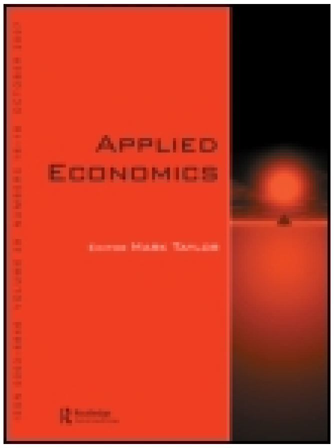

Applied Economics
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/raec20

Job preferences of students and new graduates in nursing

Denise Doiron[a], Jane Hall[b], Patricia Kenny[b] & Deborah J. Street[c]

[a] School of Economics, ASB, University of New South Wales, 2052, Australia

[b] CHERE, University of Technology, Sydney, NSW 2007, Australia

[c] School of Mathematical Sciences, University of Technology, Sydney, NSW 2007, Australia

Published online: 28 Jan 2014.

To cite this article: Denise Doiron, Jane Hall, Patricia Kenny & Deborah J. Street (2014) Job preferences of students and new graduates in nursing, Applied Economics, 46:9, 924-939, DOI: 10.1080/00036846.2013.861584

To link to this article: http://dx.doi.org/10.1080/00036846.2013.861584

Applied Economics, 2014
Vol. 46, No. 9, 924–939, http://dx.doi.org/10.1080/00036846.2013.861584

Job preferences of students and new graduates in nursing

Denise Doiron[a], Jane Hall[b], Patricia Kenny[b] and Deborah J. Street[c,*]

[a]*School of Economics, ASB, University of New South Wales, 2052, Australia*
[b]*CHERE, University of Technology, Sydney, NSW 2007, Australia*
[c]*School of Mathematical Sciences, University of Technology, Sydney, NSW 2007, Australia*

This article investigates the preferences of student and newly graduated nurses for pecuniary and nonpecuniary aspects of nursing jobs. It is the first study applying methods based on discrete choice experiments to a developed country nursing workforce. It is also the first to focus on the transition through university training and into work. This is particularly important as junior nurses have the lowest retention levels in the profession. We sample 526 individuals from nursing programmes in two Australian universities. Flexible and newly developed models combining heteroscedasticity with unobserved heterogeneity in scale and preference weights are estimated. Overall, salary remains the most important feature in increasing the probability that a job will be selected. 'Supportive management/staff' and 'quality of care' follow as the most important attributes from a list of 11 nonpecuniary characteristics. However, the subset of new graduates rank 'supportive management/staff' above salary increases, emphasizing the importance of a supportive workplace in the transition from university to the workplace. We find substantial preference heterogeneity and some attributes, such as the opportunity for clinical rotations, are found to be attractive to some nurses while seen as negative by others. Nursing retention could be improved by designing different employment packages to appeal to these different tastes.

Keywords: nursing workforce; heterogeneous job preferences; discrete choice experiment; generalized multinomial logit model

JEL Classification: J21; J62

I. Introduction

Policy-makers face a challenge in recruiting and retaining the largest component of the professional health workforce, namely nurses. An adequate supply of appropriately qualified nurses is needed to ensure that health services can be provided and that care is of the appropriate quality (Aiken *et al.*, 2002; Heinz, 2004; Needleman and Hassmiller, 2009). Population ageing and the growing incidence of disability are increasing the demand for nurses (Oulton, 2006). At the same time, the expansion of nursing roles in primary care, chronic disease management and preventive services is an important component of reforms aimed at improving the efficiency and affordability of health systems (Productivity Commission, 2005; Rother and Lavizzo-Mourey, 2009). Yet there are problems in supply, at the same time as the demands on the workforce are increasing. Many countries already face a shortage, which will be

*Corresponding author. E-mail: deborah.street@uts.edu.au

exacerbated by the ageing of the current workforce with many working nurses moving into retirement in the next decade. Nursing is still predominantly a female occupation with relatively low pay rates; and younger females with dependents or those simply changing careers have higher attrition rates (Nooney *et al*., 2010). The typical policy response is to increase pay rates and/or train more nurses. But the evidence shows that both these responses will have limited effect. Nursing wage elasticities are relatively low; indeed many nurses leave nursing for jobs with lower wages (Shields, 2004). Increasing supply through the provision of additional training places requires student nurses to complete their course, and then move into, and stay in, the nursing workforce. However, attrition rates among young and newly registered nurses are high (Aiken *et al*., 2002; Sochalski, 2002; Barron and West, 2005; Fochsen *et al*., 2006; Doiron *et al*., 2008) and there is evidence that the transition from student to registered nurse can be particularly stressful (Casey *et al*., 2004).

There is a growing body of evidence that nonpecuniary factors are significant in improving nursing retention (Shields and Ward, 2001). Hours and flexibility are important; for example, the choice of part-time or full-time work (Di Tommaso *et al*., 2009; Zeytinoglu *et al*., 2011). Workload is also a factor; stress and high workloads (Zeytinoglu *et al*., 2006; Coomber and Barriball, 2007) and management responsibilities (Frijters *et al*., 2007) have a negative effect on retention. Other nonpecuniary factors contributing to the decision or intent to leave a nursing job have included a lack of autonomy, work relationship problems (Estryn-Behar *et al*., 2010) and supervisory relationships in particular (Coomber and Barriball, 2007). Finally, supportive work environments (Zeytinoglu *et al*., 2011) and having opportunities for further training (Frijters *et al*., 2007) improve retention. There is also evidence of heterogeneity in retention across nurses (Frijters *et al*., 2007; Cunich and Whelan, 2010), but beyond the suggestion that working conditions should accommodate the needs of women with young families, there has been little investigation of this aspect of nursing labour supply (Doiron *et al*., 2008; Cunich and Whelan, 2010) and therefore little to guide policy-makers or hospital managers in addressing this issue.

The available data sets for studying the nursing labour force, primarily general household surveys or registration data, do not contain sufficiently rich information to allow for detailed study of the range of factors that influence job satisfaction and heterogeneity in preferences. While surveys enable the researcher to collect more detailed individual data, they are often limited by the range of jobs and job characteristics currently in place, particularly where pay rates and other conditions are set centrally. Stated preference techniques have become an increasingly popular approach to overcome the lack of revealed preference data. Perhaps surprisingly, given the widespread popularity of discrete choice experiments (DCEs) in health economics, and the acceptance of this approach in labour economics, there are few applications to job preferences of health workers.[1] A few of these include nurses, all of them are set in a developing country context and investigate nurses' willingness to take jobs in rural locations (Penn-Kekana *et al*., 2005; Mangham and Hanson, 2008; Blaauw *et al*., 2010). All studies demonstrate the importance of wages and nonpecuniary benefits, including the opportunity for further education and training, adequate equipment and infrastructure. However, findings from developing countries and the choice between rural and urban jobs is unlikely to transfer to the context of developed economies and nursing jobs.

This study uses the DCE approach to investigate nurses' preferences for both pecuniary and nonpecuniary aspects of nursing jobs. Australian data are used; hence, this article provides the first evidence on stated preferences over nursing jobs based on DCEs in a developed country. The hypothetical jobs are representative of those found in hospitals for new graduate nurses. We do not focus on a specific job characteristic but elicit preference rankings over a relatively large number of characteristics reflecting the complexity of modern nursing jobs and previously identified as important in the literature. We assess the strength of preferences using both predicted choice probabilities and willingness-to-pay (WTP) estimates.

A second novel aspect of the study is the focus on nurses through their training and transition from education to the workforce. As already noted, nurses in these years are especially vulnerable to attrition. The experiences in the early years of training and working as a nurse and the stress entailed, may well influence motivation and preferences over different job attributes. As is common in other developed countries, nurse education in Australia is based in universities (Robinson and Griffiths, 2007). It is likely that young students choose nurse training without experience on hospital wards and so have little idea of what it feels like to work as a nurse. Nursing courses include classroom learning, simulated experiences in laboratory tutorials and clinical placements in hospitals where they observe and practice nursing work in a structured and supervised way. In the analysis, we distinguish job preferences of nursing students according to the year in the programme and graduation status. We are interested in seeing if the job preferences differ as the students have more experience with what nurses actually do.

Alongside the growing use of DCEs, there has been increasing attention to the appropriateness of the methods, both for survey design and for the analysis of data. The standard use of multinomial logit (MNL) models has been

[1] For example, see the survey by Lagarde and Blaauw (2009). For a more recent example involving GPs, see (Sivey *et al*., 2010).

overtaken by the use of the mixed logit (MXL) models to better account for heterogeneity in preferences across individuals (Keane and Wasi, 2013). While the importance of preference heterogeneity can be considered well established, recent contributions also point to the importance of scale heterogeneity; that is, differences across individuals in utility variance, often interpreted as an individual's uncertainty over preferences. The generalized multinomial logit (GMNL) model has been developed to address both scale and preference heterogeneity (Fiebig *et al*., 2010, Keane and Wasi, 2013). Indeed, Fiebig *et al*. (2010) conclude that scale heterogeneity is relatively more important where decisions are complex.

This study also contributes to the literature by implementing recently developed econometric models. Specifically, we compare results from standard MNL and MXL models to the newly developed GMNL model. Our large sample size is important in this respect. In addition, our DCE is designed to elicit best–worst choice information and we estimate rank-ordered and heteroscedastic versions of the MNL, MXL and GMNL models. The latter approach is new to the literature.

The rest of the article is set as follows: Section II describes the DCE and the development of attributes; Section III gives a fuller description of the sample and data collection; Section IV reports the model specification and tests; Section V discusses the results; Section VI reports the results for preferences by time in programme; and Section VII concludes.

II. The Choice Experiment

Theoretically larger choice sets (scenarios) give more information than do smaller ones but of course considering a large number of options at one time is cognitively demanding. We design a choice experiment in which respondents are shown a scenario of three hypothetical jobs described in terms of different levels of the same attributes and labelled Job A, Job B and Job C. Respondents are asked which they think is the best job and which they think is the worst job. Each respondent is asked this question for eight different scenarios.

The hypothetical jobs focus on the first job as a registered nurse. The job attributes are based on the literature describing job characteristics associated with nurse retention (Naude and McCabe, 2005; Hayes *et al*., 2006; Hogan *et al*., 2007). The attributes and levels are presented in Table 1; the experiment includes 12 attributes, 11 with two levels and one (salary) with four levels. The attributes are appropriate in the context of an entry-level job in a new graduate programme. In particular, job options are limited to hospitals, as almost all new graduates are employed in hospitals which offer a 'new graduates programme'. The 12 attributes cover salary and nonpecuniary aspects including those likely to be relevant to new graduates, for example clinical rotations, i.e. the opportunity to spend a period of time in different clinical specialties. The attributes were tested in a pilot study with 60 second-year nursing students. The pilot study feedback indicated that respondents generally found the scenarios to be understandable and appropriate. In the DCE, attributes were represented by a shortened name and each choice set had a link to an explanatory glossary; see Table 1.

As Lagarde and Blaauw (2009) note, the construction of DCEs can be complex. We briefly describe the construction of the DCE used in this project; more details may be found in Street *et al*. (2005) Street and Burgess (2007). The choice sets were constructed by determining an initial set of 16 jobs. These jobs were chosen so that both levels of the attributes with two levels appeared eight times each, and all salary levels appeared four times each. Further, in order to be able to say something about the effects of each attribute independently of other attributes, for any two attributes all possible pairs of levels appeared equally often. Thus for each salary level, each of the other attributes appeared twice at one possible level and twice at the other possible level. For any two-level attribute (e.g., clinical rotations) every other two-level attribute appeared four times with one of its possible levels and four times with the other of its possible levels, for each of the possible levels of clinical rotations (say). A set of jobs constructed with levels balanced in this way is called a resolution 3 fractional factorial design. To obtain the other two jobs in each choice set we create two further resolution 3 fractional factorial designs. We do this systematically by specifying how the levels in the first set of 16 jobs is to be changed to get the levels in the second set of 16 jobs, and the levels in the third set of 16 jobs. These rules for the systematic changes are referred to as the generators of the other two designs. The generators were chosen so that the resulting set of 16 choice sets of size 3 would be D-optimal for the estimation of main effects under the null hypothesis that all of the coefficients in the multinomial choice model are equal to 0. We constructed two sets of 16 choice sets using this technique, employing two different resolution 3 fractions so that the jobs in the DCE covered a larger proportion of the sample space. These two sets of 16 choice sets were each subdivided into two versions of 8 choice sets and respondents were randomized to one of the 4 versions. A sample choice set of three hypothetical jobs is shown in Fig. 1. The choice sets are available upon request. We chose choice sets of size 3 so that respondents could indicate both the best and the worst jobs in each choice set as best–worst judgements are argued to be both easier tasks for respondents and a means of obtaining more information compared to the standard approach that asks for the preferred choice only (Flynn *et al*., 2007; Vermeulen *et al*., 2010).

Table 1. Attributes and levels for the discrete choice experiment and associated model variable names

Glossary definition of attribute	Attribute name	Levels	Variable
The type of hospital where the new graduate programme is located.**	Location	Private hospital; Public hospital	Public hosp
The number of rotations to different clinical areas	Clinical rotations	None; Three	3 rotations
Whether the new graduate programme offers full-time and part-time positions, or full-time only	Work hours	Fulltime only; Part-time or fulltime	Flex hours
The flexibility of the rostering system in accommodating requests	Rostering	Inflexible, does not allow requests; Flexible, usually accommodating requests	Flex rost
The hospital's reputation regarding staffing levels	Staffing levels	Frequently short of staff; Usually well-staffed	Well staff
The hospital's reputation regarding the workplace culture in terms of support from management and staff	Workplace culture	Unsupportive management and staff; Supportive management and staff	Suppmgt
The hospital's reputation regarding the physical work environment in terms of equipment and appearance	Physical environment	Poorly equipped and maintained facility; Well-equipped and maintained facility	Well equip
The hospital's reputation regarding whether nurses are encouraged and supported in professional development and career progression.	Professional development and progression	No encouragement for nurses; Nurses encouraged	Encourage
The parking facilities	Parking	Limited; Abundant and safe	Parking
The hospital's reputation regarding the responsibility given to nurses, relative to their qualifications and experience	Responsibility	Too much responsibility; Appropriate responsibility	App resp
The hospital's reputation regarding the quality of patient care	Quality of care	Poor; Excellent	Excell care
The gross weekly salary	Salary*	$800; $950; $1100; $1250	Salary

Notes: * Modelled as a continuous variable.
** The levels were selected because, in Australia, most clinical nurses work in public or private hospitals (Australian Institute of Health and Welfare, 2012b).
Public hospitals are owned and managed by state governments and provide approximately two-thirds of hospital beds. The remainder are provided by private hospitals which are owned and managed by private for-profit companies or not-for-profit nongovernment organizations (Australian Institute of Health and Welfare, 2012a).

III. Sample Description

To become a registered nurse in Australia, students must complete a three-year, university-based degree. Our sample was recruited from the Bachelor of Nursing (BN) degree student enrolment during the period of 2008–2010 at two Australian universities; one located in a major city, the University of Technology Sydney (UTS), and the other located in a regional centre, the University of New England (UNE). The sample consists of nursing students in each year of the course, and new graduates (within 12 months of completing their university course).[2] At each university, student intake includes school-leavers, mature age entry and other nursing workers seeking to upgrade their qualifications. Therefore, the sample covers a range of age groups, stages of household formation and exposure to nursing work.

Recruitment and data collection procedures are reported in greater detail in (Kenny *et al*., 2012). Recruiting strategies included making presentations at lectures, a recruitment desk outside lecture rooms and at student information events, and other methods of publicizing the study on campus. Willing students were registered by email, post or in person, and once registered, students were emailed invitations to complete the online survey. The 526 survey respondents (100 from UNE and 426 from UTS) represent 18% of the BN enrolment at both universities during the recruitment period (19%UNE, 18%UTS) and 43% of the emailed survey invitations. Comparable cohort studies have reported similar response rates; for example 6–10% of undergraduate nursing students at participating universities were recruited to the Nurses and Midwives e-Cohort (Turner *et al*., 2009) and 41% of young women responded to the invitation to complete the baseline survey for the Australian Longitudinal Study on Women's Health (Brown *et al*., 1999).

Our sample was similar to the BN student enrolment at the time of recruitment in terms of gender; 89% of the

[2] Three respondents were between 12 and 16 months of completing their university degree.

There are jobs available in three programs for new graduates which have the following characteristics:

To review the features of jobs, please click here.

Choice 7 of 8			
Features of Job	**Job A**	**Job B**	**Job C**
Location	Private hospital	Private hospital	Public hospital
Clinical rotations	None	None	Three
Work hours	Fulltime only	Part-time or fulltime	Fulltime only
Rostering	Inflexible, does not allow requests	Flexible, usually accommodating requests	Inflexible, does not allow requests
Staffing levels	Usually well-staffed	Frequently short of staff	Usually well-staffed
Workplace culture	Unsupportive management and staff	Unsupportive management and staff	Supportive management and staff
Physical environment	Poorly equipped and maintained facility	Poorly equipped and maintained facility	Well equipped and maintained facility
Professional development and progression	No encouragement for nurses	Nurses encouraged	No encouragement for nurses
Parking	Abundant and safe	Limited	Abundant and safe
Responsibility	Appropriate responsibility	Appropriate responsibility	Too much responsibility
Quality of care	Poor	Excellent	Excellent
Weekly salary	$1,250	$800	$1,100
Considering these three jobs:			
Which would you MOST like to get?	○ Job A	○ Job B	○ Job C
Which would you LEAST like to get?	○ Job A	○ Job B	○ Job C

>>

Fig. 1. Sample choice set with three hypothetical jobs

sample was female compared to 85% at the two recruitment universities and 87% nationally. There were some differences in the age and language distribution relative to nursing students at the recruitment universities. Specifically, young students entering the BN from secondary education were overrepresented in our sample (51% versus 35%) along with students who speak English at home (83% versus 77%).[3]

This work is part of a broader longitudinal study of nurses' training and job choices. The analysis in this article is based on the first wave of the survey as these are the only data available to date. The data come from an online survey completed between September 2009 and September 2010, and the analysis focuses on job preferences derived from responses to the DCE component of the survey. The research was conducted in accordance with the Australian Government's National Statement on Ethical Conduct of Human Research and was approved by the research ethics committees at both universities.

Nearly 14% of the respondents had graduated at the time of survey completion. The majority of respondents were females, born in Australia, aged less than 25 years and reported their health as 'very good' or 'excellent'. Almost one-third of the sample lived with a spouse or partner and 16% had dependent children; 49% were still living with their parents all or part of the time. While 65% of the sample had paid work, 35% were employed in health care. Of the 72 graduates, 50 (69%) were employed as a nurse, 11 (15%) were employed in another occupation and 11 (15%) were not in the paid workforce. Among the 454 current students, 63% were employed and 30% were employed as an enrolled nurse or assistant in nursing.

IV. Model Specification and Selection

In this section of the article we discuss the various econometric models used to estimate the preference parameters

[3] As described below, we investigate the variation in preference parameters based on observable characteristics. With a few exceptions mentioned in the text, the preference estimates are stable with respect to personal characteristics. Therefore, we find that qualitative and quantitative results are not affected when samples are reweighted to account for these differences in the distributions of personal characteristics. We cannot test for selection bias due to unobservables but the lack of impact from a large set of observables leads us to doubt that this potential source of bias is important in this case.

and their performance given our analysis sample. The results are interpreted and discussed for selected models in the following section of the article.

As a robustness check, we estimate several models. In the following, we describe these models explicitly starting from the more restrictive specifications and moving to more flexible approaches. Although the models are not new to this article, some are not well-known while others have been recently developed. The underlying model is the random utility model (RUM) as developed in Marschak (1960) and McFadden (1981) among others:

$$U_{ij} = x_{ij}'\beta + \varepsilon_{ij}^0 \tag{1}$$

where U_{ij} denotes the utility associated with an alternative or choice j for person i (the dependence on the scenario is suppressed), x is a vector of observable characteristics (including an alternative-specific constant), β is a vector of associated utility weights (we discuss heterogeneous coefficients below) and ε^0 is a component of utility unobserved by the researcher. The variance of ε_{ij}^0, denoted σ^2, is not identified in this model and the estimated parameters β are in fact scaled versions of the true underlying utility weights $\breve{\beta}$: $\beta = \breve{\beta}/\sigma$. This is the well-known scaling problem as discussed by Louviere and co-authors in various works (for example, see Ohler *et al*., 2000).

The most common model of the stochastic process assumes that ε_{ij}^0 is independent across i and j and is distributed according to an extreme type I (or Gumbel) distribution. This leads to the MNL. The left-most column of results in Table 2 presents MNL coefficients for the sample of 12 624 observations involving 526 individuals. Specification tests conducted on the MNL and other models showed that a linear function of salary did not capture preference weights adequately while a concave function could not be rejected in favour of an unrestricted function of the four salary levels. Thus, ln[salary] is used in all specifications presented below to capture this concave relationship.[4] We note that in this context, alternative-specific constants do not have a natural interpretation since A, B and C are merely labels. We have included these constants to allow for the use of heuristics in decision-making. They are significantly different from zero in some but not all specifications and more importantly, overall results are not affected. We discuss this issue further below.

In the models with heterogeneous utility parameters, the utility function becomes

$$U_{ij} = x_{ij}'\beta_i + \varepsilon_{ij}^1 = x_{ij}'(\tilde{\beta} + \eta_i) + \varepsilon_{ij}^1 \tag{2}$$

Table 2. Multinomial logit (MNL) and mixed logit (MXL) models. SEs in parentheses

	Models		
		MXL	
	MNL	Mean	SD
Salary	1.550***	2.883***	2.828***
	(0.095)	(0.241)	(0.287)
Suppmgt	1.044***	1.946***	1.381***
	(0.049)	(0.151)	(0.145)
Excell care	0.832***	1.438***	1.321***
	(0.050]	(0.120)	(0.119)
App resp	0.475***	0.961***	1.024***
	(0.048)	(0.105)	(0.137)
Flex rost	0.542***	0.912***	0.851***
	(0.042)	(0.090)	(0.125)
Encourage	0.519***	0.822***	0.611***
	(0.045)	(0.083)	(0.146)
Well equip	0.374***	0.713***	0.622***
	(0.039)	(0.084)	(0.157)
Well staff	0.400***	0.683***	0.549***
	(0.037)	(0.075)	(0.127)
Public hosp	0.241***	0.441***	0.748***
	(0.040)	(0.076)	(0.147)
3 rotations	0.205***	0.375***	0.795***
	(0.040)	(0.077)	(0.132)
Flex hours	0.128***	0.210***	0.578***
	(0.035)	(0.062)	(0.141)
Parking	0.064*	0.101	0.421**
	(0.038)	(0.061)	(0.179)
Job B Cst	0.131***	0.369***	0.117
	(0.044)	(0.095)	(0.467)
Job A Cst	0.008	0.244**	0.300
	(0.046)	(0.096)	(0.205)
Sample Size	12 624	12 624	
Likelihood	−3492.546	−3287.217	
AIC	7013.091	6630.433	
BIC	7117.298	6838.847	

Notes: MNL refers to a multinomial logit and MXL to a mixed logitmodel. For the simulations, 10 000 Halton draws are made after burning the initial 43 draws. The coefficient on salary measures, the change in utility caused by moving from a job with a weekly salary of 800 to a job with a weekly salary of 1250. The SEs are robust to arbitrary heteroscedasticity and to correlations across observations from the same individuals. LLikelihood indicates a pseudo log likelihood for the MNL and simulated log likelihoods for the mixed logit, AIC refers to the Akaike information criterion and BIC to the Bayesian information criterion. *** indicates that the parameter is significantly different from zero at a 1% level of confidence, ** at 5% and * at 10%.

where $\tilde{\beta}$ denotes the population mean of β and η_i is the variation from the mean for person i. The MXL is derived from this model under the assumption that the ε_{ij}^1's are independently drawn from the Gumbel distribution.

[4] A quadratic function performed slightly better than the log transformation but the differences were quantitatively unimportant and other coefficients were not affected. We chose the log function due to the simplification it affords when manipulating and interpreting results; specifically, the willingness-to-pay measure (Equation 6) has a unique value given a base salary and a coefficient vector.

Following most of the literature we assume that the mixing distribution is normal: $\beta_i \sim MVN(\tilde{\beta}, \Sigma)$.[5] The method of maximum simulated likelihood is used.[6] We present estimates where correlations in utility weights across attributes are set at zero but all elements of β_i are random.[7]

The right-most columns of Table 2 present estimates for the means and the SDs of the vector β_i based on 10000 replications.[8] The means of the distribution of attribute weights are all significantly different from zero at a 1% level of significance except for 'abundant parking'. This follows patterns in the MNL with fixed utility weights. All SDs for the attribute weights are significantly different from zero, an indication of heterogeneity in the utility weights across individuals. The higher estimated mean coefficients in the MXL are to be expected as the unexplained component is likely to have a smaller variance. In terms of relative importance, however, there is not much difference in the utility weights. Finally, a comparison of the AIC and Bayesian information criterion (BIC) measures also supports the use of the mixed model over the MNL.

The left-most columns of Table 3 provide results for the GMNL model. In the GMNL, the distributional assumptions along with the panel dimension of the data are used to identify parameters of the distribution of the scaling factor as well as how it interacts with the utility weights. This model can be seen as a generalization of the MXL in which the variance of the error term is heterogeneous across individuals. Specifically, the utility function in the GMNL is written as follows:

$$\begin{aligned} U_{ij} &= x'_{ij}(\varsigma_i\tilde{\beta} + \eta_i^*) + \varepsilon_{ij}^2 \\ &= x'_{ij}(\varsigma_i\tilde{\beta} + \gamma\eta_i + (1-\gamma)\varsigma_i\eta_i) + \varepsilon_{ij}^2 \end{aligned} \tag{3}$$

where ς_i is an individual-specific scalar, scaling the $\tilde{\beta}$ vector up or down and γ is a parameter that allows η_i to be scaled up by ς_i (when $\gamma = 0$) or to vary independently (when $\gamma = 1$). In practice, ς_i is assumed to follow the lognormal distribution, $\ln(\varsigma_i) \sim N(\bar{\varsigma}, \tau^2)$ with $\bar{\varsigma}$ normalized to $-\tau^2/2$. The lognormal distribution is chosen because it has positive support (Keane and Wasi, 2013). Also following Keane and Wasi (2013), γ is an unrestricted parameter. Thus effectively the GMNL model includes two additional parameters to those in the MXL model.

From a technical perspective, the GMNL model can be seen as allowing for correlation across the random utility weights (Hess and Rose, 2012), but the correlation is restricted. For example, a general correlation structure for the multivariate normal mixing distribution in our case would involve 66 additional parameters as compared to the two coefficients added by the GMNL. Not only does the GMNL model allow for correlations in a tractable way, its specification has a natural interpretation in terms of scale heterogeneity. ς_i can be interpreted as an individual-specific scale parameter (or the inverse of an individual specific error variance) which is drawn from a common distribution parametrized by τ. The interpretation of γ is perhaps less obvious; it determines the extent to which the SDs of the distributions of the utility weights are scaled relative to the means. If $\gamma <(>) 0$, as the means are scaled up by ς_i, the SDs are scaled by a larger (smaller) amount, and if $\gamma > 1$, the SDs are scaled down.[9]

Initial estimates of the GMNL model[10] with unrestricted γ yielded an estimate of γ equal to 0.099 with a SE of 0.175. Hence, there is no support for differential scaling of the mean and the heterogeneous component of the attribute weight and an interpretation of scale heterogeneity is in some sense natural. Detailed results from the model with an unrestricted γ are available from the authors. The left-most columns of Table 3 present estimates for the GMNL model with γ fixed at 0. The estimated SD of the log of the scaling factor, τ, is highly significant. From Equation 3, we see that this means that ς_i is drawn from a distribution with a nonzero variance; this can be interpreted as evidence of heterogeneity in the scale factor. The simulated likelihood is improved in GMNL relative to the MXL as are both AIC and BIC statistics. In terms of the qualitative results, the GMNL yields means and SDs that are higher than their MXL counterparts but the ranking across attributes is unchanged.

The models presented so far are based on the 'best' choice among the three alternative jobs. As described above, respondents were asked to choose the best and worst jobs and hence provide a ranking of the set of three alternatives. The most common model used to analyse this type of data is the rank-ordered logit (ROL) (Beggs *et al.*,

[5] There is some debate over the form of the mixing distribution. The work in Train (2008) shows that the independent normal mixture performs well compared to various nonparametric mixing distributions while Greene and Hensher (2003) suggest that estimates may be quite sensitive to specific assumptions in alternative specifications. For these reasons, we use the normal for all parameters.
[6] Stata version 11 is used; also, Halton draws are taken and 43 initial draws are burned (Train, 2009).
[7] See Train (2009), pp. 140–1 for a discussion of the difficulty in identifying correlations in models with this many attributes.
[8] Parameters stabilized after 10000 replications in the sense that no estimate varied by more than 10% and/or one SD.
[9] We should add that some researchers argue that emphasis on scale heterogeneity over preference heterogeneity may be misguided (Greene and Hensher, 2010). We do not take a stand on this but use the more general GMNL model as a robustness check against the more standard MXL.
[10] The GMNL model is estimated using Stata and the code written by Gu *et al.* (2013).

Table 3. Generalized multinomial logit and heteroscedastic rank-ordered GMNL. SEs in parentheses

	Models			
	GMNL		HROGMNL	
	Mean	SD	Mean	SD
Salary	4.281***	4.073***	2.999***	3.360***
	(0.819)	(0.783)	(0.321)	(0.354)
Suppmgt	2.869***	1.808***	1.993***	1.482***
	(0.528)	(0.377)	(0.198)	(0.159)
Excell care	2.100***	1.741***	1.621***	1.400***
	(0.403)	(0.366)	(0.162)	(0.145)
App resp	1.363***	1.242***	0.972***	0.845***
	(0.265)	(0.317)	(0.111)	(0.128)
Flex rost	1.359***	1.140***	0.941***	1.038***
	(0.274)	(0.280)	(0.112)	(0.125)
Encourage	1.255***	0.846***	0.935***	0.788***
	(0.253)	(0.262)	(0.103)	(0.115)
Well equip	1.055***	0.814***	0.785***	0.506***
	(0.215)	(0.251)	(0.093)	(0.133)
Well staff	1.052***	0.745***	0.819***	0.722***
	(0.215)	(0.263)	(0.094)	(0.107)
Public hosp	0.618***	0.925***	0.307***	0.450***
	(0.146)	(0.234)	(0.064)	(0.154)
3 rotations	0.544***	1.071***	0.371***	0.511***
	(0.144)	(0.247)	(0.069)	(0.115)
Flex hours	0.322***	0.767***	0.207***	0.672***
	(0.107)	(0.249)	(0.063)	(0.107)
Parking	0.159	0.652***	0.159***	0.476***
	(0.098)	(0.251)	(0.061)	(0.114)
Job B Cst	0.364***	0.216	0.118**	0.406***
	(0.099)	(0.183)	(0.052)	(0.075)
Job A Cst	0.242**	0.352**	0.012	0.227**
	(0.100)	(0.170)	(0.048)	(0.112)
δ			−0.021	
			(0.100)	
τ	0.690***		0.712***	
	(0.151)		(0.084)	
Sample Size	12 624		21 040	
SLLikelihood	−3278.283		−5751.731	
AIC	6614.566		11 563.462	
BIC	6830.423		11 802.087	

Notes: GMNL refers to a generalized mixed logit model and HROGMNL refers to a heteroscedastic rank-ordered generalized mixed logit model. For the simulations, 10 000 Halton draws are made after burning the initial 43 draws. The coefficient on salary measures the change in utility caused by moving from a job with a weekly salary of 800 to a job with a weekly salary of 1250. SLLikelihood indicates a simulated log likelihood, AIC refers to the Akaike information criterion and BIC to the Bayesian information criterion. For the HROGMNL, the BIC is calculated using a ranking as an observation. *** indicates that the parameter is significantly different from zero at a 1% level of confidence, and ** at 5%. Both models have $\gamma = 0$.

1981). It can be derived by assuming that the individual chooses best and worst jobs based on a ranking of utility levels such as those presented in Equations 1–3.[11] The efficiency gained with rank-ordered data depends on the assumption of constant preference parameters over the ranking of alternatives. Some have argued that while the utility weights may remain constant over choices in a single ranking, the variance of the error is likely to increase as one is asked to rank less preferred alternatives. One can relax the assumption of constant variance or scale across the ranking by modelling shifts in the coefficients across the decision nodes. This leads to the heteroscedastic version of the rank-ordered model developed in Hausman and Ruud (1987). We estimated rank-ordered and heteroscedastic rank-ordered versions of the MNL, MXL and GMNL models. Only the latter set of

[11] A different model of best–worst choices is the maxdiff as developed by Jordan Louviere; for example see Marley *et al*. (2008)

estimates are presented. Other results are available upon request.

In the rank-ordered heteroscedastic version of the GMNL (HROGMNL) model, ς_i is assumed to follow the lognormal distribution, $\ln(\varsigma_i) \sim N(\bar{\varsigma}, \tau^2)$, where $\bar{\varsigma} = -\tau^2/2 + \delta^* S$, S equals to 0 for the first decision node and 1 for the second choice in the ranking. In other words, δ measures the shift in the mean of $\ln(\varsigma)$ as respondents move from their best to their second best choice, a shift which is assumed to be common to all individuals. This is the first estimation of such models that we are aware of.

The estimates for the HROGMNL model are presented in the right-most columns of Table 3. Based on previous results, γ is fixed at zero. It is interesting that after allowing for individual heterogeneity in means and scaling, there is no evidence of a shift in the scaling factor across choice nodes; i.e., δ is small and insignificant. This was not the case for the heteroscedastic MNL where the shift in the scale parameter was significant statistically and quantitatively. We also note that although there is evidence of heterogeneity in the job-specific constants (the SDs are significantly different from zero at a 1% level of significance) their means are small and insignificant at 1% in this model. In other words, with enough flexibility in the modelling of preference heterogeneity, the ASCs have become negligible on average. Most importantly, the means and SDs of the attribute weights are very similar in this model compared to the previous estimations.

V. Interpretation of Estimation Results

Two sets of figures are computed from the estimation results to make the preference parameters easier to interpret: predicted probabilities of job choice and WTP measures. The two measures are each based on the full vector of parameter estimates but as detailed below, they differ in the specific questions they answer. WTP is the most popular method used to present results from models such as these. However, some researchers argue against this measure because of the high sensitivity to the marginal utility of money. Also, the WTP measure does not allow a comparison of salary and nonsalary attributes. For these reasons, we present both measures.

The predicted probabilities answer the following question: 'What is the change in the predicted probability of choosing a job Z instead of another job Y if the only difference in the two jobs lies in the level of attribute k?' For the MNL this can be written as

$$Prob\{U_Z > U_Y\} = Prob\{x'_Z\beta - x'_Y\beta > \varepsilon^0_Y - \varepsilon^0_Z\} = \left(\frac{e^{\beta_k}}{1+e^{\beta_k}}\right) \quad (4)$$

where it is assumed that the jobs differ only in the attribute k and that this attribute shifts by one unit (we discuss the shift in salary below). The base job Y is defined as the worst possibility in the sense that all attributes are set at their least preferred levels as defined by the mean attribute weight. The resulting predicted probabilities will be greater than 0.5 since the βs are greater than 0. (The predicted probability will be equal to 0.5 if the attribute is unimportant ($\beta_k = 0$) and hence the choice is completely random.)[12]

Table 4 presents predicted probabilities for the main models in our analysis. The figures in the table measure the predicted probability of accepting a job in which the corresponding attribute has shifted to its preferred level (based on mean attribute weights) and all other job attributes are held fixed at their base level. For the salary, the shift is from 800 to 1250 dollars per week; all other attributes are binary and the shift is from zero to one. All

Table 4. Predicted probabilities of job choice by attribute, various models

	MNL	MXL	GMNL	HROGMNL
Salary	0.774	0.908	0.968	0.915
Suppmgt	0.740	0.875	0.946	0.880
Excell care	0.697	0.808	0.891	0.835
App resp	0.617	0.723	0.796	0.726
Flex rost	0.632	0.713	0.796	0.719
Encourage	0.627	0.695	0.778	0.718
Well equip	0.592	0.671	0.742	0.687
Well staff	0.599	0.664	0.741	0.694
Public hosp	0.560	0.608	0.650	0.576
3 rotations	0.551	0.593	0.633	0.592
Flex hours	0.532	0.552	0.580	0.551
Parking	0.516	0.525	0.540	0.540

Notes: Figures measure predicted probabilities of job choice (relative to the base job) when the attribute is set to its preferred level, all other attributes remaining at their base level. The base job is one with all attributes set at their least preferred levels. For the salary the shift is from 800 to 1250 dollars per week, for all other attributes the shift is from zero to one. MNL refers to a multinomial logit, MXL to a mixed logit, GMNL to a generalized multinomial logit and HROGMNL to a heteroscedastic rank-ordered generalized multinomial logit. All predicted probabilities are significantly different from 0.5 at the 1% level except for those corresponding to 'Parking'; for the latter only the probability in the HROGMNL is significantly different to 0.5 at 1% level.

[12] When the coefficients are random and normally distributed, the predicted probability has a logit-normal distribution. The mean of this distribution has no analytical solution in general but the median is well-defined and equal to the logistic function evaluated at the mean $\tilde{\beta}_k$. Hence, although technically the statistic is different, qualitative interpretations are similar.

predicted probabilities are significantly different from 0.5 at the 1% level except for those corresponding to ‘Parking’.

Since we are using log salary as our attribute, we can recover the utility gain for a 1% increase in salary by dividing the numbers reported in the tables for salary by 56.25. (The increase from 800 to 1250 corresponds to a 56.25% increase in salary.) For example, for the MXL and GMNL models, the increase in utility for a 10% increase in salary is 0.5125 and 0.761, respectively. Salary has the highest effect on the predicted probability; when salary shifts from 800 to 1250, an individual is almost sure to choose the new job over the old one (the probability is over 90% in all models except for the MNL model where the probability is 77%). Only an extreme value for the unobserved components of utility would lead to a preference for the original job. We can form roughly four groups of attributes based on their importance: salary, supportive management/staff and quality of care; appropriate responsibility, flexible rostering, professional development and progression; well equipped and well-staffed premises; and public hospital, 3 rotations, flexible hours and abundant parking. The ranking across these groups is robust across all models; indeed, the ranking within the groups is also the same across models with only a few exceptions.

An alternative approach transforms utility weights into dollar values; specifically, WTP measures are constructed as marginal rates of substitution (MRS) between an attribute and a monetary attribute, in our case salary. This statistic answers the following question: ‘What is the loss in salary that would keep utility constant when one attribute, say k, is shifted to its preferred level, all other attributes remaining unchanged?’[13] The preferred level is based on the mean attribute weight. Denote the coefficient on ln[salary] as β_s and change attribute k from 0 to 1:

$$\Delta U = 0 \Rightarrow \beta_k + \beta_s \ \ln(m \times salary) = \beta_s \ \ln(salary) \tag{5}$$

where m is the proportion of the salary which is retained and which guarantees constant utility. Measuring the loss in salary in dollars from the base of 800 yields

$$WTP = 800 \times (1 - m) = 800 \times (1 - e^{-\beta_k/\beta_s}) \tag{6}$$

When coefficients are fixed, it is straightforward to derive estimates for WTP by using point estimates for β. In the MXL and its extensions, the attribute weights are normally distributed variables and their ratio will have a Gaussian ratio distribution. For general parameter values, the mean of this ratio is not well-defined. The median exists for all values of the parameters of the distributions and it has a well-defined probability density function (pdf); however, in general the pdf does not have a closed form representation and must be simulated.

Figure 2 presents the 25th percentile, the median and the 75th percentile for the simulated WTP distributions based on the GMNL model.[14] The ranking of attributes is the same as that discussed above when presenting predicted probabilities. It is also very similar across specifications. This is not surprising; since the parameter estimates are very similar across approaches, so are the WTP figures. At the median values, greater importance is given to supportive management/staff and quality of patient care relative to other characteristics such as flexible hours and the opportunity for several clinical rotations. These results are consistent qualitatively with the findings of research into nurse retention and job satisfaction in Europe and North America. Managerial or peer support have been associated with intentions to leave the profession among nurses in Canada (Zeytinoglu *et al*., 2011) and Hungary (Ujvarine *et al*., 2011); poor staffing and work environment (including lack of managerial support and promotion of care quality) were associated with job dissatisfaction and intentions to leave among nurses from 12 European countries and the United States (Aiken *et al*., 2012); and opportunities for career progression have been found to be important among British nurses (Shields and Ward, 2001) and Hungarian nurses (Ujvarine *et al*., 2011).

Not only do our results support the findings of these previous studies in identifying important characteristics of nursing jobs but the DCE provides us with quantitative estimates of the strength of preferences in regards to these characteristics. At the median, supportive management/staff is valued roughly twice as much as appropriate responsibility, flexible rostering and encouragement for professional development. In comparison, excellent patient care is valued roughly 50% more than these attributes. Working in a well-equipped hospital is valued somewhat less at around 40% of the value placed on supportive management/staff, similar to what is found for working in a well-staffed hospital. What this means is that policies to support recruitment and retention of new graduate nurses should focus on building a culture of support in the workplace, particularly for inexperienced, new entrants. It should also focus on providing the resources and processes to ensure a high quality of patient care as well as improved salary levels, which should be prioritized over less-valued attributes such as the physical

[13] We are using willingness-to-pay in a restricted sense; this experiment does not yield welfare measures that can be applied in arbitrary situations since they do not allow for a nurse's choice to move out of nursing jobs altogether (see Lancsar and Savage (2004) for more details).

[14] The distributions are simulated with 100 000 replications.

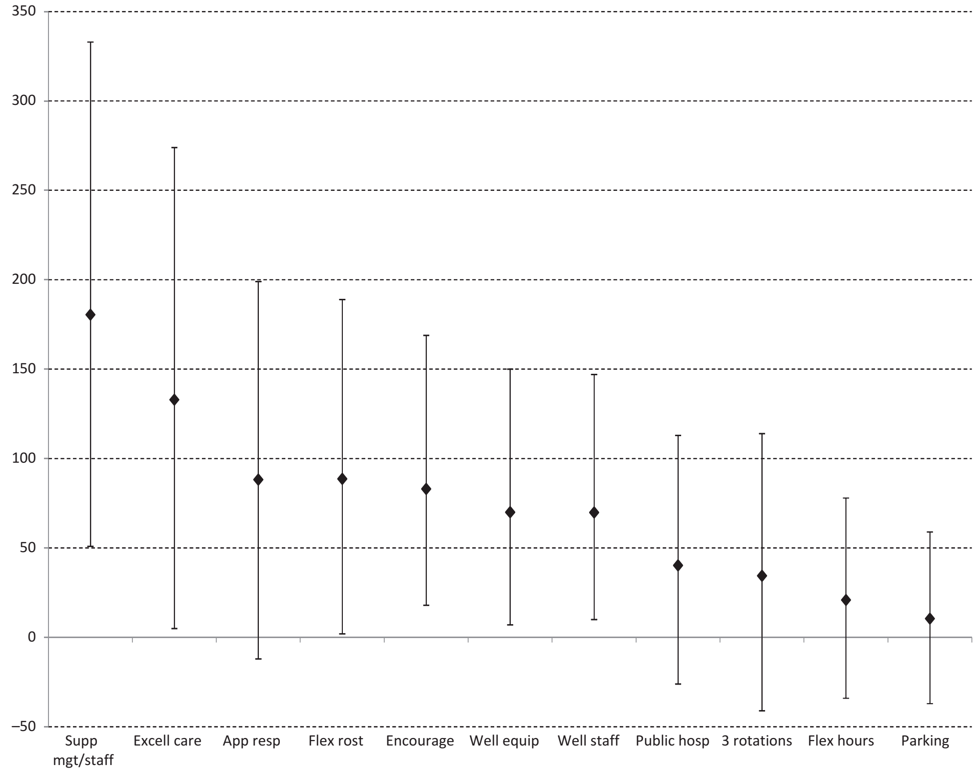

Fig. 2. Quantiles of the willingness-to-pay (WTP) distribution
Notes: For each attribute, the median and the interquartile range of the simulated WTP distributions are shown. WTP figures represent marginal rates of substitution (in absolute value) between the attributes and salary and should be compared to a base salary of $800 per week. The distribution of the WTP measure is simulated with 100 000 replications.

work environment or parking facilities. While the WTP estimates provide an alternative means of comparing the strength of preferences for job attributes, these estimates should not be interpreted literally as the amount that nurses would forgo in earnings for an improvement in job conditions. Rather, they provide a means of comparing the cost of improvements with their relative value to employees, and hence in determining the most efficient retention strategies.

Another aspect of the findings from the DCE deals with results on the heterogeneity in preferences. The simulated WTP distributions show a large amount of dispersion in the weights placed on job characteristics. This reflects the estimated SDs around mean attribute weights. For the first seven attributes (salary to well staffed), the ratio of the mean to the interquartile range is generally ≥ 0.5 while the figure for the remaining attributes is normally ≤ 0.25. Interestingly, the first group of attributes have clear better and worse levels; for example, a higher salary is always better, excellent care is better than low quality of care, and so on. Our respondents may have different strengths of preferences, but a well-equipped hospital is generally preferred to a poorly equipped one. In contrast, the characteristics in the second group do not have clear better or worse levels. With these attributes, individuals have quite divergent preferences, with some seeing them as a positive contribution to utility while others consider the same attribute as having a negative impact. For example, 3 rotations will be positive for those nurses who wish to experience a variety of clinical areas; but equally it will have a negative impact for those nurses who are already certain they want to work in one field of nursing. This preference diversity, as opposed to strength, is an important issue to be considered in designing policies to improve retention.

VI. Job Preferences and Time in the Programme

In this section of the article, we investigate if relative weights placed on job attributes differ with the progression through the programme of study and the initial post-graduation experience with the workplace. Without panel data we cannot control unobserved individual characteristics that may differ across the subsamples by year of programme (including those due to student attrition from the education programme); nevertheless, since our cross-section data spans the whole length of the programme of study, we can investigate the possibility of systematic differences in attribute weights for individuals at different levels in the programme.[15]

We construct dummy variables to represent the respondent's year in the programme. In total, there are four groups comprising as first year, second year, third year (including any fourth year) and graduates. The distribution of the 526 individuals is as follows: 183 (35%) are in the first year, 137 (26%) are in the second year, 134 (25%) are in the third year and 72 (14%) are new graduates. We estimate a MXL, where all attribute weights are heterogeneous across agents and where the means of the distributions shift across years in the programme. Detailed estimates are not shown to save space, but tests show that mean attribute weights are jointly significantly different across years. Tests on individual attributes show that equality of mean attribute weights across the four groups of respondents is rejected for three attributes (based on a 5% level of significance): three rotations (p-value of 0.010), flexible rostering (p-value of 0.008) and quality of care (p-value of 0.0327). In addition, several shifts in mean attributes are individually significantly different from zero.

Predicted probabilities of job choice and WTP measures are provided in Table 5.[16] We present figures for year 1 and shifts in the figures for subsequent years. We also present the ranking of the mean attribute weights for year 1 students and graduates to show that shifts occur in the relative ranking of the attributes as well as in the magnitude of the attribute weights.

Briefly, graduates place more weight on three rotations and flexible hours and less weight on quality of

Table 5. Predicted probabilities of job choice and willingness-to-pay for job attributes, variation by year in programme

	Rank	Value	Differences from year 1			Rank
	Year 1	Year 1	Year 2	Year 3	Graduate	Graduate
a) Predicted probabilities:						
Salary	1	0.892†††	0.055	0.037	0.002	2
Suppmgt	2	0.871†††	0.04	–0.006	0.031	1
Excell care	3	0.857†††	–0.043	–0.075**	–0.114**	4
App resp	4	0.750†††	–0.018	–0.074*	0.018	3
Flex rost	7	0.657†††	0.109***	0.116***	0.029	7
Encourage	5	0.700†††	0.011	0.019	–0.016	8
Well equip	6	0.662†††	0.078*	–0.004	–0.041	11
Well staff	8	0.656†††	0.008	0.027	0.061	5
Public hosp	9	0.599†††	0.003	0.025	0.039	9
3 rotations	12	0.519††	0.080*	0.117**	0.176***	6
Flex hours	11	0.524	0.020	0.052	0.105**	10
Parking	10	0.551††	–0.058	–0.019	–0.030	12
b) Willingness-to-pay:						
Suppmgt	1	265.635***	–24.321	–46.064	31.120	1
Excell care	2	251.806***	–88.836**	–93.215**	–93.016*	3
App resp	3	165.962***	–50.986	–70.252**	11.435	2
Flex rost	6	102.967***	31.079	50.321*	17.725	6
Encourage	4	131.513***	–27.225	–11.230	–11.712	7
Well equip	5	105.792***	13.653	–20.045	–27.493	10
Well staff	7	102.196***	–21.904	–2.285	39.073	4
Public hosp	8	65.037***	–15.579	2.015	24.330	8
3 rotations	11	12.659	35.518	61.096**	113.731***	5
Flex hours	10	16.244	5.395	25.318	67.830**	9
Parking	9	33.812*	–37.436	–16.373	–19.844	11

Notes: The value year 1 column shows the predicted probabilities and WTP figures for year 1 nursing students. Years 2 and 3 and graduate show shifts in year 1 mean attributes. Rank year 1 and rank graduate show the rankings of the probabilities and WTP figures for year 1 students and graduates, respectively. WTP measures are evaluated at the mean attribute levels. *** indicates that the parameter is significantly different from zero at a 1% level of confidence, ** at 5% and * at 10%. Similarly, ††† indicates that the parameter is significantly different from 0.5 at a 1% level of confidence, †† at 5% and ² at † 10%. Underlying SEs are computed using the delta method.

[15] Preliminary results using the first two waves of the survey that are now available suggest that the results described below are not due to nonrandom attrition. Specifically, estimates based on the balanced panel indicate that preferences change as the individuals move through stages in the programme.

[16] To simplify, WTP measures are computed at the mean of the attribute weight.

care relative to first years. Although the four highest ranked attributes do not differ between graduates and first years, the ordering within the top four does. Among graduates, supportive management/staff is ranked highest over salary and appropriate responsibility, while salary is ranked highest over supportive management/staff and quality of care among first years. The third-year group places less weight on appropriate responsibility and more weight on flexible rostering relative to first-year students. What differs as nurses move through their education and into the nursing workforce? Trainees in their later years and then graduate nurses have gained more clinical experience and insights, and are older than their first-year counterparts. Our findings suggest that this greater clinical understanding results in greater weight placed on a supportive workplace and appropriate responsibility, and that the realities of working shift work and/or changing family situations explain the stronger preference for flexible hours. While these differences do not have direct relevance for the design of job packages for registered nurses, it is important to identify the extent to which our results are influenced by year in the programme. They also point to the importance of a supportive workplace during the transition from student to registered nurse, identified by others (Rush *et al.*, 2013).

We conclude this section with a brief discussion of the results of estimations where preference parameters are allowed to differ based on other observable characteristics. Specifically, we estimate MNL allowing for variations in coefficients on the job attributes based on the university (UTS, UNE), sex, employment status (working as a nurse, working elsewhere, not working), qualifications (secondary school, nursing qualifications, higher degree, other qualifications), age (aged 21 or younger, older than 21 years), household income (under \$50 000, \$50 000 or more), presence of children under 16 years of age in the household, presence of children under 6 years of age in the household, recently arrived in Australia (arrived within 3 years of starting the BN) and English spoken at home.[17] The only jointly significant interaction effects (using a 1% significance level) are found for age and the presence of children under 6 years of age. In our sample, 8% report having children under 6 years of age in the household. On average, these individuals care more about flexible hours. Deviations from the average weights for all other attributes are individually insignificant. Forty-two per cent of our respondents are aged less than 22 years. For these, the preference weight is smaller for flexible hours and larger for abundant parking. All other shifts in attribute coefficients are individually insignificant. Detailed results are available on request.

VII. Conclusions

This article is the first study of nurses' job preferences that applies DCE method to a developed country workforce. It adds to the previous literature on stated intentions to quit, as those studies are limited to comparing the job characteristics of actual jobs with unknown alternatives. In contrast, DCEs allow the construction of a much wider range of hypothetical alternatives with defined attributes, and thus let us explore more fully how different policy options would impact attrition and retention. Our DCEs also use a greater number of job attributes than previous studies, thus increasing the realism of the choice scenarios. The choice of attributes reflects factors that have been shown to be important for nurses in various literatures and the levels of the attributes have been chosen to make the jobs realistic in the context of our sample. This article is also the first to focus on the transition through university training and into the labour force. Our sample consists of students at different stages of training and new graduate nurses. This is a particularly interesting group since junior nurses on average have the lowest retention levels in the profession. Finally, the article makes a methodological contribution in that we adapt state-of-the-art models of heterogeneity (MXL and GMNL) to best–worst information and allow for heteroscedasticity across choice nodes. The extension of the GMNL model is new to the literature. Our results remain remarkably robust across models and suggest that although there is a significant scale heterogeneity, there is no evidence of systematic shifts in scale across best–worst choices. Further, the results showing the importance of the quality of patient care and the support received from management and colleagues are consistent with international evidence on nurse retention (as discussed in Section V) supporting their relevance for health-care systems beyond Australia, especially those where registered nurses are educated in universities.

We can form roughly four groups of attributes based on their importance: salary, supportive management/staff and quality of care; appropriate responsibility, flexible rostering, professional development and progression; well-equipped and well-staffed premises; public (as opposed to private) hospital, multiple rotations, flexible hours and abundant parking. The ranking across these groups is robust across all models; indeed, the ranking within the groups is also the same across models with only a few exceptions. Based on WTP measures, we find that

[17] We look at heterogeneity based on models without unobserved heterogeneity and include the interactions for specific variables one at a time (not jointly). In this sense, we are giving the best chance to the demographics of capturing heterogeneity.

supportive management/staff is valued roughly twice as much as appropriate responsibility, flexible rostering and professional development and progression. In comparison, excellent patient care is valued roughly 50% more than these attributes. In some estimations, mean preference weights are allowed to shift across students based on the year in the programme and new graduates. We find that while preferences are similar over the transition, for nurses in their first job, supportive management/staff is valued significantly more than for student nurses. Indeed, in terms of ranked order, it is more important than salary (at normal levels). Having appropriate levels of responsibility and a greater range of training (number of rotations) are also ranked more highly by new nurses than students.

There are substantial policy implications of our work, particularly in understanding the heterogeneity of preferences for different job attributes. It is important to distinguish where nurses differ in their strength of preferences against opposing likes and dislikes. For example, everyone prefers a higher salary over a lower one but for some nurses, salary is more important than for other nurses. In contrast, nurses will disagree on whether certain characteristics are utility enhancing or reducing. For example, nursing in intensive care is attractive for some nurses but actively disliked by others. For heterogeneity in strength of preference, the relevance for policy is in identifying those job attributes which will do most to attract and retain nurses. Our findings show that salary remains an important factor in making nursing jobs attractive. Although nonpecuniary benefits are also important, policy should not ignore pay levels for nurses. Along with salaries, policies which promote a supportive workplace culture and high quality of care will also be effective in making nursing jobs more attractive. For those attributes where there are radically opposing preferences, nursing retention could be improved by designing quite different employment packages to appeal to these different tastes. Our nurses show one group who are attracted by clinical rotations, while others prefer to stay within one area. This would suggest that new graduate programmes should provide this but not require it, as some graduates will already know the specialty area in which they intend to practice or may find the transition from student easier without the need to adapt to multiple environments. Further, we see that the transition from university student to new graduate nurse is apparently a time when a supportive workplace culture and the level of responsibility make a difference, so that policies which lessen the stress and possible feelings of isolation may also be important in retaining the vulnerable group of new graduates.

It would be premature to extend our findings into firm policy and management recommendations. Our sample is predominantly nursing students, with only 14% having graduated when they completed the survey. Our results do show that job preferences change with experience; supportive management and staff culture become more important over time. We also note interesting heterogeneity in preferences which could be important in designing alternative employment packages rather than relying on uniform conditions. It will be important to see how these results change as more students enter the workforce, and junior nurses advance in years. Our study is designed as a panel and future work will report on how different nursing experiences affect preferences and retention.

Acknowledgements

We thank participants at the AHES conference 2010, workshops at UNSW and CHERE and two anonymous referees for constructive comments and suggestions. We are especially thankful to Agne Suziedelyte and Hong Il Yoo for outstanding research assistance.

Funding

This work was partially supported by the Australian Research Council [grant number DP0881205].

References

Aiken, L. H., Clarke, S. P., Sloane, D. M. *et al*. (2002) Hospital nurse staffing and patient mortality, nurse burnout, and job dissatisfaction, *Journal of American Medical Association*, **288**, 1987–93.

Aiken, L. H., Sermeus, W., Heede, K. V. *et al*. (2012) Patient safety, satisfaction, and quality of hospital care: cross sectional surveys of nurses and patients in 12 countries in Europe and the United States, *British Medical Journal*, **344**, e1717.

Australian Institute of Health and Welfare. (2012a) *Australia's Health 2012*. Australia's health series no. 13. AIHW. Available at http://www.aihw.gov.au (accessed 6 November 2013).

Australian Institute of Health and Welfare. (2012b) *Nursing and Midwifery Workforce 2011*. National health workforce series no. 2. AIHW. Available at http://www.aihw.gov.au (accessed 6 November 2013).

Barron, D. and West, E. (2005) Leaving nursing: an event-history analysis of nurses' careers, *Journal of Health Service and Research Policy*, **10**, 150–7.

Beggs, S., Cardell, S. and Hausman, J. (1981) Assessing the potential demand for electric cars, *Journal of Econometrics*, **16**, 1–19.

Blaauw, D., Erasmus, E., Pagaiya, N. *et al*. (2010) Policy interventions that attract nurses to rural areas: a multicountry discrete choice experiment, *Bulletin of the World Health Organization*, **880**, 350–6.

Brown, W. J., Bryson, L., Byles, J. E. *et al*. (1999) Women's health Australia: recruitment for a national longitudinal cohort study, *Women and Health*, **28**, 23–40.

Casey, K., Fink, R., Krugman, M. *et al*. (2004) The graduate nurse experience, *Journal of Nursing Administration*, **340**, 303–11.

Coomber, B. and Barriball, K. (2007) Impact of job satisfaction components on intent to leave and turnover for hospital-based nurses: a review of the research literature, *International Journal of Nursing Studies*, **44**, 297–314.

Cunich, M. and Whelan, S. (2010) Nurse education and the retention of registered nurses in New South Wales, *Economic Record*, **850**, 396–413.

Di Tommaso, M. L., Strøm, S. and Saether, E. M. (2009) Nurses wanted: is the job too harsh or is the wage too low?, *Journal of Health Economics*, **280**, 748–57.

Doiron, D., Hall, J. and Jones, G. (2008) Is there a crisis in nursing retention in New South Wales?, *Australia and New Zealand Health Policy*, **5**, 19.

Estryn-Behar, M., Van der Heijden, B., Fry, C. *et al*. (2010) Longitudinal analysis of personal and work-related factors associated with turnover among nurses, *Nursing Research*, **59**, 166–77.

Fiebig, D. G., Keane, M. P., Louviere, J. *et al*. (2010) The generalized multinomial logit model: accounting for scale and coefficient heterogeneity, *Marketing Science*, **29**, 393–421.

Flynn, T. N., Louviere, J. J., Peters, T. J. *et al*. (2007) Best-worst scaling: what it can do for health care research and how to do it, *Journal of Health Economics*, **260**, 171–89.

Fochsen, G., Josephson, M., Hagberg, M. *et al*. (2006) Predictors of leaving nursing care: a longitudinal study among Swedish nursing personnel, *Occupational and Environmental Medicine*, **63**, 198–201.

Frijters, P., Shields, M. A. and Price, S. W. (2007) Investigating the quitting decision of nurses: panel data evidence from the British National Health Service, *Health Economics*, **160**, 57–73.

Greene, W. and Hensher, D. (2010) Does scale heterogeneity across individuals matter? An empirical assessment of alternative logit models, *Transportation*, **370**, 413–28.

Greene, W. H. and Hensher, D. A. (2003) The mixed logit model: the state of practice, *Transportation*, **30**, 133–76.

Gu, Y., Knox, S. and Hole, A. R. (2013) Fitting the generalized multinomial logit model in Stata, *Stata Journal*, **13**, 382–97.

Hausman, J. A. and Ruud, P. A. (1987) Specifying and testing econometric models for rank-ordered data, *Journal of Econometrics*, **34**, 83–104.

Hayes, L. J., O'Brien-Pallas, L., Duffield, C. *et al*. (2006) Nurse turnover: a literature review, *International Journal of Nursing Studies*, **43**, 237–63.

Heinz, D. (2004) Hospital nurse staffing and patient outcomes: a review of current literature, *Dimensions of Critical Care Nursing*, **23**, 44–50.

Hess, R. and Rose, J. (2012) Can scale and coefficient heterogeneity be separated in random coefficients models?, *Transportation*, **39**, 1225–39.

Hogan, P., Moxham, L. and Dwyer, T. (2007) Human resource management strategies for the retention of nurses in acute care settings in hospitals in Australia, *Contemporary Nurse*, **24**, 189–99.

Keane, M. P. and Wasi, N. (2013) Comparing alternative models of heterogeneity in consumer choice behavior, *Journal of Applied Econometrics*, **28**, 1018–45.

Kenny, P., Doiron, D., Hall, J. *et al*. (2012) The training and job decisions of nurses – the first year of a longitudinal study investigating nurse recruitment and retention, CHERE Working Paper 2012/02, Centre for Health Economics Research and Evaluation, University of Technology, Sydney. Available at http://www.chere.uts.edu.au/pdf/wp2012_02.pdf (accessed 6 November 2013).

Lagarde, M. and Blaauw, D. (2009) A review of the application and contribution of discrete choice experiments to inform human resources policy interventions, *Human Resources for Health*, **7**, 62.

Lancsar, E. and Savage, E. (2004) Deriving welfare measures from discrete choice experiments: inconsistency between current methods and random utility and welfare theory, *Health Economics Letters*, **13**, 901–7.

Mangham, L. J. and Hanson, K. (2008) Employment preferences of public sector nurses in malawi: results from a discrete choice experiment, *Tropical Medicine and International Health*, **130**, 1433–41.

Marley, A. A. J., Flynn, T. N. and Louviere, J. J. (2008) Probabilistic models of set-dependent and attribute-level best-worst choice, *Journal of Mathematical Psychology*, **52**, 281–96.

Marschak, J. (1960) Binary choice constraints on random utility indications, in *Stanford Symposium on Mathematical Methods in the Social Sciences,* Arrow K. (Ed.), Stanford University Press, Stanford, CA, pp. 312–29.

McFadden, D. (1981) Econometric models of probabilistic choice, in *Structural Analysis of Discrete Data with Econometric Applications*, Manski, C. F. and McFadden, D. (Eds), The MIT Press, Cambridge, MA, pp. 198–272.

Naude, M. and McCabe, R. (2005) Magnet hospital research pilot project conducted in hospitals in Western Australia, *Contemporary Nurse*, **20**, 38–55.

Needleman, J. and Hassmiller, S. (2009) The role of nurses in improving hospital quality and efficiency: real-world results, *Health Affairs*, **28**, 625–33.

Nooney, J. G., Unruh, L. and Yore, M. M. (2010) Should I stay or should I go? Career change and labor force separation among registered nurses in the U.S., *Social Science and Medicine*, **700**, 1874–81.

Ohler, T., Le, A., Louviere, J. J. *et al*. (2000) Attribute range effects in binary response tasks, *Marketing Letters*, **11**, 249–60.

Oulton, J. A. (2006) The global nursing shortage: an overview of issues and actions, *Policy, Politics & Nursing Practice*, **7**, 34S–39S.

Penn-Kekana, L., Blaauw, D., San Tint, K. *et al*. (2005) Nursing staff dynamics and implications for maternal health provision in public health facilities in the context of HIV/AIDS, Centre for Health Policy, University of Witwatersrand Frontiers in Reproductive Health Population Council, Johannesburg. Available from http://maternalhealthtaskforce.org/component/docman/doc_details/509-nursing-staff-dynamics-and-implications-for-maternal-health-provision-in-public-health-facilities-in (accessed 15 November 2013).

Robinson, S. and Griffiths, P. (2007) *Nursing Education and Regulation: International Profiles and Perspectives*, Kings College, London.

Rother, J. and Lavizzo-Mourey, R. (2009) Addressing the nursing workforce: a critical element for health reform, *Health Affairs*, **28**, 620–4.

Rush, K. L., Adamack, M., Gordon, J. *et al*. (2013) Best practices of formal new graduate nurse transition programs: an integrative review, *International Journal of Nursing Studies*, **50**, 345–56.

Shields, M. A. (2004) Addressing nurse shortages: what can policy makers learn from the econometric evidence on nurse labour supply?, *Economic Journal*, **1140**, F464–F498.

Shields, M. A. and Ward, M. (2001) Improving nurse retention in the National Health Service in England: the impact of job satisfaction on intentions to quit, *Journal of Health Economics*, **200**, 677–701.

Sivey, P., Scott, A., Witt, J. *et al*. (2010) Why junior doctors don't want to become general practitioners: a discrete choice experiment from the Mabel longitudinal study of doctors, Working Paper 17/10, Melbourne Institute of Applied Economic and Social Research, Melbourne.

Sochalski, J. (2002) Nursing shortage redux: turning the corner on an enduring problem, *Health Affairs*, **21**, 157–64.

Street, D. J. and Burgess, L. (2007) *The Construction of Optimal Stated Choice Experiments: Theory and Methods*, Wiley, New York.

Street, D. J., Burgess, L. and Louviere, J. (2005) Quick and easy choice sets: constructing optimal and nearly optimal stated choice experiments, *International Journal of Research in Marketing*, **22**, 459–70.

The Productivity Commission. (2005) *Australia's Health Workforce: Productivity Commission Position Paper*. Productivity Commission, Canberra.

Train, K. E. (2008) EM algorithms for nonparametric estimation of mixing distributions, *Journal of Choice Modelling*, **10**, 40–69.

Train, K. E. (2009) *Discrete Choice Methods with Simulation, Second Edition*, Cambridge University Press, Cambridge.

Turner, C., Bain, C., Schluter, P. J. *et al*., and the Nurses and Midwives e-cohort Group (2009) Cohort profile: the nurses and midwives e-cohort study – a novel electronic longitudinal study, *International Journal of Epidemiology*, **38**, 53–60.

Ujvarine, A. S., Zrinyi, M., Toth, H. *et al*. (201 1) Intent to stay in nursing: internal and external migration in Hungary, *Journal of Clinical Nursing*, **20**, 882–91.

Vermeulen, B., Goos, P. and Vandebroek, M. (2010) Obtaining more information from conjoint experiments by best-worst choices, *Computational Statistics and Data Analysis*, **540**, 1426–33.

Zeytinoglu, I. U., Denton, M., Davies, S. *et al*. (2006) Retaining nurses in their employing hospitals and in the profession: effects of job preference, unpaid overtime, importance of earnings and stress, *Health Policy*, **790**, 57–72.

Zeytinoglu, I. U., Denton, M. and Plenderleith, J. M. (2011) Flexible employment and nurses' intention to leave the profession: the role of support at work, *Health Policy*, **990**, 149–57.