Unskilled Mayors and Graduate Farmers—Educational Fertility Differentials by Occupational Status and Industry in Six European Countries

Bilal F. Barakat
Rachel E. Durham

Bilal F. Barakat is a research scientist at the Wittgenstein Centre (IIASA, VID/ÖAW, WU), Vienna Institute of Demography of the Austrian Academy of Sciences.
Email: bilal.barakat@oeaw.ac.at

Rachel E. Durham is a research scientist at the Wittgenstein Centre (IIASA, VID/ÖAW, WU), Vienna University of Economics and Business.
Email: rdurham@wu.ac.at

Abstract

Understanding the relationship of education to fertility requires disentangling the potentially confounding effect of social status. We contribute to this aim by examining educational fertility differentials *within* occupational groups and industries across a broad swath of Central and Eastern Europe, specifically for Austria, Greece, Hungary, Romania, Slovenia, and Switzerland. We use the recently-released individual-level census samples from the *Integrated Public Use Microdata Series* (IPUMS). A key advantage of IPUMS is that samples are large enough to contain sizeable numbers of unusual combinations, e.g., university graduates in low-status jobs or primary school dropouts in professional categories. As an additional methodological contribution, we re-introduce to the field an alternative count model, that improves on the results of a Poisson regression by accounting for the considerable underdispersion in the data. Results show that education has a strong, direct, consistent association with fertility, net of industry and occupation, even within presumably "family-friendly" industries. Furthermore, fertility by industry and occupation yielded fairly disparate patterns. We also find that differences in fertility across countries primarily reflect country-specific compositional differences in education, industry, and occupation and interaction effects.

1 Introduction

Women's education has long been widely recognized as one of the most important factors in fertility outcomes, both at the macro and individual level (Becker 1981; Caldwell 1980; Cochrane 1979, 1983; Kasarda, Bill and West 1986), having both direct and indirect effects (Kasarda 1979). While it is known that education may reduce fertility because it exposes women to knowledge about the biological circumstances of reproduction and contraceptive practices, for reasons that are not entirely understood, even small increases in education seem to increase a woman's self-efficacy with respect to her ability to achieve her desired number of children, as well as improves her own and her offspring's health (Jejeebhoy 1995; Martin 1995). From a socialization perspective, education also exposes women to lifestyles and opportunities beyond motherhood (Lesthaeghe 1995). In sum, there is something unique and intrinsic about the process or content of education that seems to affect both reproductive behaviour and outcomes.

Yet the relationship between education and fertility behaviour is not straightforward. Educational attainment may merely reflect an individual's values and preferences regarding career and fertility (c.f. Lesthaeghe and Moors 1995; Janssen and Kalmijn 2002). If a woman does not place a high priority on family formation and having children, then other paths will likely be pursued, particularly greater career investment via higher educational attainment (Gerson 1985; Hakim 2003). But the relationship between career and fertility can work both ways: higher fertility goals may discourage career and occupational investment, but an investment in education and/or commitment to a career may lead to foregoing or postponing children (Becker 1981; Brewster and Rindfuss 2000).

Thus, while it is known that education and vocation are strongly related to fertility outcomes, the theoretical model is still surprisingly lacking. It is not definitively known whether the relationship between fertility and socioeconomic variables is primarily driven by education, or whether fertility is affected by education's close correlates, particularly income and occupational status. Furthermore, a woman's vocation is not only a product of educational attainment but is also reflective of a woman's values and preferences about family formation and childbearing. For instance, certain industries have been found to be associated with differential fertility behaviour, since each may offer varying levels flexibility and autonomy.

The goal of the current study is to more deeply investigate the relationship between education and fertility, holding occupation and industry constant. This can be accomplished due to the availability of data that allows us to examine fertility outcomes within unique combinations of education, occupational status, and industry. The *Integrated Public Use Microdata Series* (IPUMS) samples from six Central and Eastern European countries — Austria, Hungary, Greece, Romania, Slovenia, and Switzerland — contain unusually large numbers of women with relatively low educational attainment in high prestige occupations, and conversely, women with high educational attainment in low-prestige occupational groups. Fertility and education data are also available by industry of occupation. Multivariate analyses of the 1990 and 2000 samples allow us to determine the primary sources of socio-economic variation in fertility behaviour, and moreover, examining different age cohorts across waves allows us to test whether differences in fertility are a result of timing effects within cohorts. Since occupation and industry are strongly indicative of women's autonomy, social status, and income, we can test whether education drives fertility independently, or whether it primarily influences fertility through socio-economic factors and the characteristics of one's vocation.

2 Theoretical Framework

2.1 Education and Fertility

The decrease in fertility during the first demographic transition was undoubtedly driven in large part by increases in women's participation in education. A review of the relevant literature points to a number of conclusions about how education and its correlates affect fertility, yet whether educational attainment itself drives fertility preferences or behaviours, or whether other socio-economic variables mediate the relationship between these outcomes and educational attainment remains unclear. To begin with, the direction of the relationship in not one-way, i.e., fertility certainly may affect a woman's ultimate educational attainment, but the effect of education on fertility is far more powerful. The primary means by which fertility might decrease ultimate educational attainment is via its depressing effect on enrolment; however, school enrolment and the completion of schooling is more strongly predictive of birth timing and hence, ultimate fertility (Billari and Philipov 2004; Blossfield and Huinink 1991; Edwards 2002; Hoem 1986; Kravdal 1994; Liefbroer and Corijn 1999).

Second, research has demonstrated that education (whether via the process or content) has direct effects on fertility. Education determines family size preferences, since more highly educated mothers are likely to place a high value on having well-educated offspring (Axinn 1993), and in most cases this would preclude bearing a large number of children (Lesthaeghe and Meekers 1986; Martin 1995). Education also increases women's sense of self-efficacy and exposes them to lifestyles beyond motherhood (Lesthaeghe 1995), but just as importantly, it increases knowledge about contraception (Basu 2002; Jejeebhoy 1995). Basu (2002) proposed that women who are more educated are more likely to be affected by educational messages about effective contraceptive use, but also by images in the mass media presenting an idealized version of smaller, happy and healthy families.

Third, education interacts with social factors to affect fertility behaviour. Such moderators include religious participation and observance (Kasarda 1979; Lesthaeghe and Meekers 1986), region and urbanicity since geographic location determines the utility of education (Cochrane 1983), and policies designed to facilitate combining motherhood and career (Brewster and Rindfuss 2000).

But perhaps the most often-studied ways in which education affects fertility are the indirect pathways: Education affects fertility indirectly via its positive relationship with age at marriage or stable partnership formation (concerning developing countries, c.f. Martin (1995)). In industrialized countries, higher educational attainment may increase age at marriage and subsequent fertility if 'student' is perceived as being a primary role that would stand in conflict with role of mother or even wife. Further, women with higher educational attainment may delay childbirth, since if education is a career investment, a certain amount of time in the labour market would be required before leaving permanently, or even temporarily (i.e., maternity leave) to create a return on that investment (Billari and Philipov 2004; Kravdal 1994; Rindfuss, Morgan and Offutt 1996).

The opportunity cost of fertility on women's earning potential in the labour market has often been cited as reason why women with higher educational attainment have fewer children. Income earned in the labour force makes women less dependent on marriage for economic survival (Becker 1981), and from a socio-psychological perspective, the roles of motherhood may compete with those of career (Hoem and Hoem 1989; Stolzenberg and Waite 1977), especially if a significant investment in having a career was made via higher education (Liefbroer and Corijn 1999).

2.2 Occupation/Vocation and Fertility

Women with higher education generally have higher earnings potential in the labour market, but the relationship between earnings and fertility is complex, since women with higher education may have access not only to higher-paying jobs but also jobs with more autonomy and flexibility (Kravdal 2007). Occupations also differ with respect to the shape of the earnings curve over the course of one's career. Concerning how the steepness of wage profiles may impact fertility, Lappegard and Ronsen (2005) found that there seemed to be a "catching up" effect among women with higher education since they had postponed childbearing longer than women with less education, and highly educated women may, in some settings, actually experience higher rates of higher-parity births (Gerster et al. 2007). However, women working in male-dominated fields were slower to transition to a first birth, suggesting that some careers may require a longer start-up period during which childbearing is more costly than it would be later (Van Bavel 2010). Kravdal (1994) reported that women with slower-growing wage profiles seemed to enter motherhood sooner, but overall household wealth mattered as well, since women with higher earnings can more easily afford childcare.

Women's vocational choices may be endogenous to underlying preferences about family and career life and may be reflected in their chosen fields of study, occupation, or chosen industry. Women who are family-oriented may self-select into industries that are family-friendly in terms of time management and that offer flexibility in scheduling, location, or into those where seasonal or part-time work is common (Edwards 2002; Hakim 2003; Janssen and Kalmijn 2002; Lappegard 2002). While professional occupations would seem to offer the necessary autonomy to juggle the commitments of both motherhood and career, some would appear to be more amenable than others. Using Norwegian register data, Lappegard (2002) found that nurses and teachers, regardless of their educational attainment, had higher fertility than women in science and technology. Examining Austrian women, Spielauer (2005) found the same trend among teachers. Part of this effect may be due to the attractiveness of female-dominated industries (such as education, health, and social work), since these tend to include jobs where time off work for childbearing or childcare responsibilities is less problematic (Desai and Waite 1991). Jobs in healthcare and nursing also offer the possibility of non-standard shifts (Swanberg, Pitt-Catsouphis and Drescher-Burke 2005).

Edwards (2002) demonstrated that American women in higher prestige positions, i.e., professional, technical, and managerial jobs, delayed the transition to motherhood for much longer than women in nursing, teaching, sales, or service jobs and concluded that women's career ambitions played a more prominent role in fertility timing than education. Any effects of education were expressed via their preparation for a particular field of work, where women's occupation exerted a stronger effect than educational attainment. Using British data, Hakim (2003) demonstrated that fertility was less affected by social status and educational attainment directly, but instead more by their preferences regarding family and work, which were manifested in their educational attainment. These findings would suggest that any effect of education on women's fertility outcomes is merely spurious; however, if educational attainment is truly endogenous to more fundamental fertility desires and preferences, then a greater amount of variation in fertility should be explained by women's vocation than by educational attainment, whether that be their chosen industry (i.e., ones that confer lower penalties on the combination of motherhood and career or the possibility of non-standard work schedules), or occupation (i.e., those offering greater flexibility and autonomy).

2.3 National Context

Concerning the relationship between fertility, its correlates and the national context, the literature offers few reasons to expect that residence in a particular country will *per se* affect fertility. But to the extent that countries vary according to educational composition, labour market participation, market returns to education, or particular structural constraints within a country, country-level differences in fertility may be found. For instance, looking across several European countries and examining the effects of labour market activity and fertility, Adsera (2011) found that second births were significantly delayed in countries with high unemployment, and Alba, Alvarez and Carrasci (2008) demonstrated that greater opportunities for full-time employment were positively related to conception events. Fertility may also be affected by the ease with which one enters the labour market after education completion, as well as job stability (Scherer 2005). In former Communist regimes, such as Hungary, Romania, and Slovenia, fertility may have been affected by the change in returns to education after the regime collapse (c.f., (Billari and Philipov 2004)). However, any change in fertility trends would largely have been driven by overall levels of educational attainment and labour market participation. For example, Speder (2006) found that postponement among higher educated women increased sharply after the regime change due to increases in opportunity costs.

Research has also examined the relationship between fertility and country-specific policies directed towards the combination of work and motherhood. Concerning motherhood and labour market participation, Hantrais (1997) notes that monetary leave allowances enacted across many EU countries have yielded inconsistent results. In some countries, these policies have caused fertility levels to decline more slowly than they might have otherwise, but in others they have yielded little effect. Sobotka (2009) notes that in Austria, policies are designed to incentivize women to stay home to raise young children and are not directed toward promoting child-care options, making the combination of career and motherhood difficult and magnifying the opportunity costs of fertility. Perhaps the inconsistency in effects of such policies stems from the fact that nation-specific policies do not affect fertility to as great an extent as would occupational returns to education, in that achieving the desired effects via such policies may depend upon the feasibility of a single-income household or whether such allowances can compensate for women's lost income via their occupation or industry. Moreover, fertility intentions across many European countries have declined in recent years (), and with this uniformity in intention, one would not expect to find a great deal of variation in fertility rates by country. One explanation for this uniform drop may be a homogeneous rise in educational levels among women across Central and Eastern Europe (Akin and Vlad 2007).

2.4 Inferential Challenges and Strategy

There is in general a pronounced (if not necessarily monotonic) association between education and fertility, but the above discussion highlights the existence of two related but distinct issues. The first concerns the direction of causality. Education may lead women to have fewer children, through a number of pathways, however it may also be the case that low fertility intentions leads women to seek more education. The second concerns the question to what extent the link between education and fertility is direct or indirect.

The data exploited in the present study is deep in some dimensions, but limited in others. In particular, its cross-sectional nature, where both the predictors and outcomes of a decades-long process are observed only after the fact, makes causal conclusions impossible. As a corollary, the term "effect" carries no causal connotation in the following, but is used only in the sense of "effect on the predicted value".

However, robust associational findings possess their own merit. They are a necessary step towards identifying potential causal pathways that are promising subjects for more elaborate, but also more narrowly focused, research, — and to rule out others. The contribution of the present study consists in systematically disentangling the direct and indirect associations between fertility on the one hand, and education, occupation, and industry on the other, in a cross-country comparison.

Many of the potential channels through which educational attainment and fertility are thought to be linked involve other characteristics, such as income, occupation, or social status, for example. This is true of tangibles such as the opportunity cost of children, but also of many latent characteristics such as agency or autonomy. The difficulty this creates for research emerges from the fact that education level and occupation, for example, are highly correlated. Moreover, what variation there is in the relationship between education and occupation is concentrated in a narrow range. Some senior managers have tertiary degrees and others only high-school degrees, but very few have failed to complete primary schooling. This problem potentially persists even in settings where a "natural experiment" (such as a change in compulsory education laws) partly redresses the question of causation. In addition, natural experiments and instrumental variables at best isolate the causal effect of a particular education level.

The present study is based on samples large enough to contain sizeable counts of rare combinations, such as university graduates in elementary jobs or primary school drop-outs in the 'Legislators, Senior Officials and Managers' category. As a result, we are able to examine fertility gradients across virtually the entire range of education levels *within* occupation groups, and to perform statistical inference on these effects. Moreover, this is done in a way that is comparable across several different countries. The aim is to be able to interpret the estimated direct education effect as one that is connected with intrinsic properties of highly educated women, as opposed to their higher average incomes, or social status, which are assumed to be related more to occupation than educational attainment *per se*.

An additional problem that is tackled here is the dearth of appropriate stochastic models for birth counts. The variability in the number of children around a given average is poorly approximated by the Poisson assumption of equidispersion. Moreover, *both* overdispersion and underdispersion occur in empirical fertility data, ruling out staples such as the negative binomial model, which are limited to modelling the former. In this study, a lesser-known, but appropriately flexible, count model is employed to address this issue, since the data at hand display a pronounced amount of underdispersion.

3 Data and Methods

3.1 Data Source and Transformations

The present analysis is based on individual-level census samples from Austria, Greece, Hungary, Romania, Slovenia, and Switzerland, from the two most recent rounds of censuses (around 2000 and 1990). This selection is based on the availability of the crucial variables of interest in the samples. These data have been harmonized and made available for research use by the IPUMS project (Minnesota Population Center 2011). See Table 3.1 for details of the included samples. Person weights are constant within each sample and therefore ignored.

Table 3.1 Census Samples Utilized

Country	Year	Fraction (%)	Females 40–49
Austria	1991	10	50,370
	2001	10	57,230
Greece	1991	10	58,764
	2001	10	72,493
Hungary	1990	5	35,425
	2001	5	39,211
Romania	1992	10	135,293
	2002	10	153,413
Slovenia	2002	10	13,517
Switzerland	2000	5	24,716

Complete birth, educational or occupational histories are not available. Instead, two measures at census time are used as fertility indicators: the number of children ever born (CEB), and the percent childless (PCL). As a result, the analysis focuses on women aged 40 to 49 years, whose birth histories can be assumed to be essentially complete. A secondary analysis also uses both measures for women aged 35-44 and 25-34 in the earlier census wave to allow for a pseudo-cohort analysis.

For industry and occupation, both original, unharmonized, national census measures and variables were harmonized by IPUMS International. This means the national categories have been mapped to common categories as carefully as possible to achieve comparability.

The harmonized codes for occupational status conform to the International Standard Classification of Occupations (ISCO) scheme for 1988. There are ten main harmonized categories, in addition to four kinds of non-response or inapplicability. For increased clarity in our presentation, these are grouped into seven categories for purposes of the present analysis. Our aggregated occupation groups are shown in Table 3.2.

Table 3.2 Aggregated Occupation Categories

Label	**Categories**
L,SO&M	Legislators, senior officials & managers
Prof	Professionals
T&AP	Technicians and associate professionals
CSWSMS	Clerks / Service workers and shop and market sales
C&RTW	Crafts and related trades workers
PMOAE	Plant & machine operators and assemblers / Elementary occ.
SA&FW	Skilled agricultural and fishery workers

The harmonized codes for industry contain fifteen main harmonized categories, in addition to four kinds of non-response or inapplicability. For increased clarity in our presentation, these are grouped into nine categories for purposes of the present analysis. Our aggregated industry categories are displayed in Table 3.3.

The harmonized education variable does not necessarily reflect any particular country's definition of the various levels of schooling in terms of terminology or the number of years of schooling. "[It] is an attempt to merge [...] samples that provide degrees, ones that provide actual years of schooling, and those that have some of both" (Minnesota Population Center 2011). The resulting measure is "largely comparable across countries". The four resulting education levels are: *less than primary completed, primary completed,*

Table 3.3 Aggregated Industry Categories

Label	Categories
A,F&F	Agriculture, fishing and forestry
H&R	Hotels and restaurants
CM	Construction/Mining
Manuf	Manufacturing
UTC	Utilities/Transportation & communications
E&HSW	Education / Health and social work
W&RT	Wholesale and retail trade
PA&D	Public administration and defense
Service	Financial services & insurance/Real estate & business services

secondary completed, university completed, in addition to *unknown* and *not in universe* responses.

Table 3.4 Education Categories and Austrian Mapping

Label	Category	Austrian categories
<P	less than primary	n.a.
P	primary completed	Compulsory (lower) secondary
S	secondary completed	Higher general secondary
		Higher technical & vocational secondary
		Intermediate technical & vocational
		Apprenticeship training
T	tertiary completed	Technical or vocational course
		(Academic) Intermediate degrees
		University, college

The harmonized variable was unavailable for Austria. Thus, the Austrian unharmonized categories were mapped to the four levels above according to the scheme in Table 3.4, taking into account country knowledge and the IPUMS documentation on how the mapping was performed for the other five countries in this study. No persons were recorded as having failed to complete compulsory primary schooling in the Austrian sample.

An important caveat is that all three of these variables (industry, occupation, education) capture the situation *at the time of the census*. In other words, with respect to tertiary education in particular, but also with respect to occupational and industry classification of the *current* job, it is likely that a share of the completed fertility was realized while the women were in different categories. This is discussed more fully below. Because of the snapshot nature of the occupation variable, an analysis of the "not in universe" category, which corresponds to those women who cannot be assigned an occupation because they are not economically active, has not been performed. Labour force participation at a particular point in time is likely to be somewhat less informative about past labour force participation than current occupation is of past occupation.

The rationale for the present study is that the sizes of different education-occupation or education-industry cells differ by several orders of magnitude. The Spearman rank order correlation between the education and occupation classes is approximately 0.65.

3.2 Model Specifications and Estimation Techniques

Across all model alternatives below, the mean is assumed to be a log-linear function of the predictors:

$$\log \lambda_i = \mu + \sum_{m=0}^{M} \beta^{(m)}_{j_i^m}$$

Here, $\beta^{(m)}$ is the vector of coefficients of batch m, and j_i^m is the appropriate index for individual i. For example, if education represents the first batch of coefficients, β^1 would be the set of coefficients corresponding to the education categories, j_i^1 would be the index of the education category of individual i, and $\beta^1_{j_i^1}$ the corresponding coefficient.

Unless mentioned otherwise, the probability model connecting the predicted value λ_i with the observed counts is a Poisson count model:

$$y_i \sim Poisson(\lambda_i)$$

As is well known, the Poisson specification constrains the counts to be equidispersed, i.e. the variance to equal the mean. As a matter of fact, the observed counts tend to be *under*dispersed, which also explains why individual stochastic variation is not modelled with an individual error term in the (log-)linear predictor, because such an approach will on the contrary always result in overdispersion. To avoid any error associated with the violation of the assumption of equidispersion, the final model is specified as a Gamma count model. To our knowledge this is the only one of a very limited number of generalized dispersion count models that can model both overdispersion and underdispersion that was developed specifically with applications to fertility (among other things) in mind (Winkelmann 1995, 1996). The interpretation is that waiting times between births are assumed to follow a Gamma distribution (rather than an exponential distribution, as in the Poisson model). Depending on the parameters of this Gamma distribution, the hazard can be modelled to increase or decrease as a function of the waiting time, corresponding to underdispersion and overdispersion respectively.

Specifically, the Gamma count model takes the following form:

$$P(y_i = n) = G(\alpha n, \beta T) - G(\alpha(n+1), \beta T)$$

for $n = 0, 1, 2, \ldots$, where $G(\alpha k, \beta T)$ is the regularized lower incomplete Gamma function.

$$G(\alpha k, \beta T) = \frac{1}{\Gamma(n\alpha)} \int_0^{\beta T} u^{n\alpha - 1} \exp^{-u} \, \mathrm{d}u$$

We have $\alpha, \beta \in \mathbf{R}^+$ and $G(0, \beta T) \equiv 1$ by assumption. In our setting, $T = 1$ may be assumed without loss of generality. Then α is the dispersion factor, $\frac{\alpha}{\beta}$ is the mean waiting time between births, and asymptotically, $\frac{\beta}{\alpha}$ is the expected count. For $\alpha = 1$, the model reduces to the special case of Poisson counts. The regression is for the waiting times, so that we may equate the linear predictor with $\lambda_i = \frac{\alpha}{\beta_i}$ Accessible derivations for the Gamma Count model are provided in Winkelmann (2008).

Model estimation was performed within a Bayesian framework with vague or weakly informative priors. Priors for coefficients in the linear predictor term $\log \lambda_i$ are normally distributed as follows.

$$\mu, \alpha \sim N(0, 1)$$

$$\beta_j^{(m)} \sim N(0, \sigma_m^2)$$

$$\sigma_m \sim \text{half-Cauchy}(0.5)$$

For the scale hyperparameters, weakly informative prior distributions are assumed, specifically half-Cauchy distributions, as recommended by both Gelman and Hill (2007) and Polson and Scott (2011). The scale hyperparameters of these half-Cauchy priors are set to 0.5. While this is much smaller in magnitude than the "rule of thumb" suggestions of 25, it is more appropriate for our predictor on a log scale. In fact, arguably it is still too large. The 5 percent and 95 percent quantiles of the half-Cauchy with scale parameter equal to 0.5 are approximately 0.04 and 6.35 respectively. With a standard deviation of 6.35 as a hyperparameter for the prior of a regression coefficient, a value of 10 on the log scale would still be comfortably within two standard deviations of the (zero) mean. In other words, even multiplicative effects of a more than 10,000-fold increase in the number of children are granted a non-negligible prior probability. The fact that no reasonable effect size is excluded a priori by this specification justifies its characterization as merely "weakly informative".

In terms of estimation algorithm, the models in sections 4.2 and 4.3 are estimated using a fast Integrated Nested Laplace Approximation (INLA) algorithm, within the Bayesian framework outlined above. This technique is approximate, but several orders of magnitude faster than traditional Markov Chain Monte Carlo (MCMC) sampling. For exponential family specifications, which includes the Poisson distribution, the accuracy of the approximation is generally considered fairly good (Rue, Martino and Chopin 2009). In the definitive model, no high-order interactions are included; As a result, MCMC sampling becomes feasible. The starting points for the algorithm were provided by the INLA estimates of a regular Poisson model containing the same set of predictors. The results presented in Section 4.4 are based on sampling chains of 8,000 samples (after thinning) after reaching approximate stationarity according to Geweke's criterion.

4 Analysis and Results

Following a descriptive overview of the main apparent patterns in the data, a sequence of inferential models is fitted to investigate different aspects and refine the final specification. Section 4.2 fits a model containing all two-, three-, and four-way interactions between education, occupation, industry, and country to observations from the year 2000 census round, in order to identify the major direct and interaction effects to retain in later models. Section 4.3 includes only this reduced set of predictors, but interacts them with time and is fitted to data from both census rounds in order to examine the presence of changes over time. The changes are found to be minor, justifying the pooling of the two census rounds (the age groups are restricted to 40–49, so there is overlap between observed cohorts at the two points in time, which are ten years apart) in the final model. This makes use of the tractability of the more compact set of predictors to generalize the dispersion structure of the probability model.

4.1 Univariate

Figures 4.1, 4.3, and 4.4 display the level of our fertility indicator CEB by level of education, industry, and occupation group respectively, in each case for women aged 45-54 during the 2000-2002 census round. With regard to education, the expected pattern of a negative univariate association of education with fertility is on the whole confirmed, with some important caveats. Firstly, there are large differences in the size of the education differential between countries, ranging from less than 1 child in Greece to almost 2 children in Romania. This greater differential is actually achieved at both ends: the highest education group has a *lower* average CEB than the other countries, and the lowest group has a higher average CEB. In Austria the education differential between the top and bottom education level is smallest among the countries.

Similar observations can be made with respect to childlessness (c.f. Figure 4.2). For education levels of primary and up, more education is consist-

Figure 4.1 Children ever born (CEB) by education (women aged 40–49, 2000-2002 census round)

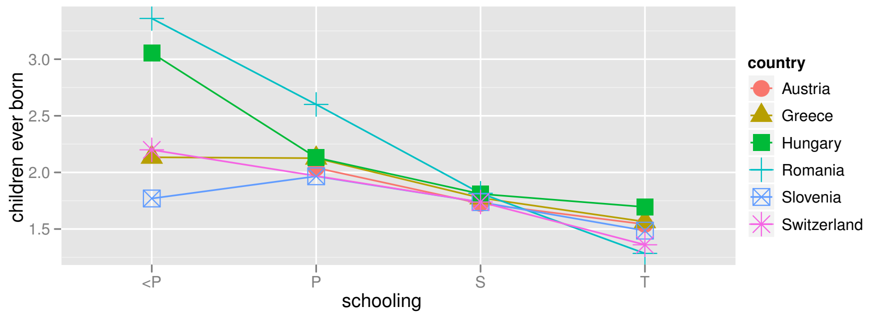
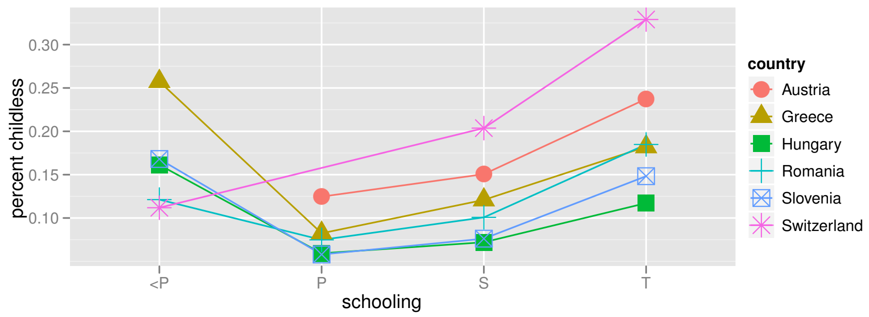


Figure 4.2 Percent childless (PCL) by education (women aged 40–49, 2000-2002 census round)



ently associated with a higher proportion of childless women in all countries. There is strikingly less variation in the gradient between countries than for CEB. Together with the fact that the differences in childlessness (of less than 10 percentage points) fall short of explaining the differences in CEB, this suggests that the effect of education is stronger at higher birth parities.

The seemingly anomalous pattern of the “less than primary” group being associated with particularly high childlessness, counter to the education effect at higher levels, may be a result of negative selection bias, as this category is probably more likely to include the permanently institutionalized and other categories of women with a relatively low chance of attracting a partner.

Figure 4.3 Children ever born (CEB) by industry (women aged 40–49, 2000-2002 census round)

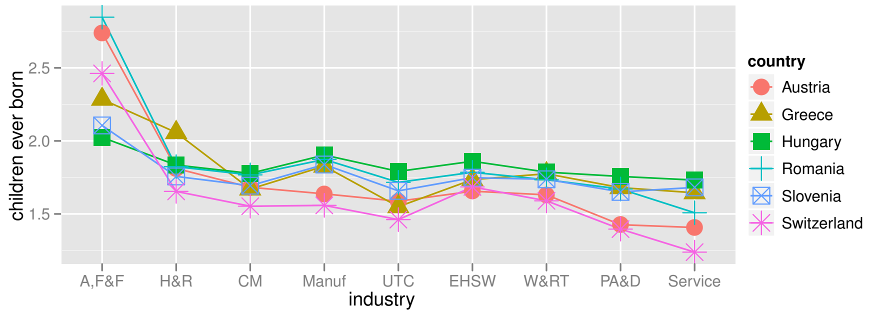
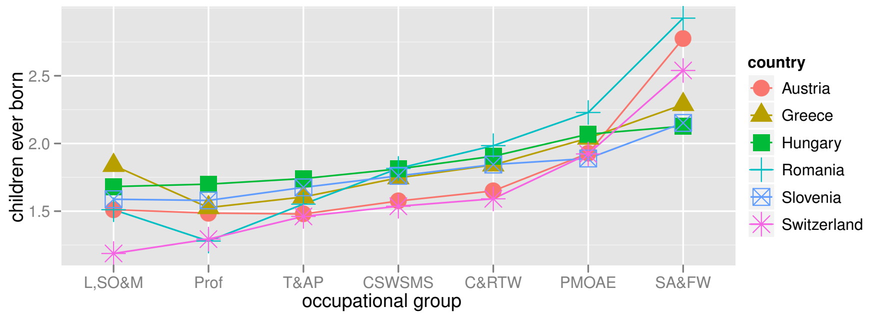


Figure 4.4 Children ever born (CEB) by occupation (women aged 40–49, 2000-2002 census round)

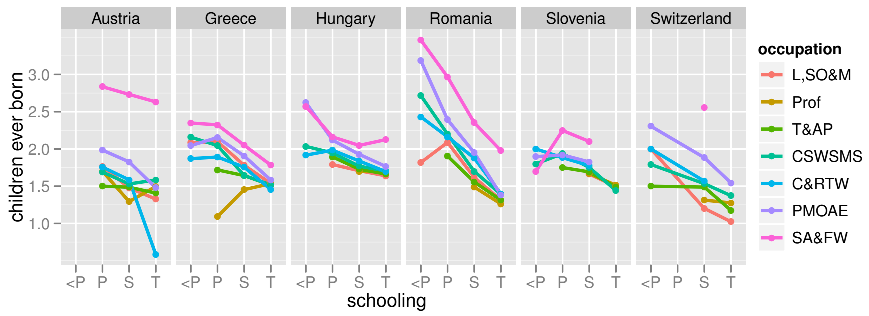


With respect to CEB and industry group (4.3), there is relatively little variation as compared to education levels and occupation groups, even between what might *a priori* be considered extremes, such as Education, Health and Social Work on the one hand and Construction and Mining on the other. The sole exception is Agriculture, Fishing and Forestry, which can easily be attributed to a natural rural bias of these industries. When it comes to differences in CEB by occupation (4.4), again the outlier SA&FW is likely to be biased upwards because of its rural predominance.

4.2 Interactions

A key objective of the present study is to analyse education differentials *within* occupation groups and industries.

Figure 4.5 Children ever born (CEB) by education and occupation (women aged 40–49, 2000-2002 census round)



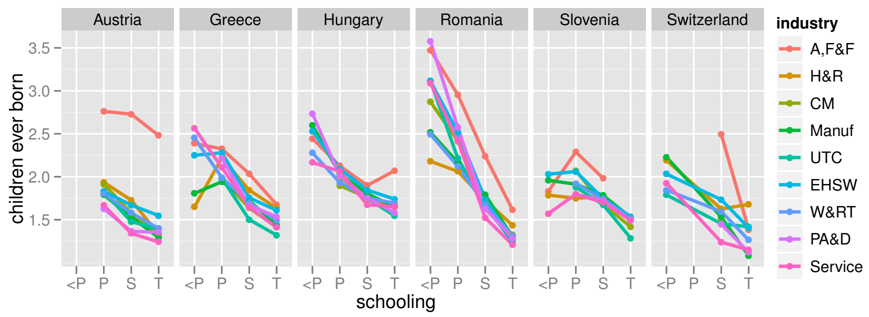
There is a fairly consistent pattern, evident in Figure 4.5, that in all countries, within the vast majority of occupation groups, there is an educational fertility differential in the expected direction, suggesting that education has a depressing effect on number of children independently of social status, income, and other factors normally associated with occupational status. In these and the following graphs, education-occupation dyads with fewer than 20 individuals have not been plotted. For many country-occupation combinations, every step up the education scale is associated with fewer children. In quite a few cases however, the difference between completed secondary and completed tertiary is marginal, non-existent or negative, especially in Austria. Nevertheless, even in these cases fertility is higher for the lowest than the highest education category in virtually all country-occupation combinations.

The same patterns were confirmed when we examined the extremes of the parity distribution — childlessness and those with four or more children (results not shown).

For childlessness (PCL), this is, however, only true with respect to the three upper education levels. In our data women with less than primary education exhibit some of the highest levels of childlessness, but this is not surprising. Since in all six countries the completion of primary school is expected universally, among those who did not we would expect to find a disproportionate number of those with health or other issues that at the same time reduces their attractiveness as partners.

The high-parity measure of the proportion with four children or more does not suffer from this distortion, and again shows a more consistent pattern. Apart from the occupation categories with extremely low shares of high-parity women (namely Professionals and Technicians & Associate Professionals), there is on the whole a consistent education differential in the expected direction in all countries.

Figure 4.6 Children ever born (CEB) by education and industry (women aged 40–54, 2000-2002 census round)



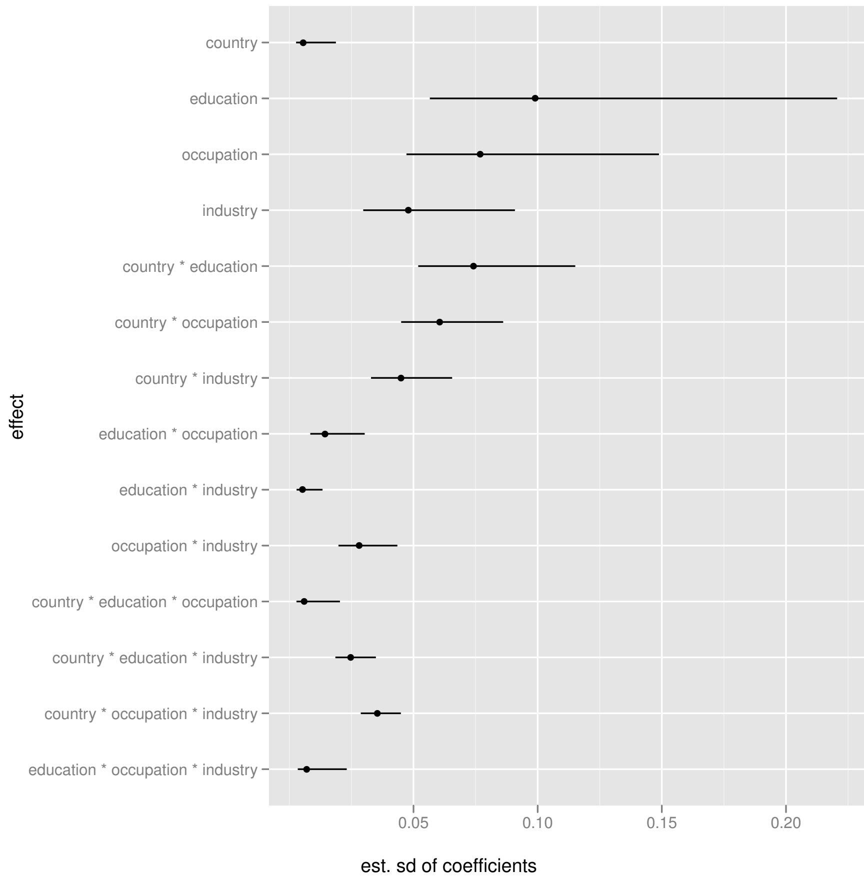
The above analysis was repeated with education level by industry (see Figure 4.6). Compared to occupation, the industry of work appears to have relatively little effect on education differentials. In particular, education differentials universally persist even in industries intuitively associated with a selection bias towards "child-friendly" women, namely in education and health and social work. Considering PCL did not change these conclusions. The results for the industry variable suggest that, perhaps unsurprisingly, industry of work at time of survey is a poor proxy for field of study, since the latter has been shown to matter to fertility (Lappegard and Ronsen 2005).

In order to maximize the number of observations even for rare combinations, this preliminary model is fitted to the CEB of all women aged 35–54 in the data from the year 2000 census round. While necessary at this stage for reasons of sample size, and although initial analyses indicated that women aged 35–39 have essentially completed their fertility, there remains a valid concern about mixing the experience of different cohorts here. Accordingly, once the most important predictors are identified, subsequent models are limited to a narrower age range.

Here and in the following sections, because interactions are included that contain dozens or hundreds of dummy variables for specific combinations of predictors, the discussion of results focuses on the amount of variation accounted for by different batches of coefficients (all occupation dummies, for example, or all occupation $\times$ education dummies). This tactic can be considered to be a form of Bayesian ANOVA (Gelman 2005). It is similar in spirit to classical ANOVA, but can deal with heavily unbalanced data, and can straightforwardly be applied to arbitrary models. The size of the standard deviation for each batch of coefficients indicates its relative importance as a source of variation in the outcome. For our data, the estimated standard deviations are displayed in Figure 4.7.

While the uncertainty surrounding the estimates is substantial, as evidenced by the interval widths, there is a clear indication that education and occupation are the most important sources of variation in CEB independently of country. This is not to say, however, that their effects are not mediated by the country setting, which they are. In fact, the interactions of country with education and occupation respectively are the next most influential sets of predictors.

Figure 4.7 ANOVA display for model regressing CEB on all interactions of education, occupation, industry, and country. Point: posterior median, Bar: posterior 80 percent interval

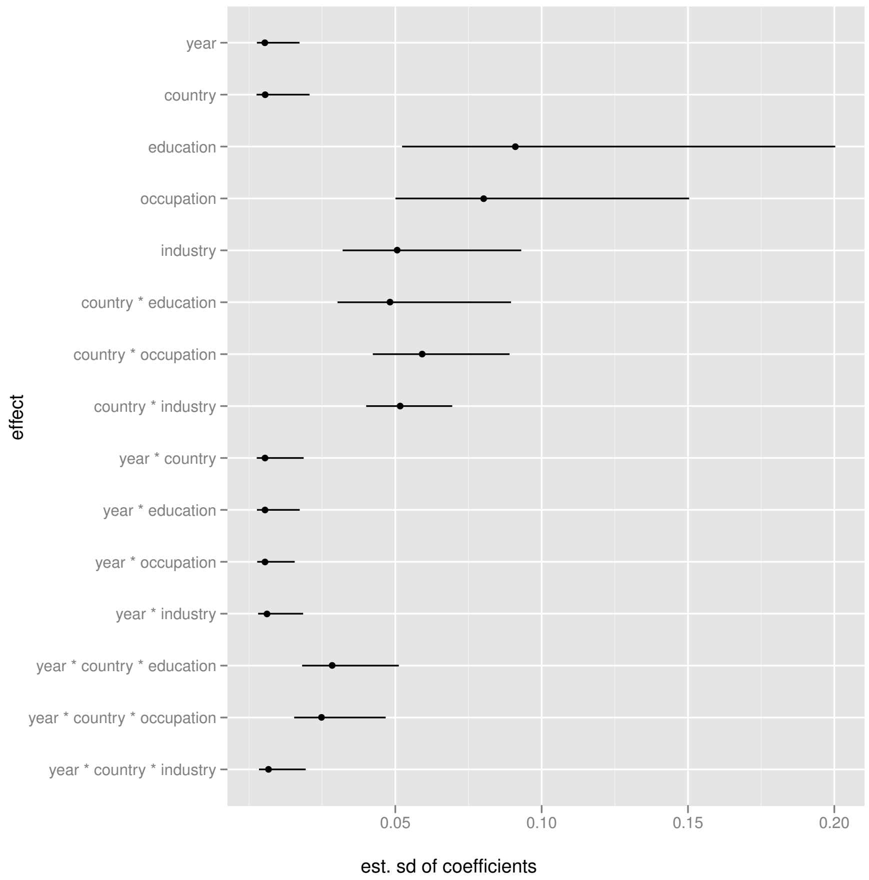


Next is the main industry effect, followed by its country interaction. The other interactions are minor compared to the main effects above.

Strikingly, the *direct* country effect is estimated to be slight. The influence of the country setting on the level of completed cohort fertility is estimated to occur almost entirely through the interaction with education, occupation, and industry, and, to a lesser extent, their two-way interactions involving industry. In other words, observed differences between the study countries in completed fertility level can, for these cohorts, be attributed almost entirely to differences in composition and its consequences. While the countries in question do form an almost contiguous block in central and south-eastern Europe, their diversity along dimensions not covered in the present analysis is enormous, if we compare Switzerland with Greece, say. This suggests the result may well hold more generally.

These insights could only be obtained jointly using the IPUMS data. While national panel or register datasets are indispensable for sophisticated causal estimation, sources as rich as those used by Kravdal (2007), for example, are not available for a large number of countries, much less in a harmonized form. Studies extracting the maximum of information from such national sources are therefore unable to estimate country effects in a cross-country model. Also, comparative international datasets resulting from coordinated surveys never reach a sufficient sample size to reliably estimate the three- and four-way interactions in the above model, because most of the rare cells would remain unobserved. The country $\times$ education $\times$ occupation $\times$ industry interaction component, for example, consists of over 1000 indicator variables. The result concerning negligible interactions is reassuring, therefore, because it confirms that little information is lost using simpler data, and indicates which effects and interactions are important to include in simpler models. The latter also benefits the remainder of the present study. Subsequent models will focus on the six most influential variance components seen here, in other words: education, occupation, industry, and their country interactions.

Figure 4.8 ANOVA display for model regressing CEB on main effects and interactions, including time. Point: posterior median, Bar: posterior 80 percent interval



4.3 Changes Over Time

The reduced set of predictors established above allows us to include time as an additional dimension, without the results becoming too unwieldy.

The results (4.8) suggest that, while the unconditional education status is likely to be the single greatest source of variation in CEB among the women in the study, the association between education and fertility is at the same time the most context-dependent. It is estimated to have the strongest interaction with the country indicator among all the country interactions, but also the greatest change over time in this interaction. The association between education and completed cohort fertility depends strongly on context, but is always considerable.

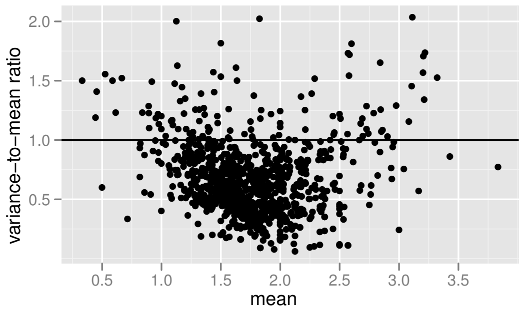
Mirroring the absence of an independent country effect (which is confirmed here), there is no secular time trend. In other words, the change in average CEB between the earlier and later census rounds is entirely due to changes in the composition of the education, occupation, and industry of employment of women.

There is virtually no change over time in the direct association of CEB with country, education, occupation, or industry. The higher-order interactions involving time are also marginal, with the exception of the above-mentioned time change in the country $\times$ education interaction. Even this, however, remains smaller than the six dominant components selected earlier. Accordingly, the same selection is retained for further analysis.

4.4 Generalized Dispersion

Based on the absence of significant changes over time, observations from both census rounds are pooled in the final analysis. The above results are all based on a standard Poisson count model for the number of births. Because the Poisson assumption of equidispersion, in other words, that the variance equals the mean, is clearly violated in the data (c.f. Figure 4.9), an alternative Gamma count model with flexible dispersion was fitted (see Section 3.2). In particular, a bivariate comparison of the variance and mean of the CEB measure for the different groups defined by predictor variables shows strong *under*dispersion, especially in the range of moderate fertility. Such a pattern is consistent with intentional behaviour targeted towards specific parities, such as 1 or 2 children, as might be expected in the countries under study.

Figure 4.9 Variance-to-mean ratio of CEB by country, education, occupation, and industry (excluding groups smaller than ten to avoid unstable variance estimates)



The more general specification does make a difference in the estimation results, but without affecting the substantive conclusions. The variance-to-mean ratio is estimated to be approximately 0.62, a considerable amount of underdispersion. Figures 4.10 and 4.11 display the estimated parameters for the education and occupation categories. While the education effect is attenuated, vindicating the effort to take dispersion into account, there remains a clear and consistent education differential after controlling for occupation and industry, as well as country context. The magnitude of the largest negative and positive effects of education and occupation respectively are broadly similar. The smaller magnitude of occupation as a variance component is mainly due to the fact that the fertility differences between the three highest occupation categories specifically are minimal. For further study, where occupation serves purely as a control variable rather than a factor of intrinsic interest, it might therefore be appropriate to collapse these groups.

Relative Model Fit

Comparing the Deviance Information Criterion (DIC) for these two specifications suggests that the model accounting for dispersion provides a small, but

Figure 4.10 Education effects on predicted fertility in generalized dispersion count model (inverse effect on Gamma waiting time between births). Point: posterior median, Thick bar: posterior 50 percent interval, Thin bar: posterior 90 percent interval

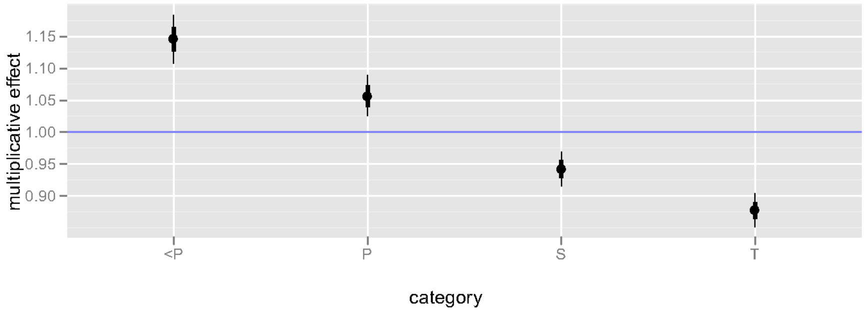
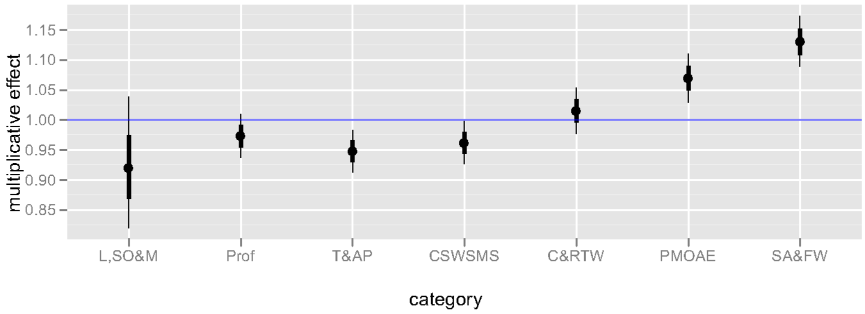


Figure 4.11 Occupation effects on predicted fertility in generalized dispersion count model (inverse effect on Gamma waiting time between births). Point: posterior median, Thick bar: posterior 50 percent interval, Thin bar: posterior 90 percent interval



measurably better fit to the data. A more telling indication of the good model fit is provided by performing posterior predictive checks.

A first check consists of comparing the average predicted values with the empirical averages for each combination of the predictors. When weighted by the number of observations for each combination, the Spearmann rank correlation coefficient for the predicted and observed average CEB is 0.98. Unweighted, i.e. treating each combination as equally important prediction targets and ignoring the different amounts of uncertainty in the true average, the correlation naturally drops, but remains at a high value. Specifically, the Spearman rank correlation between the predicted and observed group averages of CEB is 0.68. The good model fit is unsurprising, since the full interaction model from Section 4.2 had sufficiently many degrees of freedom to achieve a perfect fit, and the analysis showed that the predictors retained capture most of the variation in that model.

At the above level of analysis, the values for the Poisson model are virtually identical, at 0.98 and 0.68. However, the benefit of more careful modelling is evident when comparing indicators of interest other than the mean, such as the predicted share of childlessness or high parities. To make these comparisons, samples the size of the empirical sample were drawn across the posterior parameter distributions and the count distribution conditional on the parameters. With respect to childlessness, the observed value in the overall sample is 0.12. The share predicted by the Gamma count model is 0.11, a much closer fit than the 0.16 implied by a Poisson model. At the other end of the parity distribution, the observed share of high parities of four children or more was 0.1. Again, the Gamma model with 0.11 resembles this more closely than the Poisson model with 0.14. The relative shares *within* the range of between one and three children are still not fit sufficiently well by either model (although the Gamma model again performs better) to conclude that further work in modelling fertility distributions is not required. Nevertheless, the Gamma model represents a clear improvement over the Poisson model in replicating the empirical distribution of CEB, and is accordingly preferred for inference.

5 Discussion

An understanding of the interplay between education and occupation in influencing fertility is vital to assessing the likely impact of long-term labour market trends on demographic change. One of the obstacles to research in this area is the fact that education and occupational status tend to be highly correlated. Disentangling their associations with fertility is therefore an important task, orthogonal to the question of the direction of causation. At the same time, doing so does provide some hints as to likely causal mechanisms.

Based on the analysis presented here, it can be asserted that among the populations investigated here, fertility varies by education level, net of occupation level and industry, in a manner that is broadly consistent across a number of countries. Moreover, these direct education differences are often stronger than those of the other two factors.

One interpretation of this result is that it is inconsistent with attempting to explain educational fertility differentials as being largely driven by income effects, since we would expect income to be determined by occupation rather than educational attainment *per se*. Since both education and occupation affect social status both directly and indirectly, the implication is that using one or the other as a proxy for social status in the analysis of fertility will be incomplete at best, or worse misleading. An interesting aspect might be to investigate the effect of the *(mis)match* of education and occupation on fertility. The match between the two has been found to possess some explanatory power with regard to migration (Quinn and Rubb 2005). Intuitively, it seems plausible that a mismatch could result either in investing time and effort into further training or job search, both of which may negatively impact fertility intentions, or, conversely, indicate a change of mind towards higher fertility resulting in a career change towards a more parent-friendly occupation.

Interestingly, the analysis shows that differences between the countries in the study in terms of completed cohort fertility can be explained almost entirely through composition effects and contextual interactions with education, occupation, and industry. This insight complicates future cross-country analysis, because it means that, even once data have been harmonized, adding only independent country effects represents a misspecification.

A warning that can be extracted from the analysis is the need to carefully distinguish between industry and occupation. As is evident from many of the results above, the conclusions along these two dimensions differ substantially. This issue needs to be carefully considered when disaggregating education by "subject of study", since some disciplines (such as law) tend to predetermine an occupational category, while others, especially vocational training, may instead guide towards particular industries (say, tourism).

While allowing rare combinations such as senior managers with primary schooling to be observed, the data is limited in other ways. Only the two most recent census rounds made the requisite variables available for a significant number of European countries. As such, real longitudinal analysis is impossible. As already mentioned, the nature of the fertility indicators means completed fertility can only be estimated for relatively high ages, and as such the conclusions are not very current. More serious is the fact that no data is available on proximate determinants of fertility to control for indirect effects of education or occupational status through contraceptive use, for example. In the absence of such data, it remains unclear whether education works through behavioural changes or attitudinal/agency changes. An open question that cannot be conclusively answered based on the data at hand is to what extent the association between higher education and lower fertility is caused by selection, reflecting the fact that early and high fertility may represent an obstacle to attaining further education. In the generation of women examined, the occupation of the husband no doubt also plays a major role in determining social status. In principle, the IPUMS data allows this information to be linked in, paving the way for further analysis in this direction.

References

Adsera, Alicia. 2011. The interplay of employment uncertainty and education in explaining second births in Europe. *Demographic Research* 25 (16): 513–544.

Akin, Mustafa Seref, and Valerica Vlad. 2007. No Eastern and Central European left behind: A cross country regression for fertility, human capital and market economy. *Journal of International Development* 19 (7): 963–974.

Alba, Alfonso, Gema Alvarez and Raquel Carrasci. 2008. On the estimation of the effect of labour participation on fertility. *Spanish Economic Review* 11:1–22.

Axinn, William G. 1993. The effects of children's schooling on fertility limitation. *Population Studies* 47 (3): 481–493.

Basu, Alaka Malwade. 2002. Why does education lead to lower fertility? a critical review of some of the possibilities. *World Development* 30 (10): 1779–1790.

Becker, Gary. 1981. *A treatise on the family.* Cambridge Mass.: Harvard University Press.

Billari, Francesco C., and Dimiter Philipov. 2004. Women's education and entry into a first union: A Simultaneous-Hazard comparative analysis of Central and Eastern Europe. *Vienna Yearbook of Population Research* (Vienna, Austria) 2:91–110.

Blossfield, H.P., and J. Huinink. 1991. Human capital investments or norms of role transition? How women's schooling and career affect the process of family formation. *American Journal of Sociology* 97:143–168.

Brewster, Karin L., and Ronald R. Rindfuss. 2000. Fertility and women's employment in industrialized nations. *Annual Review of Sociology* 26:271–296.

Caldwell, John C. 1980. Mass education as a determinant of the timing of fertility decline. *Population and Development Review* 6 (2): 225–255.

Cochrane, Susan Hill. 1979. *Fertility and education: what do we really know?* Baltimore, MD: Johns Hopkins University Press.

———. 1983. Effects of education and urbanization on fertility. In *Determinants of fertility in developing countries.* Edited by R.A. Bulatao & R.D. Lee. New York: Academic Press.

Desai, Sonalde, and Linda J. Waite. 1991. Women's employment during pregnancy and after the first birth: occupational characteristics and work commitment. *American Sociological Review* 57 (4): 551–566.

Edwards, Mark Evan. 2002. Education and occupations: reexamining the conventional wisdom about later first births among American mothers. *Sociological Forum* 17 (3): 423–443.

Gelman, A. 2005. Analysis of variance: why it is more important than ever. *Annals of Statistics* 33 (1): 1–31.

Gelman, A., and J. Hill. 2007. *Data analysis using regression and multilevel/hierarchical models.* Cambridge University Press.

Gerson, Kathleen. 1985. *Hard choices: how women decide about work, career and motherhood.* Berkeley, CA: University of California Press.

Gerster, M., N. Keiding, L.B. Knudsen and K. Strandberg-Larsen. 2007. Education and second birth rates in Denmark, 1981-1994. *Demographic Research* 17 (8): 181–202.

Hakim, Catherine. 2003. A new approach to explaining fertility patterns: preference theory. *Population and Development Review* 29 (3): 349–374.

Hantrais, Linda. 1997. Exploring relationships between social policy and changing family forms within the European Union. *European Journal of Population* 13 (4): 339–379.

Hoem, Britta, and Jan M. Hoem. 1989. The impact of women's employment on 2nd and 3rd births in modern Sweden. *Population Studies* 43:47–63.

Hoem, Jan M. 1986. The impact of education on modern family-union initiation. *European Journal of Population* 2:113–133.

Janssen, M., and M. Kalmijn. 2002. Investments in family life: the impact of value orientations and patterns of consumption, production and reproduction in married and cohabiting couples. In *Meaning and choice: value orientations and life course decisions,* 129–159. Edited by Ron Lesthaeghe. The Hague, Brussels: NIDI/CBGS Publications.

Jejeebhoy, Shireen J. 1995. *Women's education, autonomy and reproductive behavior: experience from developing countries.* Oxford: Clarendon Press.

Kasarda, J.D., J.O.G. Bill and K. West. 1986. *Status enhancement and fertility: reproductive responses to social mobility and educational opportunity.* New York: Academic Press.

Kasarda, John D. 1979. How female education reduces fertility: models and needed research. *Mid-American Review of Sociology* 4 (1): 1–22.

Kravdal, Oystein. 1994. The importance of economic activity, economic potential and economic resources for the timing of first births in Norway. *Population Studies* 48:249–267.

———. 2007. Effects of current education on 2nd and 3rd birth rates among Norwegian men and women born in 1964. *Demographic Research* 17 (9): 211–246.

Lappegard, Trude. 2002. *Educational attainment and fertility patterns among Norwegian women.* Statistics Norway, Department of Statistics.

Lappegard, Trude, and Marit Ronsen. 2005. The multifaceted impact of education on entry into motherhood. *European Journal of Population* 21:31–49.

Lesthaeghe, Ron J. 1995. The second demographic transition in Western countries: An interpretation. In *Gender and family change in industrialized countries,* 17–62. Edited by K.O. Mason and A.M. Jensen. Oxford: Clarendon Press.

Lesthaeghe, Ron J., and Dominique Meekers. 1986. Value changes and the dimensions of familism in the European Community. *European Journal of Population* 2 (3/4): 225–268.

Lesthaeghe, Ron J., and G. Moors. 1995. Living arrangements, Socio-Economic position and values among young adults: A pattern description for Belgium, France, the Netherlands, and Western Germany, 1990. In *Europe's population in the 1990s.* Edited by D. Coleman. Oxford: Oxford University Press.

Liefbroer, Aart C., and Martine Corijn. 1999. Who, what, shere, and when? specifying the impact of educational attainment and labor force participation on family formation. *European Journal of Population* 15:45–75.

Martin, Teresa Castro. 1995. Women's education and fertility: results from 26 Demographic and Health Surveys. *Studies in Family Planning* 26 (4): 187–202.

Minnesota Population Center. 2011. *Integrated public use microdata series, international: version 6.1.* Minneapolis: University of Minnesota.

Polson, N. G., and J. G. Scott. 2011. On the half-Cauchy prior for a global scale parameter. *ArXiv e-prints,* no. 1104.4937.

Quinn, Michael A., and Stephen Rubb. 2005. The importance of Education-Occupation matching in migration decisions. *Demography* 42 (1): 153–167.

Rindfuss, Ronald R, S. Philip Morgan and Kate Offutt. 1996. Education and the changing age pattern of fertility: 1963-1989. *Demography* 33 (3): 277–290.

Rue, H., S. Martino and N. Chopin. 2009. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the Royal Statistical Society: Series B (statistical methodology)* 71 (2): 319–392.

Scherer, Stefani. 2005. Patterns of labour-market entry – long wait or career instability? An empirical comparison of Italy, Great Britain, and West Germany. *European Sociological Review* 21 (5): 427–440.

Sobotka, Tomas. 2009. Sub-Replacement fertility intentions in Austria. *European Journal of Population* 25:387–412.

Speder, Zsolt. 2006. Rudiments of recent fertility decline in Hungary: postponement, educational differences, and outcomes of changing partnership forms. *Demographic Research* 15 (8): 253–288.

Spielauer, Martin. 2005. Concentration of reproduction in Austria: general trends and differentials by educational attainment and Urban-Rural setting. *Vienna Yearbook of Population Research* (Vienna, Austria) 3:171–195.

Stolzenberg, Ross M., and Linda J. Waite. 1977. Age, fertility expectations and plans for employment. *American Sociological Review* 42 (5): 769–783.

Swanberg, Jennifer E., Marcie Pitt-Catsouphis and Krista Drescher-Burke. 2005. A question of justice : disparities in employees' access to flexible schedule arrangements. *Journal of Family Issues* 26:866–895.

Van Bavel, Jan. 2010. Postponement of motherhood in europe: the impact of expected earnings, gender composition, and family attitudes. *Demography* 47 (2): 439–458.

Winkelmann, R. 1995. Duration dependence and dispersion in count-data models. *Journal of Business & Economic Statistics* 13:467–474.

Winkelmann, R. 1996. A count data model for gamma waiting times. *Statistical Papers* 37 (2): 177–187.

———. 2008. *Econometric analysis of count data.* Springer Verlag.