal of Economic History* 58(3): 779-802.

Komlos, John. 2003. “An Anthropometric History of Early-Modern France.” *European Review of Economic History* 7: 159-189.

Komlos, John, 2004. “How to (and How Not to) Analyze Deficient Height Samples – an Introduction.” *Historical Methods* 37(4): 160-173.

Komlos, John and Bjorn Alecke. 1996. “The Economics of Antebellum Slave Heights Reconsidered.” *Journal of Interdisciplinary History* 26(3): 437-457.

Komlos, John and Peter Coclanis. 1997. “On the Puzzling Cycle in the Biological Standard of Living: The Case of Antebellum Georgia.” *Explorations in Economic History* 34(4): 433-459.

Komlos, John and Helmut Küchenhoff. 2012. “The Diminution of the Physical Stature of the English Male Population in the Eighteenth Century.” *Cliometrica* 6(1): 45-62.

Kotz, Samuel, N. Balakrishnan and Norman L. Johnson. 2000. *Continuous Multivariate Distributions*, Volume 1, (second edition). New York: John Wiley & Sons, Inc.

Lamm, Doron. 1988. “British Soldiers of the First World War: Creation of a Representative Sample.” *Historical Social Research/Historische Sozialforschung* 13(4): 55-98.

Lang, Stefan and Marco Sunder. 2003. “Non-parametric regression with BayesX: A Flexible Estimation of Trends in Human Physical Stature in 19th Century America.” *Economics and Human Biology* 1(1): 77-89.

MacKinnon, Mary. 1986. “Poverty and Policy: The English Poor Law, 1860-1910.” *Journal of Economic History* 46(2): 500-502.

Margo, Robert A. and Richard H. Steckel. 1982. “The Heights of American Slaves.” *Social Science History* 6(4): 516-38.

Margo, Robert A. and Richard H. Steckel. 1983. “Heights of Native-Born Whites during the Antebellum Period.” *Journal of Economic History* 43(1): 167-174.

McLynn, Frank. 1989. *Crime and Punishment in Eighteenth-Century England*. London and New York: Routledge.

Meier, Paul. 1982. “Estimating Historical Heights: Comment.” *Journal of the American Statistical Association* 77(378): 296-297.

Mokyr, Joel and Cormac Ó Gráda. 1994. “The Heights of the British and the Irish c. 1800-1815: Evidence from Recruits to the East India Company’s Army.” In Stature, Living Standards, and Economic Development: Essays in Anthropometric History, 39-59. Edited by John Komlos. Chicago and London: University of Chicago Press.

Mokyr, Joel and Cormac Ó Gráda, 1996. “Height and health in the United Kingdom 1815-1860: evidence from the East India Company army.” *Explorations in Economic History* 33(2): 141-168.

Murray, John. 1997. “Standards of the Present for People of the Past: Height, Wight, and Mortality among Men of Amherst College.” *Journal of Economic History* 57(3): 585-606.

Nicholas, Stephen and Deborah Oxley. 1996. “Living Standards of Women in England and Wales, 1785-1815: Evidence from Newgate Prison Records.” *Economic History Review* 49(3): 591-599.

Nicholas, Stephen and Richard H. Steckel, 1991. “Heights and living standards of English workers during the Early Years of Industrialization, 1770-1815.” *Journal of Economic History* 51(4): 937-57.

Nicholas, Stephen and Richard H. Steckel, 1997. “Tall but Poor: Living Standards of Men and Women in pre-Famine Ireland.” *The Journal of European Economic History* 26(1): 105-134.

Ó Gráda, Cormac. 1991. “Heights in Tipperary in the 1840s: Evidence from the Prison Records.” *Irish Economic and Social History* 18: 24-33.

Ó Gráda, Cormac. 1996. “Anthropometric History: What’s In It for Ireland?” *Historie & Mesure* 11(1-2): 139-166.

Oxley, Deborah. 2003. “’The Seat of Death and Terror’: Urbanization, Stunting, and Smallpox.” *Economic History Review*, New Series 56(4): 623-656.

Persico, Nicola, Andrew Postlewaite, and Dan Silverman. 2005. “The Effect of Adolescent Experience on Labor Market Outcomes: The Case of Height.” *Journal of Political Economy* 112, 1019-1053.

Pritchett, Jonathan B. 1997. “The Interregional Slave Trade and the Selection of Slaves for the New Orleans Market.” *Journal of Interdisciplinary History* 28(1): 57-85.

Pritchett, Jonathan B. and Richard M. Chamberlain. 1993. “Selection in the Market for Slaves: New Orleans, 1830-1860.” *Quarterly Journal of Economics* 108(2): 461-473.

Pritchett, Jonathan B. and Herman Freudenberger. 1992. “A Peculiar Sample: The Selection of Slaves for the New Orleans Market.” *Journal of Economic History* 52(1): 109-127.

Razzell, Peter. 1998. “Did Smallpox Reduce Height?” *Economic History Review*, New Series 51(2): 351-359.

Rees, R., John Komlos, Ngo. V. Long, and Ulrich Woitek. 2003. “Optimal Food Allocation in a Slave Economy.” *Journal of Population Economics* 16(1): 21-36.

Riggs, Paul. 1994. “The Standard of Living in Scotland, 1800-1850.” In *Stature, Living Standards, and Economic Development: Essays in Anthropometric History*, 60-75. Edited by John Komlos. Chicago and London: University of Chicago Press.

Robert, C.P. and G. Casella. 2004. *Monte Carlo Statistical Methods* (second edition). New York: Springer-Verlag.

Roy, A. D. 1951. “Some Thoughts on the Distribution of Earnings.” *Oxford Economic Papers* 3(2): 135-146.

Sandberg, Lars G., and Richard H. Steckel, 1988. “Overpopulation and Malnutrition Rediscovered: Hard Times in 19th-Century Sweden.” *Explorations in Economic History* 25:1-19.

Sandberg, Lars G. and Richard H. Steckel, 1997. “Was Industrialization Hazardous to your Health? Not in Sweden!” in Steckel and Floud (1997).

Schultz, T. Paul. 2002. “Wage Gains Associated with Height as a Form of Health Human Capital.” *American Economic Review* 92(2): 349-353.

Simon, Curtis H., and John T. Warner. 2008. “Youth Attitudes and Military Recruiting During Operation Iraqi Freedom.” Working paper, Clemson University.

Sokoloff, Kenneth and Georgia Villaflor. 1982. "The Early Achievement of Modern Stature in America." *Social Science History* 6(4): 453-481.

Spiers, E. M. 1980. *The Army and Society, 1815-1914*. London: Longman.

Staub, Kaspar, Frank J. Rühli, Barry Bogin, Ulrich Woitek, and Christian Pfister. 2011. "Eduoard Mallet's Early and Almost Forgotten Study of the Average Height of Genevan Conscripts in 1835." *Economics and Human Biology* 9(): 438-442.

Steckel, Richard H. 1979. "Slave Height Profiles from Coastwise Manifests." *Explorations in Economic History* 16(4): 363-380.

Steckel, Richard H. 1983. "Height and Per Capita Income." *Historical Methods* 16(1): 1-7.

Steckel, Richard H. 1986. "A Peculiar Population: The Nutrition, Health, and Mortality of American Slaves from Childhood to Maturity." *Journal of Economic History* 46(3): 721-741.

Steckel, Richard H. 1987. "Growth Depression and Recovery: The Remarkable Case of American Slaves." *Annals of Human Biology* 14: 111-132.

Steckel, Richard H., 1995. "Stature and the Standard of Living." *The Journal of Economic Literature* 33(4): 1903-1940.

Steckel, Richard H., 1998. "Strategic Ideas in the Rise of the New Anthropometric History and their Implications for Interdisciplinary Research." *The Journal of Economic History* 58(3): 803-821.

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

Steckel, Richard H. and Roderick Floud, eds, 1997. *Health and welfare during Industrialization.* Chicago: University of Chicago Press.

Steckel, Richard H. and Joseph M. Prince, 2001. "Tallest in the World: Native Americans of the Great Plains in the Nineteenth Century." *American Economic Review* 91(1): 287-294.

Sunder, Marco. 2004. "The Height of Tennessee Convicts: Another Piece of the 'Antebellum Puzzle.'" *Economics and Human Biology* 2(1): 75-86.

Sunder, Marco. 2011. "Upward and Onward: High-society American Women Eluded the Antebellum Puzzle." *Economics and Human Biology* 9(2): 165-171.

Sunder, Marco and Ulrich Woitek. 2005. "Boom, Bust, and the Human Body: Further Evidence on the Relationship between Heights and Business Cycles." *Economics and Human Biology 3*(3): 450-466.

Trussell, James and Richard H. Steckel. 1978. "The Age of Slaves at Menarche and their First Birth." *Journal of Interdisciplinary History* 8(4): 477-505.

Voth, Hans-Joachim. 2003. "Living Standards during the Industrial Revolution: An Economist's Guide." *American Economic Review* 93(2): 221-226.

Voth, Hans-Joachim and Timothy Leunig. 1996. "Did Smallpox Reduce Height? Stature and the Standard of Living in London, 1770-1873." *Economic History Review*, New Series 49(3): 541-560.

Wachter, Kenneth W. and James Trussell, 1982a. "Estimating Historical Heights." *Journal of the American Statistical Association* 77(378):279-293

Wachter, Kenneth W. and James Trussell, 1982b. "Estimating Historical Heights: Rejoinder." *Journal of the American Statistical Association* 77(378):301-303.

Warner, John T., and Beth J. Asch. 2001. "The Record and Prospects of the All-Volunteer Military in the United States." *Journal of Economic Perspectives* 15(2): 169-192.

Warner, John T., Curtis J, Simon, and Deborah M. Payne. 2000. " Enlistment Supply in the 1990s: A Study of the Navy College Fund and Other Enlistment Incentive Programs." Draft Final Report for the Office of the Undersecretary of Defense for Personnel and Readiness. Arlington, VA: Defense Manpower Data Center.

Weil, David N. 2007. "Accounting for the Effect of Health on Economic Growth." *Quarterly Journal of Economics* 122(3): 1265-1306.

Weir, David R., 1997. "Economic Welfare and Physical Well-Being in France, 1750-1990." In Steckel and Floud (1997)

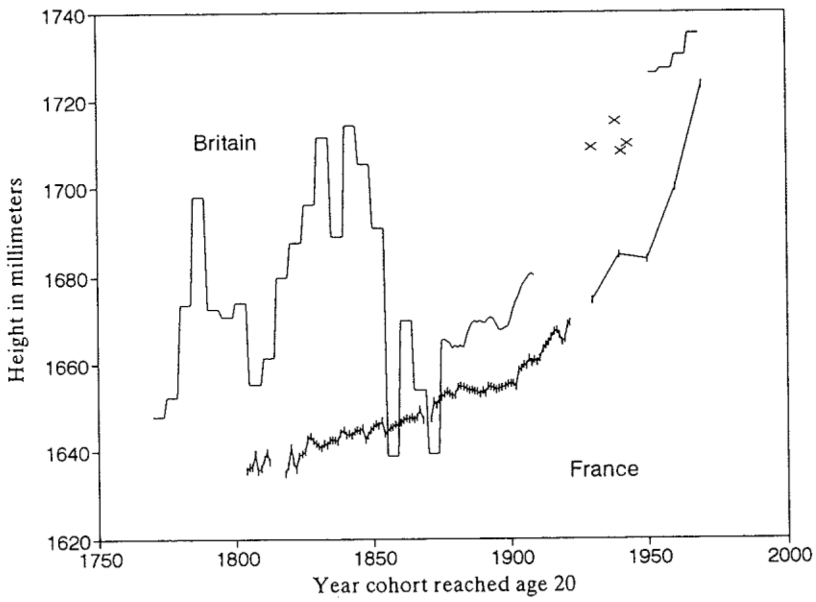
Weiss, Thomas J. 1992. "U.S. Labor Force Estimates and Economic Growth, 1800-1860." In *American Economic Growth and Standards of Living before the Civil War*, 19-78. Edited by Robert E. Gallman and John J. Wallis. Chicago: University of Chicago Press.

Whitwell, Greg, Christine de Souza, and Stephen Nicholas, 1997. "Height, Health, and Economic Growth in Australia, 1860-1940." In Steckel and Floud (1997).

Williams, Karel, 1981. *From Pauperism to Poverty*. London: Routledge & Kegan Paul.

Williamson, Jeffrey G. 1984. "Why Was British Growth So Slow during the Industrial Revolution?" *Journal of Economic History* 44(3): 687-712.

Figure 1: Male height at age 20 in France and Britain, 1770-1980



Source: Weir (1997, p. 175).

Figure 2: Effect of civilian return to height

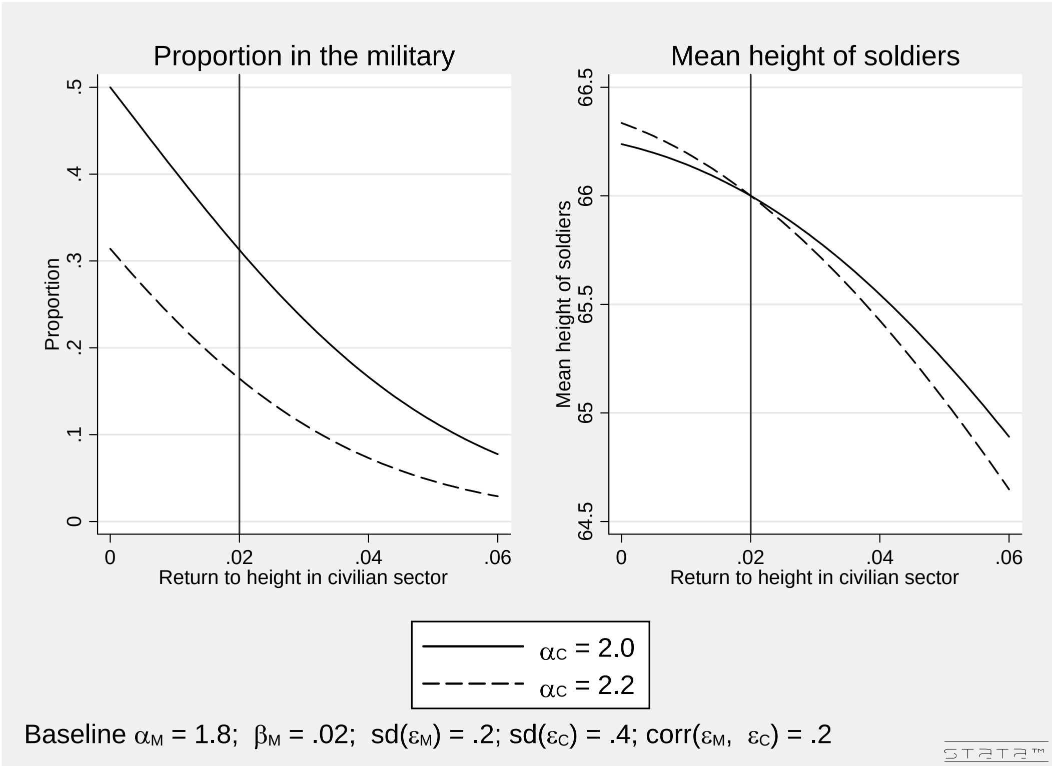


Figure 3: Heights distributions for simulated population

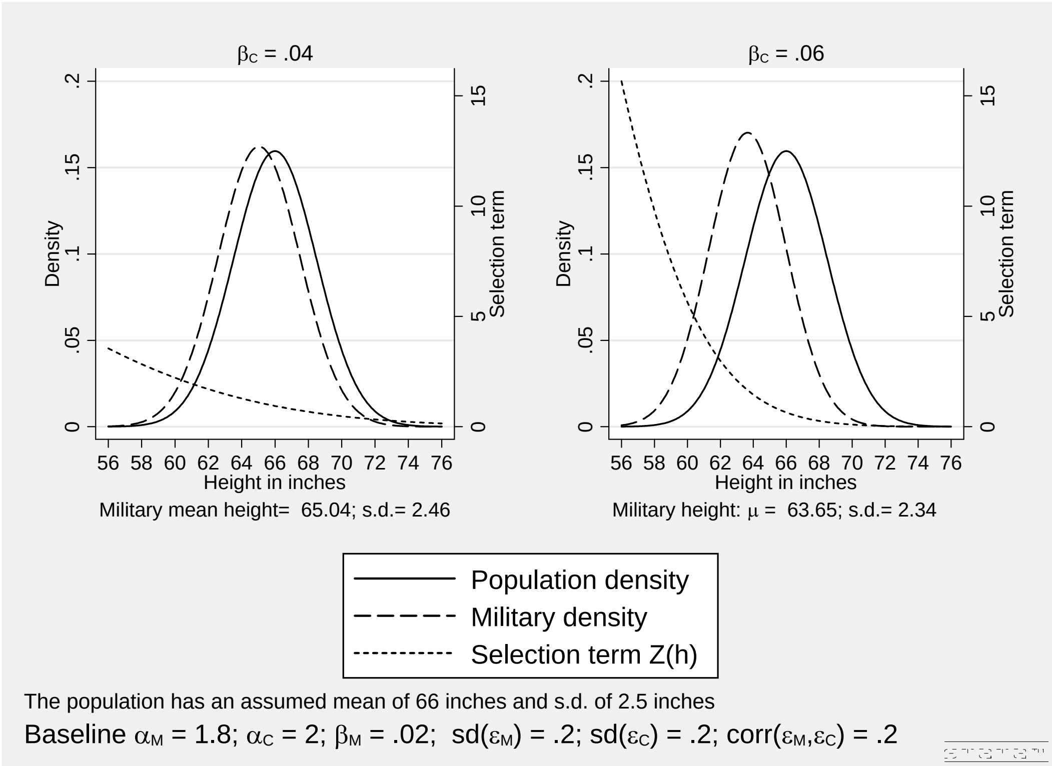


Figure 4: Tests for normality with selected samples

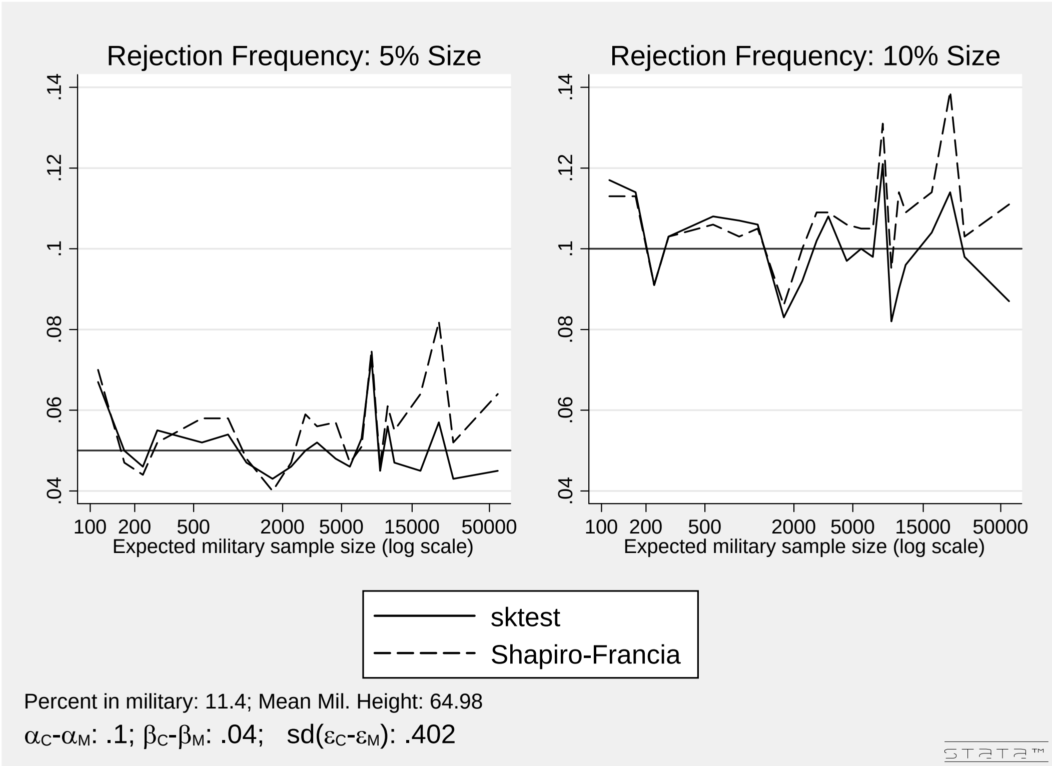


Figure 5: Rejection sample compared to Z(h) derived from Roy model

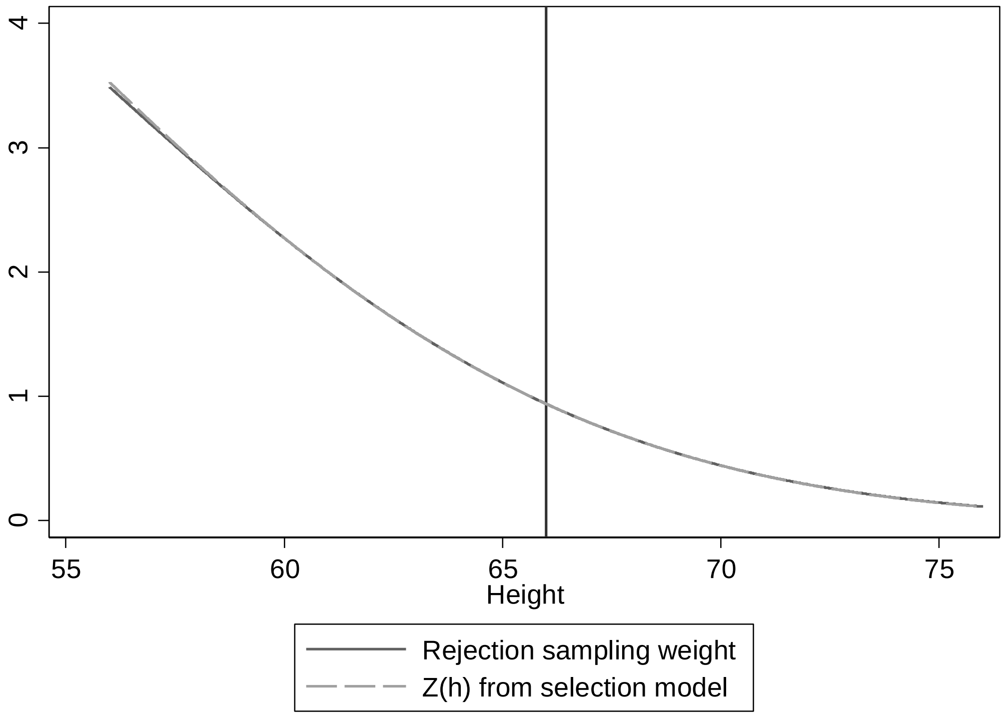


Figure 6: Many ways to get "identical" military samples

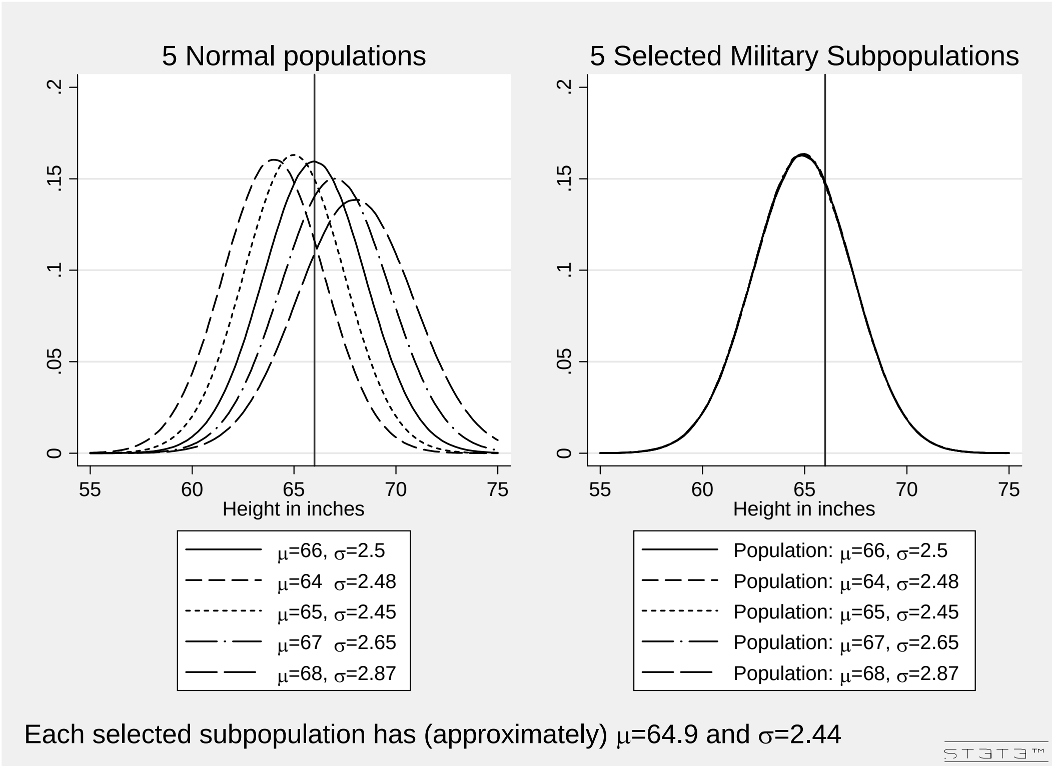


Table 1: Summary of main symbols in the occupational choice model, and their assumed values in simulation model

Symbol	Meaning
h	Height
N	Number of observations in simulated population
w_C	Potential civilian pay
w_M	Potential military pay
μ	Mean of heights in population
σ	Standard deviation of heights in population
α_C	Intercept for civilian log-wage equation at height 56 inches
α_M	Intercept for military log-wage equation at height 56 inches
β_C	Slope (return to one inch in height) for civilian log-wage equation
β_M	Slope (return to one inch in height) for military log-wage equation
ε_C	Skill (unobserved) affecting civilian log-wage equation
ε_M	Skill (unobserved) affecting military and civilian log-wages
γ_C	Return to military skill in civilian sector log-wage equation
γ_M	Return to military skill in military sector log-wage equation
δ_C	Return to civilian skill in civilian sector log-wage equation
τ_C	Taste for the civilian sector
τ_M	Taste for the military sector
σ_{CM}	Correlation of ($\varepsilon_{M,}\ \varepsilon_C$)

Table 2

	Model 1 (baseline)	Model 2	Model 3	Model 4	Model 5	Model 6 (no height selection)
Population mean	66	64	65	67	68	64.9
Population s.d.	2.5	2.48	2.45	2.65	2.87	2.44
α_M	0.4					
β_M	0.02					0
α_C	0.6	1.26	0.96	0.25	-0.07	0.99
β_C	0.06	-0.013	0.024	0.092	0.12	0
ε_M	0.1					
ε_C	0.4					
Military mean height	64.9	64.9	64.9	64.9	64.9	64.9
Military sd height	2.44	2.44	2.44	2.44	2.44	2.44
Fraction in military	0.079	0.079	0.077	0.077	0.075	0.076

Table 3

Percent of Times Each Model Selected as Best Fitting Model

(Model 1 is the true model)

	Sample Size						
	500	1,000	2,500	5,000	10,000	25,000	50,000
Number of samples	3146	1573	629	314	157	62	31
Model							
1	2.26	2.54	3.02	6.05	7.64	8.06	9.68
2	21.23	19.96	20.51	20.7	21.66	20.97	9.68
3	17.48	19.14	19.71	21.02	20.38	24.19	29.03
4	0.45	1.02	2.7	4.78	8.28	12.9	16.13
5	39.89	38.65	34.34	28.66	23.57	16.13	9.68
6	18.69	18.69	19.71	18.79	18.47	17.74	25.81
	100	100	99.99	100	100	99.99	100.01

Table 4: Summary of OLS estimates of height, using Army data

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Ages included	22-27	22-27	22-27	22-27	22-27	22-27	23-27	23-27	23-27
Heights included (inches)	65-72	65-72	65-72	67-72	67-72	67-72	67-72	67-72	67-72
Test: all age dummies zero	18.21 (0.0) [5, 13752]	13.23 (0.0) [5,13748]		2.84 (0.01) [5,9108]	3.40 (0.0) [9, 9104]		3.01 (.02) [4,6098]	3.38 (.01) [4,6094]	
Test: all recruitment-year dummies zero			6.76 (0.0) [118, 13639]			2.79 (0.0) [118,8995]			5.24 (0.0) [117,5985]
Test: all "macro" variables zero		4.79 (0.0) [5,13748]			1.03 (.39) [4,9104]			2.74 (.03) [4,6094]	
Test: ages and macro variables jointly zero		12.10 (0.0) [9,13748]			2.10 (.03) [9,9104]			2.93 (0.0) [8,6094]	
Number of observations used	13895			9224			6232		

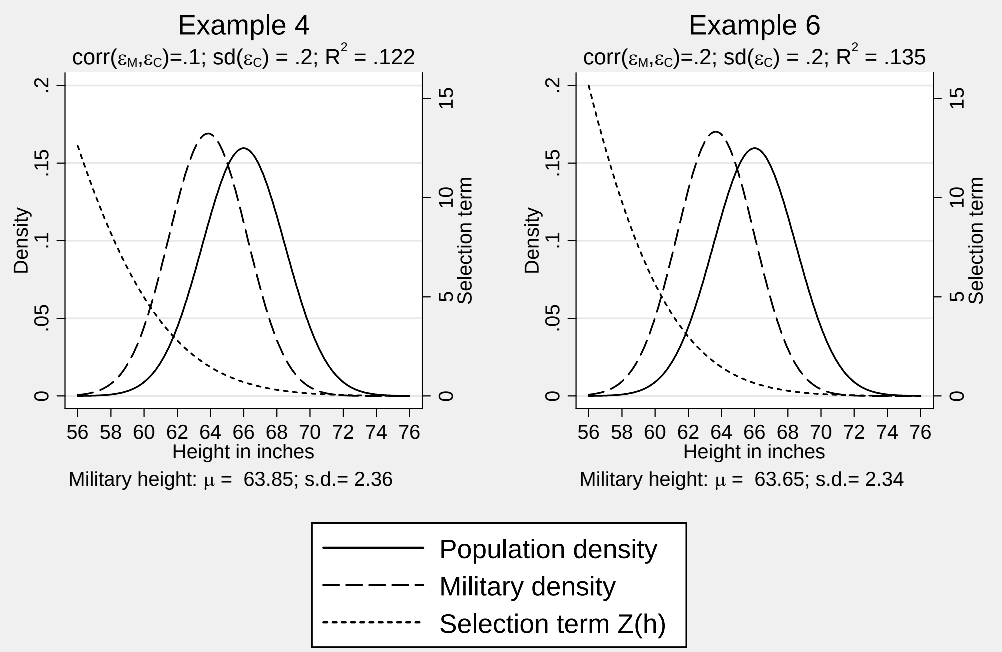
Note: Figures reported are F-statistics for the null hypothesis that the variables noted are collectively zero, with p-values in parentheses. Figures in [] are the degrees of freedom associated with the test. The regressions use robust standard errors. The full regression estimates are reported in Table E.2

Table 5: Summary of OLS estimates of height, using AMD data

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Estimation model	OLS	OLS	OLS	OLS	OLS	OLS	RSMLE	RSMLE
Ages included	22-25	22-25	22-25	22-25	22-25	22-25	22-25	22-25
Heights included (inches)	65-72	65-72	65-72	67-72	67-72	67-72	67-72	67-72
Test: all age dummies zero	35.73 (0.0) [3, 186408]	22.81 (0.0) [3, 186403]		5.16 (0.0) [3,102165]	2.68 (0.05) [3,102165]		6.40 (.09) [3]	
Test: all recruitment-year dummies zero			11.73 (0.0) [32,186379]			3.12 (0.0) [32,102141]		91.31 (0.0) [32]
Test: all "macro" variables zero		14.6 (0.0) [5,186403]			5.44 (0.0) [5,102165]		26.14 (0.0) [5]	
Test: ages and macro variables jointly zero		22.83 (0.0) [8,186403]			5.38 (0.0) [8,102165]		36.35 (0.0) [8]	
Number of observations used								

Note: Columns (1) through (7) pertain to OLS estimates; columns (8) and (9) are from RSMLE models. Figures reported are F-statistics for OLS and Wald statistics for RSMLE, with p-values in parentheses. Figure in square brackets [] is the degrees of freedom for the test. The regressions use robust standard errors. The full regression estimates are reported in Tables E.3 and E.4

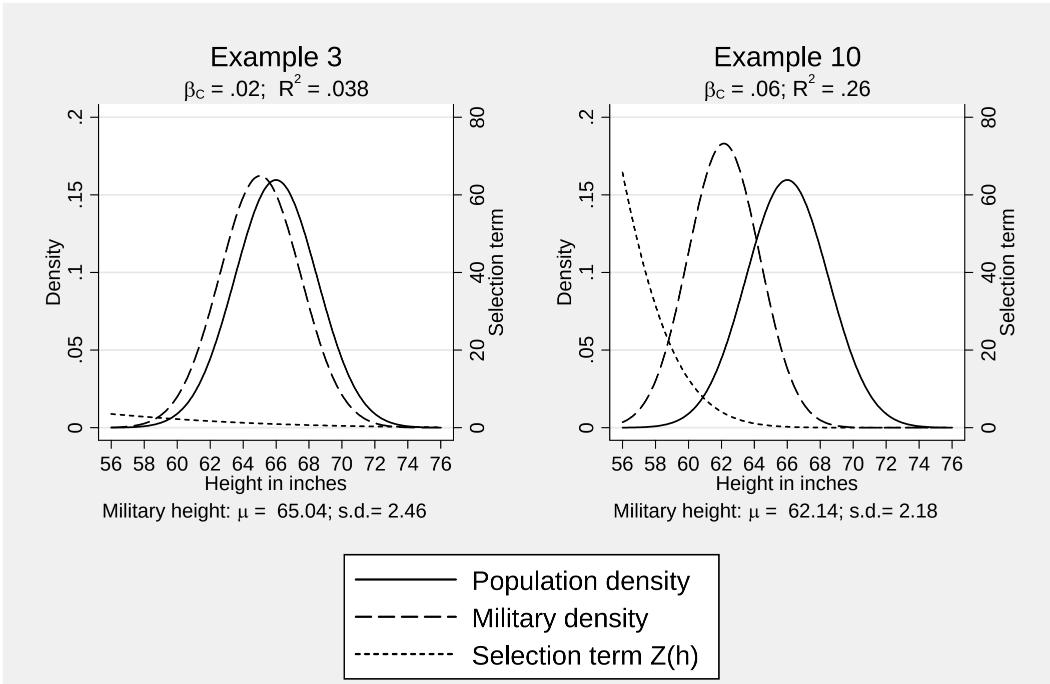
Figure B.1: Implications of wage-shock correlations



The population has an assumed mean of 66 inches and s.d. of 2.5 inches. For examples refer to Table B.1.

Baseline $\alpha_M = 1.8$; $\alpha_C = 2$; $\beta_M = 0$; $\beta_C = .04$; $sd\varepsilon_M = .2$

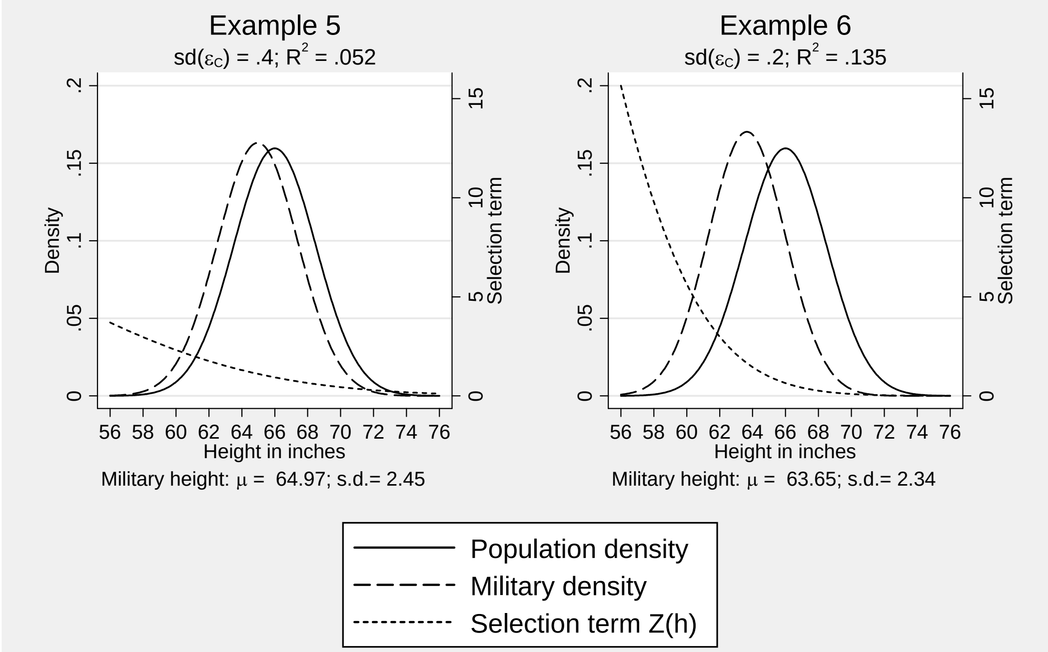
Figure B.2: Implications of return-to-heights differentials



The population has an assumed mean of 66 inches and s.d. of 2.5 inches. For examples refer to Table B.1.

Baseline $\alpha_M = 1.8$; $\alpha_C = 2$; $\beta_M = 0$; $corr(\varepsilon_M,\varepsilon_C)=.2$; $sd(\varepsilon_M) = .2$; $sd(\varepsilon_C) = .2$

Figure B.3: Implications of civilian wage shocks



The population has an assumed mean of 66 inches and s.d. of 2.5 inches. For examples refer to Table B.1.

Baseline $\alpha_M = 1.8$; $\alpha_C = 2$; $\beta_M = 0$; $\beta_C = .04$; $corr(\varepsilon_M,\varepsilon_C)=.1$; $sd(\varepsilon_M) = .2$

Figure C.1: Tests for normality with extreme selection

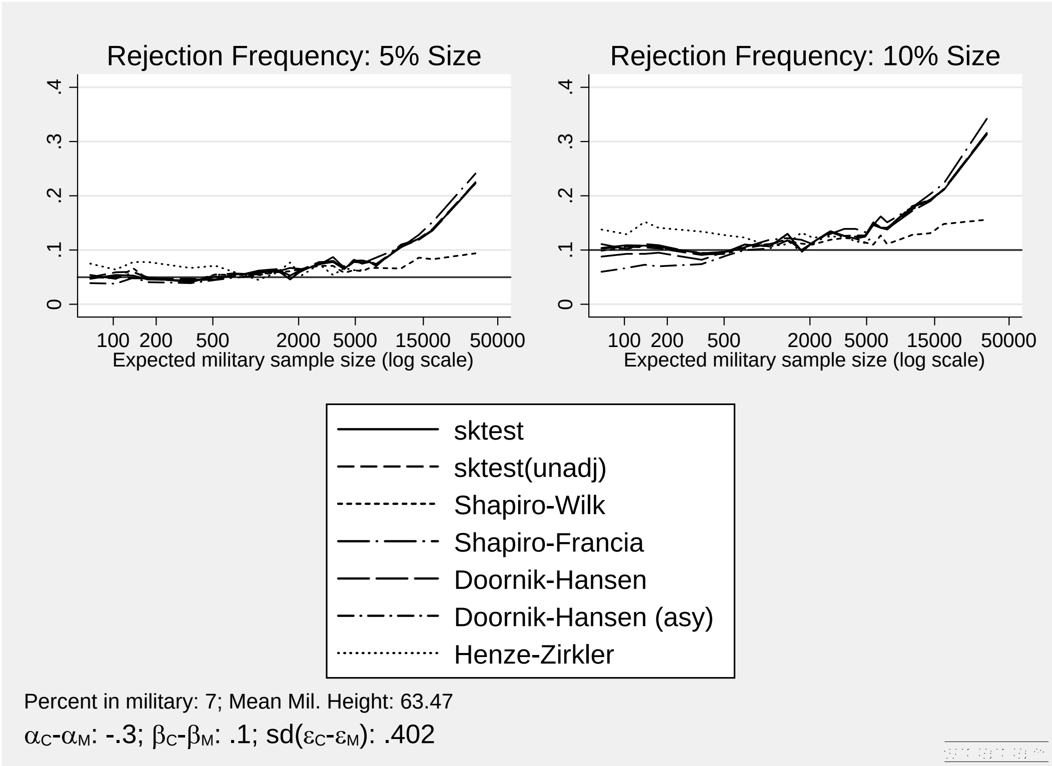
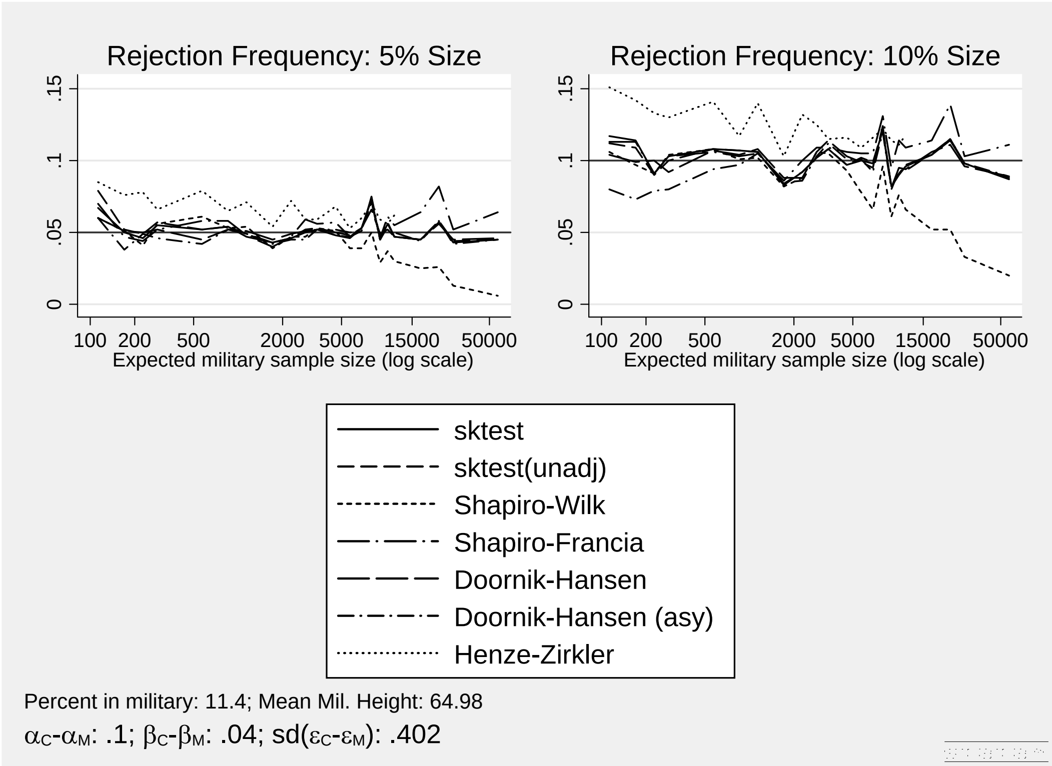


Figure C.2: Tests for normality with moderate selection



Appendix Table B.1: Example values of differences-in-returns R^2

Example	β_C	$Corr(\varepsilon_C,\varepsilon_M)$	$sd(\varepsilon_M)$	$sd(\varepsilon_C)$	R^2
1	.02	.1	.2	.2	.0335
2	.02	.1	.2	.4	.0134
3	.02	.2	.2	.2	.0376
4	.04	.1	.2	.2	.1219
5	.04	.1	.2	.4	.0515
6	.04	.2	.2	.2	.1351
7	.04	.2	.2	.4	.0562
8	.06	.1	.2	.2	.2381
9	.06	.1	.2	.4	.1089
10	.06	.2	.2	.2	.2601

All examples assume $\beta_M=0$

Table D.1: Identification of selection with additional information

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
		RMSLE model		Conditional on Military Selection (COMS) model					
	Truth (evaluated at observed sample)	Without selection controls	With selection controls in mean height	250 iterations, start at true values	250 iterations, start at true values	250 iterations, start at various values	250 iterations, start at true values, except rho=-0.1	250 iterations, start at true values, except rho=-0.30	Unconditional Selection Model (two iterations, start at true values)
Mean height	66.00	64.92	64.90	67.02	67.85	68.09	64.80	62.01	66.03
Height shift effect	1.00	0.95	0.95	1.01	0.97	0.91	0.95	1.57	1.00
Height effect in selection	0.094	NA	NA	0.139	0.034	-0.051	-0.132	-0.510	0.093
Pred. frac. prefering military	0.0815	NA	NA	1.04E-06	1.0100E-06	1.1083E-03	0.2637	0.0031	0.0814865
Pred. frac. in mil. (prefer mil & H>=min ht)	0.0502	NA		6.05E-07	6.2700E-07	7.1300E-04	0.1638	0.0023	0.0502208
Selection and height error correlation	0.2357	NA	NA	0.1697	0.2357	0.3645	-0.0146	-0.3259	0.2409
Correlation estimated or constrained	constrained	NA	NA	estimated	constrained	estimated	estimated	estimated	estimated
Convergence? (yes or no)	Evaluated at truth	yes	yes	no	yes	no	no	no	yes
log-likelihood conditional on military sample	-1816557.024	-1816775.594	-1816555.059	-1816555.098	-1816555.076	-1816555.049	-1816553.853	-1816554.119	
log-likelihood in entire sample (N=20,000,000)	-5722678.54								-5722676.178

Table E.1: Descriptive statistics for Army sub-samples used in Table 4

	Columns (1)-(3)				Columns (4)-(6)				Columns (7)-(9)			
	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max
Height	67.45	1.71	65.00	72.00	68.39	1.28	67.00	72.00	68.38	1.29	67.00	72.00
War	0.43	0.50	0.00	1.00	0.45	0.50	0.00	1.00	0.49	0.50	0.00	1.00
Defence	2.14	1.02	0.37	5.02	1.99	1.02	0.37	5.02	2.05	1.03	0.37	5.02
GDP	47.65	35.60	10.37	116.30	41.86	33.66	10.37	116.30	41.40	34.06	10.37	116.30
Wage	27.89	9.43	13.42	50.92	26.09	9.23	13.42	50.92	26.17	9.33	13.42	50.92
Age at recruitment	23.58	1.50	22.00	27.00	23.52	1.46	22.00	27.00	24.26	1.23	23.00	27.00
Year recruited	1825.75	34.01	1760.00	1879.00	1819.48	34.02	1760.00	1879.00	1819.01	34.05	1760.00	1879.00
Birth year	1802.16	34.16	1733.00	1857.00	1795.96	34.14	1733.00	1857.00	1794.74	34.17	1733.00	1856.00

Table E.2: OLS estimates for model summarized in Table 4

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Col. in Table 4	c1	c2	c3	c4	c5	c6	c7	c8	c9
War		-0.151			0.0194			0.295	
		(0.0750)			(0.0765)			(0.102)	
Defence		-0.0715			0.0141			-0.0487	
		(0.0315)			(0.0325)			(0.0411)	
GDP		0.00285			0.0203			0.0101	
		(0.00995)			(0.0106)			(0.0147)	
Wages		0.000280			-4.31e-05			0.0260	
		(0.0161)			(0.0178)			(0.0244)	
age_23	-0.0942	-0.0927		-0.0518	-0.0714				
	(0.0400)	(0.0411)		(0.0404)	(0.0414)				
age_24	0.0539	0.0502		0.0867	0.0487		0.134	0.126	
	(0.0432)	(0.0483)		(0.0430)	(0.0479)		(0.0468)	(0.0491)	
age_25	-0.142	-0.146		-0.0210	-0.0737		0.0513	0.0256	
	(0.0530)	(0.0612)		(0.0524)	(0.0598)		(0.0558)	(0.0630)	
age_26	-0.419	-0.428		-0.0546	-0.129		0.0142	-0.0287	
	(0.0645)	(0.0787)		(0.0678)	(0.0804)		(0.0713)	(0.0852)	
age_27	-0.437	-0.448		-0.123	-0.216		-0.0643	-0.119	
	(0.0683)	(0.0863)		(0.0714)	(0.0870)		(0.0764)	(0.0977)	
ry_1760			3.593			1.555			1.645
			(1.009)			(0.891)			(1.418)
ry_1761			4.157			1.959			2.188
			(1.012)			(0.904)			(1.444)
ry_1762			3.417			1.105			1.261
			(1.006)			(0.898)			(1.471)
ry_1763			4.204			1.825			1.794
			(0.982)			(0.868)			(1.383)
ry_1764			4.230			1.868			1.544
			(0.971)			(0.857)			(1.364)
ry_1765			3.766			1.549			1.285
			(0.962)			(0.845)			(1.349)
ry_1766			4.080			1.684			1.355
			(0.958)			(0.839)			(1.351)
ry_1767			4.388			2.047			1.795
			(0.989)			(0.874)			(1.341)
ry_1768			4.147			1.925			0.969
			(0.940)			(0.832)			(1.270)
ry_1769			4.102			1.838			0.998
			(0.927)			(0.816)			(1.263)
ry_1770			2.604			0.976			-0.0577
			(0.901)			(0.801)			(1.238)
ry_1771			3.896			2.217			1.676
			(0.903)			(0.788)			(1.159)
ry_1772			3.218			1.128			0.738
			(0.799)			(0.719)			(1.070)
ry_1773			3.408			1.157			0.427
			(0.692)			(0.632)			(0.984)
ry_1774			3.842			1.867			1.339
			(0.678)			(0.623)			(1.003)
ry_1775			3.320			1.519			1.369
			(0.657)			(0.602)			(0.939)
ry_1776			3.085			1.611			2.001
			(0.647)			(0.588)			(0.938)
ry_1777			2.382			0.809			1.065
			(0.615)			(0.560)			(0.897)
ry_1778			1.170			0.303			0.463

	(0.613)	(0.556)	(0.888)
ry_1779	1.959	0.745	0.878
	(0.608)	(0.553)	(0.889)
ry_1780	1.913	0.986	1.200
	(0.608)	(0.552)	(0.871)
ry_1781	2.275	1.106	1.263
	(0.609)	(0.554)	(0.861)
ry_1782	2.115	1.208	1.536
	(0.580)	(0.530)	(0.842)
ry_1783	1.990	1.196	1.396
	(0.594)	(0.536)	(0.847)
ry_1784	2.577	1.347	1.371
	(0.535)	(0.485)	(0.786)
ry_1785	2.890	1.609	1.328
	(0.516)	(0.470)	(0.763)
ry_1786	3.327	1.940	1.881
	(0.506)	(0.469)	(0.748)
ry_1787	2.589	1.337	1.280
	(0.472)	(0.440)	(0.724)
ry_1788	2.799	1.731	1.692
	(0.461)	(0.428)	(0.720)
ry_1789	2.115	1.148	1.612
	(0.426)	(0.389)	(0.641)
ry_1790	1.850	0.968	0.717
	(0.348)	(0.329)	(0.559)
ry_1791	0.486	0.177	0.746
	(0.944)	(0.958)	(0.945)
ry_1792	0.379	-0.189	-1.100
	(0.529)	(0.474)	(0.454)
ry_1793	0.757	0.0339	0.0973
	(0.262)	(0.254)	(0.340)
ry_1794	0.00854	-0.0305	0.0976
	(0.251)	(0.241)	(0.324)
ry_1795	-0.104	-0.0794	-0.125
	(0.233)	(0.219)	(0.314)
ry_1796	-0.0430	-0.301	-0.286
	(0.216)	(0.213)	(0.286)
ry_1797	-0.0556	0.120	0.0401
	(0.234)	(0.231)	(0.312)
ry_1798	0.610	0.178	0.0574
	(0.248)	(0.241)	(0.284)
ry_1799	0.242	0.188	0.121
	(0.209)	(0.190)	(0.222)
ry_1801	-0.0223	-0.133	-0.120
	(0.182)	(0.172)	(0.211)
ry_1802	-0.337	0.0181	-0.132
	(0.277)	(0.309)	(0.306)
ry_1803	-0.515	-0.367	-0.162
	(0.188)	(0.194)	(0.263)
ry_1804	-0.339	-0.207	0.0107
	(0.237)	(0.234)	(0.318)
ry_1805	-0.677	-0.217	0.00800
	(0.209)	(0.212)	(0.295)
ry_1806	-1.071	-0.326	-0.0678
	(0.267)	(0.269)	(0.365)
ry_1807	-0.826	-0.402	-0.0340
	(0.262)	(0.252)	(0.341)
ry_1808	-0.548	-0.432	-0.251
	(0.306)	(0.297)	(0.392)
ry_1809	-0.834	-0.229	-0.227

	(0.321)	(0.320)	(0.422)
ry_1810	-1.304	-0.537	-0.435
	(0.315)	(0.308)	(0.412)
ry_1811	-1.937	-0.573	-0.387
	(0.321)	(0.348)	(0.457)
ry_1812	-1.478	-0.548	-0.590
	(0.316)	(0.316)	(0.426)
ry_1813	-1.789	-0.472	-0.411
	(0.324)	(0.332)	(0.445)
ry_1814	-1.955	-0.521	-0.613
	(0.362)	(0.370)	(0.492)
ry_1815	-1.846	-0.645	-0.755
	(0.342)	(0.346)	(0.473)
ry_1816	-1.702	-0.635	-0.917
	(0.345)	(0.357)	(0.491)
ry_1817	-1.877	-0.820	-1.057
	(0.353)	(0.364)	(0.500)
ry_1818	-1.690	-0.519	-0.786
	(0.369)	(0.373)	(0.506)
ry_1819	-0.623	-0.0706	-0.465
	(0.428)	(0.424)	(0.562)
ry_1820	-1.170	-0.144	-0.608
	(0.431)	(0.429)	(0.574)
ry_1821	-1.404	-0.325	-0.703
	(0.492)	(0.520)	(0.690)
ry_1822	-0.935	0.180	-0.264
	(0.519)	(0.556)	(0.792)
ry_1823	-1.117	-0.0540	-0.424
	(0.500)	(0.517)	(0.707)
ry_1824	-1.609	-0.363	-0.884
	(0.500)	(0.529)	(0.715)
ry_1825	-1.360	0.145	-0.489
	(0.531)	(0.556)	(0.754)
ry_1826	-1.834	-0.178	-0.901
	(0.518)	(0.547)	(0.746)
ry_1827	-1.755	-0.472	-1.370
	(0.557)	(0.584)	(0.780)
ry_1828	-0.549	0.113	-1.030
	(0.631)	(0.651)	(0.880)
ry_1829	-0.652	0.343	-0.356
	(0.822)	(0.851)	(1.548)
ry_1830	-1.224	0.629	-0.491
	(0.665)	(0.710)	(1.046)
ry_1831	-2.182	-0.0441	-1.123
	(0.602)	(0.648)	(0.898)
ry_1832	-1.684	0.215	-1.021
	(0.676)	(0.721)	(0.966)
ry_1833	-1.579	0.350	1.273
	(0.854)	(0.850)	(1.041)
ry_1834	-0.855	1.008	0.999
	(0.916)	(0.869)	(1.161)
ry_1835	-2.159	0.417	
	(1.082)	(1.063)	
ry_1836	-2.440	0.113	-0.800
	(1.087)	(1.063)	(0.532)
ry_1837	-2.809	0.00248	-1.066
	(1.111)	(1.098)	(0.875)
ry_1838	-2.729	0.117	-0.746
	(1.139)	(1.125)	(0.949)
ry_1839	-3.409	-0.239	-1.345

	(1.137)	(1.123)	(0.989)
ry_1840	-2.523	0.0906	-0.537
	(1.182)	(1.177)	(1.132)
ry_1841	-1.807	0.0979	-0.460
	(1.200)	(1.197)	(1.292)
ry_1842	-1.358	0.467	0.134
	(1.226)	(1.220)	(1.403)
ry_1843	-1.637	0.743	0.0882
	(1.344)	(1.335)	(1.540)
ry_1844	-0.432	1.118	-0.245
	(1.518)	(1.514)	(2.115)
ry_1845	-2.166	-0.308	-1.848
	(1.446)	(1.428)	(2.654)
ry_1846	-3.063	-1.402	-2.239
	(1.447)	(1.412)	(2.603)
ry_1847	-3.314	-1.386	-1.567
	(1.571)	(1.534)	(2.704)
ry_1848	-3.417	-1.762	-2.958
	(1.515)	(1.471)	(2.647)
ry_1849	-4.021	-1.647	-2.922
	(1.608)	(1.560)	(2.794)
ry_1850	-3.459	-1.372	-1.003
	(1.719)	(1.714)	(2.811)
ry_1851	-3.328	-1.810	-4.336
	(1.610)	(1.596)	(2.866)
ry_1852	-4.220	-2.147	-4.493
	(1.574)	(1.559)	(2.794)
ry_1853	-4.351	-2.146	-4.420
	(1.589)	(1.574)	(2.820)
ry_1854	-5.136	-2.321	-4.537
	(1.584)	(1.570)	(2.818)
ry_1855	-5.049	-1.998	-4.151
	(1.584)	(1.570)	(2.816)
ry_1856	-5.141	-1.880	-4.114
	(1.587)	(1.574)	(2.822)
ry_1857	-5.802	-2.260	-4.437
	(1.591)	(1.579)	(2.826)
ry_1858	-5.004	-2.158	-4.235
	(1.592)	(1.579)	(2.826)
ry_1859	-5.694	-2.361	-4.624
	(1.594)	(1.583)	(2.829)
ry_1860	-5.991	-2.461	-4.944
	(1.595)	(1.584)	(2.830)
ry_1862	-6.037	-2.630	-5.213
	(1.608)	(1.596)	(2.863)
ry_1863	-6.105	-2.529	-5.416
	(1.608)	(1.600)	(2.860)
ry_1864	-6.327	-2.813	-5.620
	(1.610)	(1.596)	(2.868)
ry_1865	-5.418	-2.279	-5.017
	(1.612)	(1.601)	(2.876)
ry_1866	-6.473	-2.493	-4.997
	(1.623)	(1.617)	(2.895)
ry_1867	-6.114	-1.995	-3.964
	(1.642)	(1.665)	(2.989)
ry_1868	-6.176	-2.460	-4.626
	(1.631)	(1.639)	(2.954)
ry_1869	-5.698	-2.004	-4.038
	(1.637)	(1.645)	(2.961)
ry_1870	-5.404	-1.620	-3.420

			(1.634)			(1.642)			(2.959)
ry_1871			-6.198			-2.041			-3.928
			(1.636)			(1.645)			(2.962)
ry_1872			-6.085			-2.198			-4.071
			(1.638)			(1.647)			(2.964)
ry_1873			-6.071			-1.882			-3.420
			(1.643)			(1.655)			(2.971)
ry_1874			-5.833			-1.753			-3.621
			(1.641)			(1.650)			(2.967)
ry_1875			-5.306			-1.581			-3.381
			(1.645)			(1.657)			(2.972)
ry_1876			-5.757			-1.655			-3.300
			(1.648)			(1.661)			(2.979)
ry_1877			-5.826			-1.771			-3.190
			(1.654)			(1.686)			(3.005)
ry_1878			-5.034			-0.923			-1.867
			(1.668)			(1.698)			(3.013)
ry_1879			-4.855			-1.035			-1.924
			(1.669)			(1.700)			(3.023)
by_1733	2.020	2.125	-3.197	0.958	1.429	-0.716	1.030	1.651	-1.338
	(0.550)	(0.617)	(1.242)	(0.551)	(0.628)	(1.161)	(0.559)	(0.701)	(1.672)
by_1734	1.729	1.846	-3.809	0.646	1.105	-1.234	0.714	1.337	-1.938
	(0.273)	(0.388)	(1.147)	(0.272)	(0.405)	(1.061)	(0.285)	(0.506)	(1.614)
by_1735	2.197	2.306	-3.108	1.216	1.654	-0.554	1.277	1.882	-1.240
	(0.426)	(0.503)	(1.165)	(0.408)	(0.499)	(1.084)	(0.417)	(0.578)	(1.622)
by_1736	1.661	1.754	-3.635	0.744	1.184	-0.984	0.820	1.407	-1.564
	(0.259)	(0.367)	(1.125)	(0.254)	(0.381)	(1.036)	(0.268)	(0.478)	(1.574)
by_1737	1.132	1.217	-4.148	0.848	1.275	-0.921	0.921	1.495	-1.470
	(0.431)	(0.498)	(1.139)	(0.279)	(0.395)	(1.034)	(0.291)	(0.482)	(1.572)
by_1738	1.536	1.616	-3.545	0.704	1.107	-0.937	0.859	1.420	-1.419
	(0.285)	(0.374)	(1.121)	(0.287)	(0.390)	(1.032)	(0.372)	(0.540)	(1.556)
by_1739	2.140	2.183	-3.255	1.222	1.651	-0.498	1.329	1.863	-0.732
	(0.258)	(0.358)	(1.096)	(0.254)	(0.372)	(1.012)	(0.288)	(0.480)	(1.537)
by_1740	1.926	1.976	-3.302	1.062	1.473	-0.527	1.383	1.904	-0.763
	(0.245)	(0.339)	(1.105)	(0.238)	(0.353)	(1.020)	(0.275)	(0.457)	(1.551)
by_1741	1.823	1.840	-3.523	0.977	1.394	-0.753	1.101	1.601	-0.928
	(0.212)	(0.318)	(1.082)	(0.206)	(0.331)	(0.995)	(0.232)	(0.431)	(1.521)
by_1742	1.404	1.408	-3.959	0.714	1.127	-1.064	0.716	1.207	-1.241
	(0.258)	(0.349)	(1.085)	(0.216)	(0.334)	(0.994)	(0.245)	(0.434)	(1.500)
by_1743	1.817	1.819	-3.434	1.007	1.422	-0.731	1.207	1.698	-0.480
	(0.257)	(0.347)	(1.072)	(0.247)	(0.355)	(0.991)	(0.312)	(0.484)	(1.462)
by_1744	1.681	1.676	-3.630	1.135	1.546	-0.672	1.268	1.749	-0.438
	(0.297)	(0.374)	(1.086)	(0.258)	(0.361)	(0.980)	(0.312)	(0.481)	(1.444)
by_1745	2.093	2.083	-3.122	1.376	1.775	-0.449	1.576	2.044	0.0994
	(0.257)	(0.340)	(1.072)	(0.241)	(0.343)	(0.986)	(0.277)	(0.440)	(1.452)
by_1746	2.178	2.167	-2.742	1.475	1.864	-0.194	1.350	1.813	0.131
	(0.261)	(0.342)	(1.051)	(0.252)	(0.349)	(0.975)	(0.297)	(0.453)	(1.448)
by_1747	1.390	1.379	-3.404	0.899	1.278	-0.743	0.776	1.235	-0.330
	(0.250)	(0.328)	(1.048)	(0.218)	(0.319)	(0.965)	(0.373)	(0.501)	(1.400)
by_1748	0.877	0.872	-3.314	0.589	0.968	-0.685	0.788	1.230	-0.997
	(0.310)	(0.378)	(0.981)	(0.265)	(0.353)	(0.926)	(0.341)	(0.472)	(1.314)
by_1749	1.664	1.676	-3.152	1.125	1.505	-0.488	1.144	1.537	-0.386
	(0.458)	(0.512)	(0.958)	(0.373)	(0.444)	(0.892)	(0.444)	(0.551)	(1.226)
by_1750	2.070	2.104	-2.480	1.535	1.914	0.234	1.728	2.102	0.192
	(0.325)	(0.385)	(0.883)	(0.271)	(0.363)	(0.819)	(0.282)	(0.442)	(1.189)
by_1751	0.916	1.047	-2.969	0.810	1.188	-0.270	0.742	0.939	-1.044
	(0.279)	(0.362)	(0.820)	(0.253)	(0.353)	(0.796)	(0.308)	(0.482)	(1.165)
by_1752	0.874	1.029	-2.858	0.391	0.759	-0.625	0.301	0.473	-1.385
	(0.211)	(0.316)	(0.799)	(0.199)	(0.322)	(0.775)	(0.218)	(0.419)	(1.148)
by_1753	0.833	1.012	-2.548	0.399	0.755	-0.405	0.370	0.537	-1.203

	(0.204)	(0.310)	(0.797)	(0.195)	(0.316)	(0.772)	(0.215)	(0.408)	(1.145)
by_1754	0.769	0.951	-2.582	0.274	0.616	-0.560	0.215	0.380	-1.276
	(0.192)	(0.297)	(0.793)	(0.184)	(0.304)	(0.768)	(0.205)	(0.396)	(1.140)
by_1755	0.536	0.724	-2.572	0.160	0.490	-0.540	0.234	0.412	-1.173
	(0.184)	(0.288)	(0.788)	(0.175)	(0.295)	(0.763)	(0.206)	(0.397)	(1.131)
by_1756	0.201	0.400	-2.605	0.262	0.586	-0.360	0.625	0.814	-1.119
	(0.214)	(0.305)	(0.791)	(0.206)	(0.313)	(0.766)	(0.285)	(0.448)	(1.126)
by_1757	0.488	0.696	-2.834	0.0793	0.403	-0.900	0.246	0.451	-1.688
	(0.208)	(0.302)	(0.778)	(0.200)	(0.309)	(0.756)	(0.249)	(0.422)	(1.105)
by_1758	0.459	0.672	-2.875	0.173	0.492	-0.943	0.340	0.554	-1.680
	(0.225)	(0.312)	(0.775)	(0.203)	(0.310)	(0.752)	(0.243)	(0.409)	(1.102)
by_1759	0.946	1.127	-2.599	0.624	0.947	-0.648	0.666	0.937	-1.437
	(0.262)	(0.333)	(0.765)	(0.244)	(0.331)	(0.748)	(0.279)	(0.420)	(1.092)
by_1760	1.081	1.244	-2.404	0.773	1.086	-0.482	1.115	1.504	-0.834
	(0.246)	(0.316)	(0.758)	(0.227)	(0.317)	(0.741)	(0.300)	(0.434)	(1.079)
by_1761	1.423	1.485	-2.425	0.897	1.225	-0.535	1.053	1.496	-0.981
	(0.227)	(0.300)	(0.730)	(0.207)	(0.299)	(0.714)	(0.228)	(0.378)	(1.044)
by_1762	1.920	1.951	-2.052	1.317	1.637	-0.172	1.643	2.082	-0.467
	(0.210)	(0.282)	(0.726)	(0.197)	(0.288)	(0.707)	(0.247)	(0.392)	(1.030)
by_1763	1.861	1.875	-2.149	1.231	1.552	-0.289	1.112	1.526	-0.918
	(0.214)	(0.289)	(0.701)	(0.200)	(0.292)	(0.693)	(0.252)	(0.397)	(1.008)
by_1764	1.750	1.730	-2.342	1.022	1.343	-0.579	0.927	1.329	-1.064
	(0.206)	(0.281)	(0.697)	(0.199)	(0.288)	(0.694)	(0.236)	(0.377)	(1.019)
by_1765	1.427	1.405	-2.360	0.734	1.046	-0.670	0.878	1.279	-1.283
	(0.189)	(0.265)	(0.681)	(0.182)	(0.274)	(0.682)	(0.269)	(0.398)	(1.007)
by_1766	1.695	1.678	-1.981	0.990	1.298	-0.478	0.475	0.861	-1.454
	(0.242)	(0.305)	(0.659)	(0.234)	(0.311)	(0.660)	(0.297)	(0.421)	(0.921)
by_1767	1.272	1.318	-1.550	0.530	0.841	-0.264	0.488	0.696	-0.508
	(0.243)	(0.309)	(0.644)	(0.235)	(0.313)	(0.636)	(0.329)	(0.452)	(0.847)
by_1768	0.428	0.560	-1.651	0.236	0.542	-0.148	-0.189	-0.0492	-0.823
	(0.233)	(0.304)	(0.574)	(0.220)	(0.303)	(0.590)	(0.246)	(0.399)	(0.790)
by_1769	0.802	0.985	-0.777	0.262	0.560	0.313	0.332	0.437	-0.264
	(0.225)	(0.303)	(0.566)	(0.214)	(0.304)	(0.584)	(0.228)	(0.376)	(0.782)
by_1770	0.624	0.821	-0.836	0.381	0.664	0.452	0.431	0.527	-0.170
	(0.212)	(0.286)	(0.559)	(0.197)	(0.284)	(0.576)	(0.216)	(0.359)	(0.774)
by_1771	0.352	0.563	-1.163	0.219	0.485	0.238	0.267	0.373	-0.300
	(0.215)	(0.283)	(0.558)	(0.207)	(0.286)	(0.578)	(0.244)	(0.370)	(0.776)
by_1772	0.247	0.481	-1.061	0.165	0.421	0.209	0.373	0.478	-0.143
	(0.216)	(0.280)	(0.552)	(0.207)	(0.281)	(0.571)	(0.239)	(0.361)	(0.767)
by_1773	0.161	0.425	-1.204	0.00800	0.264	0.00353	0.0880	0.189	-0.524
	(0.210)	(0.275)	(0.535)	(0.202)	(0.277)	(0.557)	(0.240)	(0.365)	(0.742)
by_1774	-0.0977	0.169	-1.491	-0.110	0.138	-0.0633	0.134	0.220	-0.512
	(0.216)	(0.273)	(0.538)	(0.217)	(0.281)	(0.566)	(0.273)	(0.374)	(0.752)
by_1775	0.157	0.429	-1.241	-0.0947	0.146	-0.156	-0.0770	-0.00855	-0.688
	(0.217)	(0.273)	(0.529)	(0.213)	(0.276)	(0.554)	(0.248)	(0.355)	(0.731)
by_1776	-0.0883	0.165	-1.390	-0.00703	0.234	0.0123	-0.00457	0.0426	-0.574
	(0.218)	(0.272)	(0.520)	(0.212)	(0.271)	(0.548)	(0.233)	(0.335)	(0.725)
by_1777	0.331	0.582	-0.889	0.242	0.472	0.274	0.242	0.275	-0.310
	(0.209)	(0.261)	(0.512)	(0.202)	(0.260)	(0.538)	(0.229)	(0.327)	(0.710)
by_1778	-0.273	-0.0218	-1.335	-0.265	-0.0374	-0.140	-0.282	-0.294	-0.819
	(0.190)	(0.242)	(0.495)	(0.182)	(0.240)	(0.521)	(0.214)	(0.312)	(0.687)
by_1779	0.0521	0.306	-0.851	0.0831	0.312	0.271	0.0939	0.0598	-0.469
	(0.215)	(0.261)	(0.495)	(0.211)	(0.261)	(0.523)	(0.246)	(0.333)	(0.685)
by_1780	0.276	0.528	-0.437	0.293	0.519	0.539	0.253	0.192	-0.325
	(0.217)	(0.261)	(0.490)	(0.211)	(0.260)	(0.513)	(0.212)	(0.298)	(0.675)
by_1781	-0.207	0.0454	-0.792	-0.330	-0.116	-0.0114	-0.211	-0.288	-0.771
	(0.200)	(0.242)	(0.480)	(0.184)	(0.233)	(0.500)	(0.213)	(0.298)	(0.662)
by_1782	0.162	0.421	-0.421	0.0240	0.228	0.290	0.0936	0.00269	-0.499
	(0.208)	(0.248)	(0.476)	(0.197)	(0.242)	(0.499)	(0.221)	(0.300)	(0.657)
by_1783	0.0288	0.303	-0.432	0.0469	0.240	0.358	-0.0401	-0.141	-0.484

	(0.212)	(0.250)	(0.466)	(0.211)	(0.251)	(0.497)	(0.259)	(0.331)	(0.653)
by_1784	-0.313	-0.0270	-0.595	-0.0284	0.163	0.364	0.0257	-0.0949	-0.411
	(0.214)	(0.255)	(0.455)	(0.200)	(0.243)	(0.484)	(0.225)	(0.303)	(0.638)
by_1785	-0.0952	0.200	-0.329	-0.119	0.0683	0.290	0.0980	-0.0240	-0.164
	(0.210)	(0.250)	(0.449)	(0.204)	(0.246)	(0.481)	(0.243)	(0.315)	(0.629)
by_1786	-0.123	0.189	-0.162	0.0456	0.226	0.468	0.134	-0.00419	-0.114
	(0.222)	(0.264)	(0.440)	(0.213)	(0.258)	(0.466)	(0.246)	(0.327)	(0.612)
by_1787	-0.610	-0.281	-0.404	-0.296	-0.125	0.181	-0.303	-0.461	-0.438
	(0.204)	(0.251)	(0.426)	(0.197)	(0.246)	(0.453)	(0.215)	(0.303)	(0.598)
by_1788	-0.595	-0.252	-0.214	-0.358	-0.200	0.186	-0.282	-0.452	-0.363
	(0.193)	(0.242)	(0.418)	(0.193)	(0.245)	(0.446)	(0.215)	(0.304)	(0.586)
by_1789	-0.364	-0.0244	0.116	0.0309	0.175	0.585	0.205	0.0705	0.191
	(0.199)	(0.245)	(0.416)	(0.200)	(0.250)	(0.445)	(0.231)	(0.311)	(0.580)
by_1790	-0.252	0.0726	0.161	-0.0177	0.114	0.542	0.0124	-0.0789	0.0633
	(0.206)	(0.247)	(0.411)	(0.212)	(0.254)	(0.440)	(0.246)	(0.312)	(0.564)
by_1791	-0.428	-0.140	0.0673	-0.291	-0.168	0.273	-0.162	-0.169	0.0516
	(0.188)	(0.228)	(0.398)	(0.182)	(0.227)	(0.425)	(0.217)	(0.279)	(0.551)
by_1792	-0.470	-0.248	0.0595	-0.252	-0.127	0.422	-0.179	-0.148	0.138
	(0.199)	(0.231)	(0.399)	(0.206)	(0.238)	(0.428)	(0.226)	(0.279)	(0.550)
by_1793	-0.383	-0.191	0.0988	-0.298	-0.182	0.330	-0.157	-0.0954	0.138
	(0.199)	(0.229)	(0.391)	(0.195)	(0.229)	(0.419)	(0.221)	(0.270)	(0.541)
by_1794	-0.179	-0.0656	0.231	-0.273	-0.163	0.314	-0.195	-0.171	0.0346
	(0.199)	(0.224)	(0.388)	(0.198)	(0.225)	(0.415)	(0.237)	(0.273)	(0.536)
by_1795	-0.204	-0.165	0.190	-0.108	0.000414	0.426	0.230	0.241	0.294
	(0.231)	(0.246)	(0.395)	(0.235)	(0.251)	(0.420)	(0.294)	(0.317)	(0.547)
by_1796	-0.261	-0.244	-0.0421	-0.118	-0.0217	0.254	0.459	0.480	0.441
	(0.266)	(0.275)	(0.402)	(0.254)	(0.266)	(0.430)	(0.417)	(0.426)	(0.559)
by_1797	-0.0766	-0.0643	-0.161	-0.0782	0.00307	0.108	-0.0784	-0.0600	-0.0436
	(0.265)	(0.270)	(0.328)	(0.274)	(0.281)	(0.365)	(0.302)	(0.315)	(0.447)
by_1798	0.141	0.150	0.0752	-0.0390	0.0141	0.0611	0.0378	0.0750	-0.0566
	(0.236)	(0.239)	(0.259)	(0.240)	(0.242)	(0.266)	(0.271)	(0.275)	(0.296)
by_1799	0.293	0.300	0.347	0.0304	0.0564	0.155	0.237	0.261	0.166
	(0.242)	(0.244)	(0.268)	(0.241)	(0.243)	(0.269)	(0.311)	(0.313)	(0.353)
by_1801	-0.370	-0.377	-0.116	-0.346	-0.383	-0.233	-0.188	-0.215	-0.0294
	(0.206)	(0.207)	(0.211)	(0.206)	(0.207)	(0.216)	(0.240)	(0.242)	(0.258)
by_1802	-0.134	-0.144	0.319	0.0309	-0.0380	0.235	0.239	0.181	0.485
	(0.227)	(0.229)	(0.241)	(0.234)	(0.236)	(0.250)	(0.278)	(0.282)	(0.312)
by_1803	-0.359	-0.370	-0.0161	0.0154	-0.0815	0.0777	0.0226	-0.0540	0.413
	(0.239)	(0.244)	(0.272)	(0.235)	(0.241)	(0.276)	(0.265)	(0.274)	(0.330)
by_1804	-0.332	-0.351	0.224	-0.258	-0.406	-0.119	-0.000546	-0.110	0.465
	(0.218)	(0.228)	(0.267)	(0.211)	(0.222)	(0.271)	(0.313)	(0.336)	(0.469)
by_1805	0.390	0.369	0.788	0.587	0.382	0.708	0.947	0.812	1.285
	(0.375)	(0.384)	(0.431)	(0.420)	(0.427)	(0.469)	(0.590)	(0.605)	(0.702)
by_1806	0.116	0.0833	0.431	0.249	0.00472	0.121	0.268	0.112	0.462
	(0.336)	(0.357)	(0.444)	(0.345)	(0.368)	(0.483)	(0.433)	(0.474)	(0.676)
by_1807	-0.0659	-0.104	0.591	0.00998	-0.272	-0.0764	0.0684	-0.105	0.512
	(0.280)	(0.309)	(0.441)	(0.263)	(0.298)	(0.470)	(0.292)	(0.351)	(0.607)
by_1808	-0.0541	-0.0953	0.553	-0.0624	-0.365	-0.312	-0.204	-0.379	0.211
	(0.224)	(0.265)	(0.411)	(0.208)	(0.257)	(0.453)	(0.249)	(0.325)	(0.592)
by_1809	-0.425	-0.468	0.477	-0.258	-0.585	-0.266	-0.0646	-0.216	0.101
	(0.220)	(0.268)	(0.406)	(0.232)	(0.283)	(0.452)	(0.508)	(0.547)	(0.766)
by_1810	0.335	0.298	0.724	-0.0121	-0.331	-0.313	-0.0635	-0.233	-1.945
	(0.256)	(0.297)	(0.551)	(0.243)	(0.291)	(0.562)	(0.446)	(0.496)	(0.838)
by_1811	0.625	0.583	0.850	0.351	0.0126	-0.186	1.066	0.887	-0.593
	(0.293)	(0.332)	(0.732)	(0.278)	(0.324)	(0.698)	(0.563)	(0.614)	(1.071)
by_1812	1.162	1.101	1.488	0.573	0.189	-0.119	0.299	0.0798	-0.0931
	(0.371)	(0.414)	(0.879)	(0.345)	(0.397)	(0.832)	(0.445)	(0.529)	(0.896)
by_1813	0.251	0.182	1.598	-0.269	-0.680	-0.390	-0.251	-0.543	0.107
	(0.313)	(0.367)	(0.979)	(0.315)	(0.377)	(0.945)	(0.385)	(0.475)	(1.024)
by_1814	-0.260	-0.331	1.230	-0.236	-0.666	-0.284	-0.165	-0.452	0.202

	(0.224)	(0.299)	(1.002)	(0.230)	(0.314)	(0.967)	(0.273)	(0.397)	(1.152)
by_1815	-0.612	-0.685	1.257	-0.314	-0.774	-0.227	-0.358	-0.662	0.203
	(0.226)	(0.315)	(1.028)	(0.256)	(0.347)	(1.004)	(0.307)	(0.456)	(1.230)
by_1816	-0.529	-0.605	1.306	-0.111	-0.610	-0.0593	-0.0600	-0.360	0.470
	(0.225)	(0.326)	(1.041)	(0.240)	(0.347)	(1.014)	(0.298)	(0.460)	(1.248)
by_1817	-0.415	-0.488	1.356	0.408	-0.0862	0.477	0.510	0.258	0.394
	(0.257)	(0.352)	(1.059)	(0.284)	(0.373)	(1.055)	(0.349)	(0.471)	(1.406)
by_1818	0.280	0.221	1.112	0.352	-0.113	0.180	0.524	0.229	0.175
	(0.244)	(0.330)	(1.098)	(0.225)	(0.328)	(1.081)	(0.290)	(0.416)	(1.502)
by_1819	0.351	0.296	0.767	0.158	-0.297	-0.125	0.409	0.0967	-0.367
	(0.225)	(0.308)	(1.119)	(0.209)	(0.306)	(1.103)	(0.245)	(0.386)	(1.586)
by_1820	0.321	0.270	0.450	0.186	-0.279	-0.348	0.832	0.510	0.352
	(0.267)	(0.342)	(1.150)	(0.248)	(0.334)	(1.128)	(0.797)	(0.856)	(1.852)
by_1821	0.884	0.829	1.230	0.549	0.0423	-0.121	0.476	0.163	0.752
	(0.359)	(0.429)	(1.293)	(0.344)	(0.423)	(1.280)	(0.994)	(1.084)	(2.669)
by_1822	1.503	1.439	0.972	0.828	0.273	-0.134	-0.906	-1.243	0.658
	(0.308)	(0.402)	(1.443)	(0.302)	(0.409)	(1.433)	(0.327)	(0.526)	(2.754)
by_1823	0.281	0.210	1.872	0.0486	-0.580	1.048	-0.564	-0.912	0.940
	(0.533)	(0.610)	(1.455)	(0.517)	(0.610)	(1.393)	(0.450)	(0.632)	(2.732)
by_1824	0.524	0.458	2.611	0.0406	-0.627	1.620	0.0855	-0.263	1.908
	(0.514)	(0.603)	(1.463)	(0.488)	(0.596)	(1.400)	(0.657)	(0.803)	(2.773)
by_1825	-0.132	-0.196	2.169	-0.429	-1.102	1.153	-0.246	-0.608	2.315
	(0.369)	(0.487)	(1.494)	(0.377)	(0.505)	(1.455)	(0.606)	(0.770)	(2.868)
by_1826	0.0631	0.00352	2.535	-0.0573	-0.762	1.710	0.0944	-0.285	2.886
	(0.476)	(0.580)	(1.514)	(0.468)	(0.586)	(1.450)	(0.751)	(0.892)	(2.908)
by_1827	-0.235	-0.310	2.653	-0.511	-1.239	1.431	-0.367	-0.767	3.110
	(0.368)	(0.507)	(1.514)	(0.367)	(0.530)	(1.492)	(0.429)	(0.680)	(2.902)
by_1828	-0.197	-0.162	3.056	-0.235	-1.004	1.727	-0.170	-0.725	3.505
	(0.249)	(0.435)	(1.518)	(0.236)	(0.447)	(1.494)	(0.259)	(0.586)	(2.898)
by_1829	-0.0408	-0.00229	3.293	0.238	-0.565	2.292	0.270	-0.266	3.984
	(0.237)	(0.437)	(1.517)	(0.243)	(0.460)	(1.496)	(0.269)	(0.602)	(2.910)
by_1830	-0.255	-0.220	3.418	0.0609	-0.763	2.213	0.108	-0.480	3.850
	(0.192)	(0.427)	(1.511)	(0.193)	(0.450)	(1.488)	(0.215)	(0.602)	(2.909)
by_1831	-0.567	-0.495	3.193	-0.111	-0.970	1.991	-0.0403	-0.692	3.615
	(0.186)	(0.436)	(1.514)	(0.191)	(0.460)	(1.492)	(0.221)	(0.622)	(2.910)
by_1832	-0.361	-0.284	3.539	-0.147	-1.039	1.950	0.00400	-0.663	3.573
	(0.179)	(0.444)	(1.513)	(0.177)	(0.468)	(1.491)	(0.208)	(0.642)	(2.911)
by_1833	-0.407	-0.283	3.537	0.0868	-0.838	2.122	0.190	-0.386	3.852
	(0.184)	(0.461)	(1.515)	(0.181)	(0.486)	(1.493)	(0.219)	(0.677)	(2.915)
by_1834	-0.295	-0.158	3.763	0.171	-0.798	2.226	-0.0904	-0.616	3.771
	(0.204)	(0.490)	(1.520)	(0.206)	(0.517)	(1.499)	(0.265)	(0.702)	(2.921)
by_1835	-0.690	-0.724	3.651	-0.187	-1.152	2.079	-0.104	-0.665	3.766
	(0.217)	(0.504)	(1.524)	(0.216)	(0.527)	(1.503)	(0.249)	(0.705)	(2.921)
by_1836	-0.674	-0.733	3.618	-0.405	-1.389	1.892	-0.173	-0.785	4.096
	(0.209)	(0.504)	(1.524)	(0.202)	(0.525)	(1.502)	(0.279)	(0.746)	(2.927)
by_1837	-1.037	-1.103	3.575	-0.243	-1.274	2.189	0.143	-0.467	4.709
	(0.240)	(0.537)	(1.532)	(0.308)	(0.598)	(1.525)	(0.454)	(0.858)	(2.977)
by_1838	-1.133	-1.191	3.529	-0.609	-1.684	1.907	-0.644	-1.278	4.041
	(0.200)	(0.539)	(1.530)	(0.195)	(0.567)	(1.509)	(0.222)	(0.778)	(2.951)
by_1839	-0.680	-0.740	3.964	-0.816	-1.934	1.714	-0.742	-1.393	3.877
	(0.240)	(0.572)	(1.545)	(0.214)	(0.595)	(1.519)	(0.229)	(0.798)	(2.953)
by_1840	-0.533	-0.602	4.279	-0.0146	-1.155	2.559	0.135	-0.549	4.874
	(0.290)	(0.596)	(1.550)	(0.368)	(0.670)	(1.541)	(0.411)	(0.882)	(2.969)
by_1841	-0.219	-0.300	4.368	-0.110	-1.296	2.325	-0.0394	-0.733	4.474
	(0.214)	(0.588)	(1.542)	(0.208)	(0.621)	(1.521)	(0.237)	(0.852)	(2.967)
by_1842	-0.421	-0.513	4.188	-0.283	-1.496	2.078	-0.133	-0.844	4.183
	(0.206)	(0.600)	(1.544)	(0.195)	(0.633)	(1.524)	(0.220)	(0.866)	(2.971)
by_1843	-0.387	-0.486	3.975	-0.161	-1.423	1.995	-0.0509	-0.913	3.215
	(0.205)	(0.613)	(1.548)	(0.202)	(0.641)	(1.530)	(0.341)	(0.898)	(3.030)
by_1844	-0.757	-0.858	4.076	-0.467	-1.769	1.725	-0.373	-1.227	3.174

	(0.207)	(0.628)	(1.561)	(0.221)	(0.664)	(1.562)	(0.251)	(0.900)	(3.051)
by_1845	-0.514	-0.623	4.143	-0.280	-1.636	1.795	-0.189	-1.124	3.289
	(0.203)	(0.655)	(1.566)	(0.205)	(0.696)	(1.568)	(0.228)	(0.954)	(3.045)
by_1846	-0.366	-0.488	4.195	-0.0593	-1.472	1.909	0.132	-0.882	3.149
	(0.187)	(0.677)	(1.565)	(0.186)	(0.724)	(1.565)	(0.210)	(1.009)	(3.046)
by_1847	-0.486	-0.631	3.942	-0.00727	-1.491	1.785	0.0615	-0.996	3.027
	(0.193)	(0.717)	(1.567)	(0.192)	(0.756)	(1.567)	(0.214)	(1.047)	(3.048)
by_1848	-0.545	-0.700	3.881	0.0308	-1.497	1.757	-0.0714	-1.248	3.062
	(0.188)	(0.739)	(1.568)	(0.189)	(0.785)	(1.568)	(0.230)	(1.108)	(3.050)
by_1849	-0.941	-1.109	3.838	-0.260	-1.868	1.692	-0.145	-1.360	2.901
	(0.193)	(0.770)	(1.570)	(0.218)	(0.831)	(1.575)	(0.222)	(1.129)	(3.053)
by_1850	-0.924	-1.096	3.712	-0.545	-2.196	1.344	-0.395	-1.677	2.547
	(0.194)	(0.787)	(1.572)	(0.191)	(0.839)	(1.574)	(0.229)	(1.168)	(3.056)
by_1851	-0.888	-1.065	3.621	-0.250	-1.932	1.476	0.0140	-1.289	2.911
	(0.207)	(0.804)	(1.576)	(0.228)	(0.865)	(1.581)	(0.264)	(1.195)	(3.060)
by_1852	-1.153	-1.328	3.375	-0.534	-2.249	1.184	-0.532	-1.854	2.222
	(0.194)	(0.822)	(1.576)	(0.210)	(0.881)	(1.581)	(0.258)	(1.210)	(3.064)
by_1853	-0.805	-0.976	3.546	-0.160	-1.896	1.437	-0.0981	-1.459	2.403
	(0.217)	(0.839)	(1.582)	(0.223)	(0.896)	(1.588)	(0.277)	(1.225)	(3.074)
by_1854	-0.696	-0.853	3.559	-0.0100	-1.752	1.357	0.0646	-1.351	1.908
	(0.230)	(0.856)	(1.589)	(0.253)	(0.919)	(1.614)	(0.280)	(1.251)	(3.092)
by_1855	-0.672	-0.809	3.258	-0.469	-2.203	0.682	-0.345	-1.743	0.987
	(0.199)	(0.838)	(1.596)	(0.207)	(0.895)	(1.622)	(0.236)	(1.214)	(3.106)
by_1856	-1.029	-1.162	2.683	-0.422	-2.169	0.543	-0.341	-1.734	0.924
	(0.207)	(0.853)	(1.603)	(0.224)	(0.913)	(1.628)	(0.298)	(1.235)	(3.117)
by_1857	-0.572	-0.694	3.096	-0.308	-2.052	0.731			
	(0.234)	(0.853)	(1.610)	(0.265)	(0.919)	(1.640)			
Constant	67.62	67.62	68.80	68.37	67.75	68.36	68.23	67.30	68.89
	(0.158)	(0.453)	(0.509)	(0.154)	(0.494)	(0.534)	(0.175)	(0.679)	(0.721)
Observations	13,895	13,895	13,895	9,244	9,244	9,244	6,232	6,232	6,232
R-squared	0.158	0.159	0.192	0.094	0.094	0.119	0.090	0.092	0.120

Table E.3: Descriptive statistics for AMD sub-samples used in Table 5

	Columns (1) - (3)				Columns (4) - (6)			
	Mean	SD	Min	Max	Mean	SD	Min	Max
Height	67.02	1.72	65.00	72.00	68.26	1.34	67.00	72.00
Unemployment	5.20	2.42	2.00	10.20	5.14	2.37	2.00	10.20
GDP	151.95	27.26	114.09	206.54	153.07	27.51	114.09	206.54
Defense	4.91	2.76	2.65	12.68	4.96	2.75	2.65	12.68
Indoor relief	6.32	0.53	5.60	7.80	6.34	0.55	5.60	7.80
War	0.16	0.36	0.00	1.00	0.16	0.36	0.00	1.00
Age at recruitment	23.17	1.05	22.00	25.00	23.19	1.05	22.00	25.00
Year of recruitment	1892.16	9.61	1879.00	1911.00	1892.56	9.70	1879.00	1911.00
Birth year	1868.99	9.63	1854.00	1889.00	1869.37	9.73	1854.00	1889.00

Table E.4: OLS estimates for Table 5

Col in Table 5	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
ue_feinstein		-0.0138			-0.00383			0.00333	
		(0.00299)			(0.00314)			(0.00482)	
defence		-0.00871			-0.0132			-0.0140	
		(0.00418)			(0.00434)			(0.00634)	
id_relief		-0.0602			-0.0162			-0.000760	
		(0.0332)			(0.0346)			(0.0499)	
war		-0.107			-0.0361			-0.0520	
		(0.0272)			(0.0283)			(0.0398)	
gdp		-0.00215			0.00105			0.00376	
		(0.00201)			(0.00210)			(0.00300)	
age_23	0.0457	0.0537		0.00977	0.00918				
	(0.0104)	(0.0117)		(0.0110)	(0.0124)				
age_24	0.107	0.126		0.0425	0.0423		0.0380	0.0300	
	(0.0109)	(0.0153)		(0.0114)	(0.0162)		(0.0118)	(0.0144)	
age_25	0.0928	0.144		0.0302	0.0447		0.0398	0.0309	
	(0.0143)	(0.0226)		(0.0149)	(0.0237)		(0.0155)	(0.0228)	
ry_1879			-1.126			-0.360			-0.473
			(0.117)			(0.122)			(0.218)
ry_1880			-1.161			-0.392			-0.475
			(0.116)			(0.120)			(0.216)
ry_1881			-1.297			-0.326			-0.451
			(0.115)			(0.119)			(0.214)
ry_1882			-1.088			-0.323			-0.415
			(0.113)			(0.117)			(0.210)
ry_1883			-1.122			-0.292			-0.305
			(0.112)			(0.115)			(0.207)
ry_1884			-1.061			-0.295			-0.320
			(0.110)			(0.114)			(0.204)
ry_1885			-1.078			-0.347			-0.344
			(0.109)			(0.112)			(0.202)
ry_1886			-0.917			-0.309			-0.288
			(0.107)			(0.110)			(0.199)
ry_1887			-0.737			-0.268			-0.306
			(0.105)			(0.108)			(0.195)
ry_1888			-0.699			-0.225			-0.266
			(0.102)			(0.105)			(0.187)
ry_1889			-0.589			-0.238			-0.265
			(0.0963)			(0.0986)			(0.176)
ry_1890			-0.516			-0.157			-0.213
			(0.0911)			(0.0931)			(0.167)
ry_1891			-0.587			-0.165			-0.139
			(0.0850)			(0.0874)			(0.154)
ry_1892			-0.468			-0.170			-0.185
			(0.0792)			(0.0810)			(0.145)
ry_1893			-0.452			-0.118			-0.172
			(0.0742)			(0.0758)			(0.135)
ry_1894			-0.260			-0.0239			-0.0434
			(0.0689)			(0.0701)			(0.124)
ry_1895			-0.0878			-0.0193			0.0116

Table E.4: OLS estimates for Table 5

			(0.0621)			(0.0630)			(0.110)
ry_1896			0.0235			0.0182			-0.109
			(0.0549)			(0.0559)			(0.0940)
ry_1897			0.0817			0.0945			-0.0450
			(0.0433)			(0.0448)			(0.0744)
ry_1898			0.120			0.0800			0.0957
			(0.0355)			(0.0365)			(0.0508)
ry_1899			0.0761			0.0786			0.0597
			(0.0321)			(0.0332)			(0.0413)
ry_1901			-0.0560			-0.0612			-0.0700
			(0.0307)			(0.0323)			(0.0411)
ry_1902			0.0422			-0.0293			0.00605
			(0.0339)			(0.0356)			(0.0488)
ry_1903			0.106			0.0113			0.0613
			(0.0397)			(0.0415)			(0.0610)
ry_1904			0.204			0.00685			0.0900
			(0.0470)			(0.0486)			(0.0737)
ry_1905			0.224			0.0328			0.0883
			(0.0523)			(0.0542)			(0.0832)
ry_1906			0.326			0.172			0.322
			(0.0608)			(0.0623)			(0.0973)
ry_1907			0.310			0.198			0.338
			(0.0686)			(0.0699)			(0.112)
ry_1908			0.383			0.319			0.464
			(0.0752)			(0.0769)			(0.126)
ry_1909			0.412			0.400			0.579
			(0.0834)			(0.0850)			(0.142)
ry_1910			0.592			0.506			0.711
			(0.0947)			(0.0960)			(0.160)
ry_1911			0.688			0.553			0.649
			(0.103)			(0.104)			(0.177)
by_1854	0.0266	-0.288	1.212	-0.142	-0.203	0.212	-0.120	-0.0210	0.366
	(0.0592)	(0.152)	(0.131)	(0.0615)	(0.159)	(0.136)	(0.0635)	(0.222)	(0.230)
by_1855	0.0114	-0.280	1.215	-0.0954	-0.150	0.275	-0.0632	0.0291	0.422
	(0.0324)	(0.136)	(0.121)	(0.0344)	(0.142)	(0.126)	(0.0380)	(0.205)	(0.224)
by_1856	0.0143	-0.272	1.217	-0.132	-0.190	0.231	-0.0949	-0.0158	0.371
	(0.0290)	(0.127)	(0.120)	(0.0304)	(0.133)	(0.124)	(0.0344)	(0.192)	(0.222)
by_1857	-0.0522	-0.336	1.138	-0.112	-0.173	0.227	-0.0968	-0.0348	0.354
	(0.0271)	(0.121)	(0.118)	(0.0289)	(0.127)	(0.123)	(0.0360)	(0.180)	(0.220)
by_1858	-0.144	-0.430	1.047	-0.144	-0.215	0.186	-0.0670	-0.0157	0.346
	(0.0282)	(0.114)	(0.117)	(0.0302)	(0.119)	(0.121)	(0.0387)	(0.175)	(0.217)
by_1859	-0.145	-0.412	1.041	-0.136	-0.206	0.158	-0.125	-0.0912	0.221
	(0.0284)	(0.108)	(0.115)	(0.0301)	(0.114)	(0.119)	(0.0367)	(0.164)	(0.213)
by_1860	-0.0955	-0.330	1.017	-0.122	-0.184	0.173	-0.103	-0.0784	0.216
	(0.0280)	(0.105)	(0.113)	(0.0296)	(0.110)	(0.117)	(0.0359)	(0.160)	(0.210)
by_1861	-0.130	-0.334	0.965	-0.146	-0.197	0.153	-0.117	-0.0891	0.206
	(0.0274)	(0.102)	(0.112)	(0.0290)	(0.107)	(0.116)	(0.0350)	(0.160)	(0.208)
by_1862	-0.143	-0.346	0.883	-0.0769	-0.124	0.220	-0.0729	-0.0355	0.235
	(0.0275)	(0.101)	(0.111)	(0.0296)	(0.106)	(0.114)	(0.0360)	(0.159)	(0.206)
by_1863	-0.127	-0.332	0.828	-0.0821	-0.131	0.209	-0.0300	0.00603	0.249
	(0.0283)	(0.0986)	(0.110)	(0.0303)	(0.104)	(0.113)	(0.0376)	(0.153)	(0.203)
by_1864	-0.0303	-0.260	0.797	-0.0518	-0.110	0.206	0.0367	0.0537	0.315

Table E.4: OLS estimates for Table 5

	(0.0311)	(0.0964)	(0.108)	(0.0332)	(0.101)	(0.111)	(0.0464)	(0.146)	(0.197)
by_1865	0.00153	-0.259	0.691	-0.0810	-0.162	0.139	-0.0462	-0.0428	0.204
	(0.0360)	(0.0918)	(0.104)	(0.0373)	(0.0961)	(0.107)	(0.0506)	(0.141)	(0.187)
by_1866	0.0364	-0.238	0.665	-0.0305	-0.123	0.156	0.00707	-0.00453	0.228
	(0.0377)	(0.0882)	(0.0993)	(0.0392)	(0.0924)	(0.102)	(0.0491)	(0.133)	(0.177)
by_1867	0.0773	-0.181	0.650	0.0255	-0.0666	0.197	0.0831	0.0625	0.257
	(0.0378)	(0.0840)	(0.0934)	(0.0397)	(0.0882)	(0.0960)	(0.0505)	(0.129)	(0.164)
by_1868	0.0678	-0.164	0.605	-0.0765	-0.159	0.0677	-0.0720	-0.0969	0.0907
	(0.0352)	(0.0810)	(0.0871)	(0.0364)	(0.0846)	(0.0893)	(0.0444)	(0.124)	(0.154)
by_1869	0.00628	-0.204	0.515	0.0116	-0.0659	0.137	0.0736	0.0506	0.232
	(0.0346)	(0.0773)	(0.0832)	(0.0366)	(0.0814)	(0.0852)	(0.0449)	(0.122)	(0.147)
by_1870	0.0101	-0.190	0.418	-0.0326	-0.109	0.0585	0.0402	0.0105	0.135
	(0.0336)	(0.0744)	(0.0783)	(0.0354)	(0.0779)	(0.0800)	(0.0454)	(0.114)	(0.137)
by_1871	0.0554	-0.144	0.358	0.0325	-0.0479	0.0718	0.0836	0.0385	0.103
	(0.0345)	(0.0695)	(0.0732)	(0.0363)	(0.0727)	(0.0748)	(0.0474)	(0.102)	(0.125)
by_1872	0.112	-0.0913	0.253	0.0258	-0.0638	0.0101	0.0619	0.00496	0.0932
	(0.0354)	(0.0618)	(0.0674)	(0.0364)	(0.0642)	(0.0685)	(0.0473)	(0.0922)	(0.111)
by_1873	0.136	-0.0749	0.148	0.0150	-0.0826	-0.0316	0.123	0.0505	0.172
	(0.0358)	(0.0562)	(0.0596)	(0.0370)	(0.0584)	(0.0611)	(0.0481)	(0.0835)	(0.0944)
by_1874	0.170	-0.0223	0.117	0.0331	-0.0665	-0.0484	0.0609	-0.0181	0.0256
	(0.0353)	(0.0491)	(0.0496)	(0.0363)	(0.0510)	(0.0514)	(0.0449)	(0.0734)	(0.0765)
by_1875	0.0359	-0.0929	0.0295	-0.0223	-0.0985	-0.0757	-0.0295	-0.0775	-0.0538
	(0.0290)	(0.0400)	(0.0343)	(0.0304)	(0.0419)	(0.0361)	(0.0355)	(0.0616)	(0.0553)
by_1876	0.0488	-0.0241	0.0524	0.0113	-0.0347	-0.0197	0.0340	0.0198	0.0517
	(0.0298)	(0.0333)	(0.0335)	(0.0314)	(0.0352)	(0.0353)	(0.0376)	(0.0498)	(0.0483)
by_1877	0.0716	0.0615	0.0727	0.0624	0.0484	0.0445	0.0508	0.0587	0.0673
	(0.0299)	(0.0311)	(0.0316)	(0.0314)	(0.0327)	(0.0333)	(0.0379)	(0.0423)	(0.0422)
by_1879	-0.000475	0.0257	-0.0306	-0.0521	-0.0396	-0.0462	-0.0220	-0.0565	-0.0719
	(0.0319)	(0.0342)	(0.0337)	(0.0338)	(0.0362)	(0.0357)	(0.0416)	(0.0435)	(0.0454)
by_1880	0.0313	0.0383	-0.0756	-0.114	-0.122	-0.131	-0.0759	-0.167	-0.157
	(0.0313)	(0.0375)	(0.0369)	(0.0326)	(0.0391)	(0.0385)	(0.0408)	(0.0524)	(0.0567)
by_1881	0.188	0.155	0.0166	0.0279	-0.0147	-0.0243	0.0572	-0.0678	-0.0769
	(0.0337)	(0.0451)	(0.0436)	(0.0347)	(0.0468)	(0.0450)	(0.0431)	(0.0640)	(0.0697)
by_1882	0.213	0.177	-0.0229	0.0762	0.0101	-0.0133	0.155	0.00497	-0.0696
	(0.0356)	(0.0532)	(0.0508)	(0.0366)	(0.0551)	(0.0523)	(0.0475)	(0.0774)	(0.0832)
by_1883	0.222	0.204	-0.0510	-0.00693	-0.0800	-0.159	0.0456	-0.126	-0.308
	(0.0359)	(0.0610)	(0.0575)	(0.0366)	(0.0634)	(0.0589)	(0.0465)	(0.0864)	(0.0978)
by_1884	0.208	0.209	-0.122	0.0396	-0.0340	-0.210	0.129	-0.0474	-0.295
	(0.0374)	(0.0671)	(0.0655)	(0.0385)	(0.0697)	(0.0666)	(0.0507)	(0.0952)	(0.114)
by_1885	0.199	0.222	-0.162	0.0266	-0.0436	-0.294	0.112	-0.0541	-0.425
	(0.0377)	(0.0722)	(0.0728)	(0.0385)	(0.0750)	(0.0740)	(0.0497)	(0.0949)	(0.128)
by_1886	0.237	0.252	-0.203	0.0367	-0.0343	-0.374	0.113	-0.0623	-0.526
	(0.0390)	(0.0730)	(0.0798)	(0.0400)	(0.0759)	(0.0812)	(0.0547)	(0.103)	(0.145)
by_1887	0.241	0.246	-0.271	0.0829	0.00376	-0.399	0.184	-0.00467	-0.507
	(0.0424)	(0.0795)	(0.0888)	(0.0434)	(0.0826)	(0.0903)	(0.0635)	(0.112)	(0.166)
by_1888	0.293	0.292	-0.354	0.109	0.0231	-0.450	0.239	0.0360	-0.437
	(0.0521)	(0.0881)	(0.102)	(0.0531)	(0.0914)	(0.103)	(0.0801)	(0.128)	(0.187)
by_1889	0.285	0.290	-0.437	0.204	0.115	-0.385			
	(0.0628)	(0.0988)	(0.117)	(0.0659)	(0.104)	(0.119)			
Constant	66.96	67.94	66.99	68.29	68.38	68.32	68.26	67.77	68.28
	(0.0222)	(0.436)	(0.0285)	(0.0234)	(0.457)	(0.0300)	(0.0288)	(0.643)	(0.0511)

Table E.4: OLS estimates for Table 5

Observations	186,447	186,447	186,447	102,209	102,209	102,209	67,832	67,832	67,832
R-squared	0.005	0.006	0.007	0.003	0.003	0.004	0.004	0.004	0.005

Log-likelihood: -153723.67 Wald(67): 143.37, p=0 N=102209

	Estimates of mean			Estimates of log(sigma)	
	Est	SE		Est	SE
Birth-year dummies					
by_1854	-1.081	1.9345	by_1854	0.1953	0.1785
by_1855	0.577	0.7073	by_1855	0.0406	0.0731
by_1856	0.759	0.6403	by_1856	-0.0080	0.0655
by_1857	0.229	0.6585	by_1857	0.0613	0.0651
by_1858	-0.250	0.7205	by_1858	0.0973	0.0688
by_1859	0.630	0.6213	by_1859	-0.0292	0.0665
by_1860	0.360	0.6211	by_1860	0.0189	0.0646
by_1861	0.171	0.6404	by_1861	0.0335	0.0652
by_1862	0.072	0.6387	by_1862	0.0829	0.0649
by_1863	0.575	0.6119	by_1863	0.0113	0.0675
by_1864	-0.402	0.7588	by_1864	0.1347	0.0767
by_1865	0.519	0.6369	by_1865	-0.0181	0.0783
by_1866	0.427	0.6726	by_1866	0.0066	0.0856
by_1867	-0.265	0.7687	by_1867	0.1236	0.0886
by_1868	0.114	0.6113	by_1868	0.0047	0.0793
by_1869	0.247	0.5960	by_1869	0.0236	0.0768
by_1870	0.115	0.5541	by_1870	0.0013	0.0732
by_1871	0.099	0.5383	by_1871	0.0122	0.0739
by_1872	0.424	0.4626	by_1872	-0.0782	0.0701
by_1873	-0.701	0.6632	by_1873	0.0797	0.0825
by_1874	-0.648	0.5477	by_1874	0.0659	0.0733
by_1875	-0.641	0.4532	by_1875	0.0533	0.0635
by_1876	-0.775	0.5211	by_1876	0.0992	0.0680
by_1877	0.040	0.4148	by_1877	0.0154	0.0618
by_1879	-1.884	0.7977	by_1879	0.2055	0.0846
by_1880	-1.536	0.6622	by_1880	0.1284	0.0766
by_1881	-0.890	0.5819	by_1881	0.1093	0.0746
by_1882	-0.696	0.5552	by_1882	0.0876	0.0750
by_1883	-1.337	0.6836	by_1883	0.0966	0.0829
by_1884	-1.506	0.6310	by_1884	0.0949	0.0791
by_1885	-2.337	0.7791	by_1885	0.1551	0.0872
by_1886	-2.881	0.8504	by_1886	0.1802	0.0892
by_1887	-1.541	0.6078	by_1887	-0.0172	0.0803
by_1888	-3.109	1.1049	by_1888	0.1790	0.1220
by_1889	-2.594	1.0537	by_1889	0.1502	0.1225
			_cons	0.9242	0.0464
Recruitment-year dummies					
ry_1879	-1.158	0.4600			
ry_1880	-1.295	0.4551			
ry_1881	-1.013	0.4476			

ry_1882	-1.037	0.4388
ry_1883	-0.915	0.4321
ry_1884	-0.934	0.4251
ry_1885	-1.137	0.4174
ry_1886	-1.003	0.4100
ry_1887	-0.863	0.3985
ry_1888	-0.690	0.3828
ry_1889	-0.775	0.3647
ry_1890	-0.457	0.3411
ry_1891	-0.481	0.3169
ry_1892	-0.514	0.2915
ry_1893	-0.332	0.2743
ry_1894	-0.017	0.2491
ry_1895	0.016	0.2384
ry_1896	0.080	0.2098
ry_1897	0.370	0.1693
ry_1898	0.319	0.1417
ry_1899	0.285	0.1197
ry_1901	-0.244	0.1231
ry_1902	-0.064	0.1296
ry_1903	0.040	0.1586
ry_1904	0.064	0.2018
ry_1905	0.166	0.2220
ry_1906	0.686	0.2499
ry_1907	0.786	0.2802
ry_1908	1.263	0.3115
ry_1909	1.618	0.3451
ry_1910	2.030	0.3827
ry_1911	2.142	0.4094
_cons	66.396	0.3225