

Is this your Final(s) answer?*†

Florian Ederer
Nuffield College and FREVD's

Martin Oehmke
London School of Economics and FREVD's

First Draft: February 15, 2002

Revised: December 10, 2002

Abstract

This paper reports on the form in which Oxford Finals question are posed. It criticizes the apparent lack of originality and provides both theoretical and empirical estimates to support this finding. We propose a number of solutions which, if adopted, will substantially improve examination performance and unambiguously raise welfare.

1 Introduction

The University of Oxford invests a considerable amount of time and resources in testing the ability and knowledge of its students. Contrary to many other universities, examinations (generically referred to as 'Finals') are still held towards the very end of a student's university course and final marks are often exclusively based on the performance in these examinations. Undergraduate education in economics is offered for the courses Philosophy, Politics & Economics (PPE), Economics & Management (E&M) and Modern History & Economics (MH&E) and for all of these courses student performance is only assessed and formally evaluated on the basis of final

**JEL Classification* A22, B53, Z00

†We thank Matthias Osthoff, Alan Beggs, Andrew Chesher, Simon Cowan, James Forder, Andrew Glyn, David Hendry, Geoff Jones and Valérie Lechene for their wonderful quotes that provided the basis for this paper. Thanks are also due to David Myatt and Chris Wallace (1997) who inspired us to waste our time with writing this paper and to the Financial Markets Group at the LSE whose members provided useful comments. Special thanks go to the bar staff at FREVD's on Walton Street who provided copious amounts of brain cell fuel also known as 'Long Island Iced Tea' which was instrumental in sparking some of the key 'ideas' presented in this paper. Comments by anonymous referees were intentionally disregarded and the opinions of seminar participants at MIT, Harvard, Stanford, Princeton and Chicago were not even worth the paper they were written on.

examinations. As a result, Finals take a pre-eminent position in the life of every economics undergraduate.¹

Phrasing Finals questions in such a way that allows candidates to demonstrate understanding of the subject, analytical clarity and independence of judgement and thought plays a very important part in the design of each examination paper. At the same time, the questions should facilitate the marking process by clearly distinguishing between the different qualities of answers written by candidates. While the above objectives are generally met by the sophisticated examination designer, the exam usually fails to deliver on another no less important count: originality and creativity. Oxford Finals are - not unlike many other examinations - an incredibly tedious affair for both students and examiners.

In this paper we develop a theoretical model which we subsequently use to estimate the originality of current examinations. We show that these exhibit an apparent lack of originality and hence propose solutions that will considerably improve examination creativity. More importantly though, these solutions will also lead to a general improvement in students' examination results while markers will marvel at the inspired answers produced on the examination papers. Overall welfare unambiguously rises as both groups of agents are better off following the changes.

The rest of the paper substantiates these claims. In Section 2 we discuss our benchmark model used to estimate student performance. In Section 3, we present empirical results that help us evaluate the model's predictions and provide a rationale for the solutions we propose in Section 4. Section 5 concludes.

2 The Model

Assume that candidate j takes 8 examination papers. Let overall examination performance X be denoted by

$$\mathbf{X} = (x_1^j, \dots, x_8^j) \tag{1}$$

where x_i denotes performance on a particular paper i . Let the form in which a round of Finals examinations is set be denoted by

$$\mathbf{S} = (s_1, \dots, s_8) \tag{2}$$

where s_i is a variable for the form in which questions in exam i are posed. In this context S denotes the overall originality and creativity in a complete set of Finals exams.

Student performance is marked using a grading process. This grading process is modelled as a function $F(\cdot)$ which should ideally grade each student according to her understanding of the subject, denoted by η_i^j . As the

¹This is the tutor's viewpoint and not necessarily the student's.

level of understanding η_i^j in paper i for student j is not observable to the examiner, the grading process must be based on student performance x_i^j . This implies that the grading function $F(\cdot)$ must take the form

$$F(\cdot) = F(\mathbf{X}). \quad (3)$$

Conceptually, this grading function should result in an ordinal ranking that correctly reflects the ordering of η_i^j , the level of student understanding. This ordinal ranking should be invariant to monotonic transformations. Moreover, the grading function $F(\mathbf{X})$ should allow the examiner to cardinally distinguish student performance within the bounds of the marking interval $[0, 85)$. That is to say that in addition to ranking students in the correct order, the comparison of two grades assigned through the grading function should allow to infer by how much one student's performance was better than another's.

Student performance on a particular paper depends on his or her understanding of the subject, x_i^j , the creativity of the exam-setter s_i of paper i and a random shock ε_i . Hence,

$$x_i^j = \eta_i^j + s_i + \varepsilon_i \quad (4)$$

where $\varepsilon_i \sim N(0, 1)$. At present, originality in Finals can be approximated as being exactly or close to zero: often similar questions are repeated year after year, giving students little opportunity of creatively demonstrating real understanding of the material. This leads to unnecessarily low marks for students and unnecessarily boring scripts for examiners to mark, as the following analysis demonstrates. Assume that exam originality is determined by

$$s_i = a_i + b_i e_i \quad (5)$$

where a_i denotes a subject's intrinsic originality², b_i the sensitivity of originality (and hence of exam performance) to an increase in the originality-increasing effort e_i made by the examiner. Clearly, we have $b^i > 0$ for all i . Assuming that originality-increasing effort in exam-setting is costless, the above set-up implies that examiners should be original without bound when setting exams. Yet, even when there is a cost to thinking up original exam questions, the status quo must be suboptimal. This is shown using a quadratic cost term for originality to account for the examiners effort when setting the exam. Consider an examiner's utility given by

$$U = U(x_i) - c(e_i) \quad (6)$$

where $U(x_i)$ is the utility from reading the exam paper, while $c(e_i)$ is the (quadratic) cost term. Given the very low current level of originality in

²For example, for Accounting or Marketing this term would be negative.

exams, an increase in s_i will result in a first-order increase in examiner utility through its direct effect on x_i , while the cost of increasing e_i is only second-order. This implies a first-order welfare gain to examiners through an increase in the originality in exam setting.

Similarly, students will be unambiguously better off. Their utility

$$V = F(\mathbf{X}) + K(\mathbf{S}) \tag{7}$$

is entirely determined by exam results $F(\mathbf{X})$ and the intrinsic joy $K(\mathbf{S})$ that results from taking a set of exams with original questions, and therefore must rise.

This implies that a benevolent social planner maximizing social exam welfare W , the weighted sum of student and examiner utility, given an appropriate resource constraint

$$\begin{aligned} \max_E W &= U + \alpha V \\ \text{s.t. } R &= \bar{R} \end{aligned} \tag{8}$$

would choose $E = (e_1, \dots, e_8)$ with $e_i > 0$ for all i . Moreover, the amount of originality of each student in each exam paper will depend on b_i , the sensitivity of exam performance to more original exam questions. While it would be interesting to find exact values for e_i and hence s_i in each subject, this clearly exceeds the scope of this paper whose primary purpose is to demonstrate the sub-optimality of the status quo in exam setting.³

Nonetheless, we require an authoritative proof of our result.

Lemma 1 *The status quo in examination design for Finals is sub-optimal and can easily be improved.*

Lemma 2 *Lemma 1 and Lemma 3 are right.*

Lemma 3 *Lemma 2 is wrong.*

Dilemma 1 *The concurrent existence of Lemma 1, 2, and 3 creates an inescapable dilemma.*

As Dilemma 1 shows, the concurrent existence of Lemma 2 and 3 will generate a sine-curve in truth space to the effect that Lemma 1 is sometimes right and sometimes wrong. Hence, in order to establish our theoretical result we need to turn to empirical estimates. These are discussed in the next section.

³For a more definitive treatment see any paper by Jean-Michel Grandmont.

3 Empirical Estimates

It is reasonable to assume that ability for each subject η_i is normally distributed:

$$\eta_i \sim N(\mu, \sigma^2). \quad (9)$$

We calibrate our model using pre-specified numerical values for the mean and standard deviation of the ability distribution to take account of Oxford marking rules. For values $\mu = 65, \sigma^2 = 25$ there should be about 16% 1sts, 67% 2:1s, 13% 2:2s and 4% 3rds, Pass and Fails. Even though we have absolutely no clue what the hell quantile-driven identification of structural derivatives is, we use this approach pioneered by Chesher (2001) to estimate the parameters in the relationship postulated in Section 2.

We have to use this approach as ability is not only unobservable to the examiner but also to the student - essentially nobody really knows how good or bad they are. Moreover, examination results are not publicly available except for the rather crude classifications. Hence, both η_i and x_i are unknown to the econometrician. Nonetheless, this complete lack of data does not deter us from making broad sweeping statements as some manipulation (in the truest sense of the word) of the data would lead us to exactly the same conclusion: we obtain the empirical result which we wanted to see even before we started writing this paper.

4 Proposed Solutions

Given the abysmally low creativity estimates for current Economics examinations we propose a radical reform of the status quo. We suggest changes in the phrasing of questions for the two core papers and for a select number of option papers.

4.1 Macroeconomics

- Can the Diamond (1982) Coconut Model account for the fact that every year more than 150 agents are killed by falling coconuts? Does this change the nature of the long-run search equilibrium?
- “Late return fines in the economics library need never be repaid, just like national debt.” Discuss.
- Explain why Joseph Schumpeter should be regarded as a patron saint of endogenous growth theory (Robert Solow (1994)), i.e. paraded around once a year, but forgotten on all other days.
- “Der Euro ist ein Teuro.” Discuss.

- The sole contribution of Axel Leijonhufvud to macroeconomics is the addition of yet another unpronounceable name in every article's bibliography.
- "Macroeconomics is dead. It either does not rely on microeconomic principles or is nothing more than microeconomic aggregation." Should we bury Macroeconomics for good or is there life after death?
- "Real eyes would realise that real business cycles are real lies."

4.2 Microeconomics

- "Shirking is not a crime. The no-shirking condition in tutor-student models should therefore be neglected." Do you agree?
- Consider a tropical island Robinson Crusoe economy. Would 'Coconut-Pete' survive as the only producer/consumer in this economy or does he need his congenial partner 'Taxation-Jim' to give him the right incentives?
- "The profession's focus on giving microfoundations to macroeconomics is completely misguided. What we really need is macrofoundations for microeconomics." Discuss.

4.3 Econometrics

- Show that the Slutsky matrix is blue except for frozen peas. Give an intuitive interpretation.
- Outline the direct causal link between UK prices and days of rainfall in Outer Mongolia. Table A refers.
- "Econometrics - Alchemy or Science?" Show that unless econometrics helps us to find the Philosopher's Stone it does not really matter.
- Discuss the implications for world peace of a completely cointegrated world.
- "Distance from the equator is the ultimate engine of economic growth." Using your pocket calculator, run at least 20 million regressions to confirm the validity of this statement.

4.4 Economic Theory

- Is the Walrasian auctioneer French? If so, explain why there is never excess demand for baguette.

- Show that pulling at Park End is a mixed strategy equilibrium in a Bayesian game of incomplete information. Is it evolutionary stable?
- “Mantel, Sonnenschein, Debreu: k-dimensional anarchy.” Discuss.
- “Fixed-price models are dead. Real business cycle models are only in coma.” Adjudicate.
- “The Young/Peretto/Thompson-Dinopoulos/Aghion-Howitt model (Jones 1998) is far too simplistic. At least 20 other parameters such as the degree of buddhist faith and the frequency of alien landings need to be included.” Discuss.
- Show that a drunk student’s path after a Friday night at FREVD’s follows a Brownian motion. Hence show that while the student will find Hassan’s Kebab Van with probability 1, he will need to walk an infinite distance.
- “Representative Consumer - off to the graveyard!” Does he deserve a decent burial?

4.5 Public Economics

- “What is an externality? Hmmm, interesting ...” (James Forder 2001). Explain the meaning of the statement and show that Coase always meant what he said, without actually saying it.
- Should there be a tax on seminar questions?⁴
- TAXATION (one-word essay)

4.6 International Economics

- “Extensive swimming pool experiments have shown that neither notes nor coins float. A return to the Bretton Woods system is therefore the only viable option.” Do you agree?
- Demonstrate that the policy constraints placed on monetary and fiscal authorities by Mundell’s ‘Holy Trinity’ of fixed exchange rates, independent monetary policy, and capital mobility can be mitigated by devout Christian - preferably Catholic - faith.

4.7 Economics of OECD Countries

- “The miraculous economic performance of Ireland can be explained entirely by the Irish Football Team’s success in the World Cups of 1990, 1994 and 2002.” Table B (Goldman Sachs 2002) refers.

⁴See Ederer & Oehmke (2002).

4.8 Accounting

- “Accounting is sexy.” Do you agree?⁵
- “Accounting is *weird*.” No need to discuss.⁶

4.9 Finance

- “The CAPM implies slavery. It thus already ceased to hold after the civil war and not as late as the 1960s as Fama & French (1992) claim.” Discuss.
- With increased numbers of divorces, marriage has become more easily reversible. Discuss the implications of this development in the context of real options theory.
- Show that purchasing a single London underground ticket is more valuable than a travelcard as it offers the option to wait and see before making a follow-on investment (i.e. purchasing another single ticket).

5 Concluding Remarks

Originality is hard to achieve when the overruling objective is to accurately assess student performance in a period as short as 2 weeks. Yet, as we show in this paper, both students and examiners unnecessarily have to suffer from this tragic shortfall in creativity.

Adopting our proposed changes is likely to lead to a radical increase in humorous outbursts during examinations and will hence considerably improve student performance. At the same time examiners will be unable to believe their own eyes when they see the creative answers produced by candidates on the examination papers. Overall welfare is bound to rise. We do not address the concern that our improvements may lead to ‘grade inflation’ as witnessed by so many high-quality institutions at the other side of the Atlantic. However, the abundant existence of ‘conservative examiners’ in the Rogoff (1985) sense should eliminate all such concerns. Incentive contracts for examiners as advocated by Walsh (1995) are not necessary.

Finally, we find it difficult to suggest further areas of investigation as we seem to have addressed all the questions that one could possibly ask. Nonetheless it is clear that all research effort of the whole economics profession around the planet should from this day onwards be concentrated on this topic and that our paper should be cited happily ever after.

Economics at Oxford should be fun - our paper is a step towards it.

⁵The worst answer to this question will be honoured with the Enron/WorldCom Outstanding Performance Recognition Medal.

⁶The best answer to this question will be honoured with the Arthur Andersen Creative Accounting Prize.

References

- [1] Chesher, Andrew “Quantile-driven Identification of Structural Derivatives” *IFS & UCL Working Paper*.
- [2] Diamond, Peter “Aggregate Demand Management in Search Equilibrium.” *Journal of Political Economy* 90 (October 1982): 881-894.
- [3] Ederer, Florian and Martin Oehmke “Optimal Taxes on Seminar Questions.” *FREVD’s Working Paper Series* 1 (October 2002).
- [4] Fama, Eugene F. and Kenneth R. French “The Cross-Section of Expected Stock Returns.” *Journal of Finance* 47 (June 1992): 427-465.
- [5] Goldman Sachs “World Cup 2002 Preview.” *Capitalist Oppression Press* (April 2002).
- [6] Jones, Charles I. “Growth: With or Without Scale Effects?” *American Economic Review Papers and Proceedings* 89 (May 1999): 139-144.
- [7] Myatt, David P. and Chris Wallace “Rational Frenzies and Bicycle Crashes.” *Malroy Discussion Papers* 1 (September 1997): 1-2.
- [8] Rogoff, Kenneth S. “The Optimal Degree of Commitment to an Intermediate Monetary Target.” *Quarterly Journal of Economics* 100 (November 1985): 1169-1189.
- [9] Solow, Robert M. “Perspectives on growth theory.” *Journal of Economic Perspectives* 8 (Winter 1994): 45-54.
- [10] Walsh, Carl E. “Optimal Contracts for Central Bankers.” *American Economic Review* 85 (March 1995): 150-167.