<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>Enhancing the comparability of self-rated skills-matching using anchoring vignettes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Authors(s)</strong></td>
<td>Ryan, Martin; Delaney, Liam; Harmon, Colm</td>
</tr>
<tr>
<td><strong>Publication date</strong></td>
<td>2009-01-16</td>
</tr>
<tr>
<td><strong>Conference details</strong></td>
<td>Twenty-third Annual Conference of the Irish Economic Association, Blarney, Co. Cork, 24-26 April 2009</td>
</tr>
<tr>
<td><strong>Publisher</strong></td>
<td>Irish Economic Association</td>
</tr>
<tr>
<td><strong>Item record/more information</strong></td>
<td><a href="http://hdl.handle.net/10197/2066">http://hdl.handle.net/10197/2066</a></td>
</tr>
</tbody>
</table>
Enhancing the Comparability of Self-Rated Skills-Matching Using Anchoring Vignettes

IEA Draft - 16/01/09

Martin Ryan, University College Dublin
Liam Delaney, University College Dublin
Colm Harmon, University College Dublin, IZA

This research is concerned with the skills-match between researchers’ Ph.D. training and their subsequent university employment. Self-rated skills-matching is considered in light of the anchoring vignettes technique. This technique is used to address comparability issues in survey research. It has been documented that individuals with more education and skills have the highest expectations for their jobs and careers; and are more easily disappointed. This is one reason why there may be comparability problems in self-rated skills matching. Only a few studies (at least one using objective data, a few more using self-reported data) have examined the issue of skills-matching. Furthermore, the results from objective data are somewhat problematic. This underscores the need to apply anchoring vignettes to self-reported data. Mismatch is associated with substantially lower earnings; with more comparable measures, wage penalties can be more accurately estimated.

JEL: I23, J24, C83

Keywords: skills-matching, anchoring vignettes, Ph.D. outcomes, self-rated survey questions, over-education, human capital, occupational choice, career expectations

I. Introduction

Increasing attention is being placed on the outcomes of Ph.D. education all around the world (Akerlind, 2005). In recent years Science has conducted a survey in which life scientists from the United States report their income and job satisfaction (Kennedy, 2006). The need for more data on Ph.D. graduates career outcomes in Europe is one of the key conclusions of the EUA Doctoral Programmes Project (EUA, 2006). This research is concerned with

*Please seek authors’ permission to quote. Ryan is a graduate student at the UCD School of Economics and a Ph.D. fellow at the UCD Geary Institute; he acknowledges financial support from an IRCHSS scholarship. Email: martin.ryan@ucd.ie
post-Ph.D. outcomes for researchers in Irish universities - in particular, the extent of skills-match between researchers’ Ph.D. training and their subsequent university employment. The majority of Irish Ph.D. graduates (73.3 percent) work in the non-market services sector. 13 percent work in manufacturing, 4 percent in business, finance and insurance, and 2 percent in computing and the software industry. Of the 73.3 percent working in the non-market services sector, 50 percent work in third-level education; 12 percent work in health services and 3.7 percent work in secondary-school teaching (HEA, 2007 - see table in Appendix A).\textsuperscript{1}

The Irish Strategy for Science Technology and Innovation (2006-2013) contains a commitment to double the number of Ph.D. students and postdoctoral researchers operating in Irish universities. In Ireland, successive reports of the Expert Group on Future Skills Needs state that Ireland needs more doctoral candidates, especially in technical disciplines including engineering, information technology, mathematics and the physical sciences (Harmon, Delaney and Ryan, 2007). The number of Ph.D. qualifications awarded in Ireland this year topped the 1,000 mark (HEA, 2008); almost double the number a decade ago. However, Ireland’s R-and-D manpower, in terms of full time equivalent (FTE) researchers per 10,000 labour force (49), is below the EU-15 average of 57 (Singapore Agency for Science, Technology and Research, 2008).

Using the National Science Foundation’s (NSF) ‘Survey of Doctorate Recipients’ (SDR); Bender and Heywood (2006a) examine the extent of U.S. Ph.D. graduates’ ‘skills-matching’; they report that approximately one-sixth of academics in the United States report some degree of (skills) ‘mismatch’. The question on ‘matching’ in the SDR is collected in response to a demand from the (US) National Research Council for data that shows the extent of integration between “occupational detail and academic training” i.e. skills-matching. Mismatch is associated with substantially lower earnings, lower job satisfaction and a higher rate of turnover (Bender and Heywood, 2006a).

Nordin, Persson and Rooth (2008) add to the small literature on skills-matching using an objective measure; they use microdata collected by Statistics Sweden, but are forced to drop 36 percent of their sample due to restrictions on fields of education to well-defined categories. The authors state that this approach is necessary because some fields of education (e.g. in the humanities and languages) are either vague or cannot easily be matched with any specific occupation. Also, the authors exclude a further 11 percent of their sample because of missing occupation data. One way around these problems is to use self-rated measures of skill-matching; this is the approach in this research paper. Only a few studies (at least one using objective data, a few more using self-reported data) have examined the issue of skills-matching.

\textsuperscript{1}The comparative lack of researchers in Europe (compared to the US) is particularly visible in business, with almost one-third fewer researchers employed in this sector compared to the US. Unemployment among doctoral researchers in the EU is also relatively high, while their salary remains low in comparison to other professionals with a similar education (Harmon, Delaney and Ryan, 2007).
Self-rated skills-matching is considered in light of the anchoring vignettes technique. This technique is used to address comparability issues in survey research. It has been documented that individuals with more education and skills have the highest expectations for their jobs and careers; and are more easily disappointed. This is one reason why there may be comparability problems in self-rated skills matching. It is possible that the standards for what constitutes skills-matching could differ for multiple reasons, including career expectations, inputs into the Ph.D. training process and other factors such as gender and social background. We give particular consideration to the possibility that comparability problems may occur due to different expectations for Ph.D. outcomes.

The following section of the paper discusses the literature on skills-matching, particularly for Ph.D. graduates. It also covers some of the literature on expectations related to educational outcomes. The third section introduces anchoring vignettes and discusses the implementation strategy for the method. The fourth section describes the data used in our analysis. The fifth section reports the method and preliminary results.

II. Literature

Skills-Matching

Bender and Heywood (2006a) use the SDR to examine Ph.D. graduates’ skills-matching by using the following question as an indicator: “Thinking about the relationship between your work and your education, to what extent is your work related to your doctoral degree?” The possible responses are “closely related,” “somewhat related” and “not related.”

Bender and Heywood (2006a) report that approximately one-sixth of academics in the United States report some degree of mismatch in the SDR. This mismatch is associated with substantially lower earnings, lower job satisfaction and a higher rate of turnover (Bender and Heywood, 2006).

Labour market mismatch has traditionally been thought of in the context of unemployment, as opposed to a skills mismatch after educational attainment. According to Soininen’s Ph.D. thesis (2006), the ‘matching function’ and the Beveridge curve are two extensively utilized relationships in labour economics. The Beveridge curve describes the negative, convex to origo relationship between vacancies and unemployed, which is the outcome of the matching function. The name derives from William Beveridge, an English social politician.

---

2 As mentioned in the introduction, this question on ‘matching’ in the SDR is collected in response to a demand from the (US) National Research Council for data that shows the extent of integration between “occupational detail and academic training” i.e. ‘skills-matching’.

3 The SDR is conducted every two years and is a longitudinal survey that follows recipients of research doctorates from US institutions until age 76.
who discovered the relationship in the 1930s. Being far down on the Beveridge curve indicates
that a low number of vacancies corresponds to high unemployment and being far up on the
curve indicates that a high number of vacancies corresponds to low unemployment. The
matching function was first mentioned at the end of the 1970s in the economics literature.

The modern interpretation of ‘matching’ in the economics literature is provided by Budria
and Moro-Egido (2004); these authors differentiate between three different types of educa-
tional mismatch: ‘over-qualification’, ‘incorrect qualification’, and ‘strong mismatch’. They
find that while over-qualification and incorrect qualification are not associated with lower
wages, strong mismatch carries a pay penalty that ranges from 13 percent to 27 percent.
As mentioned in the introduction, strong (or skills) mismatch after the attainment of Ph.D.
qualifications is the particular focus of the research in this paper. It should also be emphasised
that over-education is a separate issue; over-education corresponds to ‘over-qualification’ in
the Budria and Moro-Egido (2004) lexicon. There is now a fairly large literature on the
relationship between overeducation/undereducation and earnings (see Bourdet and Persson,
2008; Dolton and Vignoles, 2000; Hartog, 2000; McGuinness, 2006; Rubb, 2003; Sloane, Battu
and Seaman, 1999). The survey by Hartog (2000) concludes that the return to overeducation
is about half to two-thirds of the return to required schooling. More details on the over-
education literature are provided in Nordin, Persson and Rooth (2008).

According to Nordin, Persson and Rooth (2008), very few studies have yet focused on the
mismatch between the individual’s field of education and his/her occupation. They mention a
paper by Robst (2007) which uses data on U.S. college graduates (the NSF College Graduate
Survey); this study finds that having a major that does not match one’s work is associated
with 11 percent lower annual income. Robst (2007) uses a self-reported measure of matching,
that Nordin, Persson and Rooth (2008) suggest may be a ‘form of rationalisation of a general
feeling of disappointment with the wage and/or workplace’.

Nordin, Persson and Rooth (2008) add to the small literature on skills-matching using an
objective measure; they use microdata collected by Statistics Sweden, but are forced to drop
36 percent of their sample due to restrictions on fields of education to well-defined categories.
The authors state that this approach is necessary because some fields of education (e.g. in
the humanities and languages) are either vague or cannot easily be matched with any specific
occupation. Also, the authors exclude a further 11 percent of their sample because of missing
occupation data. Robst (2007) discusses other instances where objective measures of skills-
matching may be problematic. For example, “many college majors provide students with a
broad range of skills... that apply to different occupations. It would be difficult to develop an
algorithm for determining whether a major and a job are unrelated... individual assessments,
while perhaps subjective, are expected to provide important information.”

Nordin et al. (2008) focus on the income penalty for skills mismatch; they find that the
penalty for such mismatches is large for both men and women. They also find that it is
substantially larger than has been found for the US. Interestingly, the authors also control for
cognitive ability and find that the income penalty is not caused by ability (at least for Swedish
men). The income penalty for men decreases with work experience; and given this, Nordin et
al. (2008) suggest that education-specific skills and work experience are substitutes to some
extent. However, Nordin et al. (2008) present no evidence that mismatched individuals move
to a more matched occupation over time.

At present, the literature on skills-matching only features a few studies; at least one
using objective data, and a few more using self-reported data. One of the main issues in the
literature is that the results from objective data are somewhat problematic. In addition to
this, other literature shows that the method used to measure over/under-education affects
the results (Battu, Belfield and Sloane, 1999; Groot and Maassen van der Brink, 2000).
Groot and Maassen van der Brink (2000) find that over-education is more frequent when a
self-reported rather than an objective measure is used. While over-education is a different
issue to skills-matching, the measurement problems in the over-education literature provide
further motivation to develop a measure of self-rated skills-matching that is comparable across
individuals.

Expectations Related to Educational Outcomes

Some of the literature on expectations related to educational outcomes was mentioned
in the introduction. It was noted that individuals with more education and skills have the
highest expectations for their jobs and careers; and are more easily disappointed (Tsang and
Levin, 1985; Clark and Oswald, 1996). In addition, psychological theories of expectation
suggest that the under-utilization of skills is a potential cause of diminished job satisfaction
(Bender and Heywood, 2006a). In their paper, “Job Satisfaction of the Highly Educated”,
Bender and Heywood (2006b) report that the job satisfaction of Ph.D. level scientists in the
United States is affected by gender - female scientists report lower job satisfaction than males
in academic settings (but higher job satisfaction than males in non-academic settings). It
may be the case that female Ph.D. graduates have higher expectations for being matched to
their skill-set when they are working in the academic sector.

In addition, there is also the idea that individual expectations are raised by exposure
to higher standards/expectations in the external environment. The first evidence on this
dynamic was produced in the organisational/management literature, when researchers ex-
plitly considered that raising worker expectations might enhance motivation. Previous to
this, it was established that ‘managers who expect more, get more’ (e.g., Likert, 1961, 1967;
McGregor, 1960); this has been referred to as the Pygmalion effect.4

4Pygmalion was an ancient King who hailed from Cyprus and it is said that he carved a statue in the
image of his ideal woman. So perfect was his creation, that he could not help but fall in love with it and so
strong was the love he bore his creation that with his will and a little help from the Gods he brought the
statue to life and they were married.
was the first to form a theory where self-esteem mediates the relationship between individual expectations and performance. Subsequently, this was referred to as the Galatea effect, named after Pygmalion’s sculpture, as it involves working directly on the sculpture without the Gods (Eden and Ravid, 1982).

Babad, Inbar and Rosenthal (1982) referred to the negative impact of low expectations on achievement as the Golem effect.

We are primarily interested in the Galatea effect, that is, scenarios where high self-expectations lead to high performance. According to McNatt and Judge (2004), there are several conceptual formulations of self-expectations of performance, which include specific self-efficacy, self-confidence, and performance expectancy. It is also possible that psychometric measures such as self-esteem, or extraversion and openness (from the Big Five), are somewhat related to the Galatea effect. The first experimental evidence for the Galatea effect is provided by Eden and Ravid (1982). In this experiment, Galatea trainees were given a five-minute personal interview by a psychologist, at the end of which they were told, “You have high potential for success”. Control groups were told that they had regular potential for success or were simply given interviews without the last sentence. The results showed a substantial rise in self-expectations due to the Galatea effect.

Existing research demonstrates what may constitute a Galatea effect in the Ph.D. training process. Several years ago, the Cornell Higher Education Research Institute evaluated the Graduate Education Initiative (GEI) that provided 85 million dollars in financial support towards Ph.D. education in the United States (Ehrenberg, 2005). This initiative was criticised for failing to directly address more fundamental issues on departmental levels in the American universities, such as ensuring match between supervisor and student, or match between student and Ph.D. topic. Furthermore, the Cornell evaluation describes how generous financial support may even have induced students to persist with an unloved topic for longer than they might otherwise have done (Ehrenberg, 2005). This research underscores the importance of inputs into the Ph.D. training process; and these inputs are something that may lead to higher expectations for Ph.D. outcomes. We discuss further in the following section how inputs into Ph.D. training may raise self-expectations (i.e. a Galatea effect) and thereby increase expectations for post-Ph.D. educational outcomes.

\[5\] Economists have considered the Pygmalion and Galatea effects in relation to reference-dependent preferences; see Daido and Itoh (2007): \text{http://obata.misc.hit-u.ac.jp/~itoht/pygmalion07-0116.pdf}
III. Anchoring Vignettes: Background and Implementation Strategy

Background

The anchoring vignettes technique is presented by King, Murray, Salomon and Tandon (2004) as a method to address comparability issues in survey research (see also: King and Wand, 2007). Comparability problems occur when different groups of respondents understand and use ordinal response categories such as (1) strongly disagree, (2) disagree, (3) neutral, (4) agree, or (5) strongly agree in different ways. When one group of respondents happen to have comparatively higher standards for what constitutes the definition of (for example) “strongly agree”, then they will report systematically lower levels of agreement than another group (King and Wand, 2007). We will motivate the relevance of this technique to our research question before delving further into the literature and discussing modelling strategies.

Self-reported levels of skills-matching are not comparable because respondents interpret the survey question about self-rated matching differently; and therefore the levels of matching that they report are not comparable. To illustrate, two researchers with the exact same skills-match may rate their match differently. That is, one researcher may say that their job is “very closely related” to their Ph.D. training while another may say that their job is “closely related” to their Ph.D. training (while both have the same underlying match). One understanding of this situation could be that the two individuals simply interpret thresholds on the response scale differently; however it is also likely that there are some underlying reasons for this “differential item functioning” (DIF). For example, better inputs into the Ph.D. training process, gender-related factors, or the extent of match between Ph.D. student and supervisor could all be reasons why a given Ph.D. graduate would have higher expectations for his/her Ph.D. outcomes. To clarify, we hypothesise that the process of highly skilled (i.e. Ph.D. level) human capital accumulation may set expectations at a certain level with a related impact on how Ph.D. graduates interpret the response scale on a matching question.\(^6\) In the previous section we described how inputs into Ph.D. training may raise self-expectations (i.e. a Galatea effect) and thereby increase expectations for post-Ph.D. educational outcomes. We give particular consideration to the possibility that comparability problems in self-rated skills matching may occur due to different expectations for Ph.D. outcomes.

Another possibility is that the duration of the Ph.D. process (or time-to-degree) may be a source of comparability problems in self-rated skills matching for Ph.D. graduates. The idea here is that the more the individual has committed to the process of attaining a Ph.D., the

\(^6\)An analogue to this situation is the scenario described by Sen (2002), where one Indian state with the highest level of longevity also has the highest level of self-reported morbidity (i.e. self-reported burden of ill-health). By contrast, states with low levels of longevity have the lowest rates of self-reported morbidity in India. The implication is that health conditions are affecting health expectations and thereby affecting self-reported levels of morbidity.
more he or she will want to view the outcome of that process favourably. There is a precedent for this type of comparability-bias in the anchoring vignettes literature. Buckley (2007) has used the anchoring vignettes technique to investigate the “rose-coloured glasses” effect, which refers to parents reporting higher levels of satisfaction with a school solely or partially as a justification for the effort expended in the choice process. The analogy to ‘time-to-degree’ is about the amount of time expended in the Ph.D. process. The rose-coloured glasses idea is reminiscent of the cognitive dissonance problem (see: Cooper, 2007; Festinger, 1957).

Initially, Buckley (2007) finds that parents in charter schools evaluate their child’s school more highly and are more satisfied with many dimensions of those schools than parents with children in traditional public schools. (Charter schools are elementary or secondary schools in the United States that receive public money but have been freed from some of the rules, regulations, and statutes that apply to other public schools). After applying the anchoring vignettes technique, Buckley (2007) shows that parents who change to a charter school are actually tougher graders of the new school. In relation to self-rated skills-matching, it may be the case that Ph.D. graduates with longer time-to-degree report higher levels of matching due to a similar bias. Stock and Siegfried (2006) discuss how taking a year longer during Ph.D. training implies an annual opportunity cost.

The focal paper in the anchoring vignettes literature is King, Murray, Salomon and Tandon (2004). In this paper King et al. (2004) concentrate on ‘political efficacy’ and ‘freedom’ as examples of how respondents can interpret identical survey questions in different ways. The rationale, as described in King et al. (2004), is to “measure incomparability via respondents’ assessments, on the same scale as the self-assessments to be corrected, of hypothetical individuals described in short vignettes”. The key to the technique is to avail of the fact that the actual (though not necessarily the reported) response levels on the vignettes are fixed over individuals. As a result, heterogeneity in vignette responses indicates the incomparability that we seek to measure. To summarise succintly, the anchoring vignettes technique is used to: (i) measure incomparability by asking respondents to assess hypothetical scenarios described in short vignettes, and then (ii) correct the incomparability through simple re-codes or the use of a statistical model.

King et al. (2004) illustrate the vignette design process by presenting five vignettes related to political efficacy, which fall on an ordered scale - from most to least efficacy. Importantly, the self-rated question of interest and all response scales should follow the same format. King et al. (2004) recommend asking about self-assessment first, followed by the vignettes presented in random order. Buckley (2008) presents findings on survey context effects in the use of anchoring vignettes. Using data from a randomised survey experiment Buckley investigates whether analyses based on anchoring vignettes may be vulnerable to the introduction of “survey artifacts” due to vignette ordering or the placement of the self-assessment item relative to the vignettes. He finds several patterns of bias due to context effects, and recommends
that researchers using anchoring vignettes should consider randomisation or other methods to mitigate these problems. Hopkins and King (2008) report the results of several randomised survey experiments designed to evaluate two intended improvements for anchoring vignettes. The Hopkins and King experiments (2008) show that switching the question order so that self-assessments follow the vignettes leads to a priming of respondents to define the response scale in a similar way. Though Hopkins and King (2008) state that priming is not a bias, we avoid it as we want to examine naive self-reports of skills-matching, before we re-scale those reports with information from anchoring vignettes.⁷

Crucially, two assumptions must be made before one implements the anchoring vignettes technique. The first is response consistency, which requires that each individual uses the response categories on a given question in the same way when (i) providing a self-assessment and (ii) assessing each of the hypothetical people in the vignettes. This is a difficult assumption to test, though King et al. (2004) make a useful point about how the assumption would be violated; they state that respondents who feel inferior to hypothetical individuals would set a higher threshold for what counts as their having a “lot of say” in a political efficacy example.⁸

The second assumption is called vignette equivalence - it requires that the level of the variable represented in any one vignette is perceived by all respondents in the same way and on the same unidimensional scale, apart from random measurement error. See Van Soest, Delaney, Harmon, Kapteyn and Smith (2007) for a detailed discussion of vignette equivalence, and a test of the assumption based on objective measurements. Their finding is that vignette based corrections appear quite effective in bringing objective and subjective measures closer together (Van Soest, Delaney, Harmon, Kapteyn and Smith; 2007).

**Implementation Strategy**

Following on from the vignette assumptions, we can consider how to implement the method. Before we discuss a modelling strategy, we can first consider a simple nonparametric approach; which requires that the vignette questions be asked of all the same respondents as the self-assessments; and that there are no explanatory variables (King et al., 2004). Figure 1 (on the following page - taken from King et al.) shows one self-assessment and three vignette assessments for each of two individual survey respondents (labeled 1, on the left, and 2, in the middle). The self-assessed level of political efficacy is higher for Respondent 1 (and they agree on the ordinal ranking of the vignettes). However, the fact that Alison’s (or Jane’s or Moses’s) actual level of political efficacy is the same no matter which respondent is being asked makes it possible to make the two comparable by stretching Respondent 2’s scale so

---

⁷Hopkins and King (2008) also demonstrate that combining vignettes and self-assessments in a single direct comparison induces inconsistent and less informative responses.

⁸We are hoping that our survey respondents will not feel inferior to the hypothetical individuals described in our anchoring vignettes.
that the vignette assessments for the two respondents match. This is done in the scale on the right in Figure 1. With this adjustment, we can see that Respondent 2 has a higher level of self-reported political efficacy than Respondent 1. This comes from the fact that Respondent 1 rates herself lower than Jane, whereas Respondent 2 rates herself higher than Jane.

King et al. (2004) refer to the approach in the above diagram as literally marking and stretching rubber bands. The idea is to recode the categorical self-assessment relative to the set of vignettes. King describes a situation where all respondents order the vignettes in the same way. Given this supposition and the vignettes above in Figure 1, we can assign the recoded variable as 1 if the self-assessment is below Moses, 2 if equal to Moses, 3 if between Moses and Jane, 4 if equal to Jane, 5 if between Jane and Alison, 6 if equal to Alison, and 7 if above Alison. By this coding, the first respondent in Figure 1 is coded 3 and the second is coded 5. The resulting variable is DIF-free. More formally, below we see $y_i$, the categorical survey self-assessment for respondent $i$ ($i = 1, \ldots, n$); $z_{ij}$ is the categorical survey response for respondent $i$ on vignette $j$ ($j = 1, \ldots, J$). For respondents with identical ordinal rankings on all vignettes ($z_{i,j-1} < z_{ij}$, for all $i, j$), the DIF-corrected variable is shown below as:

\[
C_i = \begin{cases} 
1 & \text{if } y_i < z_{11}, \\
2 & \text{if } y_i = z_{11}, \\
3 & \text{if } z_{11} < y_i < z_{21}, \\
\vdots & \vdots \\
2J + 1 & \text{if } y_i > z_{J1}.
\end{cases}
\]

King et al. (2004) refer to the approach in the above diagram as literally marking and stretching rubber bands. The idea is to recode the categorical self-assessment relative to the set of vignettes. King describes a situation where all respondents order the vignettes in the same way. Given this supposition and the vignettes above in Figure 1, we can assign the recoded variable as 1 if the self-assessment is below Moses, 2 if equal to Moses, 3 if between Moses and Jane, 4 if equal to Jane, 5 if between Jane and Alison, 6 if equal to Alison, and 7 if above Alison. By this coding, the first respondent in Figure 1 is coded 3 and the second is coded 5. The resulting variable is DIF-free. More formally, below we see $y_i$, the categorical survey self-assessment for respondent $i$ ($i = 1, \ldots, n$); $z_{ij}$ is the categorical survey response for respondent $i$ on vignette $j$ ($j = 1, \ldots, J$). For respondents with identical ordinal rankings on all vignettes ($z_{i,j-1} < z_{ij}$, for all $i, j$), the DIF-corrected variable is shown below as:

\[
C_i = \begin{cases} 
1 & \text{if } y_i < z_{11}, \\
2 & \text{if } y_i = z_{11}, \\
3 & \text{if } z_{11} < y_i < z_{21}, \\
\vdots & \vdots \\
2J + 1 & \text{if } y_i > z_{J1}.
\end{cases}
\]

\[9\text{A formula for the practical implementation of re-coding (i.e. the non-parametric approach) is shown in Appendix C.}\]
The main (parametric) modelling strategy allows researchers to ask vignettes of only a random sample (or subsample) from the same population as the self-assessments. The key to this method is allowing thresholds to turn unobserved continuous (latent) variables into observed categorical responses, which vary over individuals as a function of measured explanatory variables. King et al. (2004) compare this approach to an ordered probit model - where DIF is modelled through threshold variation (with the vignettes providing the key identifying information). The vignettes enable estimation of the threshold values, and with this information the self-assessment responses can be corrected. An illustration of the parametric model is provided in the model below (taken from King et al., 2004).

The GLLAMM package (developed by Sophia Rabe-Hesketh, with Anders Skrondal and Andrew Pickles) will be used for the parametric emodelling strategy; GLLAMM was obtained from the Statistical Software Components (SSC) archive and installed into STATA as an ado-file; it is a set of programs for estimating, predicting, and simulating Generalized Linear Latent And Mixed Models.\(^{10}\) GLLAMM maximises the marginal log-likelihood using Stata’s version of the Newton Raphson Algorithm (ml with method d0). In the case of discrete random effects or factors, the marginal log-likelihood is evaluated exactly whereas numerical integration is used for continuous (multivariate) normal random effects or factors.\(^{11}\)

\[
\begin{array}{c}
\text{Actual:} \quad \theta_1, \ldots, \theta_J \\
\text{Perceived:} \quad \omega^2 \rightarrow X_1, \ldots, X_J \rightarrow \omega^2 \quad 1 \rightarrow Y_1, \ldots, Y_J \rightarrow 1 \\
\text{Reported:} \quad \nu_{ij} \rightarrow \xi_i \quad \xi_i \rightarrow \nu_{ij} \quad \nu_{ij} \rightarrow \nu_{ij} \quad 1 \rightarrow \nu_{ij} \\
\end{array}
\]

<table>
<thead>
<tr>
<th>Vignettes</th>
<th>Self-Assessments</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X_i$</td>
<td>$Y_i$</td>
</tr>
</tbody>
</table>

Note: Vignette questions are on the left, with perceived and reported but not actual levels varying over observations $i$. Self-assessment questions are on the right, with all levels varying over observations $i$. The first self-assessment question (see $y^2_i$) is tied to the vignettes by the same coefficient on the variables predicting the thresholds, $\nu_{ij}$, and to the remaining self-assessment questions by person $i$’s actual value, $\mu_i$. Each solid arrow denotes a deterministic effect, a squiggly arrow denotes the addition of normal random error, with variance indicated at the arrow’s source.

According to King et al. (2004), the optimal number of vignettes to ask, in terms of the right trade-off in bias reduction vs. survey costs, depends on the nature of DIF and what information the investigator needs. King et al. (2004) suggest that only one vignette is needed to identify a parametric model in general, but they normally advise to include more. In the

\(^{10}\) An alternative program is the Anchors package developed by Jonathan Wand, Gary King and Olivia Lau; this can be run on MS-Windows with a graphical user interface (GUI).

\(^{11}\) Two methods are available for numerical integration: quadrature or adaptive quadrature. In both cases it is essential to make sure that a sufficient number of quadrature points has been used by comparing solutions with different numbers quadrature points. In most cases adaptive quadrature will perform better than ordinary quadrature. This is particularly the case if the cluster sizes are large and the responses include (large) counts and/or continuous variables. Even where ordinary quadrature performs well, adaptive quadrature often requires fewer quadrature points making it faster.
experience of King et al. (2004), much of the benefit of the anchoring vignettes approach is realised after including the first two or three vignettes - if they are carefully chosen to be near the self-assessments. In our analysis, we use three vignettes for each self-assessment; these were carefully chosen by conducting a pilot-test. This is discussed in the following section.

IV. Data

The self-rated skills-match between researchers’ Ph.D. training and their subsequent university employment will be modelled using data from the Irish Universities Study. Initially, test data was gathered from a pilot-survey conducted with a convenience sample of 68 colleagues. (King et al., 2007 advise testing out anchoring vignettes with a pilot-study). The 68 respondents in the pilot-study filled out a web-survey in November 2008, answering nine questions intended for inclusion in the Irish Universities Study (IUS).12 Questions 4-9 and the associated results can be viewed in Appendix B. Questions 4-6 are anchoring vignettes related to mismatch; questions 7-9 are anchoring vignettes related to over-education. We do not expect over-education to be an issue for researchers with Ph.D. qualifications working in Irish universities; nonetheless we will test whether this is the case.

The anchoring vignettes that we apply to self-rated matching have a five-point response scale, as follows: “very closely related”, “closely related”, “somewhat related”, “not very related” and “not at all related”. This differs from the three-point scale in the NSF survey of US doctorate recipients (SDR); we chose to use a five-point scale based closely on the SDR question. Five-point scales allow greater detail in survey response across threshold categories; they are also more common in the anchoring vignettes literature.

Test data results from the anchoring vignettes related to matching are as follows. Questions 4, 6, 7 and 9 produce results in the expected direction; we call these the “boundary vignettes”. 86 percent of test-data respondents think that a university researcher is closely or very closely matched to Ph.D. training. 73 percent of test-data respondents think that a university researcher (definitely or probably) needs a Ph.D. qualification for their work. 95 percent of test-data respondents think that a shop worker is not matched to Ph.D. training (not very or not at all). 97 percent of test-data respondents think that a shop worker does not need a Ph.D. qualification for their work (probably or definitely not).

We are confident that we have found “boundary” employment situations that are more matched than a secondary school teacher and less matched than a secondary school teacher, respectively. Given this, we refer to employment as a secondary school teacher as the “interval” vignette. Only 3.7 percent of Irish Ph.D. graduates go into secondary school

---

12 The first 3 questions are self-rated measures; we do not interpret the results as they are clearly out of sample. Questions 4-9 are the anchoring vignettes.
teaching every year, so we argue strongly that this is not what Ph.D. students are being trained for (secondary school teachers only need a BA and H-Dip to qualify).13

The Irish Universities Study is a 3-4 year research project being carried out on behalf of the Irish Universities Association by the UCD Geary Institute and is sponsored by the Strategic Innovation Fund (SIF). A component of this study has been designed to inform policy about the health, welfare and college-satisfaction of researchers in the seven Irish universities. After testing out the anchoring vignettes in the pilot-study, they were put forward for inclusion in the contract researchers’ component of the Irish Universities Study; hereafter referred to as the Contract Researchers’ Survey (CRS). The CRS is being conducted as a web-survey, which allows immediate compilation of and access to the data from the project database. Field-work for the web-survey began at the end of December; this involves dissemination of the survey to the entire population of contract researchers by the Irish Universities Association (IUA).

So far, the survey has only been launched in three universities; and 102 respondents have filled it out. The survey is due to be launched in the remaining four universities before the end of January. In addition, the IUA plan to re-issue the survey information twice before the end of March, in addition to the original launch. The IUA estimates that the entire population of contract researchers is approximately 3,000. Given this, the CRS survey team predict a response rate of roughly 50 per cent (1,500), to be achieved by the end of March. Despite the survey being live in the field; it has been possible to begin analysis with the anchoring vignettes technique, using the preliminary sample of 102 respondents. Indeed, some preliminary results (based on the non-parametric approach) are reported in the following section. These results will be substantially augmented before the conference; in particular, parametric estimation using the GLLAMM statistical package will be included. Given the advantages of the web-survey data collection process, the analysis can be replicated at any stage that substantial numbers of respondents enter the survey.

Before proceeding to the following section, we provide some more information about the data. The question about self-rated skills-matching is not presented to survey respondents who indicate that they do not have a Ph.D. qualification. Every survey respondent is presented with survey modules about their job characteristics, demographic background and academic history. From this point onwards “independent draws” are used to assign additional survey modules. 1 in 2 survey respondents are presented with a detailed “Ph.D. module”; this contains questions about Ph.D. training, skills-matching, over-education and anchoring vignettes.

13We get the figure of 3.7 percent from the most recent First Destinations report (HEA, 2007).
V. Method and Results (Preliminary)

The main questions that will be addressed during January, February and March are as follows:

(i) Is there differential item functioning (DIF) in self-rated skills matching?

(ii) If so, what causes DIF in self-rated skills matching?

With a measure of skills-match that is comparable across individuals, we can also ask:

(iii) Is mismatch more prevalent in certain disciplines or for certain individuals?

(iv) What can we say about the role of mismatch in predicting earnings for contract researchers in Irish universities?

At this stage, we only attempt to answer the first two questions using the non-parametric approach. However, over the next two to three months we will answer all four questions using both the parametric and the non-parametric approach. First of all, we present some initial findings, based on ordered probit regression. We use self-rated skills-matching (coded 1-5) as the dependent variable, and experiment with all of the intuitively sensible predictors in the data-set (the survey has been designed with a lot of relevant information). Due to the current number of observations ($n = 36$), we do not yet present any regression output.\(^{14}\)

We experiment with a small number of predictors in each regression and have found that the following variables are insignificant: age, gender, father’s education, mother’s education, success of supervisor, previous work outside academia, having a desk during Ph.D., and doing a taught component during Ph.D. However, we find that the following variables are predictive of self-rated skills-matching: Ph.D. induction, Ph.D. feedback, Ph.D. lab facilities, training on grant-writing, and training on teaching skills. The categorical breakdown of answers to the self-rated skills-matching question are shown in the table below.

<table>
<thead>
<tr>
<th>Job Relatedness</th>
<th>Freq.</th>
<th>Percent</th>
<th>Cum.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very closely</td>
<td>12</td>
<td>33.33</td>
<td>33.33</td>
</tr>
<tr>
<td>Close enough</td>
<td>17</td>
<td>47.22</td>
<td>80.55</td>
</tr>
<tr>
<td>Somewhat</td>
<td>10</td>
<td>27.78</td>
<td>98.22</td>
</tr>
<tr>
<td>Not at all</td>
<td>1</td>
<td>2.78</td>
<td>100.00</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>36</td>
<td>100.00</td>
<td></td>
</tr>
</tbody>
</table>

The categorical breakdown of answers to the three anchoring vignettes (related to skills-matching) are shown in three tables below and on the following pages (Matching Vignettes 1, 15\).

\(^{14}\)We note that $n = 36$; this is due to 15 survey respondents not having a Ph.D. qualification. 1 in 2 of the 102 respondents were asked about self-rated skills-matching; but only if they were Ph.D. graduates.
We can see that the boundary vignettes (i.e. university researcher and shop worker) are fulfilling their purpose. The interval vignette (i.e. secondary school teacher) is skewing towards being not matched. However, the interval vignette still has the most observations at the mid-point (i.e. “somewhat matched”) on the response scale.

### Matching Vignette 1

- X has been working as a university researcher since he/she got a Ph.D. How closely is X’s job related to Ph.D. training?

- Most respondents say ‘closely’ or ‘very closely’.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Very closely</td>
<td>12</td>
<td>33.33</td>
<td>33.33</td>
<td></td>
</tr>
<tr>
<td>Closely</td>
<td>19</td>
<td>52.78</td>
<td>86.11</td>
<td></td>
</tr>
<tr>
<td>Somewhat</td>
<td>4</td>
<td>11.11</td>
<td>97.22</td>
<td></td>
</tr>
<tr>
<td>Not very</td>
<td>1</td>
<td>2.78</td>
<td>100.00</td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>36</td>
<td>100.00</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Matching Vignette 2

- X has been working as a secondary school teacher since he/she got a Ph.D. How closely is X’s job related to Ph.D. training?

- Most respondents say ‘somewhat’ to ‘not at all’.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Very closely</td>
<td>1</td>
<td>2.78</td>
<td>2.78</td>
<td></td>
</tr>
<tr>
<td>Somewhat</td>
<td>9</td>
<td>25.00</td>
<td>27.78</td>
<td></td>
</tr>
<tr>
<td>Not very</td>
<td>18</td>
<td>50.00</td>
<td>77.78</td>
<td></td>
</tr>
<tr>
<td><strong>Not at all</strong></td>
<td>8</td>
<td>22.22</td>
<td>100.00</td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>36</td>
<td>100.00</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Matching Vignette 3

- X has been working in a shop in the services sector since he/she got a Ph.D. How closely is X's job related to Ph.D. training?

- Most respondents say 'not all'.

<table>
<thead>
<tr>
<th></th>
<th>Freq.</th>
<th>Percent</th>
<th>Cum.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very closely</td>
<td>1</td>
<td>2.70</td>
<td>2.70</td>
</tr>
<tr>
<td>Somewhat</td>
<td>2</td>
<td>5.56</td>
<td>8.26</td>
</tr>
<tr>
<td>Not very</td>
<td>6</td>
<td>14.67</td>
<td>25.00</td>
</tr>
<tr>
<td>Not at all</td>
<td>27</td>
<td>75.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Total</td>
<td>36</td>
<td>100.00</td>
<td></td>
</tr>
</tbody>
</table>

The non-parametric anchoring vignettes method can be implemented with the Anchors package, developed by Jonathan Wand, Gary King and Olivia Lau (2007). First one needs to confirm the ordering of vignettes; then one can calculate $C$ (the DIF corrected variable) by using Anchors (Wand, King and Lau; 2007). Anchors assumes that the vignettes are entered into the formula in ascending order from left to right. Researchers can characterise the distribution of $C$ for the sample as a whole or for contrasts between subgroups. Unfortunately there is currently a problem installing Anchors into R; Jonathan Wand has been contacted about this problem and we have verified that we can download other statistical packages from the R-GUI menu.

Fortunately, the non-parametric approach can also be implemented manually, following the formula shown in Appendix C. We re-code the answers to the self-rated skills-matching question following this routine. The categorical breakdown of answers to the anchor-adjusted measure of skills-matching is shown in the table on the following page. Comparing this to the categorical breakdown of answers to the self-rated skills-matching question, we see that far less respondents are now classified as “closely matched” or “somewhat matched”. Instead, they are now classified as “not very matched”. This seems to suggest that some respondents are over-reporting the extent of their self-rated skills-match. This could this be due to the cognitive bias (the “rose-coloured glasses”) that we mentioned in section 2. It is not due to a Galatea (higher self-expectations) effect as this would require respondents being more closely matched after the anchor-adjustment. Interestingly, Ph.D. induction, Ph.D. feedback,

---

15Technical point: values of $C$ must be scalar (i.e. with magnitude but with no direction).
Ph.D. lab facilities, training on grant-writing, and training on teaching skills do not predict anchor-adjusted skills-matching.

<table>
<thead>
<tr>
<th>match</th>
<th>Freq.</th>
<th>Percent</th>
<th>Cum.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>15</td>
<td>41.67</td>
<td>41.67</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>5.56</td>
<td>47.22</td>
</tr>
<tr>
<td>4</td>
<td>18</td>
<td>50.00</td>
<td>97.22</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>2.78</td>
<td>100.00</td>
</tr>
<tr>
<td>Total</td>
<td>36</td>
<td>100.00</td>
<td></td>
</tr>
</tbody>
</table>

1 = very closely matched, 2 = closely matched, 3 = somewhat matched,
4 = not very matched, 5 = not at all matched
### Appendix A: HEA FDG Report 2007, p.44

#### 4.1.1 Sector of Employment by Level of Award for those Employed in the Republic of Ireland

Table 4.1 outlines the proportion of those who obtained employment in the Republic of Ireland for those who graduated in 2005. This table shows the sectors of employment of Level 8 graduates in Ireland.

The highlighted cell are the most common initial sectors of graduate employment for each level of award. Table 4.1

<table>
<thead>
<tr>
<th>Employment Sector</th>
<th>Level 8: Higher Degree University Graduate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Level 7: Ordinary Degree University Graduate</td>
</tr>
<tr>
<td></td>
<td>Level 6: Foundation Degree/University Certificate</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sector</th>
<th>Level 8</th>
<th>Level 7</th>
<th>Level 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agriculture, Forestry &amp; Fisheries</td>
<td>1.8%</td>
<td>1.9%</td>
<td>0.8%</td>
</tr>
<tr>
<td>Manufacturing Industries</td>
<td>14.4%</td>
<td>8.2%</td>
<td>12.7%</td>
</tr>
<tr>
<td>Food and Drink Industries</td>
<td>2.1%</td>
<td>3.3%</td>
<td>2.3%</td>
</tr>
<tr>
<td>Pharmaceutical &amp; Biomedical Products</td>
<td>2.0%</td>
<td>2.1%</td>
<td>0.9%</td>
</tr>
<tr>
<td>Materials, Mechanical and Electrical Engineering</td>
<td>3.0%</td>
<td>2.0%</td>
<td>1.2%</td>
</tr>
<tr>
<td>Computer and Electronic Office Equipment</td>
<td>3.1%</td>
<td>4.3%</td>
<td>1.3%</td>
</tr>
<tr>
<td>Other Industrial Activities</td>
<td>7.1%</td>
<td>7.9%</td>
<td>2.3%</td>
</tr>
<tr>
<td>Electricity, Gas and Water Supply</td>
<td>2.0%</td>
<td>0.1%</td>
<td>0.1%</td>
</tr>
<tr>
<td>Building &amp; Construction</td>
<td>8.3%</td>
<td>4.0%</td>
<td>1.2%</td>
</tr>
<tr>
<td>Distribution</td>
<td>11.4%</td>
<td>6.9%</td>
<td>2.4%</td>
</tr>
<tr>
<td>Business, Finance and Administration</td>
<td>11.8%</td>
<td>14.9%</td>
<td>29.6%</td>
</tr>
<tr>
<td>Architecture, Engineering and Technical Services</td>
<td>3.3%</td>
<td>1.7%</td>
<td>3.6%</td>
</tr>
<tr>
<td>Sales and Marketing and Sales Support</td>
<td>10.9%</td>
<td>9.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Other Business Services</td>
<td>2.0%</td>
<td>4.3%</td>
<td>3.3%</td>
</tr>
<tr>
<td>Civil and Environmental Services</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Health Services</td>
<td>3.3%</td>
<td>3.3%</td>
<td>3.3%</td>
</tr>
<tr>
<td>Researcher, Artists, Scholars</td>
<td>0.4%</td>
<td>0.4%</td>
<td>0.4%</td>
</tr>
<tr>
<td>Social and Community Services</td>
<td>1.1%</td>
<td>0.4%</td>
<td>0.4%</td>
</tr>
<tr>
<td>Professional and Vocational Services</td>
<td>1.0%</td>
<td>0.7%</td>
<td>0.4%</td>
</tr>
<tr>
<td>Others</td>
<td>1.0%</td>
<td>0.3%</td>
<td>0.3%</td>
</tr>
</tbody>
</table>

**Total Response Rate: 11.623 Response Rate: 46%**
### Appendix B: Pilot-Survey for Recent Ph.D. Graduates

#### 4. Shane has been working as a university researcher since he got a Ph.D. How closely is Shane's job related to his Ph.D. training?

<table>
<thead>
<tr>
<th>Rating</th>
<th>Very closely</th>
<th>Closely</th>
<th>Somewhat</th>
<th>Not very</th>
<th>Not at all</th>
<th>Rating Average</th>
<th>Response Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Please answer here:</td>
<td>48.4%</td>
<td>37.5%</td>
<td>12.5%</td>
<td>1.6%</td>
<td>0.0%</td>
<td>1.07</td>
<td>64</td>
</tr>
</tbody>
</table>

#### 5. Caitlin has been working as a secondary school teacher since she got a Ph.D. How closely is Caitlin's job related to her Ph.D. training?

<table>
<thead>
<tr>
<th>Rating</th>
<th>Very closely</th>
<th>Closely</th>
<th>Somewhat</th>
<th>Not very</th>
<th>Not at all</th>
<th>Rating Average</th>
<th>Response Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Please answer here:</td>
<td>0.0%</td>
<td>16.8%</td>
<td>29.7%</td>
<td>46.9%</td>
<td>4.7%</td>
<td>3.38</td>
<td>64</td>
</tr>
</tbody>
</table>

#### 6. Barry has been working in a shop in the services sector since he got a Ph.D. How closely is Barry's job related to his Ph.D. training?

<table>
<thead>
<tr>
<th>Rating</th>
<th>Very closely</th>
<th>Closely</th>
<th>Somewhat</th>
<th>Not very</th>
<th>Not at all</th>
<th>Rating Average</th>
<th>Response Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Please answer here:</td>
<td>0.0%</td>
<td>0.0%</td>
<td>5.0%</td>
<td>35.0%</td>
<td>60.0%</td>
<td>4.55</td>
<td>60</td>
</tr>
</tbody>
</table>
7. Clodagh has been working as a university researcher since she got a Ph.D. Does Clodagh need her Ph.D. qualification for her current position?

<table>
<thead>
<tr>
<th></th>
<th>Definitely yes</th>
<th>Probably yes</th>
<th>Somewhat</th>
<th>Probably not</th>
<th>Definitely not</th>
<th>Rating</th>
<th>Average</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Please answer here:</td>
<td>15.6%</td>
<td>57.8%</td>
<td>15.6%</td>
<td>7.8%</td>
<td>3.1%</td>
<td>2.25</td>
<td>64</td>
<td></td>
</tr>
</tbody>
</table>

8. Dermot has been working as a secondary school teacher since he got a Ph.D. Does Dermot need his Ph.D. qualification for his current position?

<table>
<thead>
<tr>
<th></th>
<th>Definitely yes</th>
<th>Probably yes</th>
<th>Somewhat</th>
<th>Probably not</th>
<th>Definitely not</th>
<th>Rating</th>
<th>Average</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Please answer here:</td>
<td>0.0%</td>
<td>7.8%</td>
<td>14.1%</td>
<td>39.1%</td>
<td>39.1%</td>
<td>4.09</td>
<td>64</td>
<td></td>
</tr>
</tbody>
</table>

9. Sheila has been working in a shop in the services sector since she got a Ph.D. Does Sheila need her Ph.D. qualification for her current position?

<table>
<thead>
<tr>
<th></th>
<th>Definitely yes</th>
<th>Probably yes</th>
<th>Somewhat</th>
<th>Probably not</th>
<th>Definitely not</th>
<th>Rating</th>
<th>Average</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Please answer here:</td>
<td>0.0%</td>
<td>0.0%</td>
<td>3.2%</td>
<td>27.0%</td>
<td>69.8%</td>
<td>4.67</td>
<td>63</td>
<td></td>
</tr>
</tbody>
</table>
Appendix C: Practical Implementation of Re-coding in the Non-parametric Approach

<table>
<thead>
<tr>
<th>Example</th>
<th>Survey Responses</th>
<th>1: ( y &lt; z_1 )</th>
<th>2: ( y = z_1 )</th>
<th>3: ( z_1 &lt; y &lt; z_2 )</th>
<th>4: ( y = z_2 )</th>
<th>5: ( y &gt; z_2 )</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( y &lt; z_1 &lt; z_2 )</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>( y = z_1 &lt; z_2 )</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>( z_1 &lt; y &lt; z_2 )</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>( z_1 &lt; y = z_2 )</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>( z_1 &lt; z_2 &lt; y )</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>( y &lt; z_1 = z_2 )</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>( y = z_1 = z_2 )</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>2,3,4</td>
</tr>
<tr>
<td>8</td>
<td>( z_1 = z_2 &lt; y )</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>9</td>
<td>( y &lt; z_2 &lt; z_1 )</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>( y = z_2 &lt; z_1 )</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1,2,3,4</td>
</tr>
<tr>
<td>11</td>
<td>( z_2 &lt; y &lt; z_1 )</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1,2,3,4,5</td>
</tr>
<tr>
<td>12</td>
<td>( z_2 &lt; y = z_1 )</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2,3,4,5</td>
</tr>
<tr>
<td>13</td>
<td>( z_2 &lt; z_1 &lt; y )</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>5</td>
</tr>
</tbody>
</table>

Table 1: All examples with two vignettes; this table gives calculations for the nonparametric estimator \( C \) for all possible examples (sans non-response) with two vignette responses, \( z_1 \) and \( z_2 \) (intended to be ordered as \( z_1 < z_2 \)), and a self-assessment, \( y \). From King and Wand (2007)
References


22


McIver Consulting (2004). A Model to Predict the Supply and Demand for Researchers and Research Personnel in Line with Ireland’s Strategy for Contributing to the European


