Essays in Economic Behaviour and Interaction

by

Peter Slade

A Thesis
presented to
The University of Guelph

In partial fulfilment of requirements
for the degree of
Doctor of Philosophy
in
Food, Agricultural and Resource Economics

Guelph, Ontario, Canada

©Peter Slade, August, 2015
ABSTRACT

ESSAYS IN ECONOMIC BEHAVIOUR AND INTERACTION

Peter Slade  
University of Guelph, 2015

Advisor:  
Professor John Cranfield

This thesis contains three essays. The first essay explores the phenomena of countersignaling in a model with two different audiences: insiders and outsiders. Insiders have private information about the individual (or sender) and provide the individual with a reward if they believe the individual is of high ability. When the individual takes a public action that is associated with having low ability and is still rewarded by insiders, outsiders infer that insiders have favourable private information about the individual. Hence, taking an action commonly associated with low ability may signal to outsiders that the individual is more likely to have high ability by credibly revealing the private information of insiders.

The second essay proposes two maximum score estimators for two-sided one-to-one matching models with non-transferable utility: a weighted estimator and an unweighted estimator. The weighted estimator is consistent in the number of markets, based on an identification at infinity argument. Conversely, the unweighted estimator is consistent when preferences are homogeneous. The asymptotics of the unweighted estimator are in either the number of markets or the number of individuals.

The third essay examines the impact of gender on the hiring and promotion of academic economists. Past research has shown that females in North America generally experience inferior labour market outcomes, although it remains unclear the extent to which these outcomes are due to employer discrimination or gender differences in preferences. This essay seeks to answer this question through structural models that allow for the separate estimation of employer and employee preferences. The initial hiring of PhD economists is modelled as
a two-sided one-to-one matching market with non-transferable utility. Post-hiring outcomes (job exit and promotion) are estimated using a two-sided probit model. In the initial hiring process, departments exhibit no gender bias. However, significant male bias is found in post-hiring job outcomes; males are more likely to be promoted and less likely to be fired. Significant differences are also found in gender preferences, with males being more likely to exit their job voluntarily.
ACKNOWLEDGMENTS

Looking back at my four years as a PhD student I cannot help but be humbled by the unceasing support that I have received from so many people. My advisor, John Cranfield, invested more time, energy, and thought into my progress than any student has a right to expect. The members of my advisory committee, Jessica Cao, René Kirkegaard and Alan Ker, provided wonderful comments and conversation, throughout the research process. The external members of my examining committee, Henry Thille and James Vercammen, supplied several helpful suggestions that have, I hope, led to substantial improvements in this dissertation. My graduate research was funded through a Joseph-Armand Bombardier Canada Graduate Scholarship from the Social Sciences and Humanities Research Council.

The department of Food, Agricultural and Resource Economics has been an incredible intellectual home to me throughout my graduate education. My life and learning have been enriched by the department’s students, faculty, and staff; I am especially grateful to B. Deaton, A. Cairns, G. Hailu, C. Hurst, A. Lilavanichakul, K. Nadella, K. Selves, Z. Shang, and T. Tolhurst. GianCarlo Moschini generously hosted me for a semester at Iowa State University, during the writing of this dissertation.

My parents and sister provided constant encouragement throughout my twenty-two years of formal education. My partner in crime, Jessica Rosenberg, selflessly doubled as a sounding board, proofreader, and therapist. Finally, my daughter Alexandra has been a source of boundless joy (even when regressions don’t converge), endless laughter (even when propositions seem unprovable), and unconditional love (even when revisions appear unceasing).
# Contents

1 Introduction .................................................. 1

2 Crazy like a fox: countersignaling with multiple audiences  
2.1 Introduction .................................................. 3  
2.2 Model .......................................................... 9  
2.3 Solution to the model ......................................... 15  
2.4 Uncertainty ..................................................... 28  
2.5 Conclusion ...................................................... 32

3 Maximum score estimators for matching models with non-transferable utility  
3.1 Introduction .................................................... 35  
3.2 Brief literature review ......................................... 35  
3.3 Model ........................................................... 37  
3.4 Weighted maximum score estimator ............................ 38  
3.5 Unweighted estimator .......................................... 42  
3.6 Monte-Carlo evidence .......................................... 46  
3.7 Conclusion ...................................................... 47

4 A structural model of hiring, promotion, and job exit with application to  
gender and academic hiring ....................................... 49  
4.1 Introduction .................................................... 49  
4.2 The academic labour market .................................... 54  
4.2.1 The initial hiring market .................................... 54  
4.2.2 The post-hiring market ..................................... 55  
4.2.3 Gender and the academic labour market .................. 56  
4.3 Theoretical model .............................................. 57  
4.3.1 Initial hiring decision ....................................... 58  
4.3.2 Tenure, termination, and resignation ....................... 58  
4.4 Estimation ...................................................... 62  
4.4.1 Estimating initial job placements ............................ 62  
4.4.2 Estimating post-hiring outcomes ............................ 64  
4.5 Data ............................................................ 67  
4.6 Results ........................................................ 71  
4.6.1 Hiring ....................................................... 71  
4.6.2 Promotion .................................................... 79  
4.7 Conclusion ...................................................... 84

Bibliography ..................................................... 85
A Appendix to Chapter 2
  A.1 Proofs ......................................................... 93
  A.2 Comparative statics for section 2.3 .......................... 104

B Appendix to Chapter 3 .............................................. 107
  B.1 Identification of the unweighted estimator under independent preferences 107
  B.2 Example of non-identification of the unweighted estimator under heterogeneous preferences 110

C Appendix to Chapter 4 .............................................. 112
List of Figures

2.1 Distribution of evaluation signals and critical values ........................................ 18
2.2 Illustration of $\Gamma_i^k$ ................................................................................ 19
2.3 Illustration of potential equilibrium strategies .................................................... 21
2.4 Example and non-example of countersignaling equilibrium .............................. 24
2.5 Ratio of $\frac{g(\gamma|H)}{g(\gamma|H) + g(\gamma|L)}$ for countersignaling and non-countersignaling examples 26
2.6 Two examples of strong countersignaling equilibria ........................................... 27
2.7 Countersignaling equilibrium under uncertainty ............................................... 31
4.1 Plot of hypothesized movement of $\alpha(z_{c,d_k}, t - t_{c_i}^*)$ and $\omega(z_{c,d_k}, t - t_{c_i}^*)$ .... 60
4.2 Hazard plots for job mobility ............................................................................... 72
B.1 Illustration of equation B.7 .................................................................................. 111
## List of Tables

2.1 Distribution of types .................................................. 10
3.1 Monte-Carlo results .................................................. 48
4.1 Descriptive statistics .................................................. 70
4.2 Results from the estimation using hiring data .................... 73
4.3 Differences in predictive power by gender ......................... 78
4.4 Results from the estimation of post-hiring outcomes ............ 80
4.5 Effect of a change in regressors on the probability of outcomes in model 1 (%) 81
B.1 Monte-Carlo results for independent preferences ................ 109
Chapter 1

Introduction

This dissertation is composed of three essays in the area of economic behaviour and interaction. This introduction provides the motivation for, and a cursory overview of, these essays. A more fulsome introduction can be found within the essays themselves.

The first essay interests itself in the phenomena of countersignaling. Countersignaling implies that an individual can, in some circumstances, benefit from taking an action that is commonly associated with being a low type. My own interest in countersignalling was piqued when I observed the way academics dress at conferences – it seemed that those of greater status dressed according to their intrinsic preferences for casual attire while those of lower status “dressed to impress”. As I discuss further in the essay, other examples of countersignaling can be found in workplaces, product markets, and social circles.

I offer a rationale for countersignaling that relies on multiple audiences: insiders and outsiders. Insiders have private information about the individual and provide the individual with a reward if they believe the individual is of high ability. When the individual takes a public action that is associated with having low ability and is still rewarded by insiders, outsiders infer that insiders have favourable private information about the individual. The revelation of this private information may make outsiders think the individual is actually more likely have high ability when he takes an action associated with having low ability (conditional on being rewarded by insiders). The model rationalizes the pattern of dress at conferences: academics with high status are already thought of quite highly by insiders and are therefore able to signal their status to outsiders by dressing down. Conversely, academics with low status must still signal to insiders and therefore must dress to impress.

The second essay proposes maximum score estimators for two-sided one-to-one match-
ing models with non-transferable utility. The estimators are intended for markets in which two disjoint sets of players match to a single player in the opposing set; such markets would include (heterosexual) marriage markets, labour markets and various social “markets”. Previous work on matching has largely focused on matching with transferable utility. Transferable utility implies that an individual can transfer their utility to their match. This assumption is valid in markets with profit-maximizing agents, as profits are easily transferred through cash transactions. The assumption is less tenable when agents are utility-maximizing – as any participant in the marriage market would know, the utility an individual receives from being in a relationship cannot easily be transferred to their (potential) mate.

Two maximum score estimators are proposed: a weighted estimator and an unweighted estimator. The weighted estimator is consistent in the number of markets, based on an identification at infinity argument. Conversely, the unweighted estimator is consistent when preferences are homogeneous. The asymptotics of the unweighted estimator are in either the number of markets or the number of individuals.

The third essay examines the role of gender in the hiring and promotion of PhD economists. The issue of gender in the academic labour market has been much studied over the past two decades, with most researchers finding significant gender differences in outcomes. However, it remains unclear the extent to which these outcomes are a function of employer discrimination or differences in gender preferences.

To separate the preferences of employers and employees in the labour market for PhD economists I employ structural econometric models. The initial hiring of PhD economists is modelled as a two-sided one-to-one matching market with non-transferable utility. Post-hiring outcomes (job exit and promotion) are estimated using a two-sided probit model. In the initial hiring process departments exhibit no gender bias. However, significant male bias is found in post-hiring job outcomes; males are more likely to be promoted and less likely to be terminated. Significant differences are also found in preferences across genders, with males being more likely to exit their job voluntarily.
Chapter 2

Crazy like a fox: countersignaling with multiple audiences

2.1 Introduction

A first-year undergraduate student sits down for her first class in macroeconomics. A few moments later a besuited tenured professor enters the room and proceeds to lecture in an elegant, if somewhat dry, manner. The student then hurries to her next class in microeconomics. A few moments later a sloven, crazy-haired tenured professor enters the room. While this professor’s lecture is just as informative as the first, his mannerisms and gesticulations border on insanity. The student learns from her peers that these two professors have consistently acted this way since they were hired.

What is the student to believe about the relative abilities of the two professors? The first professor conformed to the norms of the professorship, whereas the second did not. Indeed, if the student had met the second professor on the street, she would have thought him to be unintelligent in absolute terms, let alone relative to other academics. After some thought, the student decides the second professor is likely to be the better at his job. She reasons that if the second professor had always acted like a madman then his superiors would have been less likely to offer him tenure, all other things being equal. Thus, professors who act like madmen would need to be even better at their job in order to achieve tenure.

In the preceding example the seemingly crazy professor benefits from countersignaling – behavior that otherwise would have denoted him as having low ability, when combined with his status as a tenured professor, made the student think he was more likely to have
high ability. To explain this phenomenon more generally, imagine an individual whose utility depends on rewards given by insiders and his reputation among outsiders. In the preceding example the individual in question is the professor, insiders are the committee that offered him tenure, and outsiders are the students (one might also think of the professor’s colleagues, both inside and outside his institution, as the outsiders).

The individual has either high or low ability. Insiders reward the individual if they believe the probability the individual has high ability is over some critical threshold. In the example, tenure is the reward bestowed by insiders.

Insiders receive two pieces of information about the individual: a noisy evaluation signal and an observation of the individual’s action. A high ability individual is more likely to produce a higher draw of the evaluation signal, which is observed by both insiders and the individual. In the example, the evaluation signal is the professor’s publication record and teaching evaluations to date.

After observing his evaluation signal, the individual takes a binary action, either A or B. Some individuals intrinsically prefer action A, others intrinsically prefer action B. High ability individuals are more likely to intrinsically prefer action A.\(^1\) Hence, if individuals acted only on the basis of their intrinsic preference then playing action A would signal that the individual is more likely to have high ability. The individual incurs a psychic cost if he acts against his preference. In the example, action A is being well mannered and groomed.\(^2\) Neither action is intrinsically preferred by insiders or outsiders – the two audiences only care about the individual’s action insofar as it signals his ability.

After observing the individual’s action and evaluation signal, insiders decide whether or not to reward the individual. Outsiders do not observe the evaluation signal. Instead they form a belief about the individual’s ability based on the action of the individual and the reward given by insiders.

---

\(^1\)To maintain some relationship with typical job-signaling models one could think of action A as undertaking training or education.

\(^2\)Some evidence that grooming is associated with an individual having high ability is provided by French et al. (2009), who show that well groomed high school students had significantly higher grade point averages.
To gain an intuition for a countersignaling equilibrium consider the minimum evaluation signal that insiders would require in order to reward the individual, conditional the individual’s action. When the individual plays action $B$ insiders might reasonably infer that the individual intrinsically prefers playing action $B$. Given that individuals who intrinsically prefer action $B$ are less likely to be of high ability than those who intrinsically prefer action $A$, insiders will require a higher draw of the evaluation signal in order to reward the individual when they observe action $B$, as opposed to action $A$.

Recall that outsiders observe the action of the individual and the reward given by insiders, but not the evaluation signal. By playing $B$ and being rewarded by insiders, the individual is able to truthfully reveal to outsiders that he has a relatively high evaluation signal. Even though outsiders know that the individual is more likely to have low ability if he prefers action $B$, the fact that the individual can only play action $B$ if he has a high evaluation signal may be enough to make outsiders believe that the individual is more likely to be of high ability when he takes action $B$ (conditional on being rewarded by insiders).$^3$

Furthermore, if the weight the individual places on his reputation among outsiders is strong enough, the individual may find it advantageous to play $B$ even if he prefers to take action $A$. That is to say the individual may choose to act “crazy” in order to demonstrate that he can get away with it, even if his intrinsic preference is to act sanely.

This type of countersignaling can be found in a myriad of situations. The most immediate examples are labour markets in which employees signal to both their direct superiors (insiders) and their colleagues inside and outside the organization (outsiders). One can also invert the power dynamic and think of examples where leaders demonstrate their power to outsiders by acting against the established the norms of leadership (assuming leadership norms reflect the typical preference of high ability leaders). In this example subordinates are insiders who have a private evaluation signal related to the true ability of the leader. If the evaluation signal is low, the leader will have to adhere to the established norms of leadership

---

$^3$The intuition for countersignaling is akin to high prices signaling quality in a product market with informed and uninformed consumers (Bagwell and Riordan, 1991).
in order to maintain the respect of subordinates. Conversely, a leader who receives a high draw of the evaluation signal is able to act against established norms without losing the respect of subordinates. From the outsiders’ perspective, a leader who violates norms of leadership but can still command the loyalty of subordinates, may be more likely to be an intrinsically good, or high ability, leader.

In a different context one could think of the individual as a graduate student, insiders as the student’s advisory committee, and outsiders as potential employers. The student’s advisors, through their interaction with the student, receive a private signal of the student’s abilities. To be reductive, one can view academic advisors as offering a binary reward to the student: recommend or not. A student who publicly acts in a manner that is typically associated with low ability students, and still receives a recommendation, may be thought of by potential employers as more likely to have high ability.

The model can also be applied to a social circle with an alpha person. Suppose the alpha receives a draw of the individual’s true sociability and then determines whether or not the individual remains in the social circle. The individual derives utility from being in the social circle and from others’ perception of his sociability. If the individual takes a public action that makes him appear less sociable and can still remain in the circle, then others may infer that the alpha’s private information regarding the individual’s sociability is particularly favourable.

In addition to having multiple audiences, the model presented in this article differs from typical job market signaling models in three important respects. First, the individual receives a binary reward from insiders. In a labour market where insiders are the individual’s employer, this implies that the individual is either hired at a fixed salary or not hired at all. In contrast, typical job-market signaling models assume that the individual is paid his expected productivity. While standard microeconomic theory predicts that wages should be set equal to marginal product, empirical evidence reveals that wages are often based on job category and tenure as opposed to a worker’s productivity (Baker, Jensen, and Murphy,
1988; Medoff and Abraham, 1980). In other applications rewards are similarly binary: a restaurant is patronized or not, a cult leader is followed or not, two people will marry or not.

Second, insiders acquire private information about the individual. In a labour market this might imply that an employee is hired for a probationary period. During this probationary period the employer is able to acquire the private information that constitutes the evaluation signal.

Third, the individual’s action is binary. In contrast, most signaling models allow the signal to be drawn from a continuous set – canonically this continuous variable is investment in education (Spence, 1973).

The idea of countersignaling has been raised in several different fields. Previous economic research has explained countersignaling through “middle-status conformity”. Feltovich, Harbaugh, and To (2002) identify a countersignaling equilibrium within a typical job market signaling model where the individual has one of three types: low, medium, and high. Employers observe both a noisy signal of the individual’s ability and the individual’s investment in education (the indifference curves of low, medium and high types have the familiar single crossing property in education-wage space.) In a countersignaling equilibrium the individual chooses high education if he has medium ability, which serves to separate him from low ability individuals. Conversely, if the individual has high ability he is sufficiently separated from low ability individuals through the noisy productivity signal, and he can therefore make a smaller investment in education, which serves to separate him from medium ability individuals. Phillips and Zuckerman (2001) discuss a similar sociological model. The motivation for countersignaling presented herein is quite different and does not rely on three types of workers – instead it depends on multiple audiences with asymmetric information.

In evolutionary biology, Zahavi (1975) suggests that animals may have an incentive to choose a mate who possesses, what would appear to be, a handicap. The argument being that if an animal has survived in spite of their handicap, they are likely to be fitter than an
animal without the handicap. For example, male birds are often brightly coloured, making them easier prey – surviving in spite of this coloration implies the bird has other assets that are valuable to his survival. These other assets may be attractive to the opposite sex, making the handicap advantageous in mating. This biological environment is similar to the model of this article with nature playing the role of insiders and the opposite sex playing the role of outsiders.

The notion of non-conformance among high ability individuals has also been broached in the consumer psychology literature. Galinsky et al. (2008) suggest that members of high-status groups may be better able to afford the consequences of breaking social norms. Bellezza, Gino, and Keinan (2014) perform a set of experiments that show non-conformance to norms does in fact increase status among some groups. Notably, these articles posit a linear relationship between non-conformity and status, which is consistent with a model of signaling with multiple audiences, though inconsistent with the hypothesis of “middle-status conformity”.

In the organizational psychology literature it has been noted that individuals with higher status have more latitude to deviate from workplace norms. Hollander (1958) theorized that over time, and through positive performance, employees are able to amass “idiosyncratic credits”. An employee with these credits is able to act idiosyncratically without facing censure from colleagues or superiors, whereas an employee lacking in credits would be punished for similar behavior. This is consistent with insiders allowing the individual to play action $B$ if his evaluation signal is high enough. To take this theory a step further, it seems logical that an individual with idiosyncratic credits might derive utility from demonstrating to others that he possesses such credits.

This article also relates to the literature on signaling with two audiences. Previous models of signaling with two audiences are premised on each audience preferring the agent to be of a different type (Bar-Isaac and Deb, 2014; Gertner, Gibbons, and Scharfstein, 1988). For example, Frenkel (2015) examines the problem of a credit agency who would like
to maintain a reputation among credit issuers for being a lenient reviewer, while maintaining a public reputation for being a stringent reviewer. To my knowledge there is no existing literature on signaling with two audiences where both audiences share the same preference over the sender’s type. Perhaps this is because such a model may sound trivial – if both audiences have the same preference why would the individual act differently than if there was only one audience? In this article it is shown that differences in information can cause the two audiences to interpret the individual’s action in quite different ways.

Finally, this article contributes to the literature on social norms. In the solution to the model it is shown that the action that is preferred by high ability individuals (action $A$) can, endogenously, become a social norm. Under this equilibrium, individuals who prefer taking action $B$ have to play $A$ in order to appear as though they had an intrinsic preference for $A$. Similar to other models of social norms (Akerlof, 1980; Bernheim, 1994), the individual has a preference over his action and incurs a psychic cost when he acts against this preference. Unlike previous literature, norms are not upheld because the audience has an intrinsic preference for individuals who take a particular action (as is assumed by Akerlof (1980)) or for individuals who have a particular preference (as is assumed by Bernheim (1994)). Instead, norms are upheld because those who prefer conformance are more likely to have high ability. One could think of norms in this model as being upheld by statistical, as opposed to preference, discrimination.

### 2.2 Model

The model has three players: the individual, the employer (or insiders), and outsiders - to maintain consistency with the signaling literature the rest of the article uses the terminology of a labour market. The timing of the game is as follows: the individual is hired at random by the employer who has no knowledge of the individual’s ability. After the individual is hired, the individual and the employer both observe the same noisy evaluation signal, at
which point the individual takes either action $A$ or action $B$. The employer then chooses to either keep on or fire the individual on the basis of the individual’s evaluation signal and action. Finally, outsiders form a belief about the individual’s ability based on the actions of the individual and the employer.

**Actions and information**

*Individual.* The individual has a two dimensional type, the first dimension captures the individual’s ability which is high ($H$) or low ($L$) with equal probability. The second dimension captures the individual’s preference which is either $a$ or $b$. The two dimensions of the individual’s type are correlated; when the individual has high ability the probability that he has preference $a$ is $q_H$, when the individual has low ability the probability that he has preference $a$ is $q_L$, where $q_H > q_L$. Table 2.1 shows the distribution of types.

It is assumed that the individual knows his preference but not his ability. Assuming the individual knows his ability would not change the results in the base model, as the individual’s action is not conditioned on his ability. When uncertainty is introduced in section 4 this assumption becomes critical.

After the individual is hired he receives an evaluation signal, $\gamma$, drawn from $g(\gamma|i)$, $i \in \{L, H\}$.

**Assumption EP.** The ratio $g(\gamma|H)/g(\gamma|L)$ is continuous over support $[0, \infty]$. 

### Table 2.1: Distribution of types

<table>
<thead>
<tr>
<th>Ability</th>
<th>Preference</th>
<th>$a$</th>
<th>$b$</th>
<th>$1/2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H$</td>
<td>$\frac{1}{2}q_H$</td>
<td>$\frac{1}{2}(1 - q_H)$</td>
<td>$1/2$</td>
<td></td>
</tr>
<tr>
<td>$L$</td>
<td>$\frac{1}{2}q_L$</td>
<td>$\frac{1}{2}(1 - q_L)$</td>
<td>$1/2$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\frac{1}{2}(q_H + q_L)$</td>
<td>$\frac{1}{2}(1 - q_H - q_L)$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Assumption MLRP.** The ratio \( g(\gamma|H)/g(\gamma|L) \) is non-decreasing in \( \gamma \) and is strictly increasing in \( \gamma \) over the interval in which both \( g(\gamma|L) \) and \( g(\gamma|H) \) have positive support.

By assumptions EP and MLRP the individual will only receive the highest draws of \( \gamma \) if he has high ability and will only receive the lowest draws of \( \gamma \) if he has low ability. Assumption EP also ensures that \( g(\gamma|L) \) converges to zero at the supremum of its support and \( g(\gamma|H) \) converges to zero at the infimum of its support. By assumption MLRP the individual is more likely to receive a higher draw of \( \gamma \) if he has high ability.

After being hired the individual receives his evaluation signal and chooses an action \( J \in \{A, B\} \). The individual intrinsically prefers action \( A \) (\( B \)) if he has preference \( a \) (\( b \)). If the individual acts against his intrinsic preference he incurs a psychic cost of \( \delta \).

**Employer.** The employer hires the individual at random with no knowledge of the individual’s type. After observing the individual’s evaluation signal (\( \gamma \)) and the individual’s action (\( J \)), the employer takes the action \( \rho \in \{F, K\} \), where \( F \) denotes the decision to fire the employee and \( K \) denotes the decision to keep on the employee. If the individual is fired the game is over.

**Outsiders.** If the individual is kept on outsiders observe the individual’s action and form a belief regarding the individual’s ability. The beliefs of outsiders are referred to as the individual’s reputation. The individual cares about his reputation among outsiders – one could think of outsiders as the individual’s professional community which might include future collaborators or employers.

In forming their belief about the individual’s ability, outsiders observe the full structure of the game, the action of the individual, and the reward given by insiders, but not the evaluation signal. The probability the individual is of the \( i \)th ability, from the perspective of outsiders, is defined as \( P(i|K, J) \) where \( J \) denotes the individual’s action and \( K \) denotes that the individual was kept on by the employer. The existence of a countersignaling equilibrium hinges on the relative values of \( P(H|K, A) \) and \( P(H|K, B) \). In a countersignaling equilibrium \( P(H|K, B) > P(H|K, A) \); outsiders hold the individual in higher esteem if he plays \( B \).
It is assumed that outsiders do not observe the individual if he is fired. In certain contexts this assumption is quite reasonable; if an employee is fired he will not have the chance to interact with outsiders at an industry conference, if a restaurant is not patronized by informed consumers it will not stay open to serve uninformed consumers. In other contexts this assumption may be less tenable; if a schoolchild is purged from their clique (insiders), they will still be observed by other classmates (outsiders).

**Payoffs and strategies**

*Employer.* If the employer keeps on the individual the employer receives a payoff of \( \pi_i = p_i - m \), where \( m \) is the wage and \( i \) is the individual’s ability. Allow that \( \pi_H > 0 > \pi_L \); the employer derives positive profit from a high ability individual and negative profit from a low ability individual. The employer receives a payoff of zero if they fire the individual. The employer’s strategy maps from the individual’s evaluation signal and action (denoted by the double \((\gamma, J))\) to a probability\(^4\) over taking action \( K \).

*Individual.* The individual receives utility,

\[
   u(J, j, \gamma, \sigma) = \begin{cases} 
   m + \beta P(H|K, J, \sigma) - \mathbb{1}\{j \neq J\}\delta & \text{if kept on} \\
   0 & \text{if fired}
   \end{cases}
\]

where \( J \) is the individual’s action, \( j \) is the individual’s intrinsic preference, and \( \beta \) is a parameter capturing the importance the individual places on his reputation among outsiders. The term \( \sigma \) represents the strategy played by the individual. The linear specification of reputational rewards, \( \beta P(H|K, J, \sigma) \), could be replaced with an increasing function of \( P(H|K, J, \sigma) \) without compromising the main qualitative results. If the individual is hired, his utility is composed of the wage received from the employer, the reputational reward from outsiders, and a psychic loss of \( \delta \) if his action is not consistent with his intrinsic preference.

\(^4\)It shall be shown that this probability is degenerate – the employer always plays a pure strategy.
It is implicitly assumed that the individual bears no cost from acting against his preference if he is fired. The timing of the game is such that the employer takes their action immediately after observing the individual’s action – if the individual chooses to act against his preference and is fired, he will have spent no time actually acting against his preference. In some situations this may be a reasonable assumption, in other situations there may be a significant lag between the employer’s observation of the individual’s action and the employer’s own action.

The individual’s strategy is a mapping from his intrinsic preference and evaluation signal, captured in the double \((j, \gamma)\), to a probability over taking action \(A\). It is assumed that the individual will play \(A\) with the same probability for all doubles that yield identical payoffs.\(^5\) Note that by equation 2.1 the individual’s payoff is not directly affected by his ability, hence his action would not be conditioned on his true ability even if it were known.

Outsiders. As mentioned previously, outsiders have no strategic role in the game – they form a belief about the individual’s ability based on the actions of the individual and the employer.

Solution concept and off-equilibrium beliefs

The solution concept is a perfect Bayesian equilibrium. It will be possible to have equilibria in which the individual plays a certain action with zero probability, therefore off-equilibrium beliefs must be defined for both the employer and outsiders. Off-equilibrium beliefs are in the sprit of the divinity criterion (Banks and Sobel, 1987). According to the divinity criterion when an audience is presented with an off-equilibrium action they ask themselves “who is most likely to benefit from taking such an action, regardless of my response to this action?”

Off-equilibrium beliefs are required for the employer if the individual plays \(A\) or \(B\) with probability one conditional on a particular draw of \(\gamma\). Suppose in equilibrium the

\(^5\)For example, if the individual receives the same relative payoffs from the actions \(A\) and \(B\), conditional on receiving either \((a, \gamma')\) or \((a, \gamma'')\), then he will play \(A\) with the same probability conditional on either of these doubles. Without this assumption, in a mixed strategy equilibrium the individual could potentially play \(A\) with varying probabilities depending on his draw of \(\gamma\), even if his payoffs did not differ with \(\gamma\).
individual plays A with probability one conditional on receiving the evaluation signal $\gamma^*$. In this situation the employer must form a belief regarding the probability that the individual is of high ability conditional on receiving $\gamma^*$ and playing B. The individual has the most to gain from playing B when his preference is $b$. Therefore, by the divinity criterion, insiders believe that the probability the individual has high ability conditional on receiving $\gamma^*$ and playing $B$, is the same as the probability the individual has high ability conditional on receiving $\gamma^*$ and having preference $b$. The probability density of the individual receiving $\gamma^*$, having preference $b$, and being of ability $i$, is $g(\gamma^*|i)(1 - q_i)$. Hence the probability the individual has high ability conditional on receiving $\gamma^*$ and preference $b$ is,

$$ \frac{g(\gamma^*|H)(1 - q_H)}{g(\gamma^*|L)(1 - q_L) + g(\gamma^*|H)(1 - q_H)}. $$ (2.2)

By the same logic if the individual plays B with probability one conditional on receiving $\gamma^*$, then insiders will believe that the probability the individual has high ability conditional on receiving $\gamma^*$ and playing $A$, is the same as the probability the individual has high ability conditional on receiving $\gamma^*$ and having preference $a$. This probability is,

$$ \frac{g(\gamma^*|H)q_H}{g(\gamma^*|L)q_L + g(\gamma^*|H)q_H}. $$ (2.3)

Given that outsiders do not observe $\gamma$, off-equilibrium beliefs for outsiders are only needed when the individual plays $B$ or $A$ with probability one (conditional on being kept on) regardless of $\gamma$. If in equilibrium the individual plays $A$ with probability one (regardless of $\gamma$), the individual will have the strongest incentive to play $B$ when he has preference $b$ and will not be fired for playing $B$ (assumption EP ensures that there is draw of $\gamma$ for which the individual will not be fired even if he plays $B$.)

**Definition of countersignaling equilibria**

A countersignaling equilibrium can be either *weak* or *strong.*
Definition 1. A weak countersignaling equilibria exists when \( P(H|K, B, \sigma) > P(H|K, A, \sigma) \), implying the individual is more likely to have high ability when he plays \( B \).

Definition 2. A strong countersignaling equilibria exists when \( \beta \left( P(H|K, B, \sigma) - P(H|K, A, \sigma) \right) > \delta \), implying that the individual will play \( B \) if he is not fired for doing so, regardless of his preference.

The definition of a strong countersignaling equilibrium comes from setting \( u(B, a, \gamma, \sigma) > u(A, a, \gamma, \sigma) \), conditional on the individual being kept on. Returning to the motivating example, in a strong countersignaling equilibrium a professor who prefers to act sanely will also act mad (if he is not fired for doing so) as the reputational benefits of acting like a madman outweigh the psychic cost of acting against his intrinsic preference. Conversely, in a weak countersignaling equilibrium a professor only acts mad if that is his intrinsic preference.

2.3 Solution to the model

The main aim of this section is to deliver a qualitative description of the equilibria and define sufficient conditions for a countersignaling equilibrium to hold. To solve the model, first define two critical thresholds of the evaluation signal: \( \gamma_A \) and \( \gamma_B \). These thresholds will be critical to the optimal action of the employer. Note that these critical values depend only on the distribution of \( \gamma \) and the probabilities \( q_H \) and \( q_L \).

Lemma 2.1. In any equilibrium there exists two values of \( \gamma \): \( \gamma_A \) and \( \gamma_B \) (where \( \gamma_A < \gamma_B \)). When the individual receives a draw of \( \gamma < \gamma_A \) he will be fired regardless of the action he takes, when the individual receives a draw of \( \gamma > \gamma_B \) he will be kept on regardless of the action he takes.

Proof. Suppose the individual plays \( A \) with the same probability conditional on receiving \( \gamma^* \), regardless of his preference. From the perspective of the employer the probability the
individual is of the $i$th ability conditional on receiving $\gamma^*$ is $\frac{g(\gamma^*|i)}{g(\gamma^*|H) + g(\gamma^*|L)}$, and the expected profit from the individual is,

$$\frac{g(\gamma^*|H)}{g(\gamma^*|H) + g(\gamma^*|L)} \pi_H + \frac{g(\gamma^*|L)}{g(\gamma^*|H) + g(\gamma^*|L)} \pi_L.$$  

(2.4)

Define $\gamma_A$ as the value of $\gamma$ that sets the employer’s payoff equal to zero. Hence $\gamma_A$ is implicitly given as the solution to,

$$\frac{g(\gamma_A|H)}{g(\gamma_A|H)} = \frac{-\pi_L}{\pi_H}. \quad (2.5)$$

Now consider the individual’s strategy and payoff when he receives the evaluation signal $\gamma^* < \gamma_A$. First, suppose the individual plays action $A$ with the same probability regardless of his preference. By assumption MLRP the employer’s expected profit from the individual would be negative and hence the individual would be fired. Now suppose that the individual did not play action $A$ with the same probability regardless of his preference. If the employer kept on the individual conditional on the individual playing a certain action, then the individual would always play the action that resulted in his being kept on (which, again, would result in the individual being fired.) Hence in equilibrium the individual will always be fired if he receives a draw of $\gamma < \gamma_A$.

The term $\gamma_B$ represents the value of $\gamma$ at which the individual would be kept on if the employer believed he had preference $b$ with probability one. Following the same logic as equation 2.5, $\gamma_B$ is the solution to,

$$\frac{g(\gamma_B|H)(1 - q_H)}{g(\gamma_B|L)(1 - q_L)} = \frac{-\pi_L}{\pi_H}, \quad (2.6)$$

where $g(\gamma|i)(1 - q_i)$ is the likelihood the individual has a given draw of $\gamma$ and preference $b$, conditional on having the $i$th ability. Assumption MLRP ensures that the individual would be hired conditional on receiving an evaluation signal greater than $\gamma_B$ regardless of
the employer’s belief about the individual’s intrinsic preference. Assumption EP ensures that $\gamma_B$ exists and is strictly less than $\text{sup}(g(\gamma|L))$.

As a corollary to lemma 2.1, the individual must be kept on if he receives a draw of $\gamma > \gamma_A$. By equation 2.5 and assumption MLRP, when the individual has a draw of $\gamma > \gamma_A$, he must be kept on conditional on playing at least one of the actions. Therefore, in equilibrium the individual must always be kept on – if he were not then he would deviate to the action that resulted in him being kept on.

One should also note that if the individual always plays action $A$ when he has preference $a$, then the individual would be fired for playing $B$ if he has a draw of $\gamma$ below $\gamma_B$. When the individual plays $A$ with probability one conditional on having preference $a$, then playing $B$ denotes that the individual has preference $b$ – by the definition of $\gamma_B$ the individual would be fired if he had preference $b$ and a draw of $\gamma < \gamma_B$. This is not to say that the individual can never play $B$ in equilibrium if he has a draw of $\gamma < \gamma_B$, only that if he does so he must play $B$ with positive probability when he has preference $a$.

The values $\gamma_A$ and $\gamma_B$ are illustrated in figure 2.1. The individual’s draw of $\gamma$ falls into one of three sets: $(0, \gamma_A)$, $[\gamma_A, \gamma_B)$, or $[\gamma_B, 1)$ – define these three sets as $\Gamma^1$, $\Gamma^2$, and $\Gamma^3$, respectively. Given that the individual will be fired if he has a draw of $\gamma < \gamma_A$, the actions of the individual conditional on receiving $\gamma \in \Gamma^1$ will not be considered further.

Denote $\Gamma^k_i$ as the probability the individual is in the $k$th set conditional on having the $i$th ability. For example, $\Gamma^2_H$ is the probability the individual’s draw of $\gamma$ is in $[\gamma_A, \gamma_B)$ conditional on the individual being of high ability,

$$\Gamma^2_H = \int_{\gamma_A}^{\gamma_B} g(\gamma|H)d\gamma. \quad (2.7)$$

These probabilities are illustrated in figure 2.2.

To understand the importance of these probabilities to a weak countersignaling equilibrium (and to foreshadow the solution to the model) suppose that the individual plays $A$
Plot of the values $\gamma_A$ and $\gamma_B$ using the parameters $\pi_H = -\pi_L = 1$, $q_H = .6$, and $q_L = .4$. The distributions of $g(\gamma|H)$ and $g(\gamma|L)$ are beta distributions with $\alpha = \beta = 1.5$. The distribution of $g(\gamma|H)$ is shifted to the right by 0.05, and the distribution of $g(\gamma|L)$ is shifted to the left by 0.05. The solid lines represent the distribution of $\gamma$ conditional on the individual’s ability. As per equation 2.5, $\gamma_A$ is the solution to $g(\gamma_A|H)/g(\gamma_A|L) = -\pi_L/\pi_H = 1$. The dotted lines represent the distribution of $\gamma$ conditional on the individual’s ability and preference $b$. The term $\gamma_B$ is found via equation 2.6.

when he has preference $a$ regardless of $\gamma$. In this case when the individual plays $B$ (and is kept on) he is telling outsiders that his draw of $\gamma$ lies in $\Gamma^3$ (he would be fired if he played $B$ and had a draw of $\gamma \notin \Gamma^3$). Conversely, if he plays $A$ (and is kept on) his draw of $\gamma$ may lie in $\Gamma^2$ or $\Gamma^3$. Hence, if the probability the individual is of high ability when he has a draw of $\gamma$ in $\Gamma^3 \left( \frac{\Gamma^3_H}{\Gamma^3_H + \Gamma^3_L} \right)$ is significantly greater than the probability the individual is of high ability with he has a draw of $\gamma \in \Gamma^2 \left( \frac{\Gamma^2_H + \Gamma^2_L}{\Gamma^2_H + \Gamma^2_L + \Gamma^2_L} \right)$, then playing $B$ may make outsiders think the individual is more likely to have high ability. Of course one still needs to account for the fact that the individual is less likely to have high ability if he has preference $b$. 
Illustration of the probabilities $\Gamma^k_i$ using the same parameters as figure 2.1.

With the critical values $\gamma_A$ and $\gamma_B$ defined, a qualitative description of the equilibria can now be presented.

**Lemma 2.2.** An equilibrium to the game always exists. If the individual plays $A$ with probability one conditional on preference $a$, the equilibrium will be unique and characterized by one of the following strategies:

1. $\sigma_1$: Play $A$ with probability 1 regardless of preference and $\gamma$.

2. $\sigma_2$: Play $B$ with probability in $(0,1)$ conditional on $(b,\gamma \in \Gamma^3)$, otherwise play $A$.

3. $\sigma_3$: Play $B$ with probability 1 conditional on $(b,\gamma \in \Gamma^3)$, otherwise play $A$.

**Proof.** Consider the second statement in the proof (leave existence aside for the moment). If the individual plays $A$ conditional on preference $a$ then by the definition of $\gamma_B$ the employer
will fire the individual if the individual plays $B$ and has a draw of $\gamma \in \Gamma^2$ (this was discussed as a corollary to lemma 2.1). Hence, the only potential equilibrium strategies are those in which the individual plays $B$ with probability in $[0, 1]$ conditional on $(b, \gamma \in \Gamma^3)$, and plays $A$ with probability 1 otherwise – these are precisely the strategies $\sigma_1, \sigma_2$ and $\sigma_3$. The proof that these strategies are unique can be found in the appendix.

To show existence, consider the case when neither $\sigma_1, \sigma_2$ or $\sigma_3$ hold as equilibria. This can only occur if the individual would prefer to play $A$ when he has preference $a$ under $\sigma_3$ (conditional on not being fired for doing so). In this case an equilibrium can always be supported in which the individual plays $B$ conditional on having a draw of $\gamma \in \Gamma^3$ and plays $A$ conditional on a draw of $\gamma \in \Gamma^2$ – term this strategy $\sigma_4$. The individual’s reputation conditional on playing $A$ is greater under $\sigma_4$ than under $\sigma_3$; $P(H|K, A, \sigma_4) = \left(1 - q_H \right) \left(1 - q_H \right) \left(1 - q_H \right) > \left(1 - q_H \right) \left(1 - q_H \right) \left(1 - q_H \right) = P(H|K, A, \sigma_3)$. Similarly, his reputation conditional on playing $B$ is lower under $\sigma_4$ than under $\sigma_3$. Hence if the the individual preferred playing $A$ to $B$ under $\sigma_3$, he will certainly prefer $A$ to $B$ under $\sigma_4$. Furthermore, the individual would not deviate from playing $B$ when he has a draw of $\gamma \in \Gamma^2$, as he would be fired if he does.

The strategies $\sigma_1, \sigma_2$, and $\sigma_3$ are illustrated in figure 2.3. Lemma 2.2 only defines strategies conditional on the individual playing action $A$ when he has preference $a$. The purpose of this section is to examine countersignaling, and in particular to define sufficient conditions for a countersignaling equilibrium. These conditions shall be developed by ruling out equilibrium that are not countersignaling equilibria – any equilibrium in which the individual plays $B$ conditional on preference $a$ is, by definition, a countersignaling equilibrium, hence we needn’t rule it out.

**Countersignaling**

From the definition of $\sigma_1$ and $\sigma_2$, neither of these strategies can represent a countersignaling equilibrium – under these strategies the individual (weakly) prefers playing $A$ even when he
Panels (a), (b), and (c) illustrate the strategies $\sigma_1$, $\sigma_2$, and $\sigma_3$, respectively.
has preference $b$, implying there is a reputational gain to playing $A$. Conversely, $\sigma_3$ would constitute a countersignaling equilibrium if outsiders believe that the individual is more likely to be of high ability when he plays $B$, that is when $P(H|K, B, \sigma_3) > P(H|K, A, \sigma_3)$.

**Weak countersignaling.** The condition $P(H|K, B, \sigma_3) > P(H|K, A, \sigma_3)$ is, in fact, a sufficient condition for a weak countersignaling equilibrium; when this condition is met, lemma $\sigma_1$ and $\sigma_2$ cannot be upheld as equilibria because the individual would always prefer to play $B$ conditional on $(b, \gamma \in \Gamma^3)$ – the equilibrium set will include only $\sigma_3$ or equilibria in which the individual plays $B$ conditional on preference $a$ (which are, by definition, strong countersignaling equilibria).

The term $P(H|K, B, \sigma_3)$ is the probability the individual has high ability conditional on playing $B$. To define this probability in terms of the model primitives consider the following set of equalities,

$$P(H|K, B, \sigma_3) = \frac{P(B, K, H|\sigma_3)}{P(B, K|\sigma_3)} = \frac{P(B, K, H|\sigma_3)}{P(B, K, L|\sigma_3) + P(B, K, H|\sigma_3)} = \frac{(1 - q_H)\Gamma_H^3}{(1 - q_L)\Gamma_L^3 + (1 - q_H)\Gamma_H^3}. \tag{2.8}$$

The first equality in equation 2.8 comes from the definition of a conditional probability. The second equality uses the fact that the unconditional probability of an outcome (in this case $B$ and $K$) is the sum of the probability of that outcome joint with low and high ability, respectively. The third equality follows from the identity $P(B, i, K|\sigma_3) = P(B, K|i, \sigma_3)P(i)$, and the assumption that $P(H) = P(L) = 1/2$. The final equality follows from the identity $P(B, K|i, \sigma_3) = (1 - q_i)\Gamma_i^3$, which can be seen in figures 2.2 and 2.3. A similar exercise yields the probability,

$$P(H|K, A, \sigma_3) = \frac{P(A, K|H, \sigma_3)}{P(A, K|H, \sigma_3) + P(A, K|L, \sigma_3)} = \frac{\Gamma_H^2 + q_H\Gamma_H^3}{\Gamma_L^2 + q_L\Gamma_L^3 + \Gamma_H^2 + q_H\Gamma_H^3}. \tag{2.9}$$
Hence the sufficient condition for a countersignaling equilibrium is,

\[
P(H|K, B, \sigma_3) - P(H|K, A, \sigma_3) = \frac{(1 - q_H)\Gamma^3_H}{(1 - q_L)\Gamma^3_L + (1 - q_H)\Gamma^3_H} - \frac{\Gamma^2_H + q_H\Gamma^3_H}{\Gamma^2_L + q_L\Gamma^3_L + \Gamma^2_H + q_H\Gamma^3_H} > 0
\]  

(2.10)

Figure 2.4 provides two examples in which \( \sigma_3 \) can be upheld as an equilibrium – in one example \( \sigma_3 \) is a countersignaling equilibrium, in the other it is not.

Comparative statics for equation 2.10 are derived in appendix B. In the appendix it is shown that the left hand side of equation 2.10 is increasing in \( \gamma_B \) and decreasing in \( \gamma_A \). An increase in \( \gamma_B \) increases the draw of \( \gamma \) the individual requires to play \( B \). By assumption MLRP, this increases the probability the individual has high ability when he plays \( B \) under \( \sigma_3 \). By the same logic, an increase in \( \gamma_A \) increases the probability the individual has high ability conditional on playing \( A \) under \( \sigma_3 \).

The model primitives, \( p_i \), \( q_i \), and \( m \), have an ambiguous effect on equation 2.10. However, it can be stated with certainty that as \( q_H \) goes to one a countersignaling equilibrium is assured.

**Proposition 2.1.** As \( q_H \) approaches one a countersignaling equilibrium is assured.

**Proof.** See appendix.

Recall that \( q_H \) is the probability the individual has preference \( a \) conditional on having high ability. As \( q_H \) goes to one (under \( \sigma_3 \)) the power of the evaluation signal becomes very strong – playing \( B \) denotes the individual as having low ability with probability approaching one. As the signalling power of the action becomes stronger to insiders, the power of the countersignal becomes stronger to outsiders – outsiders infer that the individual must have a very favourable draw of \( \gamma \) in order for insiders to have kept him on conditional on observing \( B \). Note that as \( q_H \) goes to one, \( \gamma_B \) goes to \( \sup(\gamma|L) \) (assumption EP is critical to this result), implying that the probability the individual actually plays \( B \) under \( \sigma_3 \) goes to zero.
Panels (a) and (b) plot the probability the individual plays A and B under strategy profile $\sigma_3$ for two sets of parameters. The solid lines represent the density of $\gamma$ conditional on ability, and the dashed lines represent the density of $\gamma$ conditional on ability and preference $b$. From equation 2.8, the term $P(H|K, B, \sigma_3)$ can be found by dividing the darkest grey area in the top graph of each panel by the sum of the darkest grey areas in the top and bottom graphs. The term $P(H|K, A, \sigma_3)$ can be found in the same way, using the lighter grey area. Using the parameters in panel (a), $P(H|K, A, \sigma_3) - P(H|K, B, \sigma_3) = -0.14$; under these parameters there is a countersignaling equilibrium. Note however that $\sigma_3$ is not necessarily an equilibrium – if $\delta/\beta > 0.14$ the individual would deviate and play $B$ conditional on $(a, \gamma \in \Gamma^3)$. Conversely, in panel (b) $P(H|K, A, \sigma_3) - P(H|K, B, \sigma_3) = 0.33$; under these parameters a countersignaling equilibrium is not assured. Note again that $\sigma_3$ is not necessarily an equilibrium – if $\delta/\beta > 0.33$ then the individual would prefer to deviate to $A$ conditional on $(b, \gamma \in \Gamma^3)$.

This accords with the notion that action $A$ may be a social norm that is played with high probability while action $B$ represents a rare deviation from the norm.

Whether or not the condition in equation 2.10 is satisfied when $q_H$ is less than one depends largely on the shape of $g$. Which begs the question: what shape of $g$ will ensure a countersignaling equilibrium?

To answer this question, note that the first term in equation 2.10 increases relative
to the second term when $\Gamma^3_H$ increases relative to $\Gamma^2_H$ or when $\Gamma^2_L$ increases relative to $\Gamma^3_L$.

To think of this another way, consider the fraction $\frac{g(\gamma|H)}{g(\gamma|H) + g(\gamma|L)}$; this is the probability the individual has high ability conditional on a particular draw of $\gamma$. By assumption MLRP the fraction is increasing in $\gamma$, however there is no restriction placed on its second derivative. The inequality in equation 2.10 is more likely to be met when there is relatively little curvature in the fraction around $\gamma_A$, but relatively large curvature around $\gamma_B$. This ensures that the probability the individual has high ability is much greater when he has a draw $\gamma \in \Gamma^3$ than when he has a draw of $\gamma \in \Gamma^2$.

Consider the panels in figure 2.4 again. Panel (a) represented a weak countersignaling equilibrium, whereas panel (b) did not. Figure 2.5 plots the ratio $\frac{g(\gamma|H)}{g(\gamma|H) + g(\gamma|L)}$ for both panels of figure 2.4. In panel (a) there is significant curvature in the ratio in the immediate neighbourhood of $\gamma_B$ and very little curvature around $\gamma_A$. Conversely, in panel (b) the curvature is similar in the neighbourhood of $\gamma_A$ and $\gamma_B$.

**Strong countersignaling.** There are an infinite number of potential strong countersignaling equilibrium strategies. These strategies can be pure or mixed, and are generally not unique. Figure 2.6 illustrates two examples of strong countersignaling equilibria.

A strong countersignaling equilibrium will always exist when the individual places much higher weight on his reputation than on acting consistent with their intrinsic preference.

**Proposition 2.2.** There is a value of $\xi$ such that when $\delta/\beta < \xi$ a strong countersignaling equilibrium exists regardless of any other parameter values.

*Proof. See appendix.*

To show this proof informally, suppose the individual played the following strategy: play $B$ conditional on $\gamma \in \Gamma^3$ (regardless of type) and play $A$ otherwise. Under this strategy, the individual’s reputation is greater when he plays $B$, and hence if the individual places more weight on his reputation than on acting consistent with his preference then he will play
Figure 2.5: Ratio of $\frac{g(\gamma|H)}{g(\gamma|H)+g(\gamma|L)}$ for countersignaling and non-countersignaling examples

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure2_5}
\caption{(a) Example of countersignaling \quad (b) Non-example of countersignaling}
\end{figure}

This figure plots $\frac{g(\gamma|H)}{g(\gamma|H)+g(\gamma|L)}$ for the parameters in panels (a) and (b) of figure 2.4.

$B$ regardless of his preference if he has a draw of $\gamma \in \Gamma^3$. When the individual has a draw of $\gamma \in \Gamma^2$ he would be fired for playing $B$ and must therefore play $A$.

Proposition 2.2 showed that a strong countersignaling equilibrium always exists when the individual’s cost of acting against his preference is low compared to the value the individual places on his reputation, however it did not rule out the existence of other non-countersignaling equilibria. Without specifying an equilibrium selection mechanism, a sufficient condition for a strong countersignaling equilibrium is that neither $\sigma_1$, $\sigma_2$, nor $\sigma_3$ represent equilibria. This is assured when the individual would prefer to play $B$ (if he would not be fired for doing so) conditional on preference $a$ under $\sigma_3$,

$$u(B, a, \gamma \in \Gamma^3, \sigma_3) > u(A, a, \gamma \in \Gamma^3, \sigma_3). \quad (2.11)$$
Panels (a) and (b) show two examples of strong countersignaling equilibria. The solid lines represent the density of $\gamma$ conditional on ability, and the dashed lines represent the density of $\gamma$ conditional on ability and preference $b$. For the strategies illustrated in panel (a) $P(H|K, A) - P(H|K, B) = -0.57$, hence the strategies can be upheld as equilibria if $\delta/\beta < 0.57$. If $\delta/\beta > 0.57$, the individual would deviate from the illustrated strategy and play $A$ conditional on preference $a$ (as the cost of acting against his preference would be too high compared to the value he puts on his reputation.) Similarly, in panel (b) $P(H|K, A) - P(H|K, B) = -0.47$. If these strategies were equilibria the individual would always like to play $B$ if he would be kept on, however the employer’s off-equilibrium beliefs dictate that the individual would be fired if he deviated from the equilibrium by playing $B$ instead of $A$.

When the inequality in equation 2.11 is met, $\sigma_3$ cannot be an equilibrium as the individual would deviate if he had preference $a$ and a draw of $\gamma \in \Gamma^3$. Furthermore, the individual would prefer to play $B$ conditional on preference $b$, if he would not be fired for doing so, and hence $\sigma_1$ and $\sigma_2$ cannot be equilibria. Expanding equation 2.11 (using equations 2.8 and 2.9) yields,

$$m + \beta \frac{\Gamma^H_2(1 - q_H)}{\Gamma^H_1(1 - q_L) + \Gamma^H_2(1 - q_H)} - \delta > m + \beta \frac{\Gamma^H_1 + \Gamma^H_2(q_H)}{\Gamma^H_1 + \Gamma^H_2(q_H) + \Gamma^L_1 + \Gamma^L_2(q_L)} \quad (2.12)$$
After some rearrangement the condition can be written as,

\[
\frac{\Gamma_H^3(1 - q_H)}{\Gamma_L^3(1 - q_L)} - \frac{\Gamma_H^2 + \Gamma_L^3 q_H}{\Gamma_L^2 + \Gamma_L^3 q_L} - \frac{\delta}{\beta} > 0.
\]  (2.13)

With the exception of the third term on the left hand side, this condition is identical to the sufficient condition for a weak countersignaling equilibrium (equation 2.10). Hence, the discussion surrounding equation 2.10, including the comparative static results and the affect of the shape of \( g \), apply equally to equation 2.13.

The one difference between the equation 2.10 and equation 2.13 is the presence of \( \delta/\beta \). In keeping with the logic of proposition 2.2, a decrease in \( \delta/\beta \) unambiguously increases the left hand side of equation 2.13, enlarging the parameter space over which a strong countersignaling equilibrium is assured.

2.4 Uncertainty

In most environments an evaluation signal is not perfectly observed by the individual – a manager may believe an owner attributes the success of a firm to the manager, when in fact the owner attributes the success to favourable economic conditions; an individual may mistake a friend’s bad mood as a signal that the friend dislikes him; a restaurant may mistake the good nature of their customers as praise for their food.

When the reward given by insiders is large relative to the cost of acting against one’s preference, uncertainty makes the individual more cautious in taking action \( B \). Even if the individual thinks he has received an evaluation that is high enough to permit him to “get away with” playing \( B \), if there is some probability that he is mistaken about his evaluation signal he may choose to play \( A \) rather than risk losing his reward. This caution implies that the individual only plays \( B \) when he is very confident in his evaluation signal, which in turn increases the probability the individual has high ability conditional on playing \( B \).

To introduce uncertainty allow that the individual observes his evaluation signal with
some noise. Term the signal observed by the individual as,

\[ \tilde{\gamma} = \gamma + \epsilon \]  

(2.14)

where \( \epsilon \) is a uniformly distributed mean zero noise term with support \([-s, s]\).

The purpose of this section is to explore how the variance of \( \epsilon \) (captured in \( s \)) affects the sufficient conditions for both a weak and strong countersignaling equilibrium (analogues of equations 2.10 and 2.13). In the proposition that follows it is shown that when \( m \) is large the introduction of a small amount of uncertainty increases the parameter space over which the sufficient conditions are met. In the previous section the value of \( m \) was inconsequential, yet it seems reasonable that the size of the reward bestowed by insiders should matter. For example, if an individual is in a highly competitive and lucrative job, he should be hesitant in taking an action that could result in termination. Conversely, if the employee does not value his wage very highly (either because it is low or because he can easily find another job) he will be more likely to act consistent with his preference.

This result is modest inasmuch as it only concerns the introduction of small amounts of uncertainty. When there is a large amount of uncertainty, \( \tilde{\gamma} \) will contain little information about \( \gamma \) and hence the individual’s action will largely be a function of his draw of \( \epsilon \).

**Solution Under Uncertainty**

In the previous section the individual’s strategy mapped from his preference and draw of \( \gamma \) to a probability of taking action \( A \). In this section it is assumed that the individual does not observe \( \gamma \) and hence his action will be conditioned on \( \tilde{\gamma} \) instead of \( \gamma \). Conversely, the employer’s action remains a mapping from the individual’s evaluation signal (\( \gamma \)) and the individual’s action, to a probability over playing \( K \).

**Lemma 2.3.** When \( 2s < \gamma_B - \gamma_A \) and the individual plays \( A \) with probability 1 conditional on preference \( a \), the individual’s equilibrium strategy falls into one, and only one, of three
sets:

\[ \tilde{\sigma}_1: \text{Play A regardless of } \tilde{\gamma}. \]

\[ \tilde{\sigma}_2: \text{Play B with probability in } (0, 1) \text{ if } \tilde{\gamma} \in (\gamma_B + s, \sup(\gamma)) \text{ and A otherwise} \]

\[ \tilde{\sigma}_3: \text{Play B with probability 1 if } \tilde{\gamma} \in (\tilde{\gamma}, \sup(\gamma)) \text{ where } \tilde{\gamma} \in (\gamma_B - s, \gamma_B + s) \text{ and A otherwise.} \]

**Proof.** See appendix.

The intuition for lemma 2.3 is much the same as lemma 2.2. The difference being that under \( \tilde{\sigma}_3 \) the individual plays B when he has a draw of \( \tilde{\gamma} < \gamma_B + s \), this implies that there is some probability the individual plays B and has a draw of \( \gamma \in (\gamma_B, \gamma_A) \), resulting in the individual being fired. The individual is willing to accept some small probability of being fired in order to act consistent with his preference. In keeping with the focus of this section on small amounts of uncertainty lemma 2.3 also employs an upper bound on \( s \).

Following the intuition of the preceding section, a weak countersignaling equilibrium is assured if \( P(H|K, B, \tilde{\sigma}_3) - P(H|K, A, \tilde{\sigma}_3) > 0 \), and a strong countersignaling equilibrium is assured if \( P(H|K, B, \tilde{\sigma}_3) - P(H|K, A, \tilde{\sigma}_3) > \frac{\delta}{\beta} \). Hence, the question of interest is how \( s \) affects \( P(H|K, B, \tilde{\sigma}_3) - P(H|K, A, \tilde{\sigma}_3) \).

**Proposition 2.3.** \( P(H|K, B, \tilde{\sigma}_3) - P(H|K, A, \tilde{\sigma}_3) \) is increasing in \( s \) as \( s \to 0 \) and \( m > \beta \left( \frac{r_H^2(1-q_H)}{r_H^2(1-q_H)+r_L^2(1-q_L)} - \frac{2(r_H^2+q_Hr_H^2)}{r_H^2+q_Hr_H^2+r_H^2+q_Lr_L^2} \right) + 2\delta. \)

**Proof.** See appendix.

The proof of lemma 2.3 demonstrates that \( s \) has two effects on \( P(H|K, B, \tilde{\sigma}_3) - P(H|K, A, \tilde{\sigma}_3) \): a direct effect and an indirect effect through \( \tilde{\gamma} \). The direct effect of an increase in \( s \) is to raise the probability that the individual will be fired conditional on playing B; since the individual is most likely to be fired when he has a low draw of \( \gamma \), this increases the likelihood the individual is of high ability when he plays B. The indirect effect has been discussed throughout this section; when \( m \) is large uncertainty makes the individual
less willing to play $B$ if he has a low draw of $\tilde{\gamma}$, thereby increasing $\tilde{\gamma}$. This makes it more likely that the individual is of high ability when he plays $B$. Figure 2.7 revisits the non-example of countersignaling in figure 2.4; when a small amount of uncertainty is introduced a countersignaling equilibrium is assured.

Figure 2.7: Countersignaling equilibrium under uncertainty

The figure plots the probability the individual plays actions $A$ and $B$ under the strategy $\tilde{\sigma}_3$ for the same parameter values as panel (b) in figure 2.4, with uncertainty added ($s=.2$). The solid lines represent the density of $\gamma$ conditional on ability, and the dashed lines represent the density of $\gamma$ conditional on ability and preference $b$. The individual plays $B$ only if he has preference $b$ and a draw of $\tilde{\gamma} > \tilde{\gamma}$. There is some probability the individual is fired even if he has a draw of $\gamma > \gamma_A$; this occurs when the individual has preference $b$, a draw of $\tilde{\gamma} > \tilde{\gamma}$, and a draw of $\gamma < \gamma_B$. Uncertainty has made the individual more cautious in taking action $B$, and hence the probability that the individual has high ability conditional on playing $B$ and being kept on $(P(H|K,B,\tilde{\sigma}_3) = .82)$, is now greater than the probability that he has high ability conditional on playing $A$ and being kept on $(P(H|K,A,\tilde{\sigma}_3) = .75)$.

In most environments there will be uncertainty surrounding aspects of the model other than the individual’s observation of his evaluation. For example, some players in the game may have imperfect information related to the distribution of the evaluation signal or the probability distribution of the individual’s preference $(q_H$ and $q_L)$. The insight of this section is that when the individual values the reward from insiders much more than the reputational
or psychic costs of an action, the introduction of small amounts of uncertainty will make the individual more cautious in playing an action associated with having low ability. It seems reasonable to expect this insight to hold if uncertainty were added along other dimensions of the model.

The results of this section are conditional on the individual being unaware of his ability. If the individual was aware of his ability, a high ability individual would have a higher expected draw of $\gamma$ conditional on receiving $\tilde{\gamma}$. The individual would therefore be more confident in playing $B$ if he knew himself to be of high ability, further increasing the likelihood of a countersignaling equilibrium under uncertainty.

2.5 Conclusion

The primary contribution of this article is to offer a new explanation for countersignaling in an environment with multiple audiences. Countersignaling results from the individual demonstrating to one audience (outsiders) that he is of high ability by taking an action that reduces his evaluation in the eyes of a more informed audience (insiders). If insiders are still willing to offer a reward after observing the individual’s action, then outsiders will believe that insiders must have received favourable private information about the individual. It was further shown that introducing uncertainty can increase the likelihood of a countersignaling equilibrium by making the individual more cautious in taking an action associated with having low ability. The simple model in this paper could be extended in several different directions with potentially interesting implications.

While it was assumed that insiders place no intrinsic value on the individual’s action (insiders cared about the individual’s action only inasmuch as it signalled ability) the model could be extended to allow the individual to take an action that directly reduces insiders payoffs. For example, a worker could signal to his peers that he has the favour of his superiors by showing up late to a meeting. If the superiors fail to discipline the worker, the workers’
peers will infer that the worker must have the favour of the superiors. This could develop into a culture wherein it is the norm for individuals with high evaluations to engage in destructive behaviour to demonstrate that they can “get away with it”.

The simple model in this article could be embedded into a dynamic game in which insiders receive multiple evaluation signals over time. After each signal the employee can decide what action to take, and the employer can decide whether to fire the employee or wait to gain additional information. This type of game would more fully capture Hollander’s notion of idiosyncratic credits – which are credits that an employee amasses over time (and through good behaviour), which he can “spend” by acting idiosyncratically.

The distinction between insiders and outsiders in the model is another simplification that could be relaxed. For instance one might think of a model in which there are multiple audiences each with a separate draw of the evaluation signal and a reward to give the individual. One insider will base their evaluation of the individual on both their own evaluation signal and the reward given by other insiders, with the potential for informational cascades similar to the model of Banerjee (1992).

It was implicitly assumed that preferences were fixed, however it is possible that preferences for an action may be endogenous to the ability of individuals who take the action; if those who play $B$ are of higher ability than those who play $A$, then over time it may become fashionable to play $B$, offering an explanation fashion trends and fads.\footnote{Other explanations for fads have been based on imperfect information (Bikhchandani, Hirshleifer, and Welch, 1992) and planned obsolescence by a design monopolist (Pesendorfer, 1995).}

For example, a popular celebrity may countsignal by adopting a fashion that is decidedly unstylish. This signals to the general public (outsiders) that the celebrity is confident enough in her standing within the fashion world (insiders) that she can act against established norms. Others may join the celebrity in countsignaling if their standing among the fashionistas is likewise secure. Over time the fashion the celebrity is using to to countsignal may simply become the preferred fashion. Hence a fashion that was once a countsignal transforms into an ordinary signal. To continue to be held in high esteem by outsiders the
celebrity will need to find a new method of countersignaling – perhaps starting a new fashion trend. The countersignaling explanation of fashion cycles captures the idea of Simmel (1904) who wrote, “the fashions of the upper stratum of society are never identical with those of the lower; in fact, they are abandoned by the former as soon as the latter prepares to appropriate them.” (p. 133)
Chapter 3

Maximum score estimators for matching models with non-transferable utility

3.1 Introduction

This paper explores estimation of two-sided one-to-one matching games in which two individuals from disjoint sets form a match. Examples of such markets include (heterosexual) marriage, employment, schooling, and friendship. In these markets a researcher typically observes the characteristics of individuals who match together and wishes to recover the preferences (on both sides of the market) that rationalize the observed matchings.

Two maximum score estimators are introduced: a weighted estimator and an unweighted estimator. The weighted estimator is found to be consistent under very general preferences, though the asymptotics of the estimator are in the number of markets – not in the number of individuals in a particular market. Conversely, the unweighted estimator is found to be consistent only if preferences are homogeneous, however the asymptotics of the unweighted estimator are in either the number of individuals or the number of markets.

3.2 Brief literature review

In a previous paper on matching with non-transferable utility, Agarwal and Diamond (2014) demonstrate identification of preferences on either side of the market up to a monotonic
transformation based on the assumption of preference homogeneity (this paper requires a similar assumption for the unweighted estimator). Estimation of the model is done through the minimization of an empirical moment. Sheng (2012) studies identification and estimation of a network game, of which matching with non-transferable utility is a special case. Sheng demonstrates parametric identification of the network formation model and defines a simulation estimator by which separate utility functions can be jointly estimated.

There are a handful of existing papers that have estimated matching games with non-transferable utility using micro-data. Boyd et al. (2012) use a simulated method of moments estimator, based on minimizing the difference between the actual characteristics of matches and the characteristics of matches that predicted using the Gale-Shapley algorithm. Logan, Hoff, and Newton (2008) employ a Bayesian approach to estimation, which is based on estimating the parameters that maximize the probability of a stable matching. Sørensen (2007) also uses a Bayesian approach to estimation, by restricting the payoffs on both sides of the market, his model actually estimates a joint payoff function instead of two separate utility functions. None of these papers offer an identification argument.

The matching estimators introduced in this paper are similar in some respects to Fox (2010a,b), who shows identification of matching games with transferable utility and estimates these games using a maximum score estimator. While sharing some similarities, matching games with transferable utility are quite different from and matching games with non-transferable utility. With transferable utility the equilibrium maximizes social welfare and hence the equation of interest is a joint production function (given that utility is transferable all that matters is the sum of the two utility functions). With non-transferable utility the equations of interest are the two utility functions that govern the behaviour of both sets of agents.
3.3 Model

Consider a data set with $T \geq 1$ separate markets in which $n \geq 2$ males and females match together. Define the utility the $i$th male receives from the $k$th female in the $t$th market as,

$$v_{m_i,f_k}^t = \beta z_{m_i,f_k}^t + \epsilon_{m_i,f_k}^t$$  \hspace{1cm} (3.1)

Where $z$ is a vector of observables, $\epsilon$ represents the unobservables, and $\beta$ are parameters to be estimated. The utility the $k$th female derives from the $i$th male in the $t$th market is,

$$u_{f_k,m_i}^t = \gamma x_{f_k,m_i}^t + \eta_{f_k,m_i}^t,$$  \hspace{1cm} (3.2)

where $x$, $\eta$, and $\gamma$ are analogues of $z$, $\epsilon$, and $\beta$.

The equilibrium assumption is pairwise stability. In a pairwise stable equilibrium no unmatched male and female would prefer to match with each other instead of their equilibrium match.\footnote{Gale and Shapley (1962) show that a pairwise stable equilibrium will exist in all two-sided one-to-one matching markets, though the equilibrium is not necessarily unique. Gale and Shapley further introduce the deferred acceptance algorithm which results in either a male-optimal equilibrium or female-optimal equilibrium. A male-optimal equilibrium is preferred by all males to any other stable equilibrium.} The double $(m_i, f_k)$ denotes that the $i$th male has matched with the $k$th female.

Identification of the estimators relies on the following assumptions:

Assumption 3.1. An iid random sample of individuals is observed. The result of the matching process is a pairwise stable equilibrium.

Assumption 3.2. The parameter space is compact.

Assumption 3.3. The distributions of $\eta$ and $\epsilon$ are absolutely continuous, with an unknown finite variance, open support, and median of zero. The unobservables are uncorrelated with each other and with the observables.
Assumption 3.4. The sign of one coefficient is known.

Assumption 3.3 may be potentially restrictive as it rules out dependence between the unobservables in the two utility functions. Assumption 3.4 implies that identification relies on a sign restriction. In practice this is rarely an issue as economic theory typically dictates the sign of at least one coefficient. Previous work on one-to-one matching has also found that a sign restriction to be necessary (Agarwal and Diamond, 2014).

3.4 Weighted maximum score estimator

Identification of the weighted estimator will be shown through an identification at infinity argument similar to Tamer (2003). This argument relies on a special regressor in both the male and utility function which has unbounded support.

Assumption 3.5. There exists a special regressor in both $x$ and $z$ (term them $x^*$ and $z^*$) that is absolutely continuous, with almost everywhere support on the interval $[a, \infty]$, or on the interval $[-\infty, a]$, where $a$ is some constant.

The score of the weighted estimator is,

$$
\hat{Q}(\hat{\beta}, \hat{\gamma}) = \sum_t \sum_{m_i=1}^{n_t} \sum_{m_j=m_i+1}^{n_t} \sum_{f_k=1}^{n_t} \sum_{f_l=f_k+1}^{n_t} \frac{h(Y_{m_i,m_j,f_k,f_l})}{S} \left( q(m_i,f_l),(m_j,f_k) \mathbb{I}\{(m_i,f_l),(m_j,f_k)\} + q(m_i,f_k),(m_j,f_l) \mathbb{I}\{(m_i,f_k),(m_j,f_l)\} \right).
$$

(3.3)

where,

$$
S = \sum_t \sum_{m_i=1}^{n_t} \sum_{m_j=m_i+1}^{n_t} \sum_{f_k=1}^{n_t} h(Y_{m_i,m_j,f_k,f_l}) (\mathbb{I}\{(m_i,f_l),(m_j,f_k)\} + \mathbb{I}\{(m_i,f_k),(m_j,f_l)\})
$$

(3.4)

The summation terms in equation 3.3 loop over every possible two by two submarket in the data. A two by two submarket is defined as a market with two males and two females, denoted, $m_i$, $m_j$, $f_k$, and $f_l$, respectively. The fractional term in equation 3.3 is the “weight”
assigned to the submarket. The $q$ terms are the scores for the submarket, whose value are predicated on the predicted equilibrium in the submarket. The indicator values are equal to one if the matchings inside the indicator value are actually observed in the data. Hence a submarket contributes positively to the score if (a) it contains two actual matches, (b) the observed matches are predicted by $\hat{\beta}$ and $\hat{\gamma}$, and (c) it is accorded positive weight.

Let us first define $q$ and then turn attention to the weighting function. As was just stated, the value of $q$ depends on the equilibrium predicted in the submarket by the observable part of the utility functions $-\hat{\beta}x$ and $\hat{\gamma}z$. There are two possible matches in a submarket: $\{(m_i, f_k), (m_j, f_l)\}$ and $\{(m_i, f_l), (m_j, f_k)\}$. One or both of these matches will be predicted to be stable.

If $\{(m_i, f_k), (m_j, f_l)\}$ is predicted to be a unique equilibrium then $q_{(m_i, f_k), (m_j, f_l)}(\hat{\beta}, \hat{\gamma}) = 1$ and $q_{(m_j, f_k), (m_i, f_l)}(\hat{\beta}, \hat{\gamma}) = 0$, this occurs when the model observables predict:

a) Both $m_i$ and $f_k$ both prefer to match together: $\hat{\beta}z_{m_i, f_k} > \hat{\beta}z_{m_i, f_l}$ and $\hat{\gamma}x_{f_k, m_i} > \hat{\gamma}x_{f_k, m_j}$, or

b) Both $m_j$ and $f_l$ both prefer to match together: $\hat{\beta}z_{m_j, f_l} > \hat{\beta}z_{m_j, f_k}$ and $\hat{\gamma}x_{f_l, m_j} > \hat{\gamma}x_{f_l, m_i}$.

If multiple equilibria are predicted then $q_{(m_i, f_k), (m_j, f_l)} + q_{(m_j, f_k), (m_i, f_l)}(\hat{\beta}, \hat{\gamma}) = 1$, this occurs when males and females have “opposite preferences” – that is when the model observables predict,

c) $m_i$ prefers $f_k$, $m_j$ prefers $f_l$, $f_k$ prefers $m_j$, and $f_l$ prefers $m_i$: $\hat{\beta}z_{m_i, f_k} > \hat{\beta}z_{m_i, f_l}$, $\hat{\beta}z_{m_j, f_l} > \hat{\beta}z_{m_j, f_k}$, $\hat{\gamma}x_{f_k, m_j} > \hat{\gamma}x_{f_k, m_i}$, and $\hat{\gamma}x_{f_l, m_i} > \hat{\gamma}x_{f_l, m_j}$, or

d) $m_j$ prefers $f_l$, $m_j$ prefers $f_i$, $f_k$ prefers $m_i$, and $f_l$ prefers $m_j$: $\hat{\beta}z_{m_i, f_l} > \hat{\beta}z_{m_i, f_k}$, $\hat{\beta}z_{m_j, f_k} - \hat{\beta}z_{m_j, f_l}$, $\hat{\gamma}x_{f_k, m_i} - \hat{\gamma}x_{f_k, m_j}$, and $\hat{\gamma}x_{f_l, m_j} > \hat{\gamma}x_{f_l, m_i}$.

To see why multiple equilibria are predicted in these situations, suppose that preferences are as in (c); the matching $\{(m_i, f_k), (m_j, f_l)\}$ is stable as neither male would wish to deviate, and the matching $\{(m_j, f_k), (m_i, f_l)\}$ is stable as neither female would wish to deviate. The
condition, \( q(m_i, f_k, (m_j, f_l)) + q(m_j, f_k, (m_i, f_l))(\hat{\beta}, \hat{\gamma}) = 1 \) is purposely vague and can be imposed in many ways. For example, one could allow that when multiple equilibria are predicted \( q(m_i, f_k, (m_j, f_l))(\hat{\beta}, \hat{\gamma}) = q(m_j, f_k, (m_i, f_l))(\hat{\beta}, \hat{\gamma}) = 1/2 \). Conversely one could allow that the male (or female) optimal equilibrium will always be assigned a score of one. Exactly how this condition is imposed is irrelevant to the identification argument as, asymptotically, only markets with unique equilibria are assigned positive weight by the weighting function.

The weighting function is entirely based on the value of \( h(\Upsilon_{m_i, m_j, f_k, f_l}) \) across the submarkets. The function \( h \) is non-decreasing in its argument and such that \( \lim_{n \to \infty} h(\infty) = \infty \). The value \( \Upsilon_{m_i, m_j, f_k, f_l} \) is based on the differences of the special regressors among individuals in the submarket. Let \( \Upsilon_{m_i, m_j, f_k, f_l} = \max\{z^d_{m_i, m_j, f_k, f_l}, x^d_{m_i, m_j, f_k, f_l} \} \) where,

\[
z^d_{m_i, m_j, f_k, f_l} = \min\{|z^*_{m_i, f_k} - z^*_{m_i, f_l}|, |z^*_{m_j, f_k} - z^*_{m_j, f_l}|\}
\]

if

\[
\text{sgn}(z^*_{m_i, f_k} - z^*_{m_i, f_l}) = \text{sgn}(z^*_{m_j, f_k} - z^*_{m_j, f_l})
\]

and \( z^d_{m_i, m_j, f_k, f_l} = 0 \) otherwise. If both males prefer one female to the other on the basis of the special regressor, then \( z^d \) will equal the minimum difference in the value of the special regressors across females, from the perspective of the two males. Conversely, if the two males each prefer a different female on the basis of the special regressor then \( z^d = 0 \). Define \( x^d_{f_k, f_l} \) in a similar manner.

Identification is based on the unbounded support of \( z^d \) (and \( x^d \)). As \( z^d \) goes to infinity the probability that both males prefer the same female goes to one and hence the preferred female has her choice of which male to match with (a similar story can be told with respect to \( x^d \)). For example, imagine a matching market with PhD graduates matching to their first academic employer. The employer may care about many things including research output. It seems reasonable to assume that as the difference between two candidates research output goes to infinity (i.e. one candidate has five publications in the American Economic Review,
while the other has no publications), all departments would prefer the better published candidate, regardless of the candidates’ other traits.

The denominator of the fraction $\frac{h(\Upsilon_{m_i, m_j, f_k, f_l})}{S}$ takes the sum of $h(\Upsilon_{m_i, m_j, f_k, f_l})$ across all submarkets composed of two actual matches. thereby normalizing the score to be in the interval $[0,1]$. As $n$ goes to $\infty$ the denominator will also take on an extreme value. Therefore, as the number of individuals goes to infinity the fraction $\frac{h(\Upsilon_{m_i, m_j, f_k, f_l})}{S}$ will take on positive values only when $\Upsilon_{m_i, m_j, f_k, f_l}$ is infinite, that is to say only when one individual in the submarket faces a discrete choice. Hence one might think that we can use standard discrete choice asymptotic arguments to identify the model – unfortunately this is not possible as $n$ goes to infinity.

Consider the probability that a submarket is informative. To be informative it must be the case that one of $z^d$ or $x^d$ takes on an extreme value (giving the submarket positive weight), while the other does not. In submarkets in which both $z^d$ and $x^d$ take on extreme values any $\beta, \gamma$ will yield the same score, provided that the coefficients attached to the special regressor are of the same sign. Hence, we can think of the probability that a submarket is informative as the probability that $z^d$ is takes on an extreme value while $x^d$ does not (or vice-versa): $Pr(z^d > \xi, x^d < \xi) \cup Pr(z^d < \xi, x^d > \xi)$, for some arbitrarily large $\xi$.

To determine the asymptotic behaviour of this probability, allow that $x$ and $z$ contain both individual-specific and match-specific terms. An individual-specific characteristic is a characteristic that is specific to an individual, and all members of the opposite sex agree on the value of this characteristic. Conversely, a match-specific characteristic is specific to a particular match – members of the opposite sex will have different valuations of the same individual on the basis of match-specific characteristics. It was shown by Lee (2014) that as $n$ goes to infinity each male will (with probability arbitrarily close to one) match with a female who (a) is ranked in a quantile arbitrarily close the the man’s own quantile based on individual-specific attributes and (b) provides utility (arbitrarily close) to the maximum utility possible from match-specific observables. To make this more concrete, imagine that
individuals in a marriage market match based on wealth and intelligence (assume that these characteristics are uncorrelated). Individuals assign positive value to their partners wealth (regardless of their own) but they wish to match with someone who has the same level of intelligence as their own. As $n$ gets large individuals will match with a member of the opposite sex whose quantile rank in wealth is arbitrarily close to their own and who has identical intelligence to their own.

Consider the implications of Lee’s result for the current model. Suppose the special regressor is individual specific. In this case, for submarkets in which $z^d$ goes to infinity, $x^d$ will also go to $\infty$ – this is because individuals match with a partner who is ranked in the same quantile based on individual specific characteristics. Conversely, suppose that the special regressor is match specific – in this case as $n$ goes to infinity all individuals will find a partner who provides utility arbitrarily close to infinity on this dimension. In either event the probability $Pr(z^d > \xi, x^d < \xi) \cup Pr(z^d < \xi, x^d > \xi)$ is likely decreasing in $n$, hence as $n$ increases the number of submarkets will increase, but the number of informative submarkets may not.

Conversely, the probability, $Pr(z^d > \xi, x^d < \xi) \cup Pr(z^d < \xi, x^d > \xi)$, is unaffected by an increase in the number of markets. Hence as the number of markets increases, the number of individuals who face a discrete choice will also go to infinity. Identification then follows from the argument of Fox (2007) who studied discrete choice identification of pairwise maximum score estimation.

### 3.5 Unweighted estimator

An alternative estimator is an unweighted maximum score estimator which is equivalent to equation 3.3 without the weighting term. In this section it shall be shown that the unweighted estimator is consistent under homogeneous preferences, which implies that all males share common ranking of females – that is to say $v_{mi,fk} = v_{mj,fk}$. In many contexts this
may be a reasonable assumption – for example, academic departments rankings of potential PhD candidates are likely to be highly correlated. The unweighted estimator is shown to be consistent in either the number of markets or the number of individuals – it is also likely to be more efficient than the weighted estimator as it does not rely on an identification at infinity argument (see Khan and Tamer (2010) for a discussion of the efficiency of estimators that rely on such identification assumptions).

While identification of the unweighted estimator does not require the presence of a special regressor, it does require a slight restriction on the shape of the error terms and an assumption on the support of one of the covariates.

**Assumption 3.6.** The distributions of the error terms have an increasing monotone hazard rate.

**Assumption 3.7.** Both equations 3.1 and 3.2 contain at least one covariate that is absolutely continuous with everywhere positive support.

When preferences are homogenous one can think of males and females as being ranked on a ladder. In the unique equilibrium of the game males and females who are matched together are on the same rung of the ladder. Denote $g$ as a function that maps from an individual to their ranking by the opposite sex, for example $g(m_i) = 2$ implies that $m_i$ is preferred to one other man. In equilibrium $m_i$ and $f_k$ will be matched if and only if $g(m_i) = g(f_k)$.

The score of the unweighted estimator is identical to equation 3.3 with the weighting term removed. Using the law of iterated expectations, the probability limit of the unweighted estimator is,

\[
\lim_{n,T \to \infty} \hat{Q}(\hat{\beta}, \hat{\gamma}) = \sum_{t=1}^{T} \sum_{m_i=1}^{n_t} \sum_{m_j=m_i+1}^{n_t} \sum_{f_k=1}^{n_t} \sum_{f_l=f_k+1}^{n_t}
\begin{align*}
E_{x,z} & \left[ P(\{(m_i, f_k), (m_j, f_l)\} | \beta^*, \gamma^*) q_{(m_i, f_k), (m_j, f_l)}(\hat{\beta}, \hat{\gamma}) \right] + \\
P(\{(m_i, f_l), (m_j, f_k)\} | \beta^*, \gamma^*) q_{(m_i, f_l), (m_j, f_k)}(\hat{\beta}, \hat{\gamma}) \right].
\end{align*}
\]
where \( P(\{(m_i, f_i), (m_j, f_k)\} | \beta^*, \gamma^*) \) represents the probability that both the matchings of \((m_i, f_i)\) and \((m_j, f_k)\) are stable in equilibrium under the true parameters of the model.

When preferences are homogeneous there is a unique predicted equilibrium in each submarket – either \( q(m_i, f_i), f_k, (m_j, f_k) (\hat{\beta}, \hat{\gamma}) = 1 \) and \( q(m_i, f_i), f_k, (m_j, f_k) (\hat{\beta}, \hat{\gamma}) = 0 \) or vice-versa. Hence to show identification it need only be shown that when \( q(m_i, f_i), f_k, (m_j, f_k) (\beta^*, \gamma^*) = 1 \),

\[
P(\{(m_i, f_k), (m_j, f_i)\} | \beta^*, \gamma^*) \geq P(\{(m_i, f_i), (m_j, f_k)\} | \beta^*, \gamma^*).
\] (3.8)

To see why this maximizes the probability limit (in equation 3.7), think of \( q(m_i, f_i), f_k, (m_j, f_k) (\hat{\beta}, \hat{\gamma}) \) and \( q(m_i, f_k), (m_j, f_i) (\hat{\beta}, \hat{\gamma}) \) as weights. The probability limit is maximized when the greater of the two probabilities is given positive weight and the lesser of the two is given a weight of zero.

Under homogeneous preferences the probability that \( m_i \) and \( f_k \) are matched is the probability that \( m_i \) and \( f_k \) have the same ranking: \( \sum_{r=1}^{n} Pr(g(m_i) = r) Pr(g(f_k) = r) \). The probability of observing \( \{(m_i, f_k), (m_j, f_i)\} \) is,

\[
\sum_{r=1}^{n} \sum_{s=r}^{n} Pr(g(m_i) = r) Pr(g(f_k) = r) Pr(g(m_j) = s) Pr(g(f_i) = s) + Pr(g(m_i) = s) Pr(g(f_k) = s) Pr(g(m_j) = r) Pr(g(f_i) = r)
\] (3.9)

Hence the inequality in equation 3.8 holds if for any value \( r > s \),

\[
Pr(g(m_i) = r) Pr(g(f_k) = r) Pr(g(m_j) = s) Pr(g(f_i) = s) + Pr(g(m_i) = s) Pr(g(f_k) = s) Pr(g(m_j) = r) Pr(g(f_i) = r) > Pr(g(m_i) = r) Pr(g(f_k) = s) Pr(g(m_j) = s) Pr(g(f_i) = r) + Pr(g(m_i) = s) Pr(g(f_k) = r) Pr(g(m_j) = r) Pr(g(f_i) = s)
\] (3.10)

After some algebraic manipulation it can be shown that a necessary condition for equation
3.10 to hold is,
\[
\text{sgn} \left( \frac{\Pr(g(m_i) = r)}{\Pr(g(m_i) = s)} - \frac{\Pr(g(m_j) = r)}{\Pr(g(m_j) = s)} \right) = \text{sgn} \left( \frac{\Pr(g(f_k) = r)}{\Pr(g(f_k) = s)} - \frac{\Pr(g(f_l) = r)}{\Pr(g(f_l) = s)} \right). 
\] (3.11)

Now to show identification we need only show that the relationship in equation 3.11 holds when \( q_{m_i,f_l}(m_j,f_k)(\beta^*, \gamma^*) = 1 \). Under homogeneous preferences, \( q_{m_i,f_l}(m_j,f_k)(\beta^*, \gamma^*) = 1 \) when either, (a) \( m_i \) is preferred to \( m_j \) (\( \gamma^* x_{m_i}^t > \gamma^* x_{m_j}^t \)) and \( f_k \) is preferred to \( f_l \) (\( \beta^* z_{f_k}^t > \beta^* z_{f_l}^t \)) or (b) \( m_j \) is preferred to \( m_i \) and \( f_k \) is preferred to \( f_l \). In what follows it shall be shown that the condition in equation 3.11 holds when \( m_i \) and \( f_k \) are the preferred individuals though the proof is symmetric when \( m_j \) and \( f_l \) are preferred.

Denote the utility females receive from the \( r-1 \) ranked male as \( u^{r-1} \), hence the probability that \( m_i \) is the \( r \)th ranked male is,
\[
\Pr(g(m_i) = r) = \Pr(\eta \in (u^{r-1} - \gamma^* z_{m_i}, u^{r+1} + \gamma^* z_{m_i}|\gamma^* z_{m_i} - \gamma^* z_{m_j})) > 0. 
\] (3.12)

Using this equation, the left hand side of equation 3.11 is equal to,
\[
\text{sgn} \left( \frac{\Pr(\eta \in (u^{r-1} - \gamma^* z_{m_i}, u^{r+1} + \gamma^* z_{m_i}|\gamma^* z_{m_i} - \gamma^* z_{m_j}))}{\Pr(\eta \in (u^{s-1} - \gamma^* z_{m_i}, u^{s+1} + \gamma^* z_{m_i}|\gamma^* z_{m_i} - \gamma^* z_{m_j}))} - \frac{\Pr(\eta \in (u^{s-1} - \gamma^* z_{m_j}, u^{s+1} + \gamma^* z_{m_j}|\gamma^* z_{m_i} - \gamma^* z_{m_j}))}{\Pr(\eta \in (u^{s-1} - \gamma^* z_{m_j}, u^{s+1} + \gamma^* z_{m_j}|\gamma^* z_{m_i} - \gamma^* z_{m_j}))} \right) > 0. 
\] (3.13)

The inequality follows from assumption 3.6. Intuitively, when \( m_i \) provides more non-stochastic utility than \( m_j \), it is relatively more likely that \( m_i \) is of a higher rank than \( m_j \). By a similar logic,
\[
\text{sgn} \left( \frac{\Pr(\eta \in (v^{r-1} - \beta^* x_{f_k}, v^{r+1} + \beta^* x_{f_k}|\beta^* x_{f_k} > \beta^* x_{f_l}))}{\Pr(\eta \in (v^{s-1} - \beta^* x_{f_k}, v^{s+1} + \beta^* x_{f_k}|\beta^* x_{f_k} > \beta^* x_{f_l}))} - \frac{\Pr(\eta \in (v^{r-1} - \beta^* x_{m_j}, v^{r+1} + \beta^* x_{m_j}|\beta^* x_{f_k} > \beta^* x_{f_l}))}{\Pr(\eta \in (v^{s-1} - \beta^* x_{m_j}, v^{s+1} + \beta^* x_{m_j}|\beta^* x_{f_k} > \beta^* x_{f_l}))} \right) > 0. 
\] (3.14)

This is enough to prove that \((\beta^*, \gamma^*)\) maximize the unweighted score. To show that
these parameters yield a unique maximum, it must be shown for any set of parameters, 
\((\tilde{\beta}, \tilde{\gamma})\), that are not monotonic transformations of \((\beta^*, \gamma^*)\) or 
\((-\beta^*, -\gamma^*)\), there are values of 
x and z that occur with non-zero probability such that,

\[
q_{(m_i,f_k),(m_j,f_l)}(\beta, \gamma) \neq q_{(m_i,f_k),(m_j,f_l)}(\beta^*, \gamma^*). \tag{3.15}
\]

This is is assured by the unbounded and everywhere positive support of the one of the 
regressors (assumption 3.7). Hence the score will be different under \((\tilde{\beta}, \tilde{\gamma})\) than under 
\((\beta^*, \gamma^*)\).

In the appendix it is shown that the unweighted estimator is also identified (in the 
number of individuals) when preferences are independent. Independent preferences imply
that the utility two males (females) get from \(f_k\) (\(m_i\)) are completely independent – this seems
a rather unrealistic assumption in actual markets, hence the relegation of this proof to the
appendix. The appendix also offers an example that illustrates the identification failure of
the unweighted estimator when preferences are heterogeneous.

## 3.6 Monte-Carlo evidence

Monte-Carlo experiments were run to assess the small sample properties of the estimators.
Each experiment has two observables in the male and female utility function, the coefficient
on the first observable is normalized to one leaving a single coefficient to be estimated for
each utility function. The data is generated such that the outcome of the matching game is
the male-optimal equilibrium. Conversely the function \(q_{(m_i,f_k),(m_j,f_l)}(\beta, \gamma)\) is set equal to one
if \{(\(m_i, f_k\), (\(m_j, f_l\))\} is the female-optimal equilibrium predicted by \(\hat{\beta}, \hat{\gamma}\) (demonstrating that
the estimator does not rely on knowledge of the actual equilibrium selection mechanism).
In the weighted estimator a jump function is used for \(h(\cdot)\) such that \(h(x) = x\) if \(x > 2\) and
\(h(x) = 0\) otherwise. All experiments use 1,000 replications.

Table 3.1 presents the bias and mean integrated squared error (MISE) of the ratio of
the log of the two coefficients in each utility function – employing the ratio of the log of the

46
estimates makes the mean and MISE insensitive to the normalization. Under homogenous preferences the bias of both the weighted and unweighted estimator is close to zero across all sample sizes, furthermore the MISE is shrinking in sample size. The MISE is lower for the weighted estimator when \( n \) is small – though this difference disappears as \( n \) or \( t \) increases.

Under heterogeneous preferences the unweighted estimator appears to contain significant bias that is not decreasing in the number of individuals in a market or the number of markets. The MISE of the estimator is shrinking as sample size increases, though these gains come from a reduction in variance not bias – as sample size increases the estimator returns a more precise, though incorrect, estimate. As expected, the bias and MISE of the weighted estimator is decreasing in the number of markets though not in the number of individuals in a market.

### 3.7 Conclusion

This paper introduced two maximum score estimators for matching models with non-transferable utility. It was shown that the weighted estimator is consistent in the number of markets, using an identification at infinity argument. Conversely, the unweighted estimator is consistent in either the number of markets or in the number of individuals, though this is only true when preferences are homogeneous.

\(^2\)To make this clear note that \((\hat{n}_2,s/\hat{n}_1,s - n_2^n/n_1^n)^2 \neq (\hat{n}_1,s/\hat{n}_2,s - n_1^n/n_2^n)^2\), but \((\ln n_2,s/\ln n_1,s - \ln n_2^n/\ln n_1^n)^2 = (\ln n_1,s/\ln n_2,s - \ln n_1^n/\ln n_2^n)^2\)
Table 3.1: Monte-Carlo results

<table>
<thead>
<tr>
<th>Markets</th>
<th>Players</th>
<th>Weighted estimator</th>
<th>Unweighted estimator</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>( \hat{\beta}_2/\beta_1 )</td>
<td>( \hat{\gamma}_2/\gamma_1 )</td>
</tr>
<tr>
<td>Homogeneous preferences(^2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>50</td>
<td>-0.02 (0.77)</td>
<td>-0.03 (0.72)</td>
</tr>
<tr>
<td>1</td>
<td>100</td>
<td>0.00 (0.51)</td>
<td>0.01 (0.49)</td>
</tr>
<tr>
<td>1</td>
<td>200</td>
<td>-0.02 (0.33)</td>
<td>-0.01 (0.31)</td>
</tr>
<tr>
<td>10</td>
<td>25</td>
<td>-0.01 (0.38)</td>
<td>-0.01 (0.39)</td>
</tr>
<tr>
<td>100</td>
<td>25</td>
<td>0.00 (0.14)</td>
<td>0.00 (0.13)</td>
</tr>
<tr>
<td>Full model(^3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>25</td>
<td>0.28 (0.92)</td>
<td>0.22 (0.91)</td>
</tr>
<tr>
<td>1</td>
<td>50</td>
<td>0.29 (0.80)</td>
<td>0.26 (0.79)</td>
</tr>
<tr>
<td>1</td>
<td>100</td>
<td>0.29 (0.69)</td>
<td>0.31 (0.68)</td>
</tr>
<tr>
<td>1</td>
<td>200</td>
<td>0.37 (0.60)</td>
<td>0.38 (0.59)</td>
</tr>
<tr>
<td>10</td>
<td>25</td>
<td>0.18 (0.59)</td>
<td>0.10 (0.58)</td>
</tr>
<tr>
<td>100</td>
<td>25</td>
<td>0.03 (0.27)</td>
<td>-0.02 (0.25)</td>
</tr>
</tbody>
</table>

1 Mean of the log of the estimated ratio (true value is 0) is displayed in the table with the mean integrated square error of the log of the estimated ratio in parenthesis.

2 Under homogeneous preferences the two (individual-specific) observables are drawn from N(0,1) (a normal distribution with mean zero and variance of one) and the unobservable is drawn from N(0,2).

3 Under heterogeneous preferences the first regressor is individual specific and drawn from N(0,1). The second regressor is match specific and drawn from N(0,1), there is a correlation of 0.3 between the match-specific observable in the male utility function and the match-specific observable in the female utility function. There is one match specific and one individual specific unobservable each drawn from N(0,1).
Chapter 4

A structural model of hiring, promotion, and job exit with application to gender and academic hiring

4.1 Introduction

Females in North America generally experience inferior labour market outcomes. In 2010 the ratio of female to male full time wages was 72% (Goldin, 2014) with a sizeable wage gap remaining after controlling for observable differences.\(^1\) Furthermore, research in several industries has found a gender bias in both hiring\(^2\) and promotion.\(^3\) This has led many to conclude that females face significant labour market discrimination. Yet, others have countered that gender differences in labour market outcomes may be attributable, at least

---

\(^1\) Blau and Kahn (2006) contend that controlling for observables such as education, experience, race, industry, union status and occupation, accounts for only 60% of the gender gap in pay.

\(^2\) Several field experiments have been conducted in which mock job applications are sent to employers. These field experiments tend to find a male-bias in hiring for high-status positions (Petit, 2007; Firth, 1982; Neumark, Bank, and Nort, 1996; Steinpreis, Anders, and Ritzke, 1999) and a bias towards the dominant gender in highly gendered positions, such as secretaries and car mechanics (Weichselbaumer, 2004; Glick, Zion, and Nelson, 1988). This bias is mediated somewhat by personality (Weichselbaumer, 2004; Glick, Zion, and Nelson, 1988), age (Petit, 2007), and the gender diversity of decision-makers (Gorman, 2005).

\(^3\) A male-bias in promotion has been found in both organization-level studies (Pudney and Shields, 2000; Jones and Makepeace, 1996; Joy, 1998; Groot and Brink, 1996; Cannings, 1988; DiPrete and Soule, 1988), and studies using broader survey data (McCue, 1996; Gjerde, 2002; Blau and Devaro, 2007; Cobb-Clark, 2001). A seemingly smaller number of studies have found the opposite result (Booth, Francesconi, and Frank, 2003; Hersch and Viscusi, 1996; Powell and Butterfield, 1994) or no difference in promotion by gender (Lewis, 1986).

Lazaer and Rosen (1990) show that differential promotion rates for women could be attributed to the likelihood of females taking time off (or specializing in household activities), rather than an innate bias on the part of employers. Empirically, it has been shown that the male bias is reduced when training is controlled for (Gjerde, 2002) and is generally lower in more senior jobs, where woman have already signalled their specialization in workplace activities (DiPrete and Soule, 1988).
in part, to gender differences in preferences (Konrad et al., 2000).\footnote{Reasons for these differences in preferences may relate to differences in household roles (Becker, 1985; Waldfogel, 1998), aversion to competition (Croson and Gneezy, 2009), and desire for work-life balance (Emslie and Hunt, 2009). Conversely, aspects of the workplace itself may cause differences in preferences. Workers of different genders may differ in their social isolation (Ohlott, Ruderman, and McCauley, 1994), opportunities for collaboration (McDowell, Singell Jr, and Stater, 2006), and happiness in male dominated workplaces (Clark, 1997).}

Determining whether differential gender outcomes are owed to employer discrimination or differences in preferences is complicated by the fact that one typically observes the outcome of a joint decision making process. For example, one can observe that an individual was hired but one cannot observe the choice set that faced the individual or the employer. This paper studies the hiring and promotion of PhD economists using structural models that are capable of jointly estimating the preferences of both employees and employers using data on observed outcomes.

The initial hiring decision is modelled as a two-sided matching game with non-transferable utility and is estimated semi-parametrically using a weighted maximum score estimator. Post-hiring job outcomes (job exit and promotion) are estimated using a two-sided probit estimator. Identification of the two-sided probit model is based on the differential movements of professors before and after tenure.

In the initial hiring decision there is no evidence of gender bias on the part of departments, however individuals’ preferences are affected by their gender. Females are found to have a stronger preference for departments ranked highly in their area of research, whereas men exhibit stronger preferences for departments that are ranked highly overall. Significant gender differences are found in the post-hiring data – males are both more likely to be promoted and more likely to leave their job voluntarily.

These results contribute to the ongoing debate concerning the role of gender within the economics profession. Interest in the standing of women in economics is largely driven by the gender imbalance within the profession. According to annual surveys of the Committee on the Status of Women in the Economics Profession, females accounted for only 11% of full professors in PhD granting economics departments in 2010, up from 7% in 2000, and 4% in
It has been suggested that the lack of female representation in economics has led to a bias in research, lower research productivity, and a path dependency by which a lack of female role models dissuades female students from entering the profession.

Previous research in the labour market for academic economists has found a gender bias in both hiring and promotion. Kolpin and Singell (1996), using a sample of American Economic Association (AEA) members spanning 1973 to 1989, find that higher ranked departments are less likely to hire women. Similarly, McMillen and Singer (1994), employing a sample of AEA members in 1989, find that women are less likely to be employed in top 50 economics departments. Interestingly, while these two papers both use a discrete choice model they differ significantly in their motivation. McMillen and Singer (1994) motivate their econometric approach through a random utility model in which the job candidate chooses their initial placement. Conversely, Kolpin and Singell (1996) suggest that Universities choose who they hire. There is good reason for the inelegance with which the discrete choice model is motivated and its results interpreted. Simply put, it is the wrong model. The initial placement of a PhD graduate is the result of a discrete choice by both the job candidate and the university.

Promotion within the economics profession was most recently studied by Ginther and Kahn (2004), who employ two different datasets to analyze the probability that economics professors in North America attain tenure within ten years of receiving their PhD. They find females were significantly less likely to have attained tenure within ten years of their initial hiring, even after controlling for research productivity and the quality of the hiring school. However, Ginther and Kahn (2004) also fail to account for the two-sided nature of these

---

5The percentage of females among assistant professors was 28% in 2010, 21% in 2000, and 17% in 1990. Note that these percentages come from voluntary surveys and are therefore subject to self-reporting bias.

6Ferber (1995) argues the lack of interest in topics such as household economics and women’s entry into the labour market are indicative of this bias.

7Page (2007) suggests that environments with higher levels of diversity result in more effective group decision making and higher creative output.

8Evidence for this proposition is mixed. Hale and Regev (2011) show that women are more likely to enter graduate studies in a department that has a higher proportion of female professors. Conversely, both Dynan and Rouse (1997) and Jensen and Owen (2001) find that having a female instructor for introductory economics courses does not encourage females to major in economics.
outcomes – individuals may fail to receive tenure because they are denied tenure by their institution or because they voluntarily leave their institution.

Other authors have examined gender gaps in academic pay, with a general agreement that males receive greater remuneration than their female colleagues. Some investigators find this gap to be explained by differences in academic rank (Ginther and Hayes, 2003; Ward, 2001), while others find a within rank pay gap (Blackaby, Booth, and Frank, 2005).

In contrast to these empirical findings, recent literature has raised the possibility of a preference for female candidates within academia. Friedman (1998) suggests that while females were once the subject of discrimination, “the pendulum has probably swung too far so that men are the ones currently being discriminated against” [p. 199]. Ginther and Kahn (2004) raise the possibility that gender discrimination in promotion may be the result of affirmative action policies at the time of hiring that are not carried through in the tenure and promotion phase.

While gender differences in labour outcomes are often attributed to an employer bias, evidence from sociology and behavioural economics shows that gender plays a role in shaping an individual’s preferences. Women in patriarchal societies are relatively more risk averse (Croson and Gneezy, 2009) and tend to gravitate to jobs that are less competitive, emphasize inter-personal skills, and offer a greater work-life balance (Konrad et al., 2000). If these phenomena generalized to the academic job market one might expect women to exhibit a stronger preference for job security (i.e. a higher probability of attaining tenure), interaction with students and faculty, and a less competitive environment. In a survey of PhD students in economics who were near graduation, Barbezat (1992) found that females did indeed exhibit a stronger preference for employment in liberal arts colleges, while males exhibited a stronger preference for employment in top economics or business departments.9

9While assumed exogenous in this paper, these preferences may be a function of the current distribution of female faculty and gender stereotypes within the profession. Konrad et al. (2000, p 594) note that, “Changing one’s preferences to conform to alternatives that are realistically available reduces the tension caused by frustrated desires.” Conversely, differences in gender preferences may be attributable to biological and/or sociological processes.
This paper also contributes to the literature on structural modelling in labour markets. The initial hiring market is modelled as a two-sided one-to-one matching market with non-transferable utility. It shall be argued that the process by which candidates match to departments results in a pairwise-stable equilibrium. Pairwise stability implies that no candidate and department who are not matched in equilibrium would prefer matching with each other instead of their observed partners. Using this equilibrium assumption the parameters of each utility function can be separately identified up to a sign and location restriction. Estimation is done by a weighted maximum score estimator whose properties were explored in the preceding chapter.

Post-hiring job outcomes are modelled using a two-sided probit model. After being hired an individual can experience one of three outcomes in any given year: (1) the individual stays in the same tenure status with the same department, (2) the individual is granted tenure, or (3) the individual moves to a new department. When an individual is untenured, the outcome is a function of the joint decision of the department and the candidate. An individual can choose to either stay in the department or leave the department. Similarly, the department can choose to terminate the individual, offer the individual tenure, or keep the individual in their current position.

Identification of the choice functions comes, in part, from the institution of tenure. Once an individual is granted tenure the department can no longer terminate the individual or grant tenure again. Conversely, the individual still has the binary choice of staying in the department or leaving the department. The difference in the probability a tenured professor exits their job relative to the probability an untenured professor exits their job is used to identify the probability that an individual leaves their position of their own volition. Non-random selection into the treatment (tenured) group, is accounted for by allowing for correlation in the unobservables across time.
4.2 The academic labour market

The market for academic economists is organized by strong institutional norms, some of which are common in academia, others of which are particular to the discipline. This section provides an overview of these norms, discussing in turn, the organization of the initial hiring market for recent PhD graduates, the organization of the post-hiring labour market, and the role of gender in the academic labour market.

4.2.1 The initial hiring market

In North America, most PhD graduates in economics are hired in a brief window that centres around the annual meeting of the American Economic Association (AEA) in January. The typical pattern of hiring begins in the months previous to the meeting when jobs are advertised in *Job Openings for Economists* with a deadline for applications approximately a month before the AEA meetings. Preliminary interviews are conducted during the meetings and select candidates (between three and five) are invited for campus visits between January and March. A job offer is then made to the preferred candidate between February and April (Cawley, 2011).

The outcome of this hiring process is likely to approximate a pairwise stable equilibrium wherein no candidate and department who are not matched in equilibrium would prefer to match with each other instead of their equilibrium match. The assumption of pairwise stability is justifiable as hiring takes place in a relatively brief interval of time, hence typical labour market frictions caused by temporal differences in employer hiring are minimized. Further, job candidates can both acquire and transmit information at a relatively low cost – in a 1996 survey Stock, Alston, and Milkman (2000) found that candidates sent out an average of 76 applications, in a 1997 survey List (2000) found that candidates sent an average of 41 applications. Departments can also gain more information about applicants through a brief preliminary interview during the AEA meetings. List (2000) found the mean number
of interviews per job seeker was 6 with a standard deviation of 6.6.

Typically when an offer is made to a candidate they are given ten to fourteen days to respond. One would expect that if a candidate had a reasonable expectation of an offer from a preferred department they would communicate with the preferred department to ensure the opportunity to counteroffer (Cawley, 2011).

PhD graduates may also be hired outside of academia, either by choice or because they are not able to find an academic position. These individuals are not considered in this paper.

4.2.2 The post-hiring market

A newly hired faculty member carries the rank of assistant professor. After some time as an assistant professor an individual is considered for tenure. Tenure is the most defining and controversial characteristic of the academic labour market as it affords professors strong job security; according to a joint statement of American Association of University Professors and the Association of American Colleges, “‘adequate cause’ in [tenured] faculty dismissal proceedings should be restricted to (a) demonstrated incompetence or dishonesty in teaching or research, (b) substantial and manifest neglect of duty, and (c) personal conduct which substantially impairs the individual’s fulfillment of his institutional responsibilities. The burden of proof in establishing cause for dismissal rests upon the administration.” (American Association of University Professors and Association of American Colleges, 1973, p 75)\(^{10}\)

While the exact number of tenured faculty who are fired for cause is unknown, estimates put the number at 50-75 per year for the American professoriate as a whole (Mooney, 1994; Lublin, 2005). A large fraction of these dismissals are due to moral turpitude such as violation of professional ethics. Given that economists are a small fraction of the professoriate, it would seem a rare occurrence for an economics professor to lose a tenured position for cause, especially for reasons unrelated to professional ethics.

\(^{10}\)More recent guidelines from the American Association of University Professors continue to uphold this sentiment (American Association of University Professors, 2013).
In research institutions tenure decisions are largely based on a professor’s research output, though teaching and university service are also considered. The decision of whether or not to recommend tenure is made by a committee primarily composed of faculty from within the department who generally solicit the opinion of a peer from outside the institution. The recommendation of the tenure committee is then given to the university administration who ultimately grant tenure to an individual (Hohn and Shore, 1998). In the data, approximately 80% of professors who receive tenure are granted tenure between their fifth and tenth years with the institution.

An important feature of academia is the up-or-out system – individuals who do not receive tenure must leave the institution (Siow, 1995). Assistant professors have periodic reviews during which their research record is scrutinized – typically these reviews occur on an annual basis (Hohn and Shore, 1998). After such a review, a professor may form the belief that they will likely be denied tenure and may therefore leave the institution preemptively should they be able to find a job where they are more likely to be granted tenure.

4.2.3 Gender and the academic labour market

Prior to the 1960s women faced systematic discrimination in the academic labour market (Rai and Critzer, 2000) and many studies have found continued gender discrimination (McMillen and Singer, 1994; Kolpin and Singell, 1996; Ginther and Kahn, 2004; Blackaby, Booth, and Frank, 2005). Within the field of economics positive steps to redress this bias have been advocated by professional associations, prominent members of the profession, and university administrators.

Although the US government has not extended affirmative action laws to universities (as they have for government contractors), civil rights laws put the burden on the employer to prove the non-existence of discriminatory policies in the face of a gender imbalance in

---

11 As an example, the American Economics Association created the Committee on the Status of Woman in the Economics Profession in 1971 with a mission “to eliminate discrimination against women, and to redress the low representation of women, in the economics profession.” (http://www.aeaweb.org/committees/cswe/mission.php).
hiring. To avoid this type of legal action it is in the interest of universities to ensure the
gender composition of their employees is reflective of the larger workforce, or, at least, to
maintain policies that ensure no accusation of bias could be levelled (Rai and Critzer, 2000).
Since federal laws prevent discrimination based upon gender, universities do not set quotas
for the number of women that must be hired nor do they give overt preference to female
candidates. Instead, hiring committees are often tasked with seeking out female candidates to
ensure gender balance among the candidates who receive an interview; List (2000) shows that
female PhD graduates have significantly more initial interviews than their male counterparts.

4.3 Theoretical model

The job market is defined as a game involving two disjoint sets of players: departments and
individuals. The set of departments in the $t$th time period is $D_t = \{d_{t1}, d_{t2}, ..., d_{tD_t}\}$, while
the set of individuals is $C_t = \{c_{t1}, c_{t2}, ..., c_{tC_t}\}$. Each player in the game has preferences over
players in the opposing set.\footnote{In most universities the choice of who to hire is, in theory, made by university administrators on the
recommendation of a committee composed primarily of department faculty. In practice the decision typically
rests with the department, as department faculty are better equipped to evaluate candidates in their field.
Nonetheless, in some situations the input of administration is non-trivial and is potentially reflected in the
utility function being maximized. For simplicity, the institutional decision makers are simply referred to as
the department.}

The utility the $i$th individual receives from the $k$th department in time $t$ is,

$$u_{c_i,d_k}^t = \beta z_{c_i,d_k}^t + \epsilon_{c_i,d_k}^t,$$ (4.1)

where $z$ and $\epsilon$ refer to observables and unobservables, respectively. Similarly, the utility the
$k$th department derives from the $i$th individual in time $t$ is,

$$u_{d_k,c_i}^t = \gamma x_{d_k,c_i}^t + \eta_{d_k,c_i}^t.$$ (4.2)

Equations 4.1 and 4.2 implicitly assume that utility is non-transferable. In many labour
markets utility can be transferred through wages. However, in the present context it is doubtful that wages serve this function. Ehrenberg, Pieper, and Willis (1998) note that starting salaries in PhD granting economic departments are fairly homogenous, perhaps due to institutional norms, strong labour unions, or department collusion. Furthermore, hiring and promotion decisions are typically made, in part or in whole, by a committee of faculty from within the department. The committee’s incentive is to increase the reputation of their own department – a lower quality candidate could not transfer utility to these individuals by accepting a lower salary.

4.3.1 Initial hiring decision

The process by which PhD graduates are hired is a one-to-one matching process. This matching process takes place in temporally separated markets, indexed by $t$; the set of departments and individuals active in $t$ are $\mathcal{D}^t$ and $\mathcal{C}^t$, respectively. The double $(c_i, d_k)$ indicates that $c_i$ and $d_k$ are matched in the data, similarly $\mu$ is a matching function that maps from an individual to their match – $\mu(c_i) = d_k$ is equivalent to $(c_i, d_k)$. As previously mentioned the solution concept employed is pairwise stability.

4.3.2 Tenure, termination, and resignation

After a department and individual have chosen to match together they continue to receive a flow of utility from their match and based on their utility they may choose alter their relationship. In each period after being hired and before receiving tenure, an individual can either (1) move to a new department (or leave academia), (2) remain untenured at the same department, or (3) gain tenure at their current department. Prior to an individual receiving tenure it is not possible to determine if a job exit is voluntary or not.

---

13Some departments may hire more than one individual in a particular period, hence the matching process is actually many-to-one. Nonetheless, if the department’s hiring decisions are separable, then the many-to-one and the one-to-one model is equivalent.
Department. In each period a department receives utility from their employee consistent with equation 4.2. If the individual has not yet received tenure the department may either keep the employee on without tenure, terminate the employee, or offer the employee tenure. The term “terminate” is used quite broadly to encompass any reason that an employee may leave against their wishes. The assumption that a department has the freedom to terminate an employee at may be somewhat questionable as some untenured academics hold multi-year contracts. Nonetheless, the department’s ability to signal the outcome of an individual’s tenure decision reduces the violence of this assumption, as this signal may induce the individual to pre-emptively change jobs. If an individual has been granted tenure they cannot be terminated or granted tenure again, thus the department has no decision to make about their employment or tenure status.\footnote{The department may be able to make other decisions regarding the terms of an individual’s employment, such as their salary or promotion to full professor - these decisions are not considered here.}

The action of the department is determined by the utility a department receives from the individual relative to the thresholds $\alpha$ and $\omega$. The department will terminate the individual if the utility they receive from the individual falls below the threshold $\alpha(z_{c,d_k}, t - t^*_{c_i})$, where $t$ is the current period and $t^*_{c_i}$ is the period in which the $i$th individual was hired. Conversely, the department will offer the individual tenure if the utility they receive from the individual lies above $\omega(z_{c,d_k}, t - t^*_{c_i})$. Note that the thresholds are a function of department specific characteristics – one would expect higher ranked departments to have higher thresholds.

The conditioning of $\alpha$ and $\omega$ on the time since an individual was hired captures the “up or out” aspect of academia. Most academics who achieve tenure do so sometime between their fifth and tenth years in a department, hence it is expected that $\alpha$ is relatively low in the initial years after an individual is hired and increases over time while $\omega$ follows the opposite pattern. Figure 1 captures the hypothesized movement of $\alpha$ and $\omega$ over time.

This model may appear somewhat rigid in certain labour markets. For example, opportunities for promotion are not always available in an organization and therefore an individual
may provide their employer with utility above $\omega$ and not receive a promotion. Furthermore, if an employer has market power they may have no incentive to offer a promotion even if the employee provides utility above $\omega$. In the academic market however, faculty reviews are generally quite codified and are likely to obey the more rigid structure of this model.

The actual process of tenure review differs from institution to institution. In some institutions tenure decisions are taken after a set amount of time, in other departments candidates are able to precipitate tenure decisions at a time of their choosing within a certain timeframe. Furthermore in some institutions the individual receives only one consideration for tenure, and if turned down will lose their position at the University. In other schools individuals can reapply if they are denied tenure (again within a limited timeframe.) These institutional processes may affect time to tenure, and they are therefore assumed to differ randomly from school to school – if this were not the case then the results may be biased. For example, if higher ranked schools were more likely to have a particular process for tenure then there may be bias in the estimated impact of school ranking on tenure.

*Individual.* After being hired the individual can resign their position and move to a new department (or outside academia) of their own volition regardless of their tenure
status. If the $i$th individual receives an offer from the $k$th department they will move to that department if

$$v_{ci,d_k}^t - mc > v_{ci,m(c_i)}^t;$$  \hfil (4.3)

where $mc$ is a moving cost. Of course the decision to resign and take a new position relies on an outside offer being given. An ideal data set would contain information on all the offers an individual received regardless of whether they were accepted or not – such data rarely exists and does not in the current application. One potential estimation strategy is to assume pairwise stability, as in the initial hiring market. Imposing such assumptions is problematic for a few reasons. First, the addition of the moving cost ($mc$) adds frictions to the model that would present a challenge for estimation. Second, it is unclear what departments and employees are in the market in any particular period. Third, many individuals move to positions outside of academia which is outside the scope of this analysis.

As equation 4.3 cannot be estimated directly, a reduced form equation is used to capture voluntary job exits. According to this reduced form equation an individual leaves their department if

$$g(x_{d,c_i}, z_{c_i,m(c_i)}; b) + e_{c_i,m(c_i)} > 0.$$  \hfil (4.4)

Note that $g$ captures the utility the $i$th individual receives from the observables of the department in which they are hired (through $z_{c_i,m(c_i)}$). It also implicitly captures the utility an individual could attain from an outside offer (through $x_{d,c_i}$) – if an individual has a positive draw of $x_{d,c_i}$ she will be more attractive to all departments and presumably would attract better outside offers. The error term, $e$, includes the unobserved utility an outside department receives from an individual (which influences the probability that they will offer the individual a job), the unobserved utility an individual receives from their own department, and the utility an individual receives from a department that makes them an outside offer.

The two-sided probit model is therefore semi-structural. Estimation relies on equation 4.2, a structural equation describing the department’s choice function and equation 4.4, a
reduced form equation describing the individual’s choice.

4.4 Estimation

The problem of simultaneously estimating equations 4.1 and 4.2 (in the initial hiring model), or equations 4.2 and 4.4 (in the post-hiring model) is immediately obvious – individual decisions are not observed; instead one observes the outcome of a joint decision making process. Nonetheless, it is shown in this section that the parameters in each equation can be identified and estimated.

4.4.1 Estimating initial job placements

The utility that individuals and departments receive from each other in the initial job market is estimated using a weighted maximum score estimator. The details of this estimator are described in the preceding chapter. The consistency of the estimator relies on the following assumptions.

Assumption 4.1. An iid random sampling of matches is observed. The result of the matching process is a pairwise stable equilibrium.

Assumption 4.2. The parameter space is compact.

Assumption 4.3. The distributions of $\eta$, and $\epsilon$ are absolutely continuous, with an unknown finite variance, open support, and median of zero conditional on any values of the observables.

Assumption 4.4. There exists at least one regressor in $x$ and $z$ (defined as $x^*$ and $z^*$) that are absolutely continuous with almost everywhere support on the interval $[a, \infty]$, or on the interval $[-\infty, a]$, where $a$ is some constant.

Assumption 4.5. The sign of one coefficient is known.
As in the preceding chapter the objective function of the weighted maximum score estimator is

\[
\hat{Q}(\hat{\beta}, \hat{\gamma}) = \sum_{t}^{T} \sum_{c_i=1}^{n_i} \sum_{c_j=c_i+1}^{n_i} \sum_{d_k=1}^{n_i} \sum_{d_t=1}^{n_t} h(Y_{c_i,c_j,d_k,d_t}) \frac{S}{S} \left( q_{(c_i,d_t), (c_j,d_k)}(\hat{\beta}, \hat{\gamma}) \mathbb{1}\{(c_i,d_t), (c_j,d_k)\} + q_{(c_i,d_k), (c_j,d_t)}(\hat{\beta}, \hat{\gamma}) \mathbb{1}\{(c_i,d_k), (c_j,d_t)\} \right),
\]

where

\[
S = \sum_{t}^{T} \sum_{c_i=1}^{n_i} \sum_{c_j=c_i+1}^{n_i} \sum_{d_k=1}^{n_t} h(Y_{c_i,c_j,d_k,d_t}) \left( \mathbb{1}\{(c_i,d_t), (c_j,d_k)\} + \mathbb{1}\{(c_i,d_k), (c_j,d_t)\} \right). \quad (4.6)
\]

The value of \(q_{(c_i,d_t), (c_j,d_k)}(\hat{\beta}, \hat{\gamma})\) depends on the predicted equilibrium in a two by two market composed of \(c_i, c_j, \mu(c_i), \) and \(\mu(c_j)\). The term is equal to 1 if the matching \(\{(c_i,d_k), (c_j,d_t)\}\) is predicted to be a unique equilibrium in the two by two market, equal to \(1/2\) multiple equilibria are predicted, and equal to 0 if \(\{(c_i,d_k), (c_j,d_t)\}\) is not predicted to be an equilibrium.

Again, as in the preceding section \(\Upsilon\) is defined as

\[
\Upsilon_{c_i,c_j,d_k,d_t} = \max\{z_{c_i,c_j,d_k,d_t}^d, x_{c_i,c_j,d_k,d_t}^d\}
\]

where

\[
z_{c_i,c_j,d_k,d_t}^d = \min\{|z_{c_i,d_k}^* - z_{c_i,d_t}^*|, |z_{c_j,d_k}^* - z_{c_j,d_t}^*|\}
\]

if

\[
\text{sgn}(z_{c_i,d_k}^* - z_{c_i,d_t}^*) = \text{sgn}(z_{c_j,d_k}^* - z_{c_j,d_t}^*)
\]

and \(z_{c_i,c_j,d_k,d_t} = 0\) otherwise. Define \(x_{\mu(c_i), \mu(c_j)}^d\) in a is similar manner. The function \(h(\cdot)\) is increasing in its argument and \(h(\infty) = \infty\), as the number of individuals across all markets goes to infinity. Once again the intuition for identification is that as \(\Upsilon\) goes to \(\infty\) the model collapses to a discrete choice model.

In the data, an individual’s publications play the role of the special regressor in the department’s utility function. Of course, no individual has an infinite number of publications, however what matters in practice is that as the difference in quality-adjusted publications
exceeds some threshold, departments will clearly prefer the better published candidate. Similarly, the department’s rank would fulfill the role of a special regressor in the individual’s utility function. Again, there are not an infinite amount of department’s (hence the difference in the special regressor cannot take on an infinite value), but as the difference in two departments’ ranking gets large it is clear that candidates would prefer to match with the higher ranked department.

The preceding chapter placed no additional restrictions on the shape of \( h(\cdot) \). In practice, the degree of convexity in \( h \) increases the weight placed on large draws of \( \Upsilon \) – making a small number of submarkets highly influential to the score. Of course these data points are influential precisely because they are submarkets in which one agent is likely to face a discrete choice. Hence an increase in convexity will reduce bias at the cost of efficiency. One might think of the functional form of \( h \) as akin to the bandwidth parameter in regression discontinuity models. In estimation a square root function is used for \( h \) – functions with greater convexity did not yield estimates that were significantly different (either statistically or economically.)

### 4.4.2 Estimating post-hiring outcomes

The post-hiring data is constructed such that the response term, \( y^t \), can take on three different values, \( y^t \in \{0, 1, 2\} \), where \( y^t = 0 \) denotes that an individual exited the department (either voluntarily or involuntarily), \( y^t = 1 \) denotes that an individual stayed in their current position at the same department, and \( y^t = 2 \) denotes that the individual was granted tenure within their department. Once an individual is granted tenure \( y \) becomes a binary variable (either \( y^t = 0 \) or \( y^t = 1 \)).

Consider the \( i \)th individual employed in the \( k \)th department. If the individual is tenured (that is if she received tenure in a period before \( t \)) the probability she leaves her department
is the probability that $g(x_{d,k}, c_i; z_{c,d_k}; b) + e_{c_i,d_k} < 0$,

$$P(y^t = 0|\text{tenured}) = P(g(x_{d,k}, c_i; z_{c,d_k}; b) < -e_{c_i,d_k}) \quad (4.9)$$

Similarly, if the individual is tenured the probability she stays in her department is,

$$P(y^t = 1|\text{tenured}) = P(g(x_{d,k}, c_i; z_{c,d_k}; b) > -e_{c_i,d_k}) \quad (4.10)$$

When an individual has not attained tenure the probability of a particular outcome depends on the choices of both the individual and the department. The probability an untenured individual leaves her department is the probability that either the department terminates the individual or the individual leaves of her own volition,

$$P(y^t = 0|\text{untenured}) = P(\alpha(z_{c,d_k}^t, t - t_s^*) - \gamma x_{d_k,c_i} > -\eta_{d_k,c_i}^t) \cup P(g(x_{d,c_i}, z_{c,d_k}; b) < -e_{c_i,d_k}) \quad (4.11)$$

The probability an untenured individual stays in her current position in the department is the probability that both the department chooses to keep the individual in her current position and the individual chooses to stay in her department,

$$P(y^t = 1|\text{untenured}) = P(\alpha(z_{c,d_k}^t, t - t_s^*) < \gamma x_{d_k,c_i} + \eta_{d_k,c_i} < \omega(z_{c,d_k}^t, t - t_s^*)) \cap P(g(x_{d,c_i}, z_{c,d_k}; b) > -e_{c_i,d_k}) \quad (4.12)$$

Finally, the probability that the individual attains tenure is the probability that both the department promotes the individual and the individual stays in her department,

$$P(y^t = 2|\text{untenured}) = P(\omega(t - t_s^*) < \gamma x_{d_k,c_i}) \cap P(g(x_{d,c_i}, z_{c,d_k}; b) < -e_{c_i,d_k}) \quad (4.13)$$

Identifying and estimating these equations requires two additional assumptions,
Assumption 4.6. The error terms $\eta_{d_k,c_i}$ and $e_{c_i,d_k}$ are normally distributed with a mean of zero and a finite variance. The error terms are autocorrelated over time and correlated with each other in each period.

Assumption 4.7. Both $\alpha(z_{c,d_k},t - t^*)$ and $\omega(z_{c,d_k},t - t^*)$ are polynomial functions and $\alpha(0,0) = 0$.

Assumption 4.6 imposes a parametric structure on the data. Define $\rho_\eta$ and $\rho_e$ as autocorrelation in each error term, and define $r$ as the correlation between the two terms. Assumption 4.7 imposes a parametric form on the thresholds. The assumption that $\alpha(0,0) = 0$ serves as a normalization.

Using data from untenured individuals (that is using only equations 4.11-4.13) the variables $b$ and $\gamma$ are identified by functional form. Of course this identification argument rests on the assumption of the joint normality of $\eta$ and $e$. Researchers have typically have preferred to have an exclusion restriction as opposed to relying on such functional form assumptions for identification. Unfortunately, no such restriction is available in the data.

Alternatively, employing data from tenured individuals (equations 4.9 and 4.10) allows for more robust identification of $b$ and $\gamma$. If it were assumed that the probability an individual voluntarily exits their department is unaffected by tenure then $b$ would be identified through typical discrete choice arguments using equations 4.9 and 4.10. With $b$ identified, equations 4.11-4.13 would identify $\gamma$ and the thresholds using standard ordered choice arguments.

However, an individual’s tenure status may well impact the probability that they voluntarily change jobs. Tenured professors may be loathe to move schools if it required giving up tenure, hence tenure may effectively reduce the number of jobs an individual would consider. Conversely, tenure may serve as a signal of quality to other departments – making

\[ P(y^1 = 0 \text{tenured}) = P(0 - \gamma_1x^1_{d,c_i} - \gamma_2x^2_{d,c_i} > -\eta_{d_k,c_i}) \cup P(b_1x^1_{d,c_i} - b_2x^2_{d,c_i} < -e^\prime_{c_i,d_k}). \] (4.14)

All that is needed to correctly identify $\gamma_1, \gamma_2, b_1, b_2$, and $r$, are five different values of $x^1$ and $x^2$ – this yields five independent equations with five unknowns.

\[ \text{Identification of the model by functional form is quite similar to identification of a bivariate probit model.} \]

To quickly show how functional form identifies the parameters assume that there are only two parameters in equations 4.2 and 4.4. Rewriting equation 4.11 yields,

15
the individual more attractive and therefore more likely to garner an outside offer. For this reason the preferred empirical models include an intercept shift for tenure in equation 4.4. With this shift included all the elements of \( b \) can be identified aside from functional form except for the intercept terms. Functional form arguments can then be used to separately identify \( \gamma \) from the tenure intercept shift.

The model is estimated using simulated maximum likelihood estimation. The likelihood function is,

\[
L(b, \gamma, \rho_e, \rho_y, r, a, w) = \sum_{t=1}^{T} \sum_{i \in \tilde{C}^t} \ln \left( P(y_{it}^t) \right),
\]

where \( \tilde{C}^t \), refers to the set of individuals that are active in the post-hiring labour market in time \( t \). The terms \( a \) and \( w \) reference the parameters in the threshold functions, \( \alpha \) and \( \omega \), respectively. The Geweke-Hajivassiliou-Keane (GHK) simulator (Keane, 1993, 1994) is used for estimation. The GHK simulator is capable of estimating the autocorrelation in the error terms (\( \rho_e \) and \( \rho_y \)), and correlation between the two error terms (\( r \)). The details of this algorithm are supplied in the appendix. Few professors are tenured before year five, or untenured after year ten, making estimation of the thresholds difficult outside of these years. For this reason estimation is conducted using data between five and ten years after an individual’s initial hiring.

4.5 Data

The names of PhD graduates in economics was taken from the annual list published in the Journal of Economic Literature between 1991 and 2004. The publication record of each graduate was found using the EconLit database. To focus attention on individuals who pursued a research career, any graduate with less than four lifetime publications was dropped from the sample. An exhaustive search was conducted to find the educational and job history of each graduate with four or more publications. This information was found in the individual’s CV, their web bio, or their department website. ProQuest’s dissertation database contained
information on the individual’s PhD advisor, and the *Journal of Economic Literature* listed the subject area of the graduate’s dissertation.

Unless stated on an individual’s CV, it is assumed that individuals who attain the rank of associate professor have achieved tenure. However, in certain institutions promotion to associate professor and the granting of tenure are separate decisions. The data reveals that certain schools have a very high number of professors who depart the school shortly after promotion to associate professor. These schools were deemed to have separate promotion and tenure decisions – employees at these schools were excluded from the post-hiring data.\textsuperscript{16}

A web search usually revealed the individual’s gender, if it did not a determination was made based on the individual’s first or second name. The individual’s publication record was found through a search of the EconLit database. Following Hilmer and Hilmer (2012), it is assumed that a paper published less than a year after an individual is hired is observable before they were hired. When analyzing subsequent job movements, the individual’s publication record includes all publications up to the end of that particular calendar year (the qualitative results are unchanged if the publication record includes only publications up to the previous year or publications up to the following year). Publications are weighted by their impact factor\textsuperscript{17} – using other metrics for publication quality, such as dummy variables for top 5, 10, and 30 ranked journals does not yield significantly different results. An index of advisor quality was created using the impact factor adjusted publications of an advisor at the time of the advisee’s graduation.

All schools that were active in any period were ranked based on their impact factor weighted count of faculty publications in top 30 journals during the period 1990-2000. The top 30 journals are the Diamond list which is commonly used in rankings of economic departments (Kalaitzidakis, Mamuneas, and Stengos, 2003).\textsuperscript{18} Separate rankings were calculated

---

\textsuperscript{16}Schools that were deemed to have separate tenure and promotion decisions were Chicago, Harvard, Northwestern, and Yale. The results relating to gender were insensitive to this exclusion.

\textsuperscript{17}Impact factors from 2000 are used for all publications.

\textsuperscript{18}It was deemed impractical to develop a faculty average number of publications, hence the ranking of the department is dependent on the size of the department. This is consistent with other work on the ranking of economic departments (Kalaitzidakis, Mamuneas, and Stengos, 2003; Coupé, 2003). For the current purpose,
based on a department’s overall publications and a department’s publications within each specific JEL subject code.

Most individuals that entered academia were hired in economics departments, though some were employed in other academic departments, such as business, public policy, and health policy. The analysis does not differentiate between different departments as there is typically a high correlation in the ranking of economics departments and other departments. Considering the subset of individuals that match into economics departments does not change the qualitative results.

Table 4.1 contains descriptive statistics of the data. Descriptive statistics are shown for the subset of individuals who match to top 200 institutions and for all individuals. The empirical analysis focuses on the subset of individuals that match to top 200 departments as these departments tend to be more research oriented and are therefore more likely to base their decisions on research productivity. Upon initially being hired women have, on average, fewer quality adjusted publications than men and are less likely to have graduated from a US undergraduate program. Five years after being initially hired the females remaining in academia have, on average, stronger publication records than men, though this difference is not statistically significant. No other variables show statistically significant differences between genders. As one might expect, individuals who match to top 200 programs have stronger publication records and, by construction, higher quality initial placements than the overall sample.

The plots in figure 4.2 document the likelihood of an individual achieving tenure or exiting their job. Panel (a) provides some evidence of the “up or out” nature of tenure – by the tenth year after initially being hired 79% of individuals who remain in academia have tenure. The percentage of females who have tenure at any point in time is lower than

---

the aggregate publications of a department may provide a better ranking than the average publication per faculty, given that individuals may have preferences for departments with larger research programs. The publication count also includes publications from faculty in non-economic departments (such as business and public policy) and non-faculty department members.
Table 4.1: Descriptive statistics$^a$

<table>
<thead>
<tr>
<th></th>
<th>All</th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Top 200</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>1248</td>
<td>982</td>
<td>266</td>
</tr>
<tr>
<td>Rank of PhD program</td>
<td>27.782 (55.595)</td>
<td>28.211 (56.363)</td>
<td>26.195 (52.736)</td>
</tr>
<tr>
<td>Advisor publications$^b$</td>
<td>6.276 (5.608)</td>
<td>6.38 (5.718)</td>
<td>5.903 (5.189)</td>
</tr>
<tr>
<td>Publications at initial hiring</td>
<td>0.312 (0.708)</td>
<td>0.321 (0.719)</td>
<td>0.278 (0.667)*</td>
</tr>
<tr>
<td>Post-doc dummy</td>
<td>0.108 (0.311)</td>
<td>0.108 (0.310)</td>
<td>0.109 (0.312)</td>
</tr>
<tr>
<td>US undergraduate degree</td>
<td>0.928 (0.258)</td>
<td>0.927 (0.261)</td>
<td>0.934 (0.248)</td>
</tr>
<tr>
<td>Rank of initial hiring school</td>
<td>60.481 (60.969)</td>
<td>60.336 (61.020)</td>
<td>61.017 (60.892)</td>
</tr>
<tr>
<td>Publications 5 years post-hiring$^c$</td>
<td>.863 (1.394)</td>
<td>.843 (1.406)</td>
<td>.929 (1.356)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>All</th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observations</td>
<td>1957</td>
<td>1532</td>
<td>425</td>
</tr>
<tr>
<td>Rank of PhD program</td>
<td>56.729 (105.433)</td>
<td>59.504 (112.294)</td>
<td>46.724 (74.969)</td>
</tr>
<tr>
<td>Advisor publications</td>
<td>5.125 (5.301)</td>
<td>5.119 (5.349)</td>
<td>5.148 (5.132)</td>
</tr>
<tr>
<td>Publications at initial hiring</td>
<td>0.250 (0.601)</td>
<td>0.260 (0.611)</td>
<td>0.216 (0.564)**</td>
</tr>
<tr>
<td>Post-doc dummy</td>
<td>0.127 (0.333)</td>
<td>0.129 (0.336)</td>
<td>0.120 (0.325)</td>
</tr>
<tr>
<td>US undergraduate degree</td>
<td>0.930 (0.256)</td>
<td>0.927 (0.260)</td>
<td>0.939 (0.239)</td>
</tr>
<tr>
<td>Rank of initial hiring school</td>
<td>309.844 (445.373)</td>
<td>312.479 (451.794)</td>
<td>300.344 (421.791)</td>
</tr>
<tr>
<td>Publications 5 years post-hiring$^c$</td>
<td>.863 (1.394)</td>
<td>.843 (1.406)</td>
<td>.929 (1.356)</td>
</tr>
</tbody>
</table>

Notes:

$^a$The mean value of each variable is reported with the standard deviation in parenthesis. A Wilcoxon test was used to determine if significant differences were found in the average values by gender; differences at the .1, .05, and .01 levels are denoted in the third column by one, two, and three asterisks, respectively.

$^b$Publications refer to impact factor weighted publications.

$^c$Publications 5 years post-hiring includes only individuals that are still in academia 5 years after their initial hiring.

that of men. Panels (b) and (c) plot the hazard rate of an individual leaving a department conditional on being untenured and tenured, respectively. The differential rates of exit between untenured and tenured faculty suggest that a high proportion of job exits among untenured faculty are involuntary.
4.6 Results

4.6.1 Hiring

Table 4.2 contains the estimation results from the initial hiring model. Recall that the initial hiring decision is modelled as a two-sided matching game and estimated using a weighted maximum score estimator. As mentioned previously a square root function is used for \( h: h(\Upsilon) = \sqrt{\Upsilon} \). Functions with greater convexity \( (h(\Upsilon) = \Upsilon^a \) for \( a = 1, 2 \) did not yield statistically or qualitatively different results. Five different models are estimated – the models differ in the data employed and the coefficients included.

As discussed in section 4.4, one coefficient in each utility function must be normalized. In the department’s utility function the coefficient attached to the rank of the candidate’s PhD school is normalized to -1. Similarly, in the candidate’s utility function the coefficient attached to the rank of the hiring school is normalized to -1. Note that schools are ranked such that the “best” school has a numeric rank of 1 and the 100th “best” school has a rank of 100. Hence a negative coefficient implies higher ranked schools are preferred.

In all models, gender is statistically insignificant in the department’s utility function. The estimates from model one imply that being a female provides an advantage that is equivalent to a 5% increase in the rank of an individual’s PhD institution. Conversely, it can be seen as the difference between a co-authored publication in the Journal of Econometrics (impact factor of 1.106) and a co-authored article in Journal of Public Economics (impact factor of 1.013). Thus it would appear that gender is of little economic significance to departments.

All the models were run including regressors that interacted gender and all other variables in both the department and individual utility function. Adding such interaction terms to the department utility function did not have a significant impact on the score.\(^{19}\) Consider a subsampling procedure was used to test the hypothesis that such terms did not add to the fit of the model. The literature on testing multivariate restrictions using subsampling is relatively sparse. Politis,

\(^{19}\)
The three panels in this figure demonstrate the promotion and job exits of professors after their initial hire. Panel (a) contains the cumulative percentage of professors that were granted tenure conditional on remaining within academia, panel (b) contains the percentage of untenured professors that leave their job in a given year, and panel (c) contains the percentage of tenured professors that leave their job in a given year.
Table 4.2: Results from the estimation using hiring data

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Department utility function</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PhD rank</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>Female dummy</td>
<td>0.05 (-0.32, 0.09)</td>
<td>0.07 (-0.28, 0.17)</td>
<td>-0.01 (-0.16, 0.26)</td>
</tr>
<tr>
<td>Publications</td>
<td>0.38 (0.06, 0.55)</td>
<td>0.36 (0.13, 0.58)</td>
<td>0.46 (0.25, 0.66)</td>
</tr>
<tr>
<td>Advisor publications</td>
<td>-</td>
<td>-</td>
<td>0.01 (0.00, 0.02)</td>
</tr>
<tr>
<td>North American undergrad</td>
<td>0.03 (-0.03, 0.53)</td>
<td>-0.02 (-0.12, 0.65)</td>
<td>0.06 (-0.22, 0.93)</td>
</tr>
<tr>
<td>Dissertation subject (dummies)</td>
<td>N</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td><strong>Candidate utility function</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>School rank</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>School rank in PhD subject</td>
<td>-0.95 (-1.33, -0.58)</td>
<td>-0.82 (-4.01, -0.56)</td>
<td>-0.71 (-2.79, -0.53)</td>
</tr>
<tr>
<td>Distance to PhD school</td>
<td>-0.21 (-0.32, 0.29)</td>
<td>-0.03 (-0.49, 0.38)</td>
<td>0.01 (-0.19, 0.50)</td>
</tr>
<tr>
<td>Distance to undergrad school</td>
<td>-0.69 (-1.13, -0.50)</td>
<td>-0.82 (-2.35, -0.48)</td>
<td>-0.59 (-1.62, -0.37)</td>
</tr>
<tr>
<td>Female x school rank in PhD subject</td>
<td>-7.57 (-9.81, -4.70)</td>
<td>-7.09 (-8.93, -2.81)</td>
<td>-8.06 (-9.78, -4.34)</td>
</tr>
<tr>
<td>Female x distance to PhD school</td>
<td>1.39 (-1.48, 1.88)</td>
<td>0.41 (-3.03, 1.81)</td>
<td>0.36 (-8.10, 0.99)</td>
</tr>
<tr>
<td>Female x distance to undergrad school</td>
<td>-1.57 (-2.60, 1.07)</td>
<td>-0.73 (-3.25, 1.07)</td>
<td>0.09 (-1.56, 2.24)</td>
</tr>
</tbody>
</table>

**Significance of gender interaction**

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Department utility function (p-value)</td>
<td>0.8</td>
<td>-</td>
<td>0.24</td>
</tr>
<tr>
<td>Candidate utility function (p-value)</td>
<td>1</td>
<td>-</td>
<td>0.99</td>
</tr>
</tbody>
</table>

**Dataset**

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Schools included</td>
<td>Top 200</td>
<td>Top 200</td>
<td>Top 200</td>
</tr>
<tr>
<td>Number of matches</td>
<td>1078</td>
<td>1078</td>
<td>714</td>
</tr>
</tbody>
</table>

Notes:

* 90% confidence intervals are in parenthesis. Confidence intervals are based on subsampling. The subsampling procedure randomly samples 80% of the matches in any given year.

* The p-value is derived from a subsampling test that the gender interaction terms are jointly equal to zero against the most likely alternative hypothesis. Further details are discussed in the paper.
Table 4.2: Results from the estimation using hiring data\textsuperscript{a}

<table>
<thead>
<tr>
<th></th>
<th>Model 4</th>
<th>Model 5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Department utility function</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PhD rank</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>Female dummy</td>
<td>0.05 (-0.18, 0.09)</td>
<td>-0.11 (-0.16, 0.09)</td>
</tr>
<tr>
<td>Publications</td>
<td>0.38 (0.14, 0.55)</td>
<td>0.53 (0.49, 0.64)</td>
</tr>
<tr>
<td>Advisor publications</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>North American undergrad</td>
<td>-</td>
<td>-0.24 (-0.47, 0.01)</td>
</tr>
<tr>
<td>Dissertation subject (dummies)</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td><strong>Candidate utility function</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>School rank</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>School rank in PhD subject</td>
<td>-0.87 (-2.54, -0.60)</td>
<td>-0.58 (-1.00, -0.46)</td>
</tr>
<tr>
<td>Distance to PhD school</td>
<td>-0.04 (-0.71, 0.03)</td>
<td>-0.30 (-0.46, -0.06)</td>
</tr>
<tr>
<td>Distance to undergrad school</td>
<td>-</td>
<td>-0.58 (-1.23, -0.53)</td>
</tr>
<tr>
<td>Female x school rank in PhD subject</td>
<td>-5.72 (-9.85, 0.49)</td>
<td>-0.13 (-7.27, 0.08)</td>
</tr>
<tr>
<td>Female x distance to PhD school</td>
<td>1.20 (-0.93, 2.11)</td>
<td>0.12 (-3.17, 0.31)</td>
</tr>
<tr>
<td>Female x distance to undergrad school</td>
<td>-</td>
<td>0.17 (-4.29, 0.71)</td>
</tr>
<tr>
<td><strong>Significance of gender interaction\textsuperscript{b}</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Department utility function (p-value)</td>
<td>0.3</td>
<td>0.29</td>
</tr>
<tr>
<td>Candidate utility function (p-value)</td>
<td>0.75</td>
<td>0.96</td>
</tr>
<tr>
<td><strong>Dataset</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schools included</td>
<td>Top 200</td>
<td>All</td>
</tr>
<tr>
<td>Number of matches</td>
<td>1218</td>
<td>1386</td>
</tr>
</tbody>
</table>

Notes:
\textsuperscript{a} 90\% confidence intervals are in parenthesis. Confidence intervals are based on subsampling. The subsampling procedure randomly samples 80\% of the matches in any given year.

\textsuperscript{b} The p-value is derived from a subsampling test that the gender interaction terms are jointly equal to zero against the most likely alternative hypothesis. Further details are discussed in the paper.
versely, adding interaction terms to the individual’s utility function did add significantly to the fit of the model.

The interaction terms show that females place a higher priority on a school’s ranking in the subject in which they conducted their PhD research (in models 4 and 5 the term is statistically insignificant). One interpretation of this result is that females place more value on the “fit” of the department, whereas males place more value on the overall reputation of the department. Given the strong correlation between a department’s overall rank and its rank in a particular subject matter, it is doubtful that this difference in preferences causes a significant difference in the number of females hired in top ranked departments. A potentially fruitful avenue for future research is the role of gender in mediating an individual’s valuation of other job characteristics. For example, women and men may differ in the utility they derive from the ideological orientation of a department or the proportion of same gender faculty in the department.

This result is in contrast to previous literature which found that males were more likely to be hired in higher ranked departments (McMillen and Singer, 1994; Kolpin and Singell, 1996). These studies typically used data from older cohorts of PhD economists, which suggests that the preferences of departments and/or individuals have changed over time (recall that previous studies are unable to determine whether gender differences in hiring are attributable to the preferences of individuals or departments.) This finding provides some evidence that policies intended to ensure gender parity have had their intended effect without causing the “pendulum to swing too far in the other direction”.

Romano, and Wolf (1999, p 53) suggest using a norm to reduce the problem to a univariate one, though the test is sensitive to the norm employed. The procedure introduced in this paper is to test the null hypothesis against all possible alternatives. Consider for example, the test that the three interaction terms in $\beta$ – define them as $\beta^{int}$ – are jointly equal to zero. The null hypothesis is, $H_0 : \beta^{int} = (0,0,0)$ and the alternative is $H_A : \beta^{int} = \{(b_1, b_2, b_3) | b_1 \neq 0 \text{ or } b_2 \neq 0 \text{ or } b_3 \neq 0\}$ implying that departments value observables in the same way regardless of the job candidate’s gender. For a given alternative, the probability that the alternative provides a better fit is found by taking $s$ subsamples and finding the proportion in which the alternative provides a higher score. To test the null against any alternative, this probability is treated as a function of the alternative parameters. The probability is then maximized with respect to these parameters – the resulting probability is the p-value of a test of the null hypothesis against the most likely alternative.
Males and females may also differ in the importance that they put on model unobservables, such as family considerations, work-life balance, or compatibility with future colleagues. Similarly, departments may rank females and males differently on unobservable characteristics such as teaching or interpersonal skills.

The importance of the unobservables can be measured by examining the probability that the estimated parameters correctly predict the outcome of each submarket in which there is two actual matches. Furthermore, one can test whether this proportion differs depending on the gender of the individuals in the submarket.

However, this type of test would not be able to determine whether differences in prediction were due to unobservables in the department utility function or unobservables in the candidate utility function. This question however can be addressed by weighting the submarkets. Consider the hypothesis that females place greater importance on unobservables relative to men. To test this hypothesis weight submarkets such that positive weight is placed on markets in which both departments prefer the same candidate based on the observables. Furthermore, let us place higher weight on submarkets where departments preferences seem to be strongest. Consider the following statistic,

$$\tau^M - \tau^F$$

where $M$ and $F$ refer to males and females, respectively. The statistic $\tau$ measures the weighted probability that a male or female actually makes the choice predicted by the model. If $\tau^M - \tau^F > 0$ then unobservables are less important to males. The statistic $\tau^M$ is calculated as,

$$\tau^M = \sum_{t=1}^{T} \sum_{c_i=1}^{C^t} \sum_{c_j=1}^{C^t} \sum_{t=1}^{T} \frac{h(\gamma^M_{c_i,c_j})}{\sum_{t=1}^{T} \sum_{c_i=1}^{C^t} \sum_{c_j=1}^{C^t} h(\gamma^M_{c_i,c_j}) q(c_i,\mu(c_i)),(c_j,\mu(c_j)))} (\beta, \gamma),$$

where $M^t$ is the set of all males in the market $t$ and,

$$\gamma^M_{c_i,c_j} = \max\{0, \min\{\gamma x_{\mu(c_i)},c_i - \gamma x_{\mu(c_j)},c_j; \gamma x_{\mu(c_i)},c_i - \gamma x_{\mu(c_j)},c_j\}\}.$$
The term \( \Upsilon_{c_i,c_j}^M \) will equal zero if \( c_i \) is not preferred by both departments based on the departments’ predicted utility. If \( c_i \) is preferred by both departments then \( \Upsilon_{c_i,c_j}^M \) will equal the minimum difference in utility the two departments receive from \( c_i \) relative to \( c_j \). Given the infinite support of \( \gamma x_{\mu(c_i),c_i} \), if \( h(\cdot) \) is an increasing function that goes to infinity when its argument does, then asymptotically positive weight will only be assigned to submarkets in which \( c_i \) provides infinitely more utility to each department.

Equation 4.17 resembles to the estimation equation 4.5. The critical difference lies in the \( \Upsilon \) terms. In the estimator \( \Upsilon_{c_i,c_j,d_k,d_l} \) was based on differences in both \( x^* \) and \( z^* \), whereas \( \Upsilon_{c_i,c_j} \) is based on the difference in \( x \) (and the predicted coefficients). To put this another way, in equation 4.5 weights markets by the probability that either a department or candidate face a discrete choice. Conversely, equation 4.17 weights markets by the probability that \( c_i \) faces a discrete choice. This allows us to (at least asymptotically) isolate how predictive the model is of the candidate’s choice. Note also that the summation terms are simplified in equation 4.17 relative to 4.5.

In small samples the function \( h(\cdot) \) will be important. Echoing the previous discussion on the weighting function in section 4, when \( h(\cdot) \) is convex more weight will be placed on submarkets in which one can be confident that \( c_i \) is actually preferred by both departments. The test itself is in fact more sensitive to the choice of functional form than the estimator is, this is because the estimator remains consistent without any weighting if preferences are homogenous (as per the discussion in the preceding chapter) – the test does not.

A similar test statistic can be developed for the relative importance of unobservables in the department utility function. Let \( \tau_{D_M} \) be the weighted probability that a department makes the choice that is predicted by the model when the submarket is composed of two men, and \( \tau_{D_F} \) be the weighted probability that a department makes the choice that is predicted when it is choosing either between a man and a women, or two women. Again let the probabilities be weighted by a function that captures the extent to which a department is preferred by both candidates, and test whether \( \tau_{D_M} - \tau_{D_F} > 0 \).
The results for each model are presented in table 4.3. Two different functions are used for \( h(\cdot) \); under the more convex function males make the predicted choice more often than females, and departments make the predicted choices more often when no female is involved in the submarket. This implies that unobservables are more important to females, and that departments place more weight on unobservables when evaluation females.

Table 4.3: Differences in predictive power by gender\(^a\)

<table>
<thead>
<tr>
<th>Model</th>
<th>( \tau^M - \tau^F )</th>
<th>( \tau^{DM} - \tau^{DF} )</th>
<th>( \tau^M - \tau^F )</th>
<th>( \tau^{DM} - \tau^{DF} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>0.014 (0.129)</td>
<td>0.028 (0.017)</td>
<td>0.015 (0.052)</td>
<td>0.041 (0.003)</td>
</tr>
<tr>
<td>Model 2</td>
<td>-0.004 (0.092)</td>
<td>0.01 (0.013)</td>
<td>0.015 (0.016)</td>
<td>0.031 (0.005)</td>
</tr>
<tr>
<td>Model 3</td>
<td>0 (0.451)</td>
<td>0.025 (0.052)</td>
<td>0.017 (0.093)</td>
<td>0.031 (0.050)</td>
</tr>
<tr>
<td>Model 4</td>
<td>0.025 (0.114)</td>
<td>0.053 (0.010)</td>
<td>0.011 (0.136)</td>
<td>0.036 (0.009)</td>
</tr>
<tr>
<td>Model 5</td>
<td>0.025 (0.114)</td>
<td>0.053 (0.010)</td>
<td>0.011 (0.136)</td>
<td>0.036 (0.009)</td>
</tr>
</tbody>
</table>

Notes:
\(^a\) p-value in parenthesis

Returning to the results in table 4.2, the other control variables in the department’s utility function enter with the expected sign. The individual’s impact factor weighted publications are found to be economically and statistically significant. To give some idea of the relative weight a school places on an individual’s publications vis-à-vis the rank of an individual’s PhD school, in model one a publication in *The Journal of Econometrics* (impact factor of 1.106), is equivalent to a 50% increase in the rank of an individual’s PhD school.

In the individual’s utility function, all models find a statistically significant preference for schools that are highly ranked in the individual’s subject of interest. Across all models,
individuals are found to prefer schools that are close to their undergraduate institution (which is a noisy proxy for their home location). The models do not agree on whether individuals have a preference for departments that are closer to one’s PhD institution – in some models this is statistically significant, in others it is not.

4.6.2 Promotion

Table 4.4 contains the estimates of equations 4.2 and 4.4 using post-hiring data. All models employ a moving average error process in both utility functions. The moving average error structure was preferred to an auto-regressive and auto-regressive moving average process through a likelihood ratio (LR) test. An LR test also found homoskedastic error terms to be preferred to models in which the variance of the error was a polynomial function of time. All models use a squared time trend for the thresholds, which was preferred to linear and cubic models by an LR test. Finally, all terms in both utility functions were interacted with the individual’s gender. LR tests found neither set of interaction terms to be jointly significant at standard levels. All the qualitative results relating to gender were robust in terms of magnitude and statistical significance across these various specifications.

Models one and two employ data from candidates who match into top 200 schools. Model one contains an intercept shift for tenure; as was previously discussed this intercept shift relies on functional form for identification – a potentially weak argument in a finite sample. However that the other coefficients in models one and two are quite similar – hence it would appear that there is no break down in identification when the intercept shift is included.

The coefficients in the department choice function have a clear structural interpretation – they represent the value a department places on the associated characteristics. In model one the coefficient attached to the female dummy is negative and significant, implying that departments have a preference for male employees in post-hiring decisions. Table 4.5 shows
Table 4.4: Results from the estimation of post-hiring outcomes\(^a\)

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Department choice</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>3.1 (0.14)***</td>
<td>3.02 (0.26)***</td>
<td>3.03 (0.25)***</td>
</tr>
<tr>
<td>Female dummy</td>
<td>-0.2 (0.07)***</td>
<td>-0.19 (0.08)**</td>
<td>-0.25 (0.07)***</td>
</tr>
<tr>
<td>PhD rank (log)</td>
<td>0.01 (0.01)</td>
<td>0.01 (0.03)</td>
<td>0.03 (0.02)</td>
</tr>
<tr>
<td>N. Amer undergrad</td>
<td>-0.03 (0.01)</td>
<td>-0.03 (0.07)</td>
<td>-0.03 (0.07)</td>
</tr>
<tr>
<td>Publications</td>
<td>2.2 (0.19)</td>
<td>2.04 (0.27)***</td>
<td>2.24 (0.26)***</td>
</tr>
<tr>
<td><strong>Individual choice</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>1.09 (0.11)***</td>
<td>1.00 (0.12)***</td>
<td>1.12 (0.10)***</td>
</tr>
<tr>
<td>Time trend</td>
<td>-0.04 (0.02)</td>
<td>0.04 (0.02)</td>
<td>0.003 (0.02)</td>
</tr>
<tr>
<td>Tenure dummy</td>
<td>0.4 (0.09)***</td>
<td>-</td>
<td>0.28 (0.10)***</td>
</tr>
<tr>
<td>Female dummy</td>
<td>0.16 (0.07)**</td>
<td>0.16 (0.09)**</td>
<td>0.10 (0.07)</td>
</tr>
<tr>
<td>School rank (log)</td>
<td>0.07 (0.02)***</td>
<td>0.08 (0.03)***</td>
<td>0.05 (0.02)**</td>
</tr>
<tr>
<td>PhD rank (log)</td>
<td>0.01 (0.02)</td>
<td>-0.01 (0.02)</td>
<td>0.01 (0.02)</td>
</tr>
<tr>
<td>N. Amer undergrad</td>
<td>0.11 (0.05)**</td>
<td>0.11 (0.06)*</td>
<td>0.10 (0.05)*</td>
</tr>
<tr>
<td>Publications</td>
<td>0.01 (0.12)</td>
<td>0.10 (0.13)</td>
<td>0.02 (0.13)</td>
</tr>
<tr>
<td><strong>Thresholds</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\omega): Intercept</td>
<td>6.22 (0.18)***</td>
<td>5.98 (0.29)***</td>
<td>6.36 (0.27)***</td>
</tr>
<tr>
<td>(\omega): time</td>
<td>-1.16 (0.06)***</td>
<td>-1.16 (0.08)***</td>
<td>-1.30 (0.07)***</td>
</tr>
<tr>
<td>(\omega): time(^2)</td>
<td>0.12 (0.01)***</td>
<td>0.12 (0.01)***</td>
<td>0.14 (0.01)***</td>
</tr>
<tr>
<td>(\alpha): time</td>
<td>0.71 (0.15)***</td>
<td>0.73 (0.11)***</td>
<td>0.78 (0.12)***</td>
</tr>
<tr>
<td>(\alpha): time(^2)</td>
<td>-0.08 (0.03)***</td>
<td>-0.07 (0.02)***</td>
<td>-0.08 (0.02)***</td>
</tr>
<tr>
<td>School rank (log)</td>
<td>-0.02 (0.01)</td>
<td>-0.01 (0.03)</td>
<td>-0.07 (0.02)***</td>
</tr>
<tr>
<td><strong>Error correlation</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\rho)</td>
<td>.001 (0.00)</td>
<td>-.30 (.18)</td>
<td>-0.52 (0.15)***</td>
</tr>
<tr>
<td>(\sigma_\eta)</td>
<td>-0.41 (0.18)**</td>
<td>-.34 (.22)</td>
<td>-0.13 (0.11)</td>
</tr>
<tr>
<td>(\sigma_\varepsilon)</td>
<td>0.64 (0.03)***</td>
<td>0.68 (0.05)***</td>
<td>0.62 (0.05)***</td>
</tr>
<tr>
<td><strong>Model information</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>-2923.66</td>
<td>-2930.65</td>
<td>-3854.8</td>
</tr>
<tr>
<td>Dataset</td>
<td>Top 200</td>
<td>Top 200</td>
<td>All</td>
</tr>
</tbody>
</table>

Notes:
\(^a\) Coefficient estimates are reported with standard errors in parenthesis. *, **, and *** denote significance at the .1, .05, and .01 level, respectively.

how the probability of being terminated, promoted, and resigning is affected by a change in one of the regressors. Holding all other regressors at their median value, females are 92% more likely to be fired\(^{20}\) and are 5% less likely to be promoted. As a comparison the effect

\(^{20}\) At the medial values the probability of being fired is quite low – 3.8%. Thus a 92% increase in the probability of being fired translates to a 3.5 percentage point increase in the probability of being fired.
of being male on the probability of being terminated between years five and ten is greater than having a co-authored publication in the American Economic Review (impact factor of 2.011).

Table 4.5: Effect of a change in regressors on the probability of outcomes in model 1 (%)

<table>
<thead>
<tr>
<th></th>
<th>Prob fired</th>
<th>Prob promoted</th>
<th>Prob resign</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>91.89</td>
<td>-4.42</td>
<td>-23.00</td>
</tr>
<tr>
<td>10% increase in school rank</td>
<td>0.34</td>
<td>-0.54</td>
<td>1.28</td>
</tr>
<tr>
<td>10% increase in rank of PhD School</td>
<td>0.33</td>
<td>-0.14</td>
<td>0.22</td>
</tr>
<tr>
<td>US undergrad degree</td>
<td>15.84</td>
<td>5.02</td>
<td>-15.28</td>
</tr>
<tr>
<td>Increase of 1 in lifetime impact factor adjusted publications</td>
<td>-70.88</td>
<td>14.97</td>
<td>-4.78</td>
</tr>
</tbody>
</table>

Notes:

a The value reported is the percentage change in the probability of an individual receiving a particular outcome sometime between the years five and ten, evaluated at the median values of all other regressors. The probability of termination, promotion and job exit between years five and ten at the median value of the regressors is 3.8%, 63.4% and 31.1%, respectively (note these outcomes are neither exclusive nor exhaustive, individuals may both receive tenure and resign, or conversely they may not receive tenure by their tenth year.) The large percentage changes in the probability of being fired are, in part, due to the relatively small probability of the event (at the median values of the regressors).

This result confirms earlier reduced form evidence that females face discrimination in academic promotion (Ginther and Hayes, 2003). It also speaks to the debate regarding gender differences in pay. There has been some disagreement in the literature as to whether gender differences in academic pay disappear after controlling for academic rank (Ward, 2001; Ginther and Hayes, 2003; Blackaby, Booth, and Frank, 2005). If such differences do disappear after rank is controlled for, one might be tempted to believe that there is no gender discrimination in academic pay. The current results suggest that differences in academic rank (and hence between rank salary differences) are gender biased.

The coefficients of the individual choice function lack a structural interpretation. A negative coefficient suggests an increase in the associated regressor will induce an individual to leave a department of their own volition. This implies that an increase in the regressor either gives an employee more utility from moving (by lowering moving costs, for example)
or provides other departments with utility, giving them an incentive to poach the individual from their current department. Across all models the female dummy is positive, suggesting that females are less likely to leave their department of their own volition – though the coefficient is not statistically significant in the third model. In model one, at the median values of the regressors, females are 23% less likely to voluntarily leave their department between five and ten years after their initial hiring. Given that departments were found to exhibit a male bias in promotion it is possible that departments are also more likely to recruit a male candidate, and hence the difference in the probability of voluntary job exit may reflect the preferences of outside departments. Conversely, it may be the case that males are more willing to move than females.

This is a notable result as previous literature has generally found that women have higher rates of job quits (Blau and Kahn, 1981; Keith and McWilliams, 1997). One rationalization of this result is that academic jobs tend to be sparser, and changing jobs often entails a significant geographic movement. Given that male professionals are more likely to be the primary income earner in the household, the household moving costs may, on average, be lower for males. The result is important as increased mobility is generally seen as influencing an individual’s earnings. Voluntary job changes are associated with an increase in earnings (Topel and Ward, 1992; Keith and McWilliams, 1999), furthermore the threat of job exit allows an employee to more effectively bargain with their current employer. The differential mobility of males and females may therefore provide some insight into the male-female wage differential that is found within academia (Ginther and Hayes, 2003; Ward, 2001; Blackaby, Booth, and Frank, 2005).

---

21 Bertrand, Goldin, and Katz (2010) show that male MBA graduates were more likely to have a spouse with fewer years of education (and hence lower income earning potential). Further they show that females with high income earning spouses make different labour market decisions than those who are the primary income earners in the household.

22 Indeed some commentators have suggested that the lack of mobility within academia provides employers with a degree of monopsony power that depresses wages (Ransom, 1993). However, an empirical study by Barbezat and Hughes (2001) finds the opposite result; using a national survey of academics the authors find the number of jobs a professor has held has no significant impact on their earnings unless the professor has held four or more jobs, in which case the effect is negative. The weakness of this analysis is that it does not discriminate between voluntary and involuntary job changes, which Bartel and Borjas (1981) demonstrate
The result is also important as it suggests that the relative rate of voluntary job exits between genders is different than the relative rate of overall job exits. Hence, studies that do not differentiate between these two type of exits may contain significant bias. In the current case, males are more likely to exit a job voluntarily, and less likely to leave involuntarily, these two effects offset each other making the overall rate of job exit appear more balanced.

A potential issue in the estimation of the thresholds is the impact of parental leave, which is more commonly taken by females. At most institutions the “tenure clock” is paused during parental leaves. To model this process an exogenous probability of females taking a year off is introduced into the model. In each year the probability that a female will take a year off is $\lambda$. When a year off is taken, the thresholds for promotion and firing become $\omega(t - t^* - 1)$ and $\alpha(t - t^* - 1)$. For the sake of simplicity, the model allows for only one leave of absence and the probability of that year of absence is constant over time. The $\lambda$ term was found to equal zero, implying that very few females take a year off before attaining tenure. This accords with the data from the Survey of Doctorate Recipients. In this survey the probability that a female PhD graduate (in any field of science or engineering) has a child within five years of her degree is 7%, while the probability of having a child between years five and ten is 4%.

Outside of the issue of gender, the estimates also contain some insight into the relative value that employers place on an employee’s characteristics in the tenure process. As expected, departments place a high value on publications – a co-authored publication in the American Economic Review increases an individual’s likelihood of achieving tenure between their fifth and tenth year by 15%. All models find that individuals with a PhD from a lower ranked department are more likely to be promoted, though the coefficient is only statistically significant in the second model.\footnote{It is possible that this reflects some degree of endogeneity; in order to be hired at a top 200 school, an individual from a lower ranked institution may require a higher draw of unobservables. If these unobservables correlated over time, then one might expect individuals from lower ranked institutions to be more likely to receive promotion.}
It is a common assumption that higher ranked schools have higher standards for granting tenure, hence one would expect high ranked schools to have higher thresholds. This is true in all models, though the effect is only statistically significant in model 3.

As mentioned before the coefficients estimated for equation 4.4 lack a structural interpretation as they reflect both the preferences of individuals and departments. The results show that individuals at higher ranked schools are more likely to leave of their own volition. One explanation for this result is that professors at higher ranked schools are more likely to be poached by other institutions. Though this story is weakened when one considers that an individual’s publications have no significant impact on their rates of voluntary job exit. Individuals who completed their undergraduate degree outside the United States, are more likely to leave a department of their own volition. An examination of the data reveals that a substantial portion of these individuals returned to the same country in which they pursued their undergraduate degree.

4.7 Conclusion

This paper analyzes the impact of gender in academic hiring and promotion. To capture the impact of gender the decision making of employers and employees, two structural models are introduced. The initial hiring decision is modelled as a two-sided matching game and estimated using a weighted maximum score estimator. Post-hiring outcomes were estimated using a two-sided probit model.

It was found that departments have no significant gender preference in their initial hiring decision, but have a significant male bias in promotion and firing decisions. These results should guide future efforts in redressing the gender imbalance in economics. It appears that recent efforts to encourage gender parity in hiring have had the desired impact without causing “reverse discrimination”. Furthermore, the results would suggest that those concerned with gender discrimination would do better to shift attention from the initial hiring decision,
to post-hiring decisions, such as tenure and promotion.

It was also demonstrated that males and females have different preferences. In the initial hiring decision females placed a higher value on the fit of their department, measured by the rank of the department in the individuals subject of interest, whereas males placed a higher value on the departments overall rank. In post-hiring outcomes it was shown that men are more likely to exit their job voluntarily, though it is unclear whether this is due to differences in preferences, or differences in outside opportunities. Regardless, the results point to the importance of accounting for the two-sided nature of labour market decisions.
Bibliography


Appendix A

Appendix to Chapter 2

A.1 Proofs

Proof of lemma 2.2 (continued). The first part of this proof is found in the body of the paper. It was left to be shown that when the individual plays $A$ conditional on preference $a$ there is a unique equilibrium (either $\sigma_1$, $\sigma_2$, or $\sigma_3$.) To prove this, define $\sigma_{1-3}(r)$ as a strategy in which the individual plays $B$ with probability $r \in [0,1]$ conditional on receiving $(b, \gamma \in \Gamma^3)$ and plays $A$ with probability 1 otherwise. Thus, $\sigma_1 = \sigma_{1-3}(0)$, $\sigma_2 = \sigma_{1-3}(x)$ where $x \in (0,1)$, and $\sigma_3 = \sigma_{1-3}(1)$.

Consider the individual’s best response to any value of $r$, conditional on receiving $(b, \gamma \in \Gamma^3)$. The individual’s payoff from playing $B$ relative to $A$ is,

\[ u(B, b, \gamma \in \Gamma^3, \sigma_{1-3}(r)) - u(A, b, \gamma \in \Gamma^3, \sigma_{1-3}(r)) = \]

\[ \beta(P(H|K, B, \sigma_{1-3}(r)) - P(H|K, A, \sigma_{1-3}(r))) + \delta. \]

(A.1)

Note that the individual is never fired when he has a draw of $\gamma \in \Gamma^3$ – hence wages are differenced out of equation A.1. Evidently, the individual plays $B$ when equation A.1 is greater than zero, plays $A$ when equation A.1 is less than zero, and is indifferent between the two actions when equation A.1 is equal to zero.

The term $P(H|K, B, \sigma_{1-3}(r))$ is the probability that the individual is of high ability conditional on playing $B$ and being kept on under the strategy $\sigma_{1-3}(r)$. Writing this
probability in terms of the model primitives,

\[ P(H|K, B, \sigma_{1-3}(r)) = \frac{P(B, K, H|\sigma_{1-3}(r))}{P(B, K|\sigma_{1-3}(r))} = \frac{P(B, K, H|\sigma_{1-3}(r))}{P(B, K, L|\sigma_{1-3}(r)) + P(B, K, H|\sigma_{1-3}(r))} = \frac{\Gamma_H^3(1-q_H)r}{\Gamma_L^3(1-q_L)r + \Gamma_H^3(1-q_H)r} \]  

(A.2)

Additional discussion surrounding the derivation of this equation can be found after equation 2.8 in the main body of the paper. The terms in the last equality are illustrated in figure 2.2 and panel (b) of figure 2.3.

The term, \( P(H|K, A, \sigma_{1-3}(r)) \), is similarly defined, hence the expression \( P(H|K, B, \sigma_{1-3}(r)) - P(H|K, A, \sigma_{1-3}(r)) \) can be written as,

\[ \frac{\Gamma_H^3(1-q_H)r}{\Gamma_L^3(1-q_L)r + \Gamma_H^3(1-q_H)r} = \frac{\Gamma_H^3(1-q_H)(1-r)}{\Gamma_L^3(1-q_L)(1-r) + \Gamma_H^3(1-q_H)(1-r).} \]  

(A.3)

The derivative of equation A.3 with respect to \( r \) (not shown here) is of the same sign as the term itself. Hence the sign of equation A.1 across the support of \( r \) falls into one of three cases:

Case 1: Equation A.1 is negative for all \( r \in [0, 1] \). In this case the individual plays \( A \) with probability 1 conditional on receiving \((b, \gamma \in \Gamma^3)\) regardless of \( r \); \( \sigma_2 \) and \( \sigma_3 \) cannot be equilibria.

Case 2: Equation A.1 is positive for all \( r \in [0, \bar{r}) \), negative for all \( r \in (\bar{r}, 1] \) and equal to zero for \( r = \bar{r} \). In this case, conditional on receiving \((b, \gamma \in \Gamma^3)\), the individual will play \( B \) with probability one when \( r < \bar{r} \) (ruling out \( \sigma_1 \) as an equilibrium), \( A \) with probability one when \( r > \bar{r} \) (ruling out \( \sigma_3 \) as an equilibrium), and a mixed strategy when \( r = \bar{r} \) (this is the strategy profile \( \sigma_3 \) and it is the unique equilibrium.)
Case 3: Equation A.1 is positive for all \( r \in [0, 1] \). In this case the individual plays \( B \) with probability one conditional on receiving \((b, \gamma \in \Gamma^3)\) regardless of \( r \); \( \sigma_1 \) and \( \sigma_2 \) cannot be equilibria.

\[ \square \]

**Proof of proposition 2.1.** To prove this proposition it suffices to show that \( \lim_{q_H \to 1} \left( P(H|K, B, \sigma_3) - P(H|K, A, \sigma_3) \right) > 0 \). This ensures that \( \sigma_1 \) and \( \sigma_2 \) are not equilibria and that \( \sigma_3 \), if it is an equilibrium, is a countermarking equilibrium.

Consider the limit of \( P(H|K, B, \sigma_3) \) (shown in equation 2.8) as \( q_H \) goes to one,

\[
\lim_{q_H \to 1} P(H|K, B, \sigma_3) = \lim_{q_H \to 1} \frac{(1 - q_H)\Gamma^3_H}{(1 - q_H)\Gamma^3_H + (1 - q_L)\Gamma^3_L}.
\]  

(A.4)

Substituting the definition of \( \Gamma^i_k \) into equation (A.4) yields,

\[
\lim_{q_H \to 1} P(H|K, B, \sigma_3) = \lim_{q_H \to 1} \frac{(1 - q_H) \int_{\gamma_B}^{\sup(\gamma|H)} g(\gamma|H)d\gamma}{(1 - q_H) \int_{\gamma_B}^{\sup(\gamma|H)} g(\gamma|H)d\gamma + (1 - q_L) \int_{\gamma_B}^{\sup(\gamma|L)} g(\gamma|L)d\gamma}.
\]  

(A.5)

From equation 2.6, \( \gamma_B \) is the solution to, \( \frac{g(\gamma_B|H)}{g(\gamma_B|L)} = \frac{\pi_H(1-q_L)}{\pi_L(1-q_H)} \). Therefore, when \( q_H \) goes to one, \( g(\gamma_B|L) = 0 \) (assumption EP ensures such a draw of \( \gamma \) exists). Further, by the quasi-concavity of \( g, \lim_{q_H \to 1} \int_{\gamma_B}^{\sup(\gamma)} g(\gamma|L)d\gamma = 0 \). Hence both the numerator and denominator on the right hand side of equation A.5 go to zero as \( q_H \) goes to one. By L’Hôpital’s rule,

\[
\lim_{q_H \to 1} P(H|K, B, \sigma_3) = \lim_{q_H \to 1} \frac{-\int_{\gamma_B}^{\sup(\gamma)} g(\gamma|H)d\gamma - \frac{\partial g(\gamma_B|H)}{\partial q_H} (1 - q_H)g(\gamma_B|H)}{-\int_{\gamma_B}^{\sup(\gamma)} g(\gamma|H)d\gamma - \frac{\partial g(\gamma_B|H)}{\partial q_H} (1 - q_H)g(\gamma_B|H) - (1 - q_L)g(\gamma_B|L)\frac{\partial g(\gamma_B|L)}{\partial q_H}}.
\]  

(A.6)

Given that \( \lim_{q_H \to 1} g(\gamma_B|L) = 0 \), the numerator and denominator on the right hand side of equation A.6 are equivalent; \( \lim_{q_H \to 1} P(H|K, B, \sigma_3) = 1 \).
Now consider the limit of \( P(H|K, A, \sigma_3) \) as \( q_H \) goes to one,

\[
\lim_{q_H \to 1} P(H|K, A, \sigma_3) = \lim_{q_H \to 1} \frac{\Gamma^2_H + q_H \Gamma^3_H}{\Gamma^2_H + q_H \Gamma^3_H + \Gamma^2_L + q_L \Gamma^3_L} < 1. \tag{A.7}
\]

The inequality follows from the fact that \( \Gamma^2_L \) is greater than zero even as \( q_H \) goes to one (as \( \gamma_A \) is always strictly less than \( \gamma_B \)), hence the numerator is less than the denominator. Combining equation A.5 and A.7,

\[
\lim_{q_H \to 1} (P(H|K, B, \sigma_3) - P(H|K, A, \sigma_3)) > 0, \tag{A.8}
\]

regardless of the value of any other parameters.

Proof of proposition 2.2. Suppose the individual uses the following strategy, play \( B \) conditional on receiving \( \gamma \in \Gamma^3 \) (regardless of preference) and play \( A \) conditional on \( \gamma \in \Gamma^2 \) (again regardless of preference). To show the conditions under which this strategy represents an equilibrium consider the individual’s incentive to deviate. If the individual receives \( \gamma \in \Gamma^2 \) he will be fired if he plays \( B \) (by the off-equilibrium beliefs of insiders), hence he has no incentive to deviate. If the individual receives \((a, \gamma \in \Gamma^3)\), his payo from playing \( B \) is \( m + \beta \frac{\Gamma^3_H}{\Gamma^3_L + \Gamma^3_H} - \delta \) and his payo from playing \( A \) is \( m + \beta \frac{\Gamma^2_H}{\Gamma^2_L + \Gamma^2_H} \). Hence there is no incentive for the individual to deviate if,

\[
\delta / \beta < \frac{\Gamma^3_H}{\Gamma^3_L + \Gamma^3_H} - \frac{\Gamma^2_H}{\Gamma^2_L + \Gamma^2_H}. \tag{A.9}
\]

The right hand side of equation A.9 is positive by assumption MLRP (the individual is more likely to have high ability if he has a draw of \( \gamma \in \Gamma^3 \) than if he has a draw of \( \gamma \in \Gamma^2 \)), therefore there exists a value of \( \delta / \beta \) that ensures the individual will play \( B \) conditional on \((a, \gamma \in \Gamma^3)\). If the inequality in equation A.9 holds then the individual will also have no incentive to deviate conditional on receiving \((b, \gamma \in \Gamma^3)\), as the individual has a stronger incentive to play \( B \) when he has preference \( b \).
Proof of lemma 2.3. Define \( \tilde{\Gamma}^3 \) as the set \((\gamma_B + s, \sup(\gamma))\); when the individual has a draw of \( \tilde{\gamma} \in \tilde{\Gamma}^3 \) he will not be fired regardless of his action or draw of \( \epsilon \). Also, define \( \tilde{\sigma}_{1-3}(r) \) as a strategy in which the individual plays \( B \) with probability \( r \in [0,1] \) conditional on receiving \((b, \tilde{\gamma} \in \tilde{\Gamma}^3)\), and \( A \) otherwise.

The proof can now follow the same argument as lemma 2.2, with one exception: under \( \tilde{\sigma}_{1-3}(1) \) the individual may wish to play \( B \) conditional on having preference \( b \) and a draw of \( \tilde{\gamma} \in (\gamma_B - s, \gamma_B + s) \). When the individual has a draw of \( \tilde{\gamma} \) in this region he will be fired with probability in \((0,1)\) if he plays \( B \), however the individual will be willing to exchange some probability of being fired in order to act consistent with his preference.

Consider the individual’s expected utility from playing \( B \) relative to playing \( A \) conditional on receiving \((b, \tilde{\gamma} \in (\gamma_B - s, \gamma_B + s))\),

\[
Eu(B, b, \tilde{\gamma} \in (\gamma_B - s, \gamma_B + s), \tilde{\sigma}_3) - Eu(A, b, \tilde{\gamma} \in (\gamma_B - s, \gamma_B + s), \tilde{\sigma}_3) = Pr(\gamma > \gamma_B | \tilde{\gamma})(m + \beta P(H|K, B, \tilde{\sigma}_3)) - (m + \beta P(H|K, A, \tilde{\sigma}_3) - \delta). \tag{A.10}
\]

The term \( Pr(\gamma > \gamma_B | \tilde{\gamma}) \) is the probability that the individual is hired conditional on playing \( B \) – by the uniform distribution of \( \epsilon \) this probability is increasing in \( \tilde{\gamma} \) and has range \((0,1)\) over the domain \((\gamma_B - s, \gamma_B + s)\). The individual will always be hired conditional on playing \( A \) based on the assumption that \( \gamma_B - s > \gamma_A + s \).

Under \( \tilde{\sigma}_3 \), \( m + \beta P(H|K, B, \tilde{\sigma}_3) < m + \beta P(H|K, A, \tilde{\sigma}_3) - \delta \); the individual derives more utility from \( B \) than from \( A \) when he has preference \( b \) and is kept on. Given the range of \( Pr(\gamma > \gamma_B | \tilde{\gamma}) \), for any particular values of \( P(H|K, A, \tilde{\sigma}_3) \) and \( P(H|K, B, \tilde{\sigma}_3) \) there will be a draw of \( \tilde{\gamma} \in (\gamma_B - s, \gamma_B + s) \) that makes the individual indifferent between playing \( A \) and playing \( B \). Hence in equilibrium (if one exists) there will be a draw of \( \tilde{\gamma} \), call it \( \tilde{\gamma}_\perp \), such that the individual plays \( B \) if, and only if, he has preference \( b \) and a draw of \( \tilde{\gamma} > \tilde{\gamma}_\perp \). The term \( \tilde{\gamma}_\perp \) is implicitly defined by setting equation A.10 equal to zero and solving for \( \tilde{\gamma} \) – after
rearranging this results in the following equation,

$$Pr(\gamma > \gamma_B|\tilde{\gamma}) = \frac{m + \beta P(H|K,A,\tilde{\sigma}_3) - \delta}{m + \beta P(H|K,B,\tilde{\sigma}_3)}.$$  \hfill (A.11)

The existence of the equilibrium is, however, complicated by the fact that the probabilities $P(H|K,B,\tilde{\sigma}_3)$ and $P(H|K,A,\tilde{\sigma}_3)$ are functions of $\tilde{\gamma}$. Since the left-hand side of equation A.11 is continuous with range $[0,1]$, a fixed point will exist if the right hand side of equation A.11 is continuous with a range in $[0,1]$.

To determine the range of the right hand side, recall that the denominator is greater than the numerator and therefore the range must lie within $(0,1)$. To show continuity, use the same argument as equation 2.8 to define $P(H|K,J,\tilde{\sigma}_3)$ as,

$$P(H|K,J,\tilde{\sigma}_3) = \frac{P(J,K|H,\tilde{\sigma}_3)}{P(J,K|H,\tilde{\sigma}_3) + P(J,K|L,\tilde{\sigma}_3)}. \hfill (A.12)$$

Using $F$ to denote the (uniform) distribution of $\epsilon$, the probability the individual plays $B$ and is kept on, conditional on having the $i$th ability is,

$$P(B,K|i,\tilde{\sigma}_3) = \int_{\gamma_B}^{\sup(\gamma|i)} g(\gamma|i)(1 - q_i)F(\gamma - \tilde{\gamma})d\gamma. \hfill (A.13)$$

Equation A.13 integrates over the region of $\gamma$ for which the individual would be kept on if he played $B$: $(\gamma_B, \sup(\gamma|i))$. The integrand is the probability density of $\gamma$ multiplied by the probability the individual has preference $b$, and the probability the individual has a draw of $\tilde{\gamma}$ greater than $\tilde{\gamma}$. The same probability for action $A$ is,

$$P(A,K|i,\tilde{\sigma}_3) = \int_{\gamma_A}^{\sup(\gamma|i)} g(\gamma|i)(1 - q_i) \left(1 - F(\gamma - \tilde{\gamma})\right) d\gamma + q_i(\Gamma_i^2 + \Gamma_i^3). \hfill (A.14)$$

The first term in equation A.14 integrates over the region of $\gamma$ for which the individual would be kept on if he played $A$: $(\gamma_A, \sup(\gamma|i))$. The integrand is the probability density of
\(\gamma\) multiplied by the probability the individual has preference \(b\), and the probability the individual has a draw of \(\tilde{\gamma}\) less than \(\tilde{\gamma}\). The second additive term represents the probability the individual has preference \(a\) and is kept on – recall the individual always plays \(A\) conditional on preference \(a\) under \(\tilde{\sigma}_3\).

Equations A.13 and A.14 are continuous functions of \(\tilde{\gamma}\) and hence a solution exists to equation A.11. Note for a given set of parameters there may be multiple solutions to equation A.11, although as \(s\) goes to zero, \(\tilde{\gamma}\) converges to \(\gamma_B\) (as \(\tilde{\gamma} \in (\gamma_B - s, \gamma_B + s)\)). \(\square\)

**Proof of proposition 2.3.** In this proof it is to be shown that when \(m > \frac{1}{3} H(1 - q_H) - \frac{2(\Gamma_H^2 + q_H \Gamma_L^2)}{\Gamma_H^2 + q_H \Gamma_H^2 + \Gamma_L^2 + q_L \Gamma_L^2} + 2\delta\),

\[
\lim_{s \to 0} \frac{d}{ds} P(H|B, J, \tilde{\sigma}_3) - P(H|B, J, \tilde{\sigma}_3) > 0. \tag{A.15}
\]

Substituting equation A.12 into equation A.15 yields,

\[
A.17 \lim_{s \to 0} \frac{d}{ds} \left( \frac{P(B, K|H, \tilde{\sigma}_3)}{P(B, K|H, \tilde{\sigma}_3) + P(B, K|\tilde{\sigma}_3)} - \frac{P(A, K|H, \tilde{\sigma}_3)}{P(A, K|H, \tilde{\sigma}_3) + P(A, K|\tilde{\sigma}_3)} \right) > 0. \tag{A.16}
\]

Note that a change in \(s\) will have two effects on equation A.17: a direct effect and an indirect effect through \(\tilde{\gamma}\). Consider the direct and indirect effects separately. Henceforth, the terms \(K\) and \(\tilde{\sigma}_3\) will be dropped from the probabilities – \(P(J|i) \equiv P(J, K|i, \tilde{\sigma}_3)\).

**Step 1: Direct effect.** The direct effect of a change in \(s\) is,

\[
\lim_{s \to 0} \left( \frac{\partial}{\partial s} P(B|H) \right) - \frac{\partial}{\partial s} P(A|H) > 0. \tag{A.17}
\]

Consider the two additive parts of equation A.17 separately.

**Step 1a: The sign of** \(\lim_{s \to 0} \frac{\partial}{\partial s} \left( \frac{P(B|H)}{P(B|L) + P(B|H)} \right)\). To determine the sign, first solve for \(\lim_{s \to 0} P(B|i)\) and \(\lim_{s \to 0} \frac{\partial}{\partial s} P(B|i)\). Rewrite equation A.13 substituting in the uniform distribution for \(F\),
noting that $F(x) = 1$ for any $x > 2s$, and that $\tilde{\gamma} \in (\gamma_B - s, \gamma_B + s)$,

$$P(B|i) = \int_{\gamma_B}^{\tilde{\gamma} + 2s} g(\gamma|i)(1 - q_i)\frac{\gamma - \tilde{\gamma}}{2s}d\gamma + \int_{\tilde{\gamma} + 2s}^{\text{sup}(\gamma)} g(\gamma|i)(1 - q_i)d\gamma$$  \hspace{1cm} (A.18)

Recalling that $\lim_{s \to 0} \tilde{\gamma} = \gamma_B$, the limit of equation A.18 as $s$ goes to zero is,

$$\lim_{s \to 0} P(B|i) = \int_{\gamma_B}^{\text{sup}(\gamma|i)} g(\gamma|i)(1 - q_i)d\gamma > 0.$$ \hspace{1cm} (A.19)

The derivative of equation A.18 with respect to $s$ is,

$$\frac{\partial P(B|i)}{\partial s} = -\frac{1}{4s^2} \int_{\gamma_B}^{\tilde{\gamma} + 2s} g(\gamma|i)(1 - q_i)(\gamma - \tilde{\gamma})d\gamma.$$ \hspace{1cm} (A.20)

Using L'Hôpital’s rule the limit of equation A.20 as $s \to 0$ is,

$$\lim_{s \to 0} \frac{\partial P(B|i)}{\partial s} = -\frac{1}{2} g(\gamma_B|i)(1 - q_i) < 0.$$  \hspace{1cm} (A.21)

By the quotient rule and the signs of equations A.19 and A.21,

$$\lim_{s \to 0} \left( \text{sgn} \left( \frac{\partial}{\partial s} \frac{P(i|H)}{P(i|L) + P(i|H)} \right) \right) = \lim_{s \to 0} \left( \text{sgn} \left( \frac{P(i|H)}{P(i|L)} - \frac{\partial P(i|H)/\partial s}{\partial P(i|L)/\partial s} \right) \right).$$ \hspace{1cm} (A.22)

Substituting in equations A.19 and A.21,

$$\lim_{s \to 0} \left( \frac{P(B|H)}{P(B|L)} - \frac{\partial P(B|H)/\partial s}{\partial P(B|L)/\partial s} \right) = 1 - q_H \left( \frac{\int_{\gamma_B}^{\text{sup}(\gamma|H)} g(\gamma|H)d\gamma}{\int_{\gamma_B}^{\text{sup}(\gamma|L)} g(\gamma|L)d\gamma} - \frac{g(\gamma_B|H)}{g(\gamma_B|L)} \right) > 0.$$ \hspace{1cm} (A.23)

The inequality stems from assumption MLRP; by this assumption $\frac{g(\gamma_1|H) + g(\gamma_2|H)}{g(\gamma_1|L) + g(\gamma_2|L)} > \frac{g(\gamma_B|H)}{g(\gamma_B|L)}$ for any $\gamma_1, \gamma_2 > \gamma_B$. By the same logic, $\frac{\int_{\gamma_B}^{\text{sup}(\gamma|H)} g(\gamma|H)d\gamma}{\int_{\gamma_B}^{\text{sup}(\gamma|L)} g(\gamma|L)d\gamma} > \frac{g(\gamma_B|H)}{g(\gamma_B|L)}$.

**Step 1b:** The sign of $\lim_{s \to 0} \frac{\partial}{\partial s} \left( \frac{P(A|H)}{P(A|L) + P(A|H)} \right)$. Once again, to determine the sign first solve
for \( \lim_{s \to 0} P(A|i) \) and \( \lim_{s \to 0} \frac{\partial}{\partial s} P(A|i) \). Rewriting equation A.14 yields,

\[
P(A|i) = \int_{\gamma_A}^{\gamma_B - s} g(\gamma|i)(1 - q_i)d\gamma + \int_{\gamma_B - s}^{\gamma_B + s} g(\gamma|i)(1 - q_i) \left( 1 - \frac{\gamma - \tilde{\gamma}}{2s} \right) d\gamma + q_i(\Gamma_i^2 + \Gamma_i^3). \tag{A.24}
\]

The limit of equation A.24 as \( s \) goes to zero is,

\[
\lim_{s \to 0} P(A|i) = \int_{\gamma_A}^{\gamma_B} g(\gamma)(1 - q_i)d\gamma + q_i(\Gamma_i^2 + \Gamma_i^3) > 0. \tag{A.25}
\]

The derivative of equation A.24 with respect to \( s \) is,

\[
\frac{\partial P(A|i)}{\partial s} = g(\gamma_B - s|i)(1 - q_i) - g(\gamma_B - s|i)(1 - q_i) \left( 1 - \frac{\gamma_B - s - \tilde{\gamma}}{2s} \right) + \]

\[
g(\gamma_B + s|i)(1 - q_i) \left( 1 - \frac{\gamma_B + s - \tilde{\gamma}}{2s} \right) + \int_{\gamma_B - s}^{\gamma_B + s} g(\gamma|i)(1 - q_i) \left( 1 + \frac{\gamma - \tilde{\gamma}}{2s^2} \right).
\tag{A.26}
\]

Now consider the limit of the equation A.26 as \( s \) goes to zero. As \( s \) approaches zero, \( \gamma_B + s = \gamma_B - s \) — imposing this equality, the first two lines on the right hand side of equation A.26 cancel out. Further the third line on the right hand side also goes to zero, as the upper and lower limits of integration are equal. Hence, \( \lim_{s \to 0} \frac{\partial P(A|i)}{\partial s} = 0 \) and therefore \( \lim_{s \to 0} \frac{\partial}{\partial s} \left( \frac{P(A|H)}{P(A|L) + P(A|H)} \right) = 0 \).

Combining the results from step 1a and step 1b shows the direct effect of \( s \) to be positive.

**Step 2: indirect effect.** The indirect effect of a change in \( s \) is,

\[
\lim_{s \to 0} \frac{\partial \tilde{\gamma}}{\partial s} \frac{d}{d\tilde{\gamma}} \left( \frac{P(B|H)}{P(B|L) + P(B|H)} - \frac{P(A|H)}{P(A|L) + P(A|H)} \right). \tag{A.27}
\]

Split the indirect effect into two parts: \( \lim_{s \to 0} \frac{d}{d\tilde{\gamma}} \left( \frac{P(B|H)}{P(B|L) + P(B|H)} - \frac{P(A|H)}{P(A|L) + P(A|H)} \right) \) and \( \lim_{s \to 0} \frac{\partial \tilde{\gamma}}{\partial s} \).

**Step 2a:** The sign of \( \lim_{s \to 0} \frac{d}{d\tilde{\gamma}} \left( \frac{P(B|H)}{P(B|L) + P(B|H)} - \frac{P(A|H)}{P(A|L) + P(A|H)} \right) \). Taking the derivative of
equation A.18 with respect to $\tilde{\gamma}$ as $s$ goes to zero yields,

$$
\lim_{s \to 0} \frac{dP(B|\tilde{\gamma})}{d\tilde{\gamma}} = -g(\gamma_B|\tilde{\gamma})(1 - q_i) < 0.
$$

(A.28)

By the quotient rule and the signs of equations A.19 and A.28,

$$
\text{sgn} \left( \frac{d}{d\tilde{\gamma}} \left( \frac{P(B|H)}{P(B|H) + P(B|L)} \right) \right) = \text{sgn} \left( \lim_{s \to 0} \left( \frac{P(B|H)}{P(B|L)} - \partial P(B|H)/\partial \tilde{\gamma} \right) \right)
$$

(A.29)

Substituting in equations A.19 and A.28,

$$
\lim_{s \to 0} \left( \frac{P(B|H)}{P(B|L)} - \frac{dP(B|H)/d\tilde{\gamma}}{dP(B|L)/d\tilde{\gamma}} \right) = \frac{1 - q_H}{1 - q_L} \left( \frac{\int_{\gamma_B}^{\sup(\gamma|H)} g(\gamma|H)d\gamma}{\int_{\gamma_B}^{\sup(\gamma|L)} g(\gamma|L)d\gamma} - \frac{g(\gamma_B|H)}{g(\gamma_B|L)} \right) > 0.
$$

(A.30)

The inequality follows by the same logic as equation A.23.

Similarly, the derivative of equation A.24 with respect to $\tilde{\gamma}$ as $s$ goes to zero yields,

$$
\lim_{s \to 0} \frac{dP(A|\tilde{\gamma})}{d\tilde{\gamma}} = g(\gamma_B|\tilde{\gamma})(1 - q_i) > 0.
$$

(A.31)

By the quotient rule and the signs of equations A.25 and A.31,

$$
\text{sgn} \left( \lim_{s \to 0} \frac{d}{d\tilde{\gamma}} \left( \frac{P(A|H)}{P(A|H) + P(A|L)} \right) \right) = \text{sgn} \left( \lim_{s \to 0} \left( \frac{dP(A|H)/d\tilde{\gamma}}{dP(A|L)/d\tilde{\gamma}} - \frac{P(A|H)}{P(A|L)} \right) \right)
$$

(A.32)

Substituting in equations A.25 and A.31,

$$
\lim_{s \to 0} \left( \frac{dP(A|H)/d\tilde{\gamma}}{dP(A|L)/d\tilde{\gamma}} - \frac{P(A|H)}{P(A|L)} \right) = \frac{(1 - q_H)g(\gamma_B|H)}{(1 - q_L)g(\gamma_B|L)} \left( \frac{\int_{\gamma_A}^{\gamma_H} g(\gamma)d\gamma + \int_{\gamma_B}^{\gamma_L} g(\gamma)q_Hd\gamma}{\int_{\gamma_A}^{\gamma_H} g(\gamma)d\gamma + \int_{\gamma_B}^{\gamma_L} g(\gamma)q_Ld\gamma} \right) < 0.
$$

(A.33)

To understand the inequality in equation A.33 note that by the definition of $\gamma_A$ and $\gamma_B$,

$$
\frac{(1 - q_H)g(\gamma_B|H)}{(1 - q_L)g(\gamma_B|L)} = \frac{g(\gamma_A|H)}{g(\gamma_A|L)}.
$$

(A.34)
Using the same logic as equation A.23, assumption MLRP ensures that the equation is negative. The signs of equations A.30 and A.33 together imply,
\[
\lim_{s \to 0} \frac{d}{ds} \left( \frac{P(B|H)}{P(B|L)+P(B|H)} - \frac{P(A|H)}{P(A|L)+P(A|H)} \right) > 0
\]

**Step 2b: The sign of** \( \lim_{s \to 0} \frac{d\gamma}{ds} \). **When there is no uncertainty in the model (when** \( s = 0 \), \( \tilde{\gamma} = \gamma_B \). To determine whether \( \tilde{\gamma} \) is increasing in \( s \), one need only analyze the incentives of the individual if he has a draw of \( \tilde{\gamma} = \gamma_B \) when the model contains an infinitesimal amount of uncertainty. If the individual would prefer to play \( A \) conditional on receiving \( \gamma_B \) then \( \tilde{\gamma} \) must lie above \( \gamma_B \) – implying \( \lim_{s \to 0} \frac{d\gamma}{ds} > 0 \).

For a small value of \( s \) the probability the individual will be kept on conditional on receiving \( \tilde{\gamma} = \gamma_B \) and playing \( B \) is one half (as there is a .5 probability that his draw of \( \gamma \) is above \( \gamma_B \)) conversely the probability the individual is kept on if he plays \( A \) is one. Hence the individual’s expected utility from playing \( B \) is
\[
\frac{1}{2} (m + \beta P(H|B, K, \tilde{\sigma}_3)), \tag{A.35}
\]
and the individual’s expected utility from playing \( A \) is
\[
m + \beta P(H|A, K, \tilde{\sigma}_3) - \delta. \tag{A.36}
\]

The individual will therefore play \( A \) conditional on receiving \( (b, \tilde{\gamma} = \gamma_B) \) if,
\[
m > \beta (P(H|B, K, \tilde{\sigma}_3) - 2P(H|A, K, \tilde{\sigma}_3)) + 2\delta. \tag{A.37}
\]

The right hand side of the equation may be less than or greater than zero – hence the condition is not necessarily binding. Equation A.37 can be written in terms of the model primitives by noting that \( \lim_{s \to 0} P(H|K, J, \tilde{\sigma}_3) = P(H|K, J, \sigma_3) \). Substituting in equations
2.8 and 2.9,

\[ m > \beta \left( \frac{\Gamma^3_H (1 - q_H)}{\Gamma^3_H (1 - q_H) + \Gamma^2_L (1 - q_L)} - \frac{2(\Gamma^2_H + q_H \Gamma^3_H)}{\Gamma^2_H + q_H \Gamma^3_H + \Gamma^2_L + q_L \Gamma^3_L} \right) + 2\delta. \]  

(A.38)

Hence if \( m \) is large enough \( \lim_{s \to 0} \frac{d\tilde{\gamma}}{ds} > 0 \). Combining the results of step 2a and step 2b, the direct effect is positive. Given that both the indirect and direct effect of an increase in \( s \) is positive, the total effect is positive.

\( \square \)

A.2 Comparative statics for section 2.3

To get comparative static results for equation 2.10 set the left hand side of the equation equal to \( \chi \) and substitute in the definition of \( \Gamma^k_i \),

\[ \chi = \frac{(1 - q_H)(1 - G(\gamma_B|H))}{(1 - q_L)(1 - G(\gamma_B|L)) + (1 - q_H)(1 - G(\gamma_B|H))} \\
- \frac{q_H(1 - G(\gamma_B|H)) + G(\gamma_B|H) - G(\gamma_A|H)}{q_H(1 - G(\gamma_B|H)) + G(\gamma_B|H) - G(\gamma_A|H) + q_L(1 - G(\gamma_B|L)) + G(\gamma_B|L) - G(\gamma_A|L)}, \]

where \( G \) is the CDF of \( \gamma \).

**A change in the value of \( \gamma_B \).** Consider the two fractional terms in \( \chi \) separately. Note that to sign the derivative of a fraction, \( N(x)/D(x) \), one need only evaluate the sign of \( N'(x)D(x) - D'(x)N(x) \). Using this identity and rearranging, the derivative of the first term in \( \chi \) taken with respect to \( \gamma_B \) has the same sign as,

\[ \frac{\{1 - G(\gamma_B|H)\}}{(1 - G(\gamma_B|L))} - \frac{g(\gamma_B|H)}{g(\gamma_B|L)} > 0. \]  

(A.40)

The inequality stems from assumption MLRP; by this assumption \( \frac{g(\gamma_1|H) + g(\gamma_2|H)}{g(\gamma_1|L) + g(\gamma_2|L)} > \frac{g(\gamma|H)}{g(\gamma|L)} \) for any \( \gamma_1, \gamma_2 > \gamma_B \). By the same logic, \( \frac{1 - G(\gamma_B|H)}{1 - G(\gamma_B|L)} > \frac{g(\gamma_B|H)}{g(\gamma_B|L)} \).
The derivative of the second term in $\chi$ with respect to $\gamma_B$ has the same sign as,

$$
\frac{q_H(1 - G(\gamma_B|H)) + G(\gamma_B|H) - G(\gamma_A|H)}{q_L(1 - G(\gamma_B|L)) + G(\gamma_B|L) - G(\gamma_A|L)} - \frac{(1 - q_H)g(\gamma_B|H)}{(1 - q_L)g(\gamma_B|L)} > 0. \tag{A.41}
$$

To understand the sign of equation A.41 note that by the definition of $\gamma_A$ and $\gamma_B$, 

$$
\frac{g(\gamma_A|H)}{g(\gamma_A|L)} = \frac{1-q_H}{1-q_L}g(\gamma_B|H) \text{ and hence the inequality once again holds by assumption MLRP. Given that both of its additive terms are increasing in $\gamma_B$, $\chi$ must be increasing in $\gamma_B$.}
$$

**A change in the value of $\gamma_A$.** The first term in $\chi$ is not a function of $\gamma_A$. The derivative of the second term in $\chi$ with respect to $\gamma_A$ is of the same sign as,

$$
\frac{g(\gamma_A|H)}{g(\gamma_A|L)} - \frac{q_H(1 - G(\gamma_B|H)) + G(\gamma_B|H) - G(\gamma_A|H)}{q_L(1 - G(\gamma_B|L)) + G(\gamma_B|L) - G(\gamma_A|L)} < 0, \tag{A.42}
$$

where the inequality holds by assumption MLRP. Hence $\chi$ is decreasing in $\gamma_A$.

**A change in the value of $q_H$.** A change in $q_H$ has two effects on $\chi$: a direct effect, and an indirect effect through $\gamma_B$. Consider the direct effect first. Recall that $\gamma_B$ is the solution to,

$$
\frac{g(\gamma_B|H)}{g(\gamma_B|L)} = \frac{\pi_L}{\pi_H} \frac{1 - q_L}{1 - q_H}. \tag{A.43}
$$

Taking the derivative of both sides of equation A.43 with respect to $q_H$ yields,

$$
\frac{d\gamma_B}{dq_H} \frac{d}{d\gamma} \left( \frac{g(\gamma_B|H)}{g(\gamma_B|L)} \right) = -\pi_L \frac{1 - q_L}{\pi_H (1 - q_H)^2}. \tag{A.44}
$$

The right hand side of equation A.44 is positive because $\pi_L$ is negative and all other terms are positive. By assumption MLRP, $\frac{d}{d\gamma} \left( \frac{g(\gamma_B|H)}{g(\gamma_B|L)} \right) > 0$ and therefore $\frac{d\gamma_B}{dq_H} > 0$. Recalling that, $\frac{\partial \chi}{\partial \gamma_B} > 0$, the indirect effect of $q_H$ is therefore positive,

$$
\frac{d\gamma_B}{dq_H} \frac{\partial \chi}{\partial \gamma_B} > 0. \tag{A.45}
$$

Now consider the direct effect. The partial derivative of the first term of $\chi$ with respect
to \( q_H \) is of the same sign as,

\[-(1 - q_L)(1 - G(\gamma_B|H))(1 - G(\gamma_B|L)) < 0. \tag{A.46} \]

The partial derivative of the second term of \( \chi \) with respect to \( q_H \) is of the same sign as,

\[-q_L(1 - G(\gamma_B|H))(1 - G(\gamma_B|L)) < 0. \tag{A.47} \]

Given that both equation A.46 and A.47 are negative, the direct effect of \( q_H \) on \( \chi \) is negative. Hence the total effect of \( q_H \) on \( \chi \) is of ambiguous sign and depends on whether the direct or indirect effect dominates.

By a similar logic the impact of \( q_L \) on \( \chi \) is ambiguous. In this case the direct effect is negative and the indirect effect (through \( \gamma_B \)) is positive.

A change in the value of \( \pi_i \). A change \( \pi_i \) affects \( \chi \) through both \( \gamma_B \) and \( \gamma_A \). One can verify that an increase in \( \pi_i \) will reduce both \( \gamma_B \) and \( \gamma_A \). An increase in the profit the employer receives from a worker of either ability will induce the employer to reduce their standards – decreasing \( \gamma_A \) and \( \gamma_B \). Hence the effect of \( \pi_i \) on \( \chi \),

\[
\frac{d\chi}{d\pi_i} = \frac{\partial \chi}{\partial \gamma_A} \frac{d\gamma_A}{d\pi_i} + \frac{\partial \chi}{\partial \gamma_B} \frac{d\gamma_B}{d\pi_i}, \tag{A.48}
\]

is of ambiguous sign as \( \frac{\partial \chi}{\partial \gamma_A} < 0 \) and \( \frac{\partial \chi}{\partial \gamma_B} > 0 \).
Appendix B

Appendix to Chapter 3

B.1 Identification of the unweighted estimator under independent preferences

The purpose of this section is to show that the unweighted estimator is consistent under independent preferences. When preferences are independent the utility one male derives from a female is strictly independent of the utility another male derives from the same female (the same is true for the utility females receive from males). As in the previous section identification hinges on the inequality,

\[ P(\{(m_i, f_k), (m_j, f_l)\} | \beta^*, \gamma^*) \geq P(\{(m_i, f_l), (m_j, f_k)\} | \beta^*, \gamma^*). \] (B.1)

To show identification it needs to be shown that the inequality holds strictly when \( q_{(m_i, f_k), (m_j, f_l)}(\beta^*, \gamma^*) = 1 \) and \( q_{(m_i, f_l), (m_j, f_k)}(\beta^*, \gamma^*) = 0 \). However, under independent preferences there is the possibility of multiple equilibria being predicted in the submarket. Recall that when multiple equilibria are predicted \( q_{(m_i, f_k), (m_j, f_l)}(\beta^*, \gamma^*) + q_{(m_i, f_l), (m_j, f_k)}(\beta^*, \gamma^*) = 1 \), but there is no restriction on the individual terms in the equation (both can be positive). Hence, it must be shown that when multiple equilibria are predicted in the submarket, equation B.1 holds as a strict equality.

We must not only deal with the possibility of multiple equilibria being predicted in a submarket, but also multiple equilibria in the data. One might think that multiple equilibria present a problem in evaluating the relative sizes of \( P(\{(m_i, f_k), (m_j, f_l)\} | \beta^*, \gamma^*) \) and...
$P(\{(m_i, f_i), (m_j, f_k)\}|\beta^*, \gamma^*)$ — this however is not the case, at least asymptotically. While multiple equilibria can be present in the data, the probability that the matchings $(m_i, f_k)$ and $(m_i, f_l)$ can both be pairwise stable in equilibrium is (asymptotically) trivial relative to the probability that just one is stable.

To formalize this concept define the utility that the $i$th male receives from his best possible match other than $f_k$ or $f_l$ as $v^*_m$, (when preferences are independent $E(v^*_m) = E(v^*_m)$). The probability that $f_k$ is in $m_i$’s choice set and provides $m_i$ utility above $v^*_m$ is $1/(n-1)$ (this is simply the unconditional probability that a female and $m_i$ are matched when $f_l$ is removed from consideration). The probability that both $f_k$ and $f_l$ are in $m_i$’s choice set and provide him utility above $v^*_m$ is $1/(n-1)^2$. Hence, $Pr(v_{m_i, f_k} > v_{m_i, f_l} > \max v^*_m) = O(Pr(v_{m_i, f_k} > \max v^*_m)) = 1/(n-1)$ as $n \to \infty$ (note that this result holds only as $n \to \infty$, not as $T \to \infty$). Therefore the probability that the matchings $(m_i, f_k)$ and $(m_j, f_l)$ are stable as $n \to \infty$ is $Pr(u_{f_k, m_i} > u^*_{f_k}) Pr(v_{m_i, f_k} > v^*_m) Pr(u_{f_l, m_j} > u^*_{f_l}) Pr(v_{m_j, f_l} > v^*_m)$. Substituting these probabilities into equation B.1 and rearranging yields,

$$\frac{Pr(u_{f_k, m_i} > u^*_{f_k}) Pr(v_{m_i, f_k} > v^*_m)}{Pr(u_{f_k, m_j} > u^*_{f_k}) Pr(v_{m_j, f_k} > v^*_m)} \geq \frac{Pr(v_{m_i, f_k} > u^*_m) Pr(u_{f_l, m_j} > u^*_{f_l})}{Pr(v_{m_j, f_l} > u^*_m) Pr(u_{f_i, m_j} > u^*_{f_l})} \tag{B.2}$$

It needs to be shown that the inequality in equation B.2 holds strictly when $\{(m_i, f_k), (m_j, f_l)\}$ is predicted to be a unique equilibrium, and holds as a strict equality when multiple equilibria are predicted.

Consider the following equation,

$$\frac{Pr(u_{f_k, m_i} > u^*_{f_k} | \bar{u}_{f_k, m_i}(\gamma^*) > \bar{u}_{f_k, m_i}(\gamma^*))}{Pr(u_{f_k, m_j} > u^*_{f_k} | \bar{u}_{f_k, m_j}(\gamma^*) > \bar{u}_{f_k, m_j}(\gamma^*))} = \frac{Pr(u_{f_l, m_i} > u^*_{f_l} | \bar{u}_{f_l, m_i}(\gamma^*) > \bar{u}_{f_l, m_i}(\gamma^*))}{Pr(u_{f_l, m_j} > u^*_{f_l} | \bar{u}_{f_l, m_j}(\gamma^*) > \bar{u}_{f_l, m_j}(\gamma^*))} > \frac{Pr(u_{f_i, m_i} > u^*_{f_i} | \bar{u}_{f_i, m_i}(\gamma^*) > \bar{u}_{f_i, m_i}(\gamma^*))}{Pr(u_{f_i, m_j} > u^*_{f_i} | \bar{u}_{f_i, m_j}(\gamma^*) > \bar{u}_{f_i, m_j}(\gamma^*))}. \tag{B.3}$$

The first equality follows from symmetry. The second equality follows from the fact that when $\bar{u}_{f_i, m_i} > \bar{u}_{f_j, m_j}$ the probability that $u_{f_i, m_i} > x$ is greater than $u_{f_j, m_j} > x$ for any $x$. By
a similar logic,

\[
\frac{Pr(v_{m_i,f_k} > v_{m_i}^* | \bar{v}_{m_i,f_k}(\beta^*) > \bar{v}_{m_i,f_i}(\beta^*))}{Pr(v_{m_i,f_i} > v_{m_i}^* | \bar{v}_{m_i,f_k}(\beta^*) > \bar{v}_{m_i,f_i}(\beta^*))} = \frac{Pr(v_{m_j,f_k} > v_{m_j}^* | \bar{v}_{m_j,f_k}(\beta^*) > \bar{v}_{m_j,f_i}(\beta^*))}{Pr(v_{m_j,f_i} > v_{m_j}^* | \bar{v}_{m_j,f_k}(\beta^*) > \bar{v}_{m_j,f_i}(\beta^*))} > 1
\]

(B.4)

By the preceding two equations, when \(q_{(m_i,f_k),(m_j,f_i)}(\beta, \gamma) = 1\) with certainty (as described in section 3.4) the inequality in equation B.2 holds as a strict inequality. Conversely, when \(q_{(m_i,f_k),(m_j,f_i)}(\beta, \gamma) \in [0, 1]\) (again, as described in section 3.4) the inequality in equation B.2 holds as a strict equality. As in the section 3.5, unique identification is assured by the support of utility functions.

Table B.1 provides the results of a Monte-Carlo experiment using independent preferences. The bias of the unweighted estimator is close to zero across all experiments. Furthermore, the MISE of the unweighted estimator is declining in the number of individuals in the market, and also in the number of markets (though we have not shown the estimator to be identified in the number of markets). The bias and MISE of the weighted estimator are both decreasing in the number of markets, though not necessarily in the number of individuals.

Table B.1: Monte-Carlo results for independent preferences

<table>
<thead>
<tr>
<th>Markets</th>
<th>Players</th>
<th>Weighted estimator</th>
<th>Unweighted estimator</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(\hat{\beta}_2/\beta_1)</td>
<td>(\hat{\gamma}_2/\gamma_1)</td>
</tr>
<tr>
<td>1</td>
<td>25</td>
<td>0.27 (0.91)</td>
<td>0.07 (1.01)</td>
</tr>
<tr>
<td>1</td>
<td>50</td>
<td>0.10 (0.88)</td>
<td>-0.05 (0.97)</td>
</tr>
<tr>
<td>1</td>
<td>100</td>
<td>-0.04 (0.64)</td>
<td>-0.08 (0.81)</td>
</tr>
<tr>
<td>1</td>
<td>200</td>
<td>-0.11 (0.48)</td>
<td>-0.07 (0.62)</td>
</tr>
<tr>
<td>10</td>
<td>25</td>
<td>0.01 (0.69)</td>
<td>-0.04 (0.79)</td>
</tr>
<tr>
<td>100</td>
<td>25</td>
<td>-0.02 (0.29)</td>
<td>0.01 (0.40)</td>
</tr>
</tbody>
</table>

\(^1\) Mean of the log of the estimated ratio (true value is 0) is displayed in the table with the mean integrated square error of the log of the estimated ratio in parenthesis. Under independent preferences the two (match-specific) observables are drawn from \(N(0,1)\), there is a correlation of 0.3 between the second observable in the male utility function and the second observable in the female utility function. The unobservable is drawn from \(N(0,2)\).
B.2 Example of non-identification of the unweighted estimator under heterogeneous preferences

When utility is neither homogeneous nor independent the unweighted estimator cannot be proven to be identified using the same arguments as in the preceding section. In sections B.1 and 3.5, the estimator was shown to be consistent if the inequality in equation 3.11 holds when \(\{(m_i, f_k), (m_j, f_l)\}\) is predicted to be an equilibrium in a two by two submarket. This section will provide a brief (graphical) example where the inequality does not hold when \(\{(m_i, f_k), (m_j, f_l)\}\) is predicted to be an equilibrium, under heterogeneous preferences.

Consider a model in which each utility function has one individual specific observable (denoted \(z_f\) and \(x_m\), respectively), one match specific observable (denoted with double subscripts), and an individual specific unobservable. Normalizing the coefficient on the individual specific term to one leaves the coefficient on the match specific term to be estimated.

Suppose that,

\[
\bar{u}_{m_i, f_k} - \bar{u}_{m_i, f_l} = z_{f_k} - z_{f_l} + \beta^*(z_{m_j, f_k} - z_{m_j, f_l}) > 0 \quad (B.5)
\]

and

\[
\bar{v}_{f_l, m_i} - \bar{v}_{f_l, m_j} = x_{m_i} - x_{m_j} + \gamma^*(x_{f_k, m_i} - x_{f_k, m_j}) > 0 \quad (B.6)
\]

From section 3.4 this implies that \(\{(m_i, f_k), (m_j, f_l)\}\) is predicted to be a unique equilibrium.

Yet, as figure B.1 shows, it can be the case that when these inequalities hold,

\[
\frac{Pr(v_{m_i, f_k} > v_{m_i}^*)}{Pr(v_{m_i, f_l} > v_{m_i}^*)} < \frac{Pr(v_{m_j, f_k} > v_{m_j}^*)}{Pr(v_{m_j, f_l} > v_{m_j}^*)}. \quad (B.7)
\]
By a similar logic it can be the case that,

\[
\frac{Pr(u_{fk,m_i} > u^*_{fk})}{Pr(u_{fk,m_j} > u^*_{fk})} < \frac{Pr(u_{fl,m_i} > u^*_{fl})}{Pr(u_{fl,m_j} > u^*_{fl})}
\]  

(B.8)

Hence, the inequality in equation 3.11 does not necessarily hold when \{(m_i, f_k), (m_j, f_l)\} is predicted to be an equilibrium.

Figure B.1: Illustration of equation B.7

The top plot shows the utility the \(i\)th male receives from the \(k\)th and \(l\)th female conditional on \(\bar{v}_{m_i,f_k}\) and \(\bar{v}_{m_i,f_l}\). The bottom plot shows the same functions for the \(j\)th male. The plots demonstrate that the \(j\)th man is more likely to want to match with the \(k\)th female relative to the \(l\)th female.
Appendix C

Appendix to Chapter 4

Details of the GHK estimation

The simulated maximum likelihood estimator used herein is described in detail by Keane (1993) and Borsh-Supan et al. (1991). Denote the sequence of inter temporal outcomes for a particular individual-department match as $Y_{ti} = \{y_{ti}, ..., y_{ti}^{T}\}$ and denote $\xi(y_{ti}^{T}, b, \gamma)$ as the space of $(e, \eta)$ that is consistent with the observed outcome, $y_{ti}^{T}$, given the parameters $b, \gamma$. For example (dropping the individual subscript),

$$\xi(y_{1} = 0, b, \gamma) = \left\{ e, \eta | \eta < \alpha(1 - t^{*}) - \gamma x^{1} \right\}, \text{ or } (e < g(z^{1}; b)) \right\} \quad \text{(C.1)}$$

For any particular set of parameters $S$ different simulations are run, with the simulated maximum likelihood calculated as follows:

1. For $s \in \{1, 2, ..., S\}$

   (a) In period $t = 1$ draw $\eta^{1}$ and $e^{1}$ from the truncated joint distribution $F_{\eta,e}(r)$ (the marginal distributions of $\eta$ and $e$ are both standard normals – $r$ is the correlation between the two terms) such that $(e, \eta) \in \xi(y_{ci}^{1} = 0, b, \gamma)$. Denote these error terms as $(e_{1}^{t=1,s}, \eta_{1}^{t=1,s})$.

   (b) In period $t = 2$ draw $\eta^{2}$ and $e^{2}$ from the joint distribution $F_{\eta,e}(r)$ such that (a) $(e, \eta) \in \xi(y_{ci}^{2}, b, \gamma)$ and (b) the error terms are correlated across periods as per the covariance parameters, $\rho_{e}$ and $\rho_{\eta}$ (this is done via a Cholesky decomposition of the covariance matrix.)
(c) Continue drawing error terms in this manner for the remaining periods.

2. The estimated probability of any particular outcome is then defined as,

\[
\hat{\text{Prob}}(y^t | Y^{t-1}, b, \gamma) = \frac{1}{S} \sum_{s=1}^{S} \Pr \left( (e^{t,s}, \eta^{t,s}) \in \xi(y^t, b, \gamma) | (e^{1,s}, \eta^{1,s}), \ldots, (e^{t-1,s}, \eta^{t-1,s}) \right),
\]

where \( w \) are importance sampling type weights which sum to 1,

\[
w \left( (e^{1,s}, \eta^{1,s}), \ldots, (e^{t-1,s}, \eta^{t-1,s}) \right) = \\
\left( \Pr \left( (e^{t-1,s}, \eta^{t-1,s}) \in \xi(y^t, b, \gamma) | (e^{1,s}, \eta^{1,s}), \ldots, (e^{t-2,s}, \eta^{t-2,s}) \right) \right) / \\
\left( \frac{1}{S} \Pr \left( (e^{t-1,r}, \eta^{t-1,r}) \in \xi(y^t, b, \gamma) | (e^{1,r}, \eta^{1,r}), \ldots, (e^{t-2,r}, \eta^{t-2,r}) \right) \right) \ldots \\
\frac{1}{S} \Pr \left( (e^{r}, \eta^{r}) \in \xi(y^t, b, \gamma) | (e^{1,r}, \eta^{1,r}) \right).
\]

3. The simulated likelihood is then

\[
L(b, \gamma, \rho_e, \rho_\eta, r, a, w) = \sum_{t=1}^{T} \sum_{c=1}^{\tilde{c}_t} \ln \left( \hat{\text{Prob}}(y^t | Y^{t-1}, b, \gamma) \right).
\]