ECONOMIC GROWTH CENTER

YALE UNIVERSITY

P.O. Box 208629

New Haven, CT 06520-8269

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

CENTER DISCUSSION PAPER NO. 993

Social Class and the Fertility Transition: A Critical Comment on the Statistical Results Reported in Simon Szreter's *Fertility, Class and Gender in Britain, 1860-1940*

Geoffrey Barnes

Yale University

Timothy W. Guinnane

Yale University

November 2010

Notes: Center discussion papers are preliminary materials circulated to stimulate discussion and critical comments.

Barnes is a student in Yale College. Guinnane is the Philip Golden Bartlett Professor of Economic History in the Department of Economics, Yale University. Address correspondence to timothy.guinnane@yale.edu. We are grateful to Simon Szreter for his comments on an earlier version of this note.

Social Class and the Fertility Transition:
A Critical Comment on the Statistical Results Reported in Simon Szreter's
Fertility, Class and Gender in Britain, 1860-1940

Geoffrey Barnes and Timothy W. Guinnane

Abstract

Simon Szreter's book *Fertility, Class, and Gender in Britain, 1860-1940* argues that social and economic class fails to explain the cross-sectional differences in marital fertility as reported in the 1911 census of England and Wales. Szreter's conclusion made the book immediately influential, and it remains so. This finding matters a great deal for debates about the causes of the European fertility decline of the nineteenth and twentieth centuries. For decades scholars have argued whether the main forces at work were ideational or social and economic. This note reports a simple re-analysis of Szreter's own data, which suggests that social class *does* explain cross-sectional differences in English marital fertility in 1911.

Keywords: Fertility transition, 1911 Census of England and Wales

JEL Codes: J13, N33

The 1996 publication of Simon Szreter's *Fertility, Class, and Gender* brought the author acclaim seldom produced by any scholar's first book. The work was widely and almost uniformly positively reviewed. David Kertzer, for example, called it "something of a tour de force, combining a historian's concern for local context, a demographer's multivariate statistical techniques for the analysis of fertility-history data, and a social theorist's engagement with epistemological issues."[1] More than one reviewer remarked that the book reports research sufficient in scope and quality for at least two monographs. One consists of a thoughtful analysis of the way the British statistical authorities thought about "class," and the way those notions shaped the taking of the census. The second is the subject of our discussion here. Szreter provides an analysis of the remarkable fertility survey undertaken as part of the 1911 Census of England and Wales. He argues that the "professional model" of class-based differential fertility held by the census officials and others at the time does not explain the actual patterns in the data. Many take Szreter to have demonstrated that socio-economic differences in fertility were not terribly significant in turn-of-the-century Britain, and that, by extension, socio-economic forces did not play a major role in the decline of fertility.[2] Only slightly less influential is Szreter's positive conclusion that Britain underwent many simultaneous fertility transitions, each reflecting participation in a "communication community" rather than social class or economic status.

The statistical heart of the book relies on fertility measures drawn from the published fertility survey. This unusual exercise was published in multiple cross-classifications, allowing later scholars to examine differences in fertility according to the occupations of both husbands and wives. Social "class" reflects an amalgamation of groups of occupations. Szreter refers to the then-prevailing theory of fertility decline and its relationship to social class as the "professional model" and concludes, based on his statistical analysis that the model "fails." We find that, on the contrary, that the professional model performs well by conventional standards . Since we rely entirely on the professional model as he

[1] *Journal of Interdisciplinary History*, 28(1), Summer 1997, p.103.
[2] We located fourteen reviews of the book in scholarly journals. As of September 2010, the book had 209 citations on Google Scholar. The only generally negative review was by Charles Tilly, *Population and Development Review* 22(3), September 1996, pp. 557-66.

describes it, and the fertility measures he reports, the only differences at issue are statistical methods and interpretation of results. Our discussion should not be taken as a rejection of *Fertility, Class and Gender* in its entirety. Much of Szreter's argument deals with the nature of fertility control, more particularly the issue of spacing versus stopping. This part of the book is not affected by the results we report here. Moreover, we do not claim to have made our own, positive contribution to understanding the British fertility transition. Such a contribution would require an even more complete exploitation of the published tabulations than Szreter attempted, and would in any case be superseded by research based on the manuscript census schedules for 1911, which will soon be available to scholars.

1. Class, occupation, and fertility

Part II of *Fertility, Class and Gender* is an intellectual and policy history of the making of the 1911 fertility survey. Szreter starts from the premise that most social scientists probably share, but few take as seriously: that the sources from which we gather our data were themselves devised and used with specific purposes in mind. That is, data collection reflects a theory of social behavior. If a census reports considerable information on religion, for example, it is likely that those responsible for its construction thought religion a powerful force. Those theories inform the categories used in the sources, as well as the selection of people for inclusion in the study.

Szreter documents a growing fear about fertility decline in late-nineteenth century Britain, especially the fear that the "lower orders" were providing an ever-larger share of each birth cohort. These developments justified the expense and trouble of the special fertility survey undertaken in connection with the census. The manner in which the census authorities decided to publish the data reflects, as Szreter stresses, their understanding of the causal mechanisms underlying fertility patterns. The census tabulations of fertility by social class reflected a shared, almost dominant understanding of differential fertility at the time: fertility decline had gone furthest among the upper classes. In discussing Szreter's argument and results, it is important to stress that the professional model implies regular gradations by

social class: Class I should have smaller families than Class II, etc. This argument is stronger than one that simply claims social Class I has fertility *different from* Class II.

The 1911 census organized the British population into 465 occupations. The social-class groupings were amalgamations of subsets of these 465 occupations. The underlying idea was really one of three classes: Class I, thought to be an "upper and middle class;" Class III, comprising skilled workmen; and Class V, the unskilled. The remaining two classes were inserted as Class II and Class IV.[3] Classes II and IV consisted of couples who did not fit together as naturally as is the case for the other three Classes, and it is not clear that they serve as anything more than intermediates. Szreter discusses several different groups of couples defined by their age at marriage and marital duration as of 1911, but in his analysis he stresses one particular group, those married at ages 20-24 in the period 1881-1885. This is the largest single age-at-marriage category. By 1911, these couples were nearing the end of their reproductive lives but were not so old that one would have to worry much about selective mortality affecting the chances that higher-fertility wives would still survive.[4]

2. Szreter's analysis

We focus on the discussion presented in Chapter Six of *Fertility Class and Gender* (entitled "A test of the coherence of the professional model of class-differential fertility decline"). The 1911 census published complete parity distributions for each marriage-age and duration category, but the occupational cross-tabulations only present the total number of couples and total births (and child deaths) in each age-duration category.[5] Thus Szreter works

[3] The system was devised by the Registrar-General, who also ran the census, and was first described in the Registrar-General's Annual Report for 1911 (published in 1913) Szreter (1996, pp. 255-256). The reference to the key source in Szreter's bibliography is not quite right; the proper full citation is in the list of references.

[4] We restrict our analysis in this note to the sub-population Szreter stresses. But we should note that age at marriage itself could reflect the influence of social class, or more broadly social and economic forces, on the desire to limit family size. Elsewhere in his book Szreter stresses precisely this point.

[5] One could use the parity distributions to look for heterogeneity in the way couples achieved smaller families, for example. Others have used the similar fertility survey from the 1911 Census of Ireland to examine questions similar to this one. Paul David, Warren Sanderson, and their collaborators have developed an entirely new statistical methodology that exploits all the information in those parity distributions for the purpose of arriving at measures of

with the mean number of children ever born (CEB). In evaluating the social-class model we need to bear in mind that the census reports cross-classification of fertility by occupation for a sub-set of occupations. The census itself counts 465 different occupations, but the fertility cross-classifications restrict themselves to 206 occupations chosen, Szreter (1996, p. 291) reports, "for a combination of their quantitative importance and their specificity of occupational description." They include 87.5 percent of couples enumerated in the census. Szreter argues that Classes I-V are graded, and thus focuses on the 176 occupations in those classes. Including Classes VI-VIII (which we discuss below) increases the total number of occupations to 195.

Szreter uses these fertility measures drawn from the 176 different occupational groups to demonstrate that the social-class model of fertility, in his word, fails. Table 1 reports descriptive statistics for mean CEB in the five classes. The table confirms that fertility was graded by social class. Yet one could suggest, as Szreter does, that there is too much variation in CEB within each class to view them as distinct groups:

> A strong and relatively unequivocal corroboration for a theory of class-differentials would be provided by a pattern in which fertility levels of the component occupations within each class proved to be reasonably tightly bunched around their class mean, demonstrating that the classes approximated to both homogenous and discrete entities. Of course, there would be some outliers in each class but a high degree of continuity in the run of occupational fertility values, without any clustering, certainly would not conform to this kind of highly positive confirmation of the class-differential model.
>
> Secondly, it might be found that there was no bunching of occupational fertility values around their class mean and correspondingly no clear separation between the graded social classes, as each one merged into the next. But it might still be the case that the fertility levels of the individual occupations followed a rank ordering more or less in accord with their graded class affiliations: class I occupations appearing on a rank-ordered list before those of class II and so on. This would be a qualified success for the professional model: class-graded rather than class-differential fertility decline.
>
> Thirdly, a still greater degree of overlap between the occupations of the adjacent graded classes might present itself. More than a merging at the edges, there might be such substantial overlapping of the fertility values recorded by many of the individual occupations drawn from different classes that some of the social classes could hardly be distinguished from each other in

---

the extent and nature of fertility regulation behavior in specific birth and marriage-age cohorts. Cohort-Parity Analysis (CPA) compares a "model" population of couples not thought to be controlling their fertility to a "target" population of fertility controllers. It yields parameters that indicate the proportion of each age-duration group that is controlling fertility, and the number of births averted by controllers. See David et al (1988), and David and Sanderson (1988, 1990).

> terms of their occupational contents, although still perhaps preserving some overall and very general correspondence with the professional model. This would suggest that the class-graded model was far from an accurate or useful summary of the occupational patterns revealed by the 1911 census (Szreter 1996, p.298-99).

Szreter refers to having examined these questions using "factor, cluster, and discriminant analysis," but he does not report any statistical results in this chapter. In what follows we describe what we think would be the most natural way to cast his criteria as testable hypotheses into a statistical framework.[6]

Szreter's judgment that the social-class model failed to explain fertility differentials rests on two main bodies of evidence. First, he plots tabulations of the mean of CEB for all observations, and argues that because the figures do not display "discontinuities," they do not suggest that each social-class has its own fertility patterns (these are his Figures 6.2 through 6.7). Our Figure 1 is a slight variant on his Figure 6.6. We divide the couples into "bins" according to their mean CEB, and then graph the number of couples in each bin.[7] Szreter is correct: while some outcomes are more common than others, we do not see five different spikes in frequency corresponding to the five different classes. However, plotting the data a different way reveals more. Figure 2 presents smoothed distributions of mean CEB separately for each class, but overlaid on the same horizontal axis to facilitate comparison. Setting aside Class IV for the moment, we see that each class *does* correspond to a distinct peak in the frequency of a given mean CEB. The class-specific peaks we see in Figure 2 are obscured in Figure 1 by the overlaps across classes. Figure 2 also shows that Class IV is different. Its mean is not much larger than Class III's, but it overlaps much of both III and V. This feature of Class IV affects all of the statistical analysis discussed below.[8]

Szreter's Tables 6.2 through 6.6 appear to report the results of a more formal approach to his question, namely, an application of linear discriminant analysis (hereafter LDA). LDA is a classification

[6] One review of *Fertility, Class and Gender* makes the puzzling claim that Szreter's analysis of the 1911 fertility survey shows that the social-class model is "… little better at predicting fertility behavior than a random selection of five groups (Carl Ipsen in *Victorian Studies* 40(1), Autumn 1996, pp. 156). Szreter makes no such claim.
[7] The bins are .25 CEB wide, as in his figures.
[8] Class IV comprises 33 very heterogeneous occupation groups, including engine drivers, domestic and other gardeners, gas works service workers, domestic coachmen, grooms, postmen, shepherds, fishermen, some manufacturing occupations, police, and "men of the navy."

procedure.[9] In our context, LDA uses the relationship between one variable (fertility) and a set of categories (class) to construct a classification rule. This classification rule seeks a separation between each class that minimizes the number of misclassifications of actual classes into predicted classes. When applying LDA, a "misclassification" consists of an observation from the true Class III, for example, that is assigned to the predicted class IV.

Table 2 reports an LDA exercise with results similar to Szreter's. (Given that there are only minor differences between what he reports and what we were able to re-create, in the following we discuss his Table 6.4, rather than our Table 2). LDA allocates 23 of the 31 groups in actual Class I to the predicted Class I. It places a further 8 observations in the wrong category, with 6 in class II and another 2 in class III. Thus 8/31 or 26 percent are misclassified. As we would expect on the basis of Figure 2, LDA does not perform as well for Classes III and IV. Fifty-nine percent of the 44 observations in Class III are assigned to the wrong predicted social class. Class IV is even worse, with 73 percent in the wrong group. A glance at Figure 2 suggests the source of the problem: for a low value of CEB such as 3.3, there is little doubt that the occupation belongs in the predicted class I. But in other cases the overlap among the class distributions is so substantial that multiple classes are plausible, if not equally likely. A CEB value of about 5.1, for example, could belong in class II, III, or IV.

Szreter then reports a weighted version of the LDA, which allows him to estimate the fraction of all *couples* that are misclassified. This important exercise reflects the different sizes of the occupational groups that form the observations. The smallest group (Class I's "scientific pursuits") has about 2400 couples in England overall, while the largest, Class V's "cartmen and carters – nonfarm" has some 181 thousand couples. Thus a misclassification in the latter has greater implications for our understanding of fertility patterns. Table 6.6 reports that about 52 percent of all couples are misclassified.[10]

[9] We assume Szreter used LDA, but cannot precisely replicate his Table 6.4 using LDA. See Eisenbeis and Avery (1972) for an explanation.

[10] Only 39 percent of cells are misclassified; that is, cells with more couples are more likely to be misclassified. The weights Szreter uses do not correspond to the number of couples in each age at marriage/period of marriage/occupational group cell reported in the fertility survey. Rather, he weights by the total number of men with that occupation in England and Wales (Szreter 1996, p.303).

Szreter views the LDA exercise as a rejection of the "professional model." We disagree. First, his standard for rejection is unusual. A statistical model can be useful even if it does not fit *perfectly.* His LDA model implies that the class variables correctly allocate about half the English population to the correct discrete category. Many social scientists would be ecstatic if their models had such good predictive powers! Second, his conclusion that the model fails relies heavily on the binary criterion of "right class/wrong class." Most misclassifications in his Table 6.4 are into neighboring social classes. Consider the 31 observations in class I. LDA assigns 23 to the predicted class I, and another 6 to the predicted class II. Thus 29/31 = 93 percent are in the correct predicted class or one of its neighbors. If we go through the table and compute the percentage of all cells that are assigned either to the right social class or to one of its neighbors, we find the fit is actually rather good. Our logic reflects Szreter's own argument that the social classes are graded. If we think Class II belongs between Class I and Class III in some meaningful scale of status, as is presumed by the entire classification structure, then it is indeed relevant that most of the misclassifications from one social class appear in the predicted version of the class either above or below it. Applying this procedure to Table 6.6, we find that 88 percent of the population is classified into either the correct class or one of its immediate neighbors. There is no formal statistical basis for this procedure that we know of, but it does show that LDA does not assign occupations randomly, even when it is "wrong."[11]

3. A different approach

The approaches Szreter took do not, in our view, survive reexamination. We disagree particularly with his idea of what constitutes a useful statistical model. In our view, however, the most surprising

[11] A variant on the LDA approach yields a better fit than Szreter reports. The *linear* version of the model (upon which we surmise Szreter relied) assumes that the distribution of mean CEB within each class is normally distributed. This assumption is violated for all classes except class I. In addition, LDA, like all linear models, is sensitive to outliers. Note, for example, that nearly all occupations in Class I have a mean CEB of between 3.5 and 4.5, but one has a mean CEB of over 5.5. An alternative but similar approach is to use the k^{th} nearest-neighbor (knn) technique, which is a nonparametric approach. This makes knn models less sensitive to a few unusual observations within a class. Experimenting with the knn model, we found a fit dramatically better than in the LDA model reported in Table 2. The fit is sensitive to the choice of k, but no choice of k<20 yields a fit as poor as the linear model's.

feature of his statistical analysis is the technique he does not mention or use, analysis of variance (ANOVA). When we attempt to determine whether a model fits a given set of data, we are not so much interested in whether it can assign outcomes to specific predicted classes as we are in asking whether most of the variation in the outcome variable (fertility) follows the predictions implied by the model. The most natural way to approach this question is to ask about sources of variation in fertility. What fraction of the variation in fertility we observe in these 176 occupational categories reflects differences *within* a social-class, and what fraction reflects differences *between* the social classes? This question suggests the use of ANOVA or its close cousin, linear regression. To the extent that the latter type of variation predominates, it would imply that social class differences provide a good account of the couple-to-couple variations in marital fertility, or that the social-class model "fits" the 1911 survey data.

ANOVA and LDA are closely related, and in fact one can view LDA as the reverse of ANOVA. In this section we are not really presenting new information as much as re-presenting it in a way that more closely corresponds to widely accepted social science notions of testing a behavioral hypothesis. Referring back to the distributions shown in Figure 2, LDA asks if it is possible to find a rule that distinguishes most observations for Class I from most observations for Class II. As is readily apparent, no such "clean" separation is possible, and, as we already know, LDA will place some Class II occupations in the predicted Class I and vice-versa. ANOVA treats that variation differently. Instead of asking whether it can classify the observations into five distinct groups, it asks whether there are in fact five distinct groups. Note that Table 1 has already shown than Classes I-V are graded. If we can establish that there are five distinct groups in a statistical sense, then we have demonstrated the validity of the professional model as Szreter describes it.

The answer is surprising in its divergence from the interpretation that Szreter gives to the LDA results: an unweighted analysis of variance (ANOVA) has an adjusted R^2 of .67. Weighting by the number of couples (as Szreter defines them) reduces that R^2 to about .65. Thus about two-thirds of all variation in fertility across couples occurs *between* social classes; in the imprecise parlance often used by

social scientists, social class "explains" about two-thirds of the variation in marital fertility.[12] A single-factor model that achieves so much hardly can be called a failure.

ANOVA also provides us with a better intuitive grasp of the problem posed by Class IV. We construct four pseudo-classes in which III and IV are combined. Using those pseudo-classes instead of the real classes, ANOVA reports an adjusted R^2, of, once again, .66. Our amalgamated scheme does not lose any explanatory power because it affects only that part of the class distribution that does not help to explain fertility in any case. Now we have a more complicated answer to the question, "In what sense does social class 'explain' fertility?" Knowing whether a couple was allocated to class I, II or V does matter, as does knowing whether it was in III or IV as opposed to one of the others. The differences between III and IV matter so little, however, that they can be combined into a single group without the loss of any explanatory power. So part of the social-class model's strength turns on the question of where we look in this hierarchy of "classes".[13]

Finally, a general way to determine whether the values in one sample are larger than the values in another is to ask whether the second distribution stochastically dominates the first. If distribution A stochastically dominates distribution B, then for any x, $F_A(x) < F_B(x)$, where F(x) denotes the cumulative distribution function. In Figure 3 we plot the (unsmoothed) cumulative distribution functions for each of the five classes. The figure shows that the distribution of mean CEB in each social class dominates that of

[12] This R^2 is estimated from observations that are themselves aggregated; all variation between couples but with an occupation group has been removed. Ordinarily this aggregation results in a R^2 much higher than is possible with individual-level data. Brown and Guinnane (2007) explain the aggregation issue in the context of the Princeton studies. As an example, consider Teitelbaum's contribution to the Princeton series, which is also on Britain. He reports (Table 7.2) a series of linear models with adjusted R^2 values ranging as high as .78. This is possible because his units of analysis are British counties. In another famous historical demographic study, the reported adjusted R^2 values rarely exceed .05 (Preston and Haines, for example, Table 4.4). Preston and Haines are working with individual-level data, and the "unexplained" portion of the variation in the dependent variable includes much inter-couple variation removed by aggregation in Szreter's approach.

[13] ANOVA corresponds to a linear regression model in which all regressors are dummies. If we estimate the model in this form, we cannot reject the null hypothesis that Class III and Class IV have the same conditional mean ($F=2.17$, with a probability of .143). This is just a restatement of the point made in the text. ANOVA could also be sensitive to parametric form. We re-estimated the model with the ranks (rather than values) of occupation-specific CEB as the dependent variables. This procedure is the Kruskal-Wallis nonparametric analogue of ANOVA. The results in ranks are nearly identical to those in levels.

the class "above" it. Class II, for example, has higher fertility in this sense than Class I. The sole exception is that involving Class IV, which by now familiarly anomalous.[14]

The only problem with the "professional model" for Classes I-V appears to be a Class IV that really is not distinct from Class III. Szreter's discussion of the way these classes were defined and used shows that to the extent there was a clear theory of class, the theory pertained to Classes I, III, and V, while Classes II and IV were afterthoughts. Class II nonetheless appears to have real explanatory power for fertility differentials. It is unfortunate that Szreter's study did not explicitly focus on the problem with Class IV. Class IV is indeed a surprising amalgam of occupations. It would have been an interesting and valuable contribution to have worked out why the composition of Class IV failed to contribute to the otherwise valid account of marital fertility variation provided by the "professional" .

In concluding this discussion of Chapter 6 we should emphasize two points. First, our discussion of the ANOVAs stresses that the classes are indeed different. *But Table 1 shows they are also graded*, in the sense Szreter stresses: the mean CEB for Class I is less than for Class II, etc. Thus the model as Szreter defines it does not fail. Second, our efforts here follow the fundamental choices Szreter makes in his Chapter 6. We already noted that the 176 occupations we use in our analysis are those stressed in his work. He does not include Classes VI, VII, and VIII in his test of the professional model because he views them as having been excluded from the professional model on *a priori* grounds. He does report mean CEB for occupations grouped under these classes, however, and it is worth briefly considering their role in the argument. Class VI comprises a diverse group of textile workers. Class VII is mostly coal miners. Class VIII is agricultural laborers. These three classes clearly do not fit the graded class scheme. The mean CEB for Class VI is 5.2 births (N=12), which would place it between Classes II and III. (Class VII's mean CEB is 7.1, (N=4); for agricultural laborers, mean CEB is 6.7 for three occupational groups). Adding these three classes to our analysis does not affect the main ANOVA model, however. Using the

[14] Class IV does stochastically dominate Class III at second order, which is a slightly weaker condition still implying that "most" of the distribution for Class IV is larger than for Class III.

full 195 occupations reflected in Classes I-VIII, we obtain an R^2 of .68.[15] Szreter (1996, p. 307) stresses that Classes VI-VIII demonstrate the professional model's inherent "incoherence," since they comprise large parts of the workforce and obviously do not fit the graded fertility scheme. Yet since they were not included in the statistical discussion reported in Chapter Six of *Fertility, Class, and Gender,* they are really not at issue in our re-analysis of his data.

4. Szreter's explanation

The heart of *Fertility, Class, and Gender* for later scholars has been the rejection of the social class model. This exercise cleared the way for Szreter's preferred alternative, which is to argue that

> The idea of a neatly socially graded single event [the fertility decline] should be replaced by the picture of many geographically and chronologically disparate processes occurring in distinct contexts and for different reasons. It is because these have formed, overall, a chronologically overlapping sequence of such changes that this has produced the illusion at the aggregate, national level of a relatively smooth unitary process. This misapprehension needs to be recognized for what it is: an elementary statistical fallacy, an artifact created by aggregation (Szreter 1996, p.364).

We have demonstrated, we believe, that there is no statistical fallacy. But the question remains: what is the basis for his interpretation?

We see it as follows: Szreter ordered all occupational groups (the observations) by the value of mean CEB. He then notes what he sees as patterns in the "misclassifications." A misclassification here amounts to an occupation from Class II, for example, that appears in the middle of occupations from Class I. He stresses two general forces in accounting for these patterns. First, couples' fertility may reflect the influence of others with whom they share close social proximity. A good example is domestic indoor servants (class III, but with an average CEB of 4.01, well below the mean CEB of 5.77 for their class and

[15] Szreter (1996, pp.306-307) discusses an earlier paper by Michael Haines that is also based on the 1911 fertility survey. Haines (1989, Table 3) reports OLS regressions of fertility on social class, and concludes that "social class was a good predictor of differential fertility." He uses Classes I-VIII and reports R^2 values similar to ours. Szreter claims that "Had Haines used his method to test only the five classes of the professional model, the proportion of variance explained would certainly have been lower, as reported in this exercise." Our results show this is not true, as we find similar fits with just Classes I-V. Haines did not stress the class ordering of fertility that Szreter sees as central to the professional model. Szreter (1996, p.307) also notes that Haines forms his observations by combing all women married in a given period, regardless of their ages. We cannot say precisely how that practice affects Haines' results, but given the similarity to our own, it does not seem to have been a source of great error.

lower even than the mean for Class I). Szreter (1996, p.315) notes that these occupations are "predominantly composed of couples who lived in the direct employ and in close social contact with the most wealthy part of society, being their personal servants." Szreter suggests that servants might have very low fertility because they spend so much time with others who have low fertility, and either learn from them the benefits of a smaller family or simply ape their social superiors. In other instances Szreter stresses geographic patterns. Thus, noting apparent differences among workers in the textile industry, he stresses that some are most common in Lancashire, while others are mostly in Yorkshire's West Riding (Szreter 1996, p.317). Lancashire and the West Riding have different fertility patterns for persons in the same class.

This discussion forms the basis for the common interpretation of his argument as asserting the fertility transition was not driven by socio-economic forces but the influence of other people in a place, whether that place is defined as a place of employment or a region. Thus what mattered was less a couple's social or economic position as its participation in a "communication community." That argument resonates with much social history and social-science over the past two decades, which has stressed the important of social interactions and various sorts of geographically-based spillover effects. Properly documented, the argument could well be the basis for important insights into the fertility transition in England and Wales.

Yet the approach used in Chapter 7 ("Multiple fertility declines in Britain") is not what social scientists usually have in mind in conducting statistical analysis, because it is entirely *ad hoc*. A more useful analytical approach would be to specify, *a priori*, a set of forces one thinks affects marital fertility, and then to classify the occupational groups according to these forces. Thus Szreter might assign to each occupation a score reflecting the degree to which people in this occupation are in close contact with members of the upper classes. One could use statistical models to test whether this quality of an occupation has any power in explaining fertility differences. Similarly, one might explicitly model the role of local forces in driving fertility. From Table 45 of Part II one could compute the mean CEB measure discussed here for the urban and rural districts of every county in England and Wales.

Unfortunately, this table does not further cross-classify fertility by occupation, so one could not construct from this table the information needed to undertake a fair test of Szreter's "local social communications hypothesis." Still, given his claim of multiple fertility declines, it he might well have made us of the available geographic information to document the existence of such declines.

The absence of such analysis is puzzling because Szreter later undertakes one exploration that has precisely the character that seems to be indicated. His Table 7.2 reports correlations between fertility and three distinct measures of employment status. Here he defines a characteristic of each occupation that he thinks affects fertility, uses other census data to construct a measure of this characteristic, and then tests this measure's relationship to measured fertility. But, unfortunately, the more widely noticed portions of Szreter's effort to support his thesis rests on the procedure described above.

5. Conclusions

Szreter's impressive first book made a significant mark in the historiography of the European fertility transition, in large measure because he claimed to have shown that a "professional model" of British fertility in 1911 did not fit the available evidence. Others expanded upon this claim to argue that he had shown socio-economic differences were not an important part of the British fertility transition. We have shown this finding to be incorrect. Using Szreter's own data and simply reworking the statistical analysis he presents as a test of the professional model, we demonstrate that by conventional criteria a five-part division of the British population by social class explains the observed variations in marital fertility extremely well. The only serious weakness in the model's power is Class IV, which appears no different from Class III. Szreter rightly stresses that the graded class scheme excludes large parts of the British population, and that many of these excluded couples (Classes VI-VIII) do not fit the professional model. But these couples are not relevant to our critique; Szreter did not include them in the analysis at issue here.

The manuscript schedules from the 1911 census of England and Wales will be available soon, and their use will support tests of more complicated multivariate hypotheses. Szreter argued that social class

did not explain fertility differentials. We have shown to the contrary that it did. But this is a one-factor model, and begs several different questions. Social class might well be correlated with some other trait that better explains fertility. More likely, it might be found that, as Szreter suggested, the influence of social class is slight once we have accounted for other causal forces such as the nature of the local community. None of these or other factors can be identified from aggregated data of the sort published by the census, which is why individual-level data are so important for the study of fertility.

What are the larger stakes in this discussion? *Fertility, Class, and Gender* occupies an important place in the received understanding of the European fertility transition. For decades scholars have been speculating about and arguing over the suspected causes of the European fertility transition. Much of this debate has been framed around Carlsson (1969)'s distinction between "innovation" on the one hand and "adjustment" on the other. Although not strictly correct, most scholars associate innovation with ideational forces and adjustment with a changing economic and social environment. Princeton University's European Fertility Project famously argued that social and economic forces played little or no role in the transition (Coale and Watkins (1986)). For a while that view came close to becoming the consensus position, but recent research has stressed the view that the Princeton findings were largely an artifact of faulty statistical methods (Brown and Guinnane (2007)).[16] Szreter's book sometimes has been unfairly portrayed as similar in argument to that advanced by the work of the Princeton Project. Nevertheless, although it is a distinct intellectual contribution, its conclusions for the case of the British transition were aligned with and appeared to reinforce the Princeton Project's position. As we have shown here, Szreter's conclusion does not survive re-analysis of his data.

[16] Guinnane (In Press) provides an overview of the economics of the historical fertility transition.

References

Brown, John C. and Timothy W. Guinnane, 2007. “Regions and Time in the European Fertility Transition: Problems in the Princeton Project’s Statistical Methodology.” *Economic History Review* 60(3): 574-595.

Carlsson, Gösta, 1969. "The Decline of Fertility: Innovation or Adjustment Process." *Population Studies* 20(2):149-174.

Census of England and Wales 1911, 1923, *Fertility of Marriage, Part 2*. (Vol XIII of the 1911 Census). London: HM Stationary Office.

Coale, Ansley J., and Susan Cotts Watkins, editors, 1986. *The Decline of Fertility in Europe*. Princeton: Princeton University Press.

David, Paul, Thomas Mroz, Warren Sanderson, Kenneth Wachter, and David Weir, 1988. "Cohort Parity Analysis: Statistical Estimates of the Extent of Fertility Control." *Demography* 25(2): 163-188.

David, Paul A., and Warren C. Sanderson, 1988. "Measuring Marital Fertility Control with CPA." *Population Index* 54(Winter): 691-731.

David, Paul A., and Warren C. Sanderson, 1990. "Cohort Parity Analysis and Fertility Transition Dynamics: Reconstructing Historical Trends in Fertility Control from a Single Ceneus." *Population Studies* 44(3): 421-445.

Eisenbeis, Robert A. and Robert B Avery, 1972. *Discriminant Analysis and Classification Procedures: Theory and Applications*. London.

Guinnane, Timothy W., In Press. “The Historical Fertility Transition: A Guide for Economists.” *Journal of Economic Literature.*

Haines, Michael R., 1989. “Social Class Differentials during Fertility Decline: England and Wales Revisted.” *Population Studies* 43(2): 305-323.

Preston, Samuel, and Michael Haines, 1991. *The Fatal Years: Child Mortality in Late Nineteenth Century America*. Princeton: Princeton University Press.

Szreter, Simon, 1996. *Fertility, class, and gender in Britain, 1860-1940*. Cambridge: Cambridge University Press.

Teitelbaum, M., 1984. *The British Fertility Decline: Demographic Transition in the Crucible of the Industrial Revolution*. Princeton: Princeton University Press.

Table 1: Descriptive statistics for AM2/81-85, by social class

Social class	Number of groups	Number of children ever-born			
		Mean	Minimum	Maximum	Inter-quartile range
I	31	4.11	3.16	5.66	.69
II	40	4.92	3.55	6.06	.88
III	44	5.77	4.01	7.04	.66
IV	33	5.97	4.49	7.03	1.04
V	28	6.56	5.76	7.29	.53

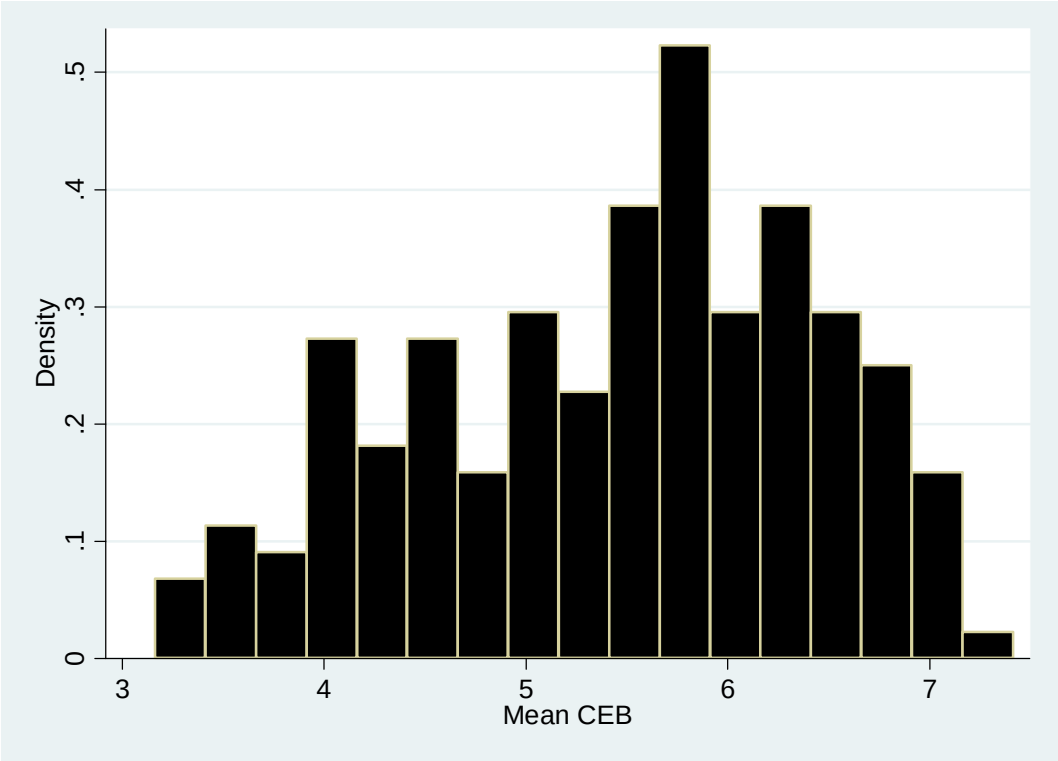
Note: Computed from the fertility measure reported in Szreter (1996, Appendix C)

Table 2: Linear discriminant analysis

	Predicted class					
True class	I	II	III	IV	V	Total
I	26	4	1	0	0	31
	83.87	12.90	3.23	0.00	0.00	100.00
II	9	19	10	2	0	40
	22.50	47.50	25.00	5.00	0.00	100.00
III	2	6	17	13	6	44
	4.55	13.64	38.64	29.55	13.64	100.00
IV	1	6	9	4	13	33
	3.03	18.18	27.27	12.12	39.39	100.00
V	0	0	2	3	23	28
	0.00	0.00	7.14	10.71	82.14	100.00
Total	38	35	39	22	42	176
	21.59	19.89	22.16	12.50	23.86	100.00

Note: The first number in each row is the number of observations, and the second is the corresponding row percentage. Source is as in Table 1.

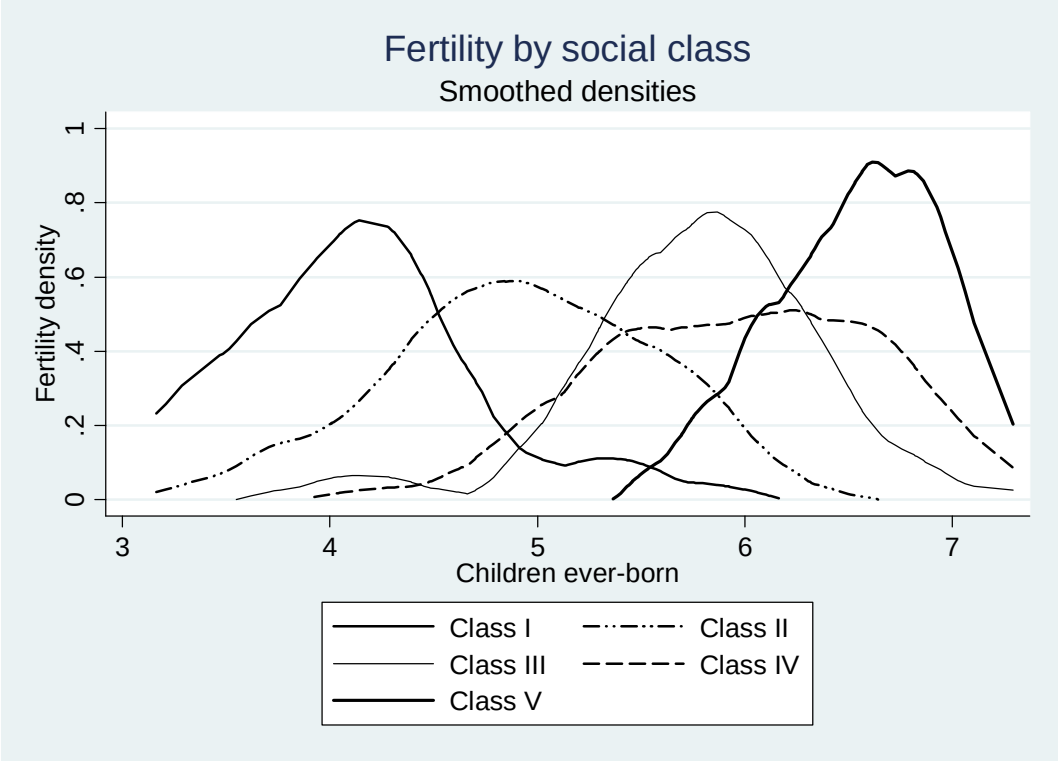
Figure 1: The distribution of occupations by mean children ever born



Note: This figure approximates Szreter (1986, Figure 6.6)

Source: Szreter (1996, Appendix C)

Figure 2



Note: The figures are kernel estimates of the density for each class. They are plotted on the same x-axis.

Source: Szreter (1986, Appendix C).

Figure 3: Cumulative Distribution Function of Mean Fertility by Social Class