Does the Trend Need a Mend? The Effects of Herding Behavior on the Returns to Higher Education

Amy K. French
The College of Wooster, akfrench16@gmail.com

Follow this and additional works at: https://openworks.wooster.edu/independentstudy

Part of the Social and Behavioral Sciences Commons

Recommended Citation
https://openworks.wooster.edu/independentstudy/6008
Does the Trend Need a Mend?
The Effects of Herding Behavior on the Returns to Higher Education

By
Amy French

Submitted in Partial Fulfillment of the Requirements of Senior
Independent Study for the Department of Business Economics at the
College of Wooster

Advised by
Dr. James Burnell
Department of Economics

March 24, 2014
Abstract

Using the human capital model, supplemented by a preference framework focusing on utility maximization, this paper examines the decision making process of the individual when considering enrollment in higher education. The study applies a theory of herding behavior to the investment decision, arguing that individuals use the decisions of others who have gone before them as signals to reduce information costs and to aid in their own decision making process. The paper further discusses the possibility of this behavior leading to irrational investment decisions. Herding behavior will likely lead to an increase in the supply of highly educated workers, perhaps to the extent that the supply exceeds the demand. As a result, it is hypothesized that this phenomenon will yield a decrease in returns on the investment in higher education. After a literature review, the hypothesis is empirically tested by examining the earnings difference between individuals who have obtained a bachelor’s degree and those who have earned higher degrees. The Oaxaca-Blinder decomposition procedure is employed to gain insight on the payoff of pursuing higher education in the face of increasing student enrollment. Although the results do not provide evidence in support of the hypothesis, a discussion on how they may still reflect irrational decisions is provided.
Acknowledgements

First and foremost, I would like to thank Dr. James Burnell, my Independent Study advisor, for his encouragement and support. I would not have been able to complete the Independent Study process without his continuous guidance and patience. I thank him for his constructive criticism and helpful advice as he tirelessly read through numerous drafts and revisions.

I would also like to express my gratitude to family and friends, always willing to celebrate or commiserate with me depending on what the situation called for. I am appreciative of their constant reassurance and support.

Last, but certainly not least, I would like to acknowledge my mother, Brenda Reynolds French. No longer with us, but never forgotten, she continues to be my source of inspiration and strength to overcome obstacles that pale in comparison to those she fought herself.
# Table of Contents

I. Introduction 6

II. A Theory of Herding Behavior and Irrational Decisions Pertaining to an Investment in Higher Education 10
   2.1 The Human Capital Model 11
   2.2 Preferences and Utility Maximization 15
   2.3 The Theory of Herding Behavior 20
      The Formal Model 22
   2.4 Distortions in Preferences 27
   2.5 Implications of Herding Behavior on the Human Capital Model 34
   2.6 Parental Influence 37
   2.7 Subsidizing 39
   2.8 Importance of Completion 41
   2.9 Incomplete Information 43
   2.10 Credit and Financial Constraints 44
   2.11 Conclusion 45

III. Literature Review 49
   3.1 Evaluating the Ex Ante Expected Returns to College 51
   3.2 Differential Pricing in Undergraduate Education: Effects on Degree Production by Field 58
   3.3 Who goes to graduate/professional school? The importance of economic fluctuations, undergraduate field, and ability 64
   3.4 How have college decisions changed over time? An application of the conditional logistic choice model 70
   3.5 “Overeducation” in the Labor Market 78
   3.6 Conclusion 85

IV. Model Specification 87
   4.1 The Graduate School Decision 88
   4.2 The Oaxaca-Blinder Decomposition Procedure 91
   4.3 Explained and Unexplained Portions in Terms of the Hypothesis 95
   4.4 The Earnings Function 97
   4.5 Data 100
   4.6 Conclusion 104

V. Results and Analysis 105
   5.1 The Basic Regression 106
   5.2 The Chow Test 110
   5.3 An Application of the Oaxaca-Blinder Decomposition Procedure 120
5.4 Controlling for Experience
   A Further Examination of the Results
5.5 Conclusion

VI. Conclusion

VII. Appendices
   Appendix A: The Basic Regression
   Appendix B: The Chow Test
   Appendix C: The Initial Oaxaca-Blinder Decomposition Procedure
   Appendix D: The Oaxaca-Blinder Decomposition Procedure Controlling for Experience

VIII. Bibliography
I. Introduction

One would be hard-pressed to find a teacher in the United States who fails to hammer into the brains of her students the value of a college education. Nonetheless, there are still those who ask if too many people are making the choice to pursue higher education. With rising tuition rates, some individuals are wondering about the true value of a college degree, begging the question of whether people are being asked to spend a great deal of money for a meager result (Wood, 2011). Compared to the past returns on earning a higher degree, the present rewards may differ substantially. This paper attempts to examine the decision making process of the individual as she considers her investment decision in college as well as the results associated with such a decision.

Nobody wants to believe that America’s higher education system could be failing to provide students with better lives. However, if it is the case that more and more individuals are engaging in an investment in higher education, it is possible that the supply of educated workers entering the labor force could exceed the demand. Institutions are not necessarily duping individuals into making poor decisions by hustling them and stockpiling their money. Rather, the nature of the problem exists within the demand for education.

Individuals may be quick to engage in the decision to enroll without considering the costs and benefits on a personal level. Instead, decision makers may be making the choice to attend simply because everyone else is doing it. This does not imply that the choice to attend college is a bad decision for everyone. On the contrary, a college degree does offer many individuals a means for obtaining a better life. Nonetheless, the push for higher education is not without consequences, either.
If more and more individuals are choosing to enroll in higher education, it is important to investigate what can be attributed as the cause of this as well as possible repercussions. It is possible that individuals observe others following a certain path, signaling to the decision maker that this is the ‘correct’ path since everybody is choosing it. However, this signal may differ from the decision maker’s own private signals if she were to do an analysis of the costs and benefits on her own instead of following the decisions of those who made the choice before her. This has the potential to alter the decision making process, implying that decisions being made currently would perhaps be different if made in another environment where the individual is not influenced by others.

It follows that an irrational decision can be made in this context if the decision results in returns that are less than initially anticipated and a situation in which the decision maker could have been better off by choosing an alternative choice. This is made even more likely if more and more students are choosing to attend college, resulting in an oversupply of educated workers and a possible decrease in returns realized by the investor.

To aid in a deeper discussion of this theory and phenomenon, Chapter II presents a basic theoretical model to analyze how individuals typically make their decision to enroll in college. The human capital model is used as the primary framework, supplemented by a model of individual preferences and utility maximization. After these basic theories are introduced as a foundation, the theory of herding behavior is introduced to explain trends in the decision to pursue higher education. This model explores how individuals often make choices based on others’ decisions. Herding is also further analyzed in the context of distorted preferences, parental influence, and the subsidization
of higher education. Other factors that pertain to the college investment decision, but are left out of traditional models, are then explored. These include the importance of completion, incomplete information that the decision maker faces, and credit and financial constraints.

Chapter III presents a review of literature, providing a presentation of relevant articles and research that have previously looked at how individuals make their decision to attend college, as well as the returns on such an investment. Five articles are outlined, focusing on empirical investigations of the college investment decision and the connections they exhibit to the way individuals process information and make decisions. The first article is useful in providing a basic method for empirically evaluating the returns to the college investment decision by employing the human capital model. Following, a look at recently implemented differential pricing policies between college majors is used to glean an insight into the subsidization of education and how individuals respond to changes in price.

Nonetheless, the decision to pursue higher education extends beyond simply a pecuniary analysis. A third article examines what types of individuals enroll in graduate and professional school, relating the decision to pursue higher education to individual characteristics that likely reflect individual preferences. Extending beyond this, the fourth article engages in a discussion of how the choice to attend school has changed over time.

Since most of the literature focuses on the decision making process, the last article is a study on overeducation with the labor market, providing, by extension, a method to indirectly measure herding. If workers hold jobs that require less schooling than they
actually possess, this may suggest an overinvestment in education, or an irrational decision to invest.

Since the topic has received little attention in the area of economics, however, this paper expands upon prior research by offering an empirical investigation to better understand the effects of herding. Chapter IV focuses on the model specification used for the empirical work. The results, supplemented by an analysis, are then provided in Chapter V. Since herding is not confined to only the undergraduate level, the empirical work focuses on the difference in earnings between individuals who have earned bachelor’s degrees versus individuals who have earned higher degrees.

The model uses the decision to pursue higher education beyond the undergraduate level to diminish the effects of herding. This is based on the idea that individuals who make this enrollment decision are more informed about the investment and less influenced by those around them. The paper hypothesizes, however, that herding still affects the returns to a graduate degree, yielding a reward that is not as significant as one may initially presume.

The Oaxaca-Blinder decomposition procedure is employed to decompose the difference in earnings exhibited by the two groups. This is used to gain a better understanding of the returns associated with earning a higher degree in the presence of an increasing educated labor supply. Although the results do not support the hypothesis, a discussion of how irrational decisions may still occur is provided. Finally, Chapter VI summarizes relevant and important findings, critiques the model, and suggests further extensions that would complement the research provided in this paper.
II. A Theory of Herding Behavior and Irrational Decisions Pertaining to an Investment in Higher Education

The choice to attend college has long been a decision faced by many individuals. This chapter provides an overview of how individuals consider their decision to enroll in higher education based on previous frameworks for analyzing the choice. Only after understanding how decisions are made in a so-called ‘neutral’ or rational environment, assuming no information costs, is it possible to contrast how choices are actually being made. Although models are useful in simplifying a scenario for understanding, the choice to pursue higher education is no doubt affected by factors that are not considered in the classical frameworks used to analyze the decision.

The traditional framework used to evaluate the choice to invest in higher education is the human capital model. This is presented as an overview of how individuals consider the pecuniary returns on their investment, suggesting that a rational individual should invest only if the net benefits exceed the cost of enrolling. This is useful in that it allows for a basic and simplistic analysis of the decision.

Next, the chapter expands on this theory by considering non-pecuniary benefits an individual may expect to receive from attending college. The theory considers a framework that incorporates the preferences of the individual and aligns her decision with utility maximization. Outlined are equilibrium conditions that will occur under utility maximization.

The remainder of the chapter presents challenges to these typical models, suggesting distortions in the decision making process and providing possible explanations for an oversupply of educated workers as well the consequences that may ensue. A simple model of herding behavior is presented and applied to the decision making
process. This framework attempts to examine how irrational decisions are made in the context of distorted signals. That is, there is a potential that individuals are led in the wrong direction as impediments of information alter individual preferences and the choices that would be made in different scenarios. The implications of this framework are then analyzed in the context of the human capital model.

The chapter also examines the affect of parental influence on a student’s decision to enroll. This provides support for the herding model by suggesting that preferences may be distorted in the context of signals that are being received. The subsidization of education by state governments also mirrors the effects of herding and distorts the signals that students receive in regards to the price and cost of obtaining education.

The remainder of the chapter is devoted to understanding the importance of completion when obtaining a degree as well as incomplete information students may hold. In addition, the costs associated with obtaining information are examined. Finally, credit and financial constraints faced by the decision maker are explored as well. Each of these components relates back to individual decision process and how decisions may be altered in the context of distorted signals and information.

2.1: The Human Capital Model

The first step in understanding how an individual would evaluate her choice to attend college is in terms of the financial gains she would receive from doing so. The basic framework and classic theory for analyzing the decision to pursue higher education is the human capital model. This theory is based on the idea that attending college is ultimately an investment in oneself (Blau, Ferber, & Winkler, 1997). Capital can be
thought of as wealth, whether said wealth may be quantified as actual money or as other assets. Thus, human capital is therefore the ‘value’ or ‘worth’ of an individual in terms of the benefits she produces in the labor market. If an individual makes the decision to pursue higher education, it is assumed that she does so because she is enhancing her ability to perform or produce labor at a higher quality and economic value. Thus, a rational individual should view the decision in this most basic sense. In its simplest form, the choice go to college or end formal education requires weighing the costs of attending college against the returns (Blau et al., 1997).

The appeal of this method is in its simplicity, as it measures only pecuniary costs and benefits of the investment. The decision necessitates comparing the expected experience-earnings profile against each level of schooling. The expected experience-earnings profile is the annual earnings that correspond to each level of labor market experience (Blau et al., 1997). A rational individual will choose to go to college only if the returns are greater than the costs. Consider Figure 2.1.1. The graph depicts the experience-earnings profiles for attending college versus entering the labor market directly after high school.

The curve $EF$ illustrates an individual’s earnings without a college degree. Point $E$ indicates the initial level of earnings, which correlates to zero years of experience. On the contrary, if this individual decides to attend college she will observe a much lower initial level of earnings. This is because she not only incurs expenses related to obtaining a degree, but also foregoes income by deciding not to work during the years she devotes to education. The expenses she incurs are direct costs that correlate to ‘negative earnings.’ These costs include expenses such as tuition and books. If the student incurs
$O_A$ dollars per year in expenses, the sum of direct costs is equivalent to the area $OAB4$, as illustrated in the diagram.

Further, the individual must consider costs beyond simply direct costs. Indirect costs represent opportunity costs associated with the investment decision. This includes the earnings that are foregone by the individual during schooling. Although the individual at hand has decided not to work during college, even those that do can still expect to experience large foregone earnings. When looking at the costs of the investment overall, the individual must look at both direct and indirect costs. The area of $EABCD$ represents the total cost of a college education.
Upon graduation, the student expects to receive earnings at the level of $C'$. Note that this lies above point $E$, the earnings she could have expected without a degree. This demonstrates the increase in productivity that the individual gains from her investment in education. Over her lifetime, the experience-earning profile is represented by $ABCG$. Therefore, although she sacrifices earnings in the beginning years of her career, by investing in her human capital, she can anticipate higher lifetime earnings (Blau et al., 1997).

The experience-earning profiles for both scenarios, entering the workforce or obtaining more education, are increasing for the majority of the employee’s working lifetime. This can be attributed to experience that the individual gains from on-the-job training. However, it is also important to point out that the profile of the college graduate rises at a faster rate. This has been attributed to empirical evidence that finds that those who have earned college degrees boast more informal on-the-job training in addition to formal training acquired in school (Blau et al., 1997).

The gross benefits of a college education are then illustrated by the difference between the expected earnings for a college graduate compared to those of a high school graduate, as shown in Figure 2.1.1. This is directly related to the length of an individual’s work life. For example, if the individual decides to work 45 years after starting college, as depicted, the gross benefits are equal to the area $DRR'$.

When analyzing the decision to attend college or not, it is necessary to compare the gross benefits to the costs. The human capital model asserts that an individual will decide to pursue higher education so long as the gross benefits outweigh the costs. That is, the area $DRR'$ must be greater than the area $EABCD$. Time preference, however,
declares that all else being equal, individuals prefer to have any given end sooner rather
than receive it later in the future. Thus, a sufficient return is required to induce
individuals to undertake such investments (Blau et al., 1997). Beyond this, however,
individuals also have different rates of return that entice them to partake in an investment.
Gross benefits must therefore also be large enough to provide the investor with an
adequate return based on their own preferences. This return should outweigh the yield of
other investments, such as profit received from investing in other assets or interest earned
through a savings account (Blau et al., 1997). Only after the benefits and returns are
analyzed in comparison to costs can the individual make her investment decision.

2.2: Preferences and Utility Maximization

Although the human capital model is valuable in its simplistic analysis, it does not
consider the individual’s preferences and the non-pecuniary benefits that she will receive
from her decision. These factors are no less important in the investment decision relative
to monetary gains. This does not discredit the human capital theory, however, but rather
adds another dimension to be analyzed.

Preference theory dictates that an individual will make her investment decision on
more than simply pecuniary measures. In this case, the individual will make her decision
based on the amount of utility the choice provides. That is, the decision maker is
concerned with the level of happiness, satisfaction, and enjoyment that the decision will
yield. In this context, her choice is the result of a process that optimizes utility.

Non-pecuniary benefits emerge both while in college and after graduation.
Students are likely to enjoy activities, such as sports for example, that they may not have
had the opportunity to participate in otherwise. They are also likely to enjoy spending time with peers, and engaging in intellectual exploration (Oreopoulos & Petronijevic, 2013). Beyond graduation, empirical evidence also suggests that, controlling for different measures of family background and income, college graduates hold jobs that provide a greater sense of accomplishment, independence, creativity, and social activity. In addition, those with more education display higher reports of health (Oreopoulos & Petronijevic, 2013). Although these benefits are difficult to quantify, they play a significant role in the decision of the individual.

When making the choice to engage in additional schooling, the individual faces a choice between consuming education \( E \) versus a composite good \( Y \). The composite good can be defined as an amalgam of other goods. By convention, it is thought of as the amount of income the consumer has left over, which allows for all other opportunities to be accounted for as the individual allocates her resources.

It follows that the decision maker’s utility is dependent upon education and the composite good. The consumer attempts to maximize utility subject to a budget constraint, the set of all bundles that exactly exhaust the consumer’s income at given prices. This can be expressed as follows:

\[
U = f(Y, E) \quad \text{subject to} \quad I = P_E E + P_Y Y
\]

Note that \( I \) indicates the income that can be allocated to consumption, \( P_E \) represents the price of education, and \( P_Y \) represents the price of the composite good.

To find the values of \( E \) and \( Y \) that produce the highest value of utility, subject to the constraint, the method of Lagrangian multipliers can be used. The more-is-better assumption implies that the best affordable bundle will lie on the budget constraint. To
determine this bundle, the constrained equation can be transformed into the following unconstrained maximization problem where \( \lambda \) is the Lagrangian multiplier:

\[
\text{Maximize } L = U(E, Y) - \lambda(P_E E + P_Y Y - I)
\]

The first-order conditions for maximizing \( L \) are obtained by taking the first partial derivatives of \( L \) with respect to \( E \), \( Y \), and \( \lambda \), and then setting these equal to zero. This is illustrated below:

1. \[
\frac{\partial L}{\partial E} = \frac{\partial U}{\partial E} - \lambda P_E = 0
\]
2. \[
\frac{\partial L}{\partial Y} = \frac{\partial U}{\partial Y} - \lambda P_Y = 0
\]
3. \[
\frac{\partial L}{\partial \lambda} = I - P_E E - P_Y Y = 0
\]

Then, to find the values of \( E \) and \( Y \) that maximize the utility function, it is necessary to solve the first order conditions for \( E \), \( Y \), and \( \lambda \). Note that the solutions for \( E \) and \( Y \) are the only quantities that are relevant for determining the maximization solution. Setting \( \lambda \) equal to zero simply guarantees that the budget constraint is satisfied.

The equilibrium condition, displayed in Figure 2.2.1, specifies that utility is maximized where the marginal rate of substitution at \( A \) is equal to the absolute value of the slope of the budget constraint. This ensures that the best affordable bundle will occur at the point of tangency between the budget constraint and indifference curve.
Following, it must be the case that:

\[ MRS = \frac{P_E}{P_Y} \]

The right-hand side of this equilibrium condition represents the opportunity cost of education in terms of the composite good. The left-hand side of the equation is the marginal rate of substitution, which conveys the rate at which the consumer is willing to exchange education for the composite good. In other words, it is the amount of additional education the consumer would need to be given in order to compensate her fully for the loss of one unit of the composite good, holding utility constant. This is the absolute value of the slope of the indifference curve, or \( \left| \frac{\Delta Y}{\Delta E} \right| \).
As an example, suppose that the slope of the budget constraint is -2. The tangency condition requires that 2 units of education would be needed to offset the benefits lost due to giving up one unit of the composite good. In general, based on preferences and utility maximization, the rational decision maker will consume education up to the point where \( MRS = \frac{P_E}{P_Y} \).

From this stipulation, as well as the first-order conditions, another equilibrium condition can be derived. Using the first partial derivatives, it is possible to derive the marginal utility of each good. The marginal utility of each good is the additional amount of utility the individual gains from consuming one more unit of the good at hand. Mathematically, the marginal utility of education and the composite good can be written as follows:

\[ m_u_E = \frac{\partial U}{\partial E} \quad \text{and} \quad m_u_Y = \frac{\partial U}{\partial Y} \]

By taking the first partial derivatives and setting them equal to zero, we already know that:

1. \( \frac{\partial U}{\partial E} - \lambda P_E = 0 \)
2. \( \frac{\partial U}{\partial Y} - \lambda P_Y = 0 \)

Redistributing allows for rewriting these equations in the following forms:

1. \( \frac{\partial U}{\partial E} = \lambda P_E \)
2. \( \frac{\partial U}{\partial Y} = \lambda P_Y \)
Then, dividing the first equation by the second equation, it follows that:

\[ \frac{\partial U / \partial E}{\partial U / \partial Y} = \frac{\lambda P_E}{\lambda P_Y} = \frac{P_E}{P_Y} \]

Thus, the equilibrium condition is as follows:

\[ \frac{m u_E}{m u_Y} = \frac{P_E}{P_Y} = MRS \]

In other words, the ratio of the marginal utilities is equal to the marginal rate of substitution of E and Y. The optimal bundle will only exist where these conditions are satisfied.

2.3: The Theory of Herding Behavior

If the pervious theories indicate how a rational individual ‘should’ assess her decision to go to college, how is it that irrational decisions stray from these models? In such cases, signals are being distorted; causing a change in behavior that would not be observed prior to the distortion. Subsequently, it is necessary to ask how individuals are actually making decisions. A likely cause for a change in behavior is herding. Herding behavior is used to explain the occurrence of a large number of people doing the same thing at the same time. This has the potential to explain why increasing amounts of individuals are making the decision to pursue higher education. In this sense, students are influenced in their own decision making based on what others are doing.

Banerjee (1992) sets forth a simple model to explain herding behavior, which can then be applied to the college investment decision. Herding behavior refers to following the decisions of everyone else, even if private information provides a signal that suggests following other alternatives. Individuals engage in this behavior because they assume the
choices of decision makers who have gone before them may reflect information that they
do not hold (Banerjee, 1992). This, therefore, reduces the information and search costs
an individual faces. The decision maker will continue her search process and information
gathering only as long the marginal benefit is greater than the marginal cost. She will
stop the search process prior to acquiring all information because it would be too costly
to continue. Herding, thus, allows for a more efficient search for the individual.
However, although herding itself may be a rational behavior, the result of the behavior
may be an irrational decision.

For illustrative purposes, a straightforward example will be used. Suppose a
population of 100 people faces a choice between two options, A and B. It is known that
the prior probabilities are 51% for A being better and 49% for B being better. Further,
each individual has a private signal indicating which option is better. Although each
individual’s signal may be correct (or incorrect) all of the signals are of the same quality.
Now, out of these 100 people, suppose that 99 have received signals that favor B. If the
one person whose signal dictates that A is the better option chooses first, clearly she will
choose Option A (Banerjee, 1992).

The second-mover will then know that the first-mover’s signal favored Option A.
Since the second-mover has an opposing signal of equal quality, the two signals cancel
each other out. This leads the second decision maker to make a choice based on prior
probabilities and choose A regardless of her own signal indicating otherwise. Not
surprisingly, the third-player faces the same choice as the second-player. Likewise, she
will also choose A based on prior probabilities. Without further analysis, it is easy to see
that the entire population will decide upon Option A, despite Option B originally being
the favored choice in the rational environment (Banerjee, 1992).

This example demonstrates how a negative externality is imposed on the entire
population. Suppose instead, however, that the second-player would have chosen to
follow her own signal. The third player would have then received information that the
second player held a signal that favored B. Thus, the third person would have chosen B
as well, and the rest of the population would have followed suit. However, since the
actual scenario results in everyone joining the ‘herd,’ a negative externality ensues, which
can be identified as the herd externality (Banerjee, 1992).

The Formal Model

Following this example, herding can be deduced into a more formal model, as
presented by Banerjee (1992) in his article, “A Simple Model of Herd Behavior.” First,
the article considers a population of size $N$, attempting to maximize a utility function
defined on the space of asset returns. It is assumed that the utility received is identical to
the monetary payout the individual acquires. Suppose, also, that there is a set of assets
indexed by numbers $[0,1]$. The $i$th asset can be denoted $a(i)$. Then, the return to the $i$th
asset to the $n$th person is notated as $z(i) \in R$. Let there be a unique $i^*$ such that $z(i^*) = z,$
where $z$ is positive. For all $i \neq i^*$, $z(i) = 0$. Naturally, every individual would want to
choose $i^*$. Although no individual knows which $i$ is $i^*$, some people do receive a signal
that suggests which asset is $i^*$. However, there is a probability that the signal is false.
Others receive no signal at all (Banerjee, 1992).
The process of choosing assets in this model is sequential. Each decision maker is informed about the choices of each individual that has moved before herself. However, she is ignorant to whether these people held signals or not. The game proceeds until every person in the population has made a decision. The players who chose $i^*$ correctly receive their payout while those who chose incorrectly receive nothing. If nobody chooses $i^*$ it is not revealed and no one gains any rewards (Banerjee, 1992). Three assumptions are made, which minimize the possibility of herding:

**Assumption A.** Whenever a decision maker has no signal and everyone else has chosen $i = 0$, she always chooses $i = 0$ (Banerjee, 1992).

**Assumption B.** When decision makers are indifferent between following their own signal and following someone else’s choice, they always follow their own signal (Banerjee, 1992).

**Assumption C.** When a decision maker is indifferent between following more than one of the previous decision makers, she chooses to follow the one who has the highest value of $i$ (Banerjee, 1992).

The first decision maker makes her choice based on whether or not she has a signal. Clearly, if she does have a signal, she will choose that option. If the has no signal, based on Assumption A, she chooses $i = 0$. The second player will follow the choice of the first player so long as she, herself, has no signal. If she does have a signal, she will follow her own, according to Assumption B (Banerjee, 1992).
The third individual then faces a choice based on the history of the game so far. If she observes that both the first and second player chose $i = 0$, she will also chose $i = 0$ so long as she has no signal. Otherwise, she will follow her own signal. In all other scenarios, if the third individual does not have a signal, she should choose to follow the person who did not choose $i = 0$. However, if both of the two players have not chosen $i = 0$ and have also not chosen the same $i$, then she will make the same decision as the individual with the highest $i$, as indicated by Assumption C (Banerjee, 1992).

It is also necessary to consider the possibility that player three does have a signal, but the two players before her both agreed upon an $i$ that was different than her signal suggested, such that $i$ was not $i = 0$. In this case she will follow the choice of those who have decided before her. If the two players had disagreed, however, she would choose to follow her own signal according to Assumption B (Banerjee, 1992).

If her signal matches the choice made by a predecessor (or both predecessors), she should obviously follow her signal because the probability of that choice being $i^*$ has increased. Obviously, the previous person would have chosen $i = 0$ if she had no signal. So the fact that both individuals received the same signal suggests that the choice is more likely to be $i^*$ (Banerjee, 1992).

Further, the same intuition indicates what should happen in any situation when several options other than $i = 0$ have been chosen, but only one of them has been chosen by two people. Consider the next decision maker in the process outlined so far and assume she does not have a signal that matches any of the options that have already been chosen. (If her signal did match someone else’s choice, she would clearly follow her own signal.) If the choice chosen by the two decision makers was not the highest $i$, both
of those individuals must have the same signal, suggesting that their choice has a higher chance of being $i^*$. Thus, the decision maker will choose this option. In the situation where the choice is the highest $i$, this option is still chosen since it cannot be ruled out that the decision makers did not have the same signal. Thus, in either case, the next person chooses the option that has been chosen by two people. In summary, this implies that once one option has been chosen by two people, the next person should always choose that option as well unless her signal matches an option previously chosen by another player. In such a case, she should follow her own signal (Banerjee, 1992).

A combination of this decision rule with Assumption C indicates that the next decision maker will face one of the following three histories:

**History I)** One option (other than $i = 0$) has been chosen by more than one person, and this is the one that has the highest $i$ (Banerjee, 1992).

**History II)** One option (other than $i = 0$) has been chosen by more than one person, and this is not the one with the highest $i$ (Banerjee, 1992).

**History III)** Two options (other than $i = 0$) have been chosen by more than one person, one of which is the one with the highest $i$ (Banerjee, 1992).

If the first history occurs, the individual will choose the option with the highest $i$. In the second and third histories, the option that is not the highest value of $i$ is the ‘correct’ option because those who chose the option with the highest $i$ most likely did so simply due to Assumption C. Thus, the option that is not the highest $i$ will be chosen and all of
the subsequent decision makers will choose this option as well (Banerjee, 1992). A tree
diagram depicted in Figure 2.3.1 summarizes the outcomes.

![Figure 2.3.1: The kth decision maker’s choice problem (k > 2) (Banerjee, 1992)](image)

This model demonstrates how individuals herd considerably. Even with
assumptions that aim to minimize herding, it is still possible to witness decision makers
who ignore their own signals to follow the decisions made by those who have gone
before them. This is the case, even if the individual is not positive that the choice is
absolutely correct. Thus, although the first two individuals will follow their signals if
they have one, not even the third person is assured to follow her own signal. Even if an
individual has chosen the correct choice, herding at an incorrect choice still may exist
(Banerjee, 1992).

So long as each individual up to k has chosen a different option, if the k + 1
decision maker does not have a signal, she will choose the option with the highest value
of \( i \). As a result, according to the model, all other players will choose that option as well, unless their signal matches the choice of an option already chosen. Thus, even if another decision maker previously chose the correct choice, inefficient herding can ensue at an incorrect option (Banerjee, 1992).

In addition, it is important to note that the equilibrium pattern of choices will be very volatile if the process was repeated several times. The outcome depends on the signals (or lack thereof) of the first few decision makers in the decision making process. These signals may be random and also incorrect. Hence, the power lies in the hands of these few decision makers rather than the quality of the information. From this, it is easy to see how a herding externality occurs when the choices made by previous individuals affect the information that subsequent decision makers have. As a result, this suggests that each individual is less responsive to her own information. In a welfare sense, although it is rational for an individual to join the herd, it reduces the probability that future decision makers will choose the correct option. Thus, in an ex ante societal point of view, the equilibrium pattern of choices may be inefficient for the population as a whole (Banerjee, 1992).

2.4: Distortions in Preferences

The theory of herding behavior is effective in showing how the decision for college can be distorted by the choices that peers and predecessors make. That is, many students substitute the decisions of others for their own information. Although this is rational in the sense that it reduces information costs for the individual, as each person engages in this behavior, they may be effectively making a less efficient decision (a
decision that is different than would be the case if they pursued their own information). This may result in a herd externality, as those who observe the decisions of those before them engage in this same behavior, choosing to do what everyone else is doing. Even more, the process is exacerbated as it is transferred to the herd. In successive rounds, as more and more people chose an inefficient or irrational choice, this option is reinforced. Hence, the likelihood that people will follow along increases over time and the population moves farther and farther away from efficient behavior (Baddeley, 2013).

This has the potential to explain why people are not learning from the mistakes of others. Since everyone is doing the same thing, the information that individuals take away and extract from the process is perceived to be of high quality, even though in reality the information may not be as good as it is initially recognized to be. This demonstrates an information cascade, characterized by individuals observing the actions of others and then increasingly choosing that same option, independent of their own private information and signals. As a result, the conclusions these individuals come to are different than those that would be made autonomously. This suggests a ‘lower’ outcome than if everyone had just chosen independently of everyone else.

Appealing more to emotion, Wood (2011) discusses this behavior as, simply, appealing to custom. It is possible that parents and students are persuaded that attending college is the best decision merely because everyone else is doing it. Citing embarrassment as a factor, many students finishing high school may feel a sense of shame in entering the work force if all of their peers are pursuing a college degree. Moreover, a sense of fear comes into play when individuals believe others will ridicule
their choice if it deviates from the norm (Wood, 2011). Even if decision makers choose an irrational choice, there is comfort in numbers.

On a similar note, a ‘sharing-the-blame’ effect arises, such that even when irrational decisions are made, it is not perceived to be as bad when others make the same mistake (Scharfstein & Stein, 1990). As a result, many people feel it is better to take the risk of being conventionally wrong, than unconventionally right (Baddeley, 2013). This is supported by prospect theory, which postulates that people are risk-seeking when it comes to a loss, but risk-averse in the region of gains. In other words, the loss resulting from disagreeing with the majority is weighted more heavily as negative than the gains are weighted as positive from agreeing with the consensus (Baddeley, 2013). However, it is also important to note that supporting consensus views does not necessarily imply those choices are wrong. Problems may arise when following the herd reflects motivations that differ from individual signals and private information, causing a distortion in behavior.

Specifically, consider the choice of a high school senior. Assume her choice falls between two options, continuing her education versus consumption of the composite good. If the decision maker is accepting other people’s behavior as a substitute for finding out information on her own, she is in effect choosing to ignore private signals she may have come across otherwise. Suppose her indifference curve is, as a result, modified from \( I \) to \( I' \) as displayed in Figure 2.4.1.
The graph shows a distortion in the individual’s preferences. The slope of the new indifference curve, $I'$, is ‘steeper,’ suggesting that the individual is now willing to give up more of the composite good than before, for a one unit increase in education. This is shown in Figure 2.4.2, which presents a comparison of the two indifference curves (illustrated in Figure 2.4.1), removing the budget constraint and all extraneous details.
Figure 2.4.2: A Comparison of Two Indifference Curves

The point of intersection between the two indifference curves is used simply for ease of comparison. The graph shows that for an identical increase in education, the individual is now willing to sacrifice more of the composite good than she was before since the decrease in the composite good for $I'$ is greater than the decrease in the composite good for $I$. In other words, the marginal rate of substitution is steeper for the new indifference curve. Hence, this implies the individual has developed a stronger preference for education and therefore places more value upon it.

Further reflecting this change in preferences, Figure 2.4.1 depicts a shift in the optimal bundle from $A$ to $A'$. Not surprisingly, the new bundle consists of an increase in the quantity of education, accompanied by a decrease in consumption of the composite good. In addition, as shown in Figure 4.2.3, with the individual’s new preferences, if she
were to resort back to consumption at bundle \( A \) (the bundle with a lower quantity of education), she would experience a lower level of utility as shown using the indifference map. This, again, illustrates her new, stronger preference for education.

![Indifference Map](image)

**Figure 4.2.3: A Comparison of Utility Using an Indifference Map**

Herding has the capability of distorting preferences by transforming the function concerning the preference for education. Originally, it was assumed that the preference for \( E \) was a function of only individual preferences, which will be denoted as \( IP \). That is, \( E = f(IP) \). However, an assessment of other individuals changes the nature of the signal that is being received, suggesting that the decision maker is not following her individual preferences only. For example, a student may have a strong preference to major in
drama, but after evaluating the choices of others, she may come to the conclusion that another decision is better because everyone else is doing it.

In many uncertain situations, individuals employ heuristics to guide their interpretation of events. This creates cognitive bias, leading to herding and path dependency, which explains a continued practice based on historical perspectives. Even if a better alternative is available, an individual may choose to follow a previous path because it is easier than following a new one and requires less information costs. While resisting peer pressure may create awkwardness, conformity with others gives psychological reassurance (Baddeley, 2013).

An example of this is the decision regarding going to grad school. Many students faced with this decision may have the mentality that graduate school is now ‘required’ to set an individual apart from her peers, especially with so many individuals now boasting undergraduate degrees. Whether this be the case or not, the fact that many students are choosing to pursue grad school signals to other students that they need to be doing the same thing, whether the costs are warranted or not.

Since the bundle being chosen is no longer a reflection of only individual preferences, the functional form of $E$ must change to reflect the individual’s preferences as well as person $X$’s choice. Therefore $E = f(IP, X)$ since the bundle is no longer just an image of the decision maker’s selection.

Further, if the decision made is based on person $X$, the choice also reflects a sense of status. In other words, choosing the ‘right’ college is also a reflection of person $X$’s choice. For example, this may mean choosing a more expensive or prestigious college. This also applies to the choice of program. Based on the herd’s point of view, some
majors are ‘wrong.’ This may be because they are not respected or because they do not provide an adequate expected future income, for example. However, just because they do not yield a high income does not mean that they are unfulfilling in regards to personal preferences. Nonetheless, herding causes a transformation in such preferences. The quality of an institution and/or program is therefore based not only upon individual preferences, but also upon how the herd assesses it.

2.5: Implications of Herding Behavior on the Human Capital Model

Herding also has the ability to affect the human capital model. To understand this in a more focused sense, it is necessary to first understand the rewards and consequences of herding in a broader sense. First, decision makers can be divided into two types: consensuals and contrarians.

Consensuals are those that follow the herd. For these consensuals, rewards are first increasing and then decreasing in \( n \): the number of people adopting the consensus hypothesis. In other words, by being among the first to pursue an original idea, the value of joining the group is initially high. As more and more people join, however, the negative externality emerges, as returns from following the herd tend to decrease. That is, the individuals are no longer able to signal that they are more talented than any of their fellow herd members. They are in effect lost in the crowd, causing a decrease in returns based on an increasing proportion of consensuals (Baddeley, 2013).

For contrarians, those who oppose popular trends, private rewards are declining in \( n \) at a decreasing rate. In terms of education, a contrarian can be thought of as someone who chooses an alternative to what everyone else is doing. This may be reflective of a
choice as simple as choosing a unique major. Since the contrarian refuses to go along with the herd while more people are joining it, those in the herd reject her behavior and the decision is viewed as disconnected from the ‘correct’ choice (Baddeley, 2013).

As a result, the contrarian will not benefit from the payoff externalities that initially exist for the consensuals. However, as the number of consensuals increases, the decision will be viewed as more pedestrian, causing the returns to the consensual to decrease. In time, the returns to the contrarian will eventually increase because she has a unique set of skills that others do not have to offer. For this reason, the rewards for contrarians fall at a decreasing rate, increasing eventually if the contrarian view evolves into a more acceptable choice (Baddeley, 2013). A graph depicting the rewards to consensuals and contrarians is depicted in Figure 2.5.1.

![Figure 2.5.1: The Rewards Associated with Consensuals and Contrarians as More Individuals Join the Herd](image_url)
An individual who chooses to follow her own signals and preferences in terms of the decision to pursue higher education may experience backlash at first. However, it is entirely possible that her choice may be a suitable choice in the long run. This is especially true if the individual made the decision because it aligned with her preferences and maximized her utility, implying she made an informed decision based on her personal characteristics in an individual context. This is not to say that every contrarian will fair well simply by choosing a path that is not the norm. The consensual and contrarian model does, however, illustrate how herding, or a trend, is the most valuable to those who join the herd early on in the phenomenon. For those who join later, the returns become less and less.

As an extension, as more individuals choose to pursue higher education, the returns may decrease if the norm becomes commonplace. As discussed previously in terms of consensuals, if the supply of educated workers entering the labor force increases, it is likely that the returns will fall. This reflects the increase in $n$. That is, if the supply becomes higher due to herding behavior, net earnings will effectively decrease because too many people will be joining the herd.

The implications of this are reflected in the fact that students make ex ante decisions based on the human capital model. Therefore, based on herding, decision makers may hold false expectations of the present value of earnings, exaggerating the true outlook. As a result, the human capital model would be altered as shown in Figure 2.5.2, based on the consequences of herding.
2.5.2: A Change in Gross Benefits Resulting from an Increase in Individuals Pursuing Higher Education

Clearly, as illustrated, the realized gross benefits are less than the expected gross benefits. This has serious ramifications because a contrast from the original model suggests that the decision would be evaluated entirely different. That is, when comparing the costs to the benefits, it may be the case that the costs do not outweigh future expected earnings if they are significantly lower than originally anticipated.

2.6: Parental Influence

Parents may also influence an individual’s college decision, thus distorting the signal that they receive. Altonji, Blom, and Meghir (2012) comment that this even includes genetic and cultural factors that are not necessarily deliberate decisions made by the parent. Further, choices made by parents that influence their child’s health, exposure
to education, and social experiences prior to high school have strong implications for the child’s decision regarding higher education (Altonji, Blom, & Meghir, 2012).

Parents also have the capability to reinforce herding since they, as well, may feel as if it is best to do what everyone else is doing. Not immune to social pressures themselves, parents are sensitive to others looking down upon them or their children if they do not steer their sons and daughters in the ‘right’ direction. It follows that parents may also face a fear of ridicule that inflames the herding decision to attend college. Assuming that parents only want what is best for their children, it is likely that they will look to see how others have made decisions before them and follow suit. They may feel as if it is not only important for their child to attend college, but to attend colleges with certain characteristics that are viewed by society as prestigious (Wood, 2011).

Further, if parents are basing their beliefs on their own experiences, they may not be up to date on the returns of such a decision. Even if the choice paid off for a parent in the past, this does not necessarily mean that the return will be the same today. In this sense, children may be following the decisions of not only their peers, but of decisions that were made a full generation ahead of them.

Thus, although students themselves may feel pressure from peers to pursue higher education, the expectation for them to attend college may be heightened by parental influence. In this scenario, we see a shift in preferences similar to those seen through herding. The individual will have a stronger preference for education, now a function of individual preferences, X’s choice, and parental influence.

Likewise, the human capital model is expected to reflect this expanding push for higher education, exhibiting decreasing returns based on an increasing supply of educated
workers. As a result, the decision maker may have come to a different conclusion regarding her investment decision if she had followed her own private signals. However, as parental pressures and herding distort her signal, she may end up making an irrational choice.

2.7: Subsidizing

For a majority of undergraduate institutions, states provide large subsidies to higher education, which allows them to provide schooling below actual costs and with little price differentiation between students and programs. It follows that herding, in regards to making the decision to invest in higher education, may be heightened as a result of subsidization if individuals perceive the cost of education to be less than it actually is. In addition, without tuition differentials, students currently participate in implicit cross-subsidization across majors which results in students paying prices that do not actually align with true instructional costs. This has the ability to distort an individual’s perception of the true cost of the degree. This is unique from an economic standpoint as profit-maximizing institutions typically charge prices based on the marginal cost and the consumer’s willingness to pay (Stange, 2013).

As an example, a science major is likely paying a lot less than an art major when comparing the price to the actual instructional costs. The implications of this are that it may deter some individuals who are very interested in art from pursuing this major, especially if they do not feel as if the costs align with future rewards. If art majors are now deciding to become science majors, the supply of science majors will increase,
exacerbating the herd. Therefore, a possible result is a distortion in the composition of
degree types, yielding a world with too many science majors and too few art majors.

Although subsidization is an attempt to enable individuals to obtain higher
lifetime earnings and boost the economy, pushing for everyone to go to college may not
necessarily result in the returns anticipated for every individual. The more students
enrolled in college, the less a college degree serves as a distinguishing signal. In other
words, a degree signals to employers that the individual is qualified and that she
possesses skills that others may not have. However, if many others also hold the same
degree, she may find it difficult to differentiate herself, essentially signaling to employers
that she has the same credentials as everyone else. That is, the individual is less likely to
boast skills that other candidates lack. As a result, the signal will most likely stand for
less and less if a bachelor’s degree is becoming more commonplace.

Herding, however, is not incorporated into the ex ante evaluation of returns. Even
if a college graduate makes $1 million dollars more over her lifetime today than she
would without the degree, this only signals the current expected earnings. This is not
representative of what actual future earnings may be in the presence of a growing and
more competitive labor force supply. This may be indicative of why more individuals are
also choosing to pursue graduate school. They may feel as if this is a necessary step to
take in order to differentiate and distinguish themselves to employers from other
candidates.

Subsidies incentivize individuals by offering a reward, essentially encouraging a
specific behavior in the opposite way a tax would (Stange, 2013). In other words, this
will lower the opportunity cost of choosing education, because $P_E$ decreases. It is
anticipated that this will yield an increase in the supply of educated laborers, perhaps to the extent that it exceeds demand. As a result, degree-holders may realize lower returns or be forced to work outside of the field in which they obtained their degree. Therefore, it is possible that subsidies are enabling students to purchase degrees that they may not use.

This does not necessarily imply that students do not receive non-pecuniary benefits from their time devoted to learning. However, it nonetheless affects individual preferences and the human capital model, mirroring the effects of herding. An individual will hold stronger preferences for education, causing a transformation in the shape of her indifference curve. Also, as more students pursue higher education, it is possible that a decrease in expected earnings may ensue, causing the individual to realize smaller gross benefits than originally anticipated. As a result, this distortion in behavior may lead to an inefficient outcome. This is especially the case if the individual would not have come to the same decision without the distortion in her signal that reflects a price that is not actually aligned with true instructional costs or the ability to repay such costs after graduation.

2.8: Importance of Completion

Completing school, not just attending school, is a major determinant of future earnings. Labor economists have specified that ‘diploma’ or ‘sheepskin’ effects exist, which imply that beyond just going to school for a few years, the actual possession of a college degree boosts one’s earnings significantly. Thus, there are significant benefits to completion over just attending college for some time. Specifically, data indicates that
those who complete some schooling but do not finish their degree only earn slightly more than workers who have obtained no further education beyond high school (Oreopoulos & Petronijevic, 2013).

Individuals also possess abilities, which are a source of uncertainty to the individual (Oreopoulos & Petronijevic, 2013). It may be the case that individuals are being influenced to pursue education, but lack the capability of completing the required work necessary for obtaining a degree. This may lead to a higher dropout rate or having to repeat classes, delaying the time until graduation and therefore, also, the rewards of the investment. This increases direct and indirect costs, such that the return of the investment is less than originally forecasted. Even more, those who are unable to complete their education will have incurred significant costs while realizing a very minimal return.

This is not to suggest that some students should not drop out. Oreopoulos and Petronijevic (2013) contend that school should be thought of as an ‘experience good.’ After individuals discover more about their true potential to succeed in college, some students should choose to let bygones be bygones if they lack an ability to do well. However, the authors also cite that most students discount the possibility they will not perform well from the very beginning (Oreopoulos & Petronijevic, 2013). Supporting the idea of an increasing preference for education, possibly by those lacking the ability to succeed, Oreopoulos and Petronijevic (2013) cite evidence that verifies students are taking longer to complete college and that completion rates have stagnated.
2.9: Incomplete Information

Many students are ill-informed about the costs and benefits associated with college (Oreopoulos & Petronijevic, 2013). As with any investment, risk is involved. Individuals are using only an idea of expected earnings to evaluate the investment decision. However, expected income and returns are not guaranteed. This relates to information costs and the search process an individual must go through. If the costs of obtaining information are too high, especially considering the complexity of financial aid, students may use other details to make the search less taxing. For this reason, high information costs and incomplete information will reinforce herding behavior. It is often the case that students place more emphasis on extraneous details that are not necessarily pertinent to the decision being made (Oreopoulos & Petronijevic, 2013). When the plethora of choices becomes overwhelming, it may be easier to focus on factors that have no relation to a long-term investment.

This illustrates the consequences of relying on unreliable signals. The more informed a student is about her investment decision, the more likely she will be able to receive a return she is satisfied with. However, it must be noted that attaining a higher degree is not an irrational decision for all individuals. The decision becomes irrational when an individual could have chosen an alternative that would have made her better off. This emphasizes how the choice must be made in a personal context. If an individual is relying on unreliable signals rather than gathering information on her own, the decision may not align with the individual’s best interests.
2.10: Credit and Financial Constraints

A fault of the human capital theory is that it inaccurately assumes that individuals encounter no credit constraints. However, research suggests that students do face borrowing and credit constraints, as those coming from wealthier families are 16 percentage points more likely to go to college than those from low-income families (Oreopoulos & Petronijevic, 2013).

An option for those who cannot borrow perfectly against their future income is to work during school. This can decrease the direct costs of education if earnings are applied to tuition costs. However, this also has the potential to delay time until graduation, thus delaying the rewards of the investment. Further, working while in school reduces the amount of time that students may devote to studying. This has the potential to decrease the quality of the skills the investor is attempting to acquire for her future performance in the labor force (Oreopoulos & Petronijevic, 2013). A graph of this scenario is presented in Figure 2.10.1.

Although the student saved a cost equal to the area of $A'AHH'$ through working, by delaying time until graduation by two years, she acquired additional costs equal to area $CB'H'ID'D$. The area labeled Gross Benefits’ indicates the new expected earnings for the individual based on this scenario. Clearly, the costs and benefits must be examined differently than originally expected.
Although individuals may take out loans to pay for the investment, this also has the ability to distort present decisions, particularly if debt aspects in the future are not carefully considered. In the short run, loans make education look less expensive. This may possibly lead to an inefficient decision if individuals are not able to pay off loans in the long run, especially if the returns they receive on their investment are not as high as initially anticipated. An individual must therefore be able to understand the long-term consequences of holding debt when making the college investment decision.

2.11: Conclusion

The human capital theory and the preference theory framework regarding utility maximization provide a foundation for understanding how individuals typically consider
the decision to investment in higher education. However, as it is possible to observe
more and more individuals making the decision to attend college, this may result in an
oversupply of educated workers, potentially affecting the returns realized by the
investors.

The theory of herding behavior explains why large bodies of people are doing the
same thing at the same time, illustrating how students are influenced in their own
decision making based on what others are doing. In an attempt to reduce information
costs, individuals are likely to substitute the decisions of others for their own information.
However, as more and more individuals choose an inefficient or irrational choice, the
process is exacerbated and the population is likely to move further away from efficient
behavior. This also has the potential to explain why individuals are not learning from
their mistakes.

This theory has implications for the human capital model as consensuals are
expected to receive decreasing returns from choosing the consensual decision. As more
students earn degrees, the signal they possess that distinguishes them as a valuable and
skilled worker effectively decreases. Because individuals make ex ante decisions,
however, individuals may hold false expectations of the value of earnings they will
actually receive.

Similarly, the signals that decision makers receive are also distorted by parents,
who may pressure their child to go to college, deeming it as the ‘right’ choice. This
phenomenon appeals to the idea of custom and the notion that individuals are deciding
upon the choices they are making simply because it is what is expected and because it is
believed to ultimately be the correct option. This, too, may result in an oversupply of educated workers, effectually lowering the returns of the investment.

Subsidizing of education by state governments also has the ability to alter the decisions made by individuals. By providing education below actual costs, this incentivizes decision makers to engage in higher education. Although this certainly has the ability to increase social welfare, an unintended consequence may be an increase in the supply of educated workers, perhaps to the extent that it exceeds demand. Again, the result may be that of decreasing rewards pertaining the investment. This may also result in degree-holders having to work in fields outside of their degree, clearly distorting the ex ante predictions regarding the human capital model and utility maximization. Without this bias in price the choice of the individual may indeed be different.

The importance of completion also suggests serious consequences for herding. As diploma or sheepskin effects have been proven to exist, it is crucial for a student to finish her degree if she expects to receive a substantial return on her investment. If herding is occurring, it may be the case that individuals who do not possess the ability to succeed in college are being influenced to acquire a degree. Thus, these students will have incurred significant costs while realizing a very minimal return, suggesting that an irrational decision was made.

Incomplete information is also expected to affect the choice of the decision maker. As many students are ill-informed about the costs and benefits associated with college, as well as the risk involved, students may opt to use other details to make the search process less burdensome. This reinforces the herding model since following the choices made by predecessors significantly reduces the costs of obtaining information. A
more informed student is more likely to make a better decision and choose the option that is the most satisfying to her, personally.

Lastly, the impact of credit and financial constraints has repercussions on the decision making process as well. A fault of the human capital theory is that it assumes individuals encounter no credit constraints. This is not the case. Students may choose to work during college to makeup for such constraints. However this may reduce the amount of time available to devote to studying and may entail a delay in the time it takes the student to graduate, thus, also delaying the rewards of the investment. This requires a different analysis of the costs and benefits that are originally made. In addition, the human capital theory does not consider the individual’s assessment of debt and her ability to pay it off in the future. It is possible that credit constraints may distort present decisions if debt aspects in the future are not considered.

The concepts outlined all relate to the decision making process of an individual concerning the choice to invest in college. Many illuminate how irrational decisions may be made if individuals are influenced by distorted signals. If herding behavior entices individuals to ignore private signals in favor of other unreliable signals, this may lead to an increase in the enrollment in higher education. As a result, the supply of educated workers may be too high, in the sense that it could exceed demand. This has serious implications for the decision maker if the rewards of the investment do not align with original expectations of future returns.
III. Literature Review

In order to extrapolate on the theory further, as well as empirically test the hypothesis that too many individuals are making the choice for higher education, it is first useful to examine previous literature. The articles reviewed provide relevant studies, models, and research pertaining to the college investment decision involving possible irrational decisions. Five articles are discussed, revealing important connections to the choices made by individuals and possible distortions in the way information is being processed.

The first article is included to provide and understand a basic method for empirically evaluating the returns to the college investment. It is useful in that it employs a basic human capital model, thus relating directly to the theory. The article is also helpful in that it considers earnings conditional on either graduating or not graduating, relating to the importance of completion. Further, the study incorporates individual characteristics to investigate the returns more holistically.

After understanding a basic way to measure the returns to education, it is possible to consider how decisions are distorted and altered. The article, “Differential Pricing in Undergraduate Education: Effects on Degree Production by Field” explores how employing varying tuition prices for undergraduates based on major has affected the decisions that are made (Stange, 2013). This tackles the cross-subsidization of education, which is predicted to alter decisions based on distortions in prices. Differential pricing, however, is able to align tuition more closely with instructional costs and the ability to repay costs post-graduation (Stange, 2013). As a result, this has the power to significantly alter the proportion of students entering different majors, as well as the
decision to attend college at all, suggesting that the possibility that irrational decisions are currently being made.

The theory so far has dealt primarily with earning an undergraduate degree, but another possible avenue to explore is the decision to attend graduate school. Analyzing this choice has its benefits in that the decision is more removed from factors that intensify herding behaviors. It is assumed that at this stage in life, decision makers are more informed about their choice and are less affected by parental influence and the decisions of peers. Bedard and Herman (2006) attempt to determine who goes to graduate/professional school based on economic fluctuations, undergraduate field, and ability. This model is appropriate in that it considers individual specific characteristics that are likely to reflect preferences and aspects that affect the choice made. Although economic fluctuations are not of the greatest concern to this paper, other variables may be substituted, such as financial and credit constraints for example. Overall, the article provides an analysis of how individuals are affected by external information beyond the rational environment.

Although it is important to consider how individuals make certain decisions at certain points in time, it is also necessary to consider how college decisions have changed over time. Long (2003) insists that this is particularly important as the college market has shifted from a local market to a more regionally and nationally integrated competitive market. In addition, the tuition price has nearly doubled in real terms over the 30 years that is investigated (Long, 2003). The author maintains that the most important aspects to consider are thus cost, quality, and distance. These are analyzed according to how their role in the decision process has evolved over time. The author also includes individual
characteristics in the decision making process, making the assessment even more valuable. Overall, the article is beneficial in that it essentially examines how individuals process information, especially in the face of a shifting market. Clearly, the number of choices available to students has expanded, creating more information for the decision maker to process. It is probable that the individual may resort to herding or rely on specific signals to aid her in the decision making process.

Although the previous articles consider the returns to education as well as the choice process, it is also helpful to think about ways in which herding may be measured. Sicherman (1991) does not address the decision making process but does provide information regarding ‘overeducation’ in the labor market. This refers to workers holding jobs that require less schooling than they actually have. Sicherman aims to explain why this discrepancy between actual and required level of schooling occurs as well as the resulting differences in returns to schooling. He considers alternative forms of human capital as well as occupational mobility. However, his theory and results also have the capability to reveal an oversupply of workers if individuals are taking lower-level jobs because they have no other choice. This would support a theory of herding behavior, especially if individuals are realizing lower returns than anticipated.

3.1: Estimating the Ex Ante Expected Returns to College

With an increasing amount of individuals making the choice for college education, it is important to understand empirically the returns that the investment provides. Hussey and Swinton (2011) provide an empirical model for estimating the ex ante returns to college, asserting that it has never been more important to be able to
interpret the expected return. This is due to the notion that a college degree is increasingly considered as a necessary credential for employment in many occupations (Hussey & Swinton, 2011).

The article is also important in the way that it considers how individuals measure and decipher the returns to education, ultimately considering how the decision to invest is made. Although the authors do not explicitly refer to the human capital theory, they are in essence measuring returns according to this model. They attempt to understand how students should estimate their expected returns based on an ex ante investment decision. Ex ante, meaning before the event, implies that individuals make a decision based on what they anticipate the return being in the future. In this sense, the decision is a prospective method of calculating costs and benefits (Hussey & Swinton, 2011).

The authors note that the attempts of previous literature to estimate the returns to a college education often make the comparison between the earnings of individuals who obtain a bachelor’s degree and the earnings of those who do not. However, as an improvement to this most basic model, Hussey and Swinton (2011) comment that a growing body of literature has attempted to isolate the effects of graduating from an institution of higher quality, based upon some proxy variable. Nonetheless, they contend that the uncertainty involved in graduating has not been explored. Therefore, since the likelihood of graduating is expected to vary depending on the type of institution one attends, the authors suspect the ex ante expected return on the investment to differ from the estimated returns conditional on graduating (Hussey & Swinton, 2011).

As indicated in Chapter II of this paper, completing school, not just attending school, is a major determinant of future earnings since ‘diploma’ or ‘sheepskin’ effects
do exist (Oreopoulos & Petronijevic, 2013). If, as suggested by the herding model, an increasing number of individuals are making the decision to invest in higher education, it is possible that those who are being influenced to attend are those who lack the ability to actually complete college, potentially incurring a significant cost without realizing the expected return. This may be especially problematic in the face of lengthier completion times and stagnating graduation rates (Oreopoulos & Petronijevic, 2013).

Hussey and Swinton’s (2011) article is unique in that it takes into consideration the importance of completion. They authors consider a highly varied college experience, as well as the variance in the net returns to college. The authors predict earnings, probabilities of graduation, and ex ante expected returns separately by quartiles of a standardized test score. Two broad classes of college selectivity are then used to measure the quality of education.

The dataset by the authors is the National Education Longitudinal Study of 1988/2000 Postsecondary Education Transcript Study Data Collection (NELS 1988/2000 and PETS). This study used a nationally representative sample of eighth-graders at the time of the first survey. A sample of these respondents was then surveyed through four follow-ups in 1990, 1992, 1994, and 2000 (Hussey & Swinton, 2011). The researchers obtained the data through an IES restricted-use data license. Individual annual earnings in 1999, as reported in the latest follow-up survey in 2000, are used to measure returns, the dependent variable. A sample of 4,970 high school graduates with non-missing earnings and covariates were used, some of whom obtained a bachelor’s degree and some of whom attended college but did not graduate (Hussey & Swinton, 2011).
The researchers then divided those who attended college into two groups, based on the college in which they were enrolled. Each was categorized as attending either a selective or nonselective institution based upon the selectivity measure provided by NELS. Using this information, the authors employed a linear regression model to estimate earnings. The dependent variable was identified as the log of annual salary in 1999 (an average of seven years from high school graduation and an average of two years from college graduation). This is denoted by \( Y_i \).

Independent or control variables \( (X_i) \) included by Hussey and Swinton (2011) are NELS test score and indicator variables for race and sex. Then, each of the models is run separately for five groups \( (j) \). These groups can be defined as high school only, some selective college, some nonselective college, nonselective bachelor’s, and selective bachelor’s. Hussey and Swinton (2011) employed this method to measure varying returns to individual characteristics across schooling levels. With the coefficient estimates \( (\beta_j) \), the authors predict the earnings for the entire sample for all five categories \( (Y_{ij}) \), regardless of the education path actually chosen \( (Y_{ij} = \beta_jX_{ij}) \).

Since some students enroll in college but fail to graduate, the likelihood of graduating \( (Pr(g_k)) \) is calculated based on the type of institution that is attended. Note that \( (k) \) represents the type of school. The authors calculate this because they assert that the probability of graduating may differ across school type (Hussey & Swinton, 2011). Hussey and Swinton (2011) state that higher quality (or more selective) schools tend to have higher graduation rates, possibly due to selection effects as well as the school environment. Estimates of the coefficients \( (\gamma_k) \) were then obtained from separate probit models based on individuals who attended colleges of each selectivity group,
conditioning on the same variables as in the earnings equations as well as additional explanatory variables \((Z_i)\). These additional explanatory variables include factors such as whether a parent has earned a bachelor’s degree or higher and whether parental income in 1991 was greater than $50,000 (Hussey & Swinton, 2011).

Hussey and Swinton (2011) represent the probability of each individual graduating from both selective and nonselective institutions by the equation: 

\[
\Pr(g_{ik}) = f(X_t, Z_t, \theta_k).
\]

Finally, to estimate the ex ante expected returns to schooling based on attending college of a particular type, the authors multiplied the predicted salary, conditional on graduating from that college type, by the probability of graduating. This was then added to the product of the predicted salary from some college and the probability of not graduating from that type of college. Thus, the equation is as follows:

\[
\Pr(g_{ik}) \times (y_{ik}|g_k = 1) + [1 - \Pr(g_{ik})] \times (y_{ik}|g_k = 0) \quad \text{(Hussey & Swinton, 2011)}.
\]

To understand the results based on the data used, the authors first consider the estimated returns from schooling for the entire sample when the probability of graduating is not taken into account. This imitates the traditional model often presented in previous literature (Hussey & Swinton, 2011). Following, the authors compare this to the estimates of expected returns that are derived from incorporating graduation rates. The results suggest positive expected returns to college for all groups, and strong returns to school selectivity (Hussey & Swinton, 2011). However, after incorporating individual’s estimated probabilities of graduating, the overall returns to college estimates decreased (Hussey & Swinton, 2011).

The authors’ next step involved comparing the expected earnings from college for individuals who obtained only a high school diploma to the earnings actually obtained by
college graduates. After the researchers ran a traditional earnings regression, the results indicate large returns to a bachelor’s degree and smaller returns to only some college (Hussey & Swinton, 2011).

Looking at the ex ante expected returns including individual-specific graduation rates, Hussey and Swinton (2011) state that the results do not support college attendance for a large portion of students in regards to attending a nonselective institution. Interestingly, students in the top and bottom quartiles of NELS test scores who attended nonselective institutions did not experience a significant difference in expected salaries compared to their actual, current earnings. However, significant increases in earnings are expected for all NELS quartiles from attending a selective college (Hussey & Swinton, 2011).

The authors do not suggest any hypotheses for why the expected salary and actual income for those in the top and bottom NELS test quartiles do not exhibit a significant difference (in regards to attending a nonselective institution). However, they do mention in their introduction and literature review the theory of “under-matching,” referencing previous studies. This applies to talented students, perhaps those in the top quartile, who opt to attend lower quality institutions with high dropout rates instead of attending higher quality schools with low dropout rates (Hussey & Swinton, 2011). It may be that proficient students suffer from an environment that does not foster growth and learning. Those in the bottom quartile perhaps lack the ability to succeed, possibly exacerbated by a below-average learning environment. Thus, the findings may imply that the quality of the school is important for completion and future returns.
Nonetheless, the results are important in that they reinforce the concept of heterogeneity in returns to the investment decision. The article stresses that it is necessary to look at ex ante expected returns (i.e. the human capital model) based on more than just the average return to earning a bachelor’s degree. There are strong implications concerning varying returns based on individual ability and personal characteristics, school selectivity, and probability of completion (Hussey & Swinton, 2011).

It is important to recognize that based on these attributes, some students may be making irrational decisions to attend college if the ex ante returns are predicted to be much higher than what is actually realized. The authors themselves assert that their results do not support college attendance for a large portion of students, specifically for those in the top and bottom NELS test quartiles looking to enroll in a nonselective institution (Hussey & Swinton, 2011). These individuals may incur high costs without a significant return to justify their decision. In this sense, it is possible that students with incomplete information, or those that face high costs in regards to obtaining information, could make an irrational decision if they are following the decisions of others.

Hussey and Swinton’s (2011) model is useful in that it provides a more complete and holistic view of the human capital model. However, it still fails to recognize how individuals consider non-pecuniary rewards in terms of their decision to pursue higher education. The authors assume that monetary returns to education are the only thing that should be considered when evaluating the investment decision. Further, the model does not distinguish between various majors and the returns associated with each.
Nonetheless, the article provides a good benchmark in measuring pecuniary returns to education and is advantageous in its simplicity. The framework is useful in that it exhibits an appropriate model for evaluating differing monetary returns to education by tailoring it to be more specific to the individual at hand.

3.2: Differential Pricing in Undergraduate Education: Effects on Degree Production by Field

In order to assess the demand for education and the individual’s decision process in regards to this investment, it is important to understand how the supply and price of education play a role. Clearly, as pointed out by the human capital model, the costs of attending school are a main component of the investment decision as they are measured against the returns. Stange (2013) points out, however, that higher education is highly unusual in the way that pricing is implemented.

Non-differentiated pricing in education creates cross-subsidization across major fields by charging students similar prices regardless of instructional costs or ability to pay post-graduation, thus affecting how students decide on what major to choose (Stange, 2013). The author examines the concept of differential pricing mainly as a means for policy-makers to alter the mix of undergraduate degrees in order to achieve the greatest return for society as a whole. He believes this method can be used as well as to encourage students to enter high-need fields (Stange, 2013). However, the article is just as important in understanding how the subsidization of education has the ability to alter the choice of the decision maker, ultimately nudging the individual towards a possible irrational decision.
Stange (2013) declares that in an attempt to provide individuals with access to high-quality postsecondary education, states have historically provided large subsidies to public institutions, allowing them to charge students a price well below actual costs and with little price variation between in-state undergraduate students within institutions. However, this is atypical for profit-maximizing firms since they are predicted to charge prices based on marginal cost and willingness to pay (Stange, 2013). In agreement with the human capital theory, the author notes that individuals weigh the long-term expected benefits of studying a particular program against the short-term costs of doing so. If the supply of available spaces in a major is perfectly inelastic, the author declares that individuals are free to choose the major for which the expected benefit is the greatest (Stange, 2013).

After implementing price differentiation, the author predicts that high prices should discourage students from entering impacted fields, holding all else constant. Although he contends that the most salient benefit is the financial return, he also acknowledges, however, that benefits include non-financial aspects associated with each major as well. These include things such as the pleasure of doing something that the individual enjoys, as well as the consumption value during college (Stange, 2013). The analysis is thus valuable in that it considers pecuniary benefits but does not ignore a preference framework that incorporates non-pecuniary benefits as well.

The author asserts that the effect of implementing differential pricing must be tested empirically as the result of charging varying prices on enrollment is ambiguous. For example, an increase in price has the possibility of deterring students if they feel as if the program has become too expensive. However, if programs receive additional revenue
from differential pricing, it is possible that they will use this additional revenue to improve quality, possibly leading to an increase in demand. For example, the department may be able to invest this increase in revenue on better professors or better resources. This has the ability to attract more students, despite the increase in price (Stange, 2013).

Another scenario to examine is an initial equilibrium of over-demand. That is, more students wish to pursue a major than there are spots available. In this case, higher prices may permit oversubscribed departments to expand the supply of education (that is, the number of available spots for students) and increase the number of admitted students (Stange, 2013). Since a wide variety of results could ensue from implementing differential pricing, it is important to investigate the implications of putting into effect such a policy. The results are important, as almost no research on how major-specific prices affect students’ major choice has been conducted (Stange, 2013). This allows for an examination of how the choice of an individual may change after a more accurate price is signaled to the decision maker.

To empirically test the outcome, the author uses data on the mix of degrees awarded by 142 large public universities from 1990 to 2010, focusing on engineering, business, and nursing. These fields were chosen because they are the most common targets for differential pricing and account for a sizable share of all undergraduate students (Stange, 2013). Of the 142 universities, 50 of these adopted differential pricing within this time period (Stange, 2013).

Although the article notes that the most common data source for tuition information is the Integrated Postsecondary Education Data System (IPEDS), Stange (2013) comments that it only publishes differentials by instate status. Thus, the author
obtained data on tuition differentials by program compiled by Glen Nelson (2008) for his doctoral dissertation “Differential tuition by undergraduate major: Its use, amount, and impact on public research universities.” The data contains incremental tuition or fees charged to different majors above base tuition (in percentage terms) for the 2007-2008 academic year at 161 public research universities. This was then narrowed to the 142 institutions of which the precise timing of differential adoption was available, 50 of which had implemented differentials for engineering, business, and nursing majors as of the 2007/2008 academic year (Stange, 2013). Since the timing of field-specific differentials was not obtained, the author assumes that differentials for all majors at an individual school were adopted at the same time.

The primary outcome examined (the dependent variable) is the share of undergraduate degrees awarded by field, which is assessed using the IPEDS Degrees and Certificates Conferred (Completions) module. Stange (2013) then calculates the fraction of bachelor’s degrees awarded in engineering, business, and nursing for each institution in each year from 1990 to 2008 overall and by sex and race.

Institutions are grouped into three selectivity categories as well (most or highly competitive, very competitive, and competitive or less competitive) to aid in understanding what types of institutions have implemented differential pricing. This is also done to control for observed (and unobserved) differences between colleges that may correlate with both major choice and the adoption of differential pricing (Stange, 2013). For example, the author indicates that institutions that adopt differential pricing tend to be slightly larger, better resourced, and more likely to be in the “very-competitive” category. Further, institutions with differential pricing tend to have more engineering and business
majors than colleges without differentials. Thus, since there are apparent differences between institutions with and without differentials, it is necessary to control for such differences (Stange, 2013). The author also takes into account how differential pricing affects the compositions of financial aid of students in impacted fields as well as individual-level data from the 1996, 2000, 2004, and 2008 waves of the National Postsecondary Student Aid Study (NPSAS).

The main strategy employed by the author is a difference-in-difference model. Based upon staggered adoption by institutions, the empirical examination compares changes in major shares between those universities that have recently adopted differential tuition pricing to changes at universities that did not alter their tuition policy during the time period. This is useful in that it isolates how the choices of individuals have altered based on changes in prices, thus eliminating other trends that may affect decisions during the time period. Basically, the model assumes that outcomes for treatment and control schools trend similarly in the absence of treatment (Stange, 2013). Thus, the difference-in-difference strategy is helpful in segregating the effect of pricing to analyze how irrational decisions may be occurring in the face of previous cross-subsidization.

Stange (2013) declares that the coefficient of interest is the change in the share of degrees granted for the three majors following the adoption of differential pricing. To estimate this coefficient, the author first documents the major share changes following each school’s adoption of differential tuition by calculating the change in the fraction in each major following the policy change. The results indicate that the majority of schools experienced a decrease in the fraction of students majoring in engineering and business but, in contrast, an increase in the fraction of students majoring in nursing when
differential pricing was introduced. This trend persists even after a number of different identification strategies and robustness checks (Stange, 2013).

To better understand how time plays a role in the decision-process, the author specifies a difference-in-difference model that permits separate effects for the immediate (0, 1, 2 years after the policy was enacted) and medium-run (3 and 4 years after) time periods. These event-study estimates suggest that any treatment effects may take 3 to 4 years to emerge as the most notable changes occurred three years after differential pricing was enacted (Stange, 2013). Specifically, the author’s results show that differential pricing for engineering is associated with a statistically significant 1.1 percentage point decrease in the share of degrees awarded in engineering after three years (on a base of 14.7 percent). For business, a 0.8 percentage point decrease in the share of degrees awarded was recorded within three years (on a base on 19.5 percent). Dissimilarly, nursing actually displayed a 0.8 percentage point increase in the share of degrees awarded (on a base on 4.4 percent), though imprecise and not significantly different from zero (Stange, 2013).

Stange (2013) concludes that these figures are associated with fairly large elasticities. For engineering and business, the implied elasticities are positive 0.51 and 0.30, respectively. The elasticity for nursing is positive and almost unity (elasticity = 1.0). Stange (2013) also reports that it does not appear that additional institutional grant aid, from an increase in prices, offsets the increased tuition for impacted majors. Women and minorities, however, have larger proportionate effects than male and white students (Stange, 2013).
This study is important in that it illustrates how implementing differentials may impact the field that students pursue. Although the author focuses only on specific majors, it is possible that such price changes may affect not only what field students choose to pursue, but also what college to attend, or even whether to attend college at all. The fact that students appear to respond to differential pricing suggests that students may be currently making irrational decisions based off of ‘imprecise’ information. If the signal sent to the student about pricing is distorted based on subsidization, the individual’s decision is expected to be different than a decision made in a ‘rational’ or neutral environment. If the price were aligned more closely with true instructional costs or the ability to repay costs after graduation, the choice would likely be altered. Thus, if a decrease in students pursuing degrees with higher prices does actually result from the implemented policy, this may suggest that herding is indeed occurring in the higher education investment decision.

3.3: Who goes to graduate/professional school? The importance of economic fluctuations, undergraduate field, and ability

Although extensive research has been conducted on the decision process of choosing whether or not to attend college, the focus has primarily been on high school students pursuing an undergraduate degree. However, the decision to attend graduate school is likely to reveal important contributing factors to the decision process as well. The choice to pursue graduate school may be the most useful decision to examine as it is more removed from factors that spur on herding behavior. In other words, by reducing other ‘noise’ in the process, it may be possible to identify how irrational decisions are made even by individuals who are more informed or less affected by others in their
decision. It is probable that at this point in the decision maker’s life, she is likely to hold more information and be less influenced by her parents or peers.

Bedard and Herman (2006) attempt to analyze this decision to attend graduate school beyond what has been done in the past. They note that a substantial amount of literature already exists that examines the impact of the business cycle on the decision of individuals to complete high school and enroll in college. The majority reports that enrollment increases when the unemployment rate rises, thus, displaying a countercyclical relationship in general (Bedard & Herman, 2006). Although the authors consider several descriptive editorial/review articles on the subject related to graduate school, they are not aware of any previous econometric study that analyzes the effect of the business cycle on advanced degree enrollment (Bedard & Herman, 2006). As they point out, the choice to attend graduate school is an important decision to consider as the cost, in terms of lost wages, is much higher than for high school or undergraduate education. Thus, it is has the potential to be more seriously affected by economic fluctuations (Bedard & Herman, 2006).

Although the article looks at enrollment as a whole in regards to economic fluctuations, it is valuable in that it may reveal or mimic ways in which prices or financial constraints affect the choice being made. As a substantial fraction of the population now holds advanced degrees, if that number continues to rise, it is critical to understand how the decision to pursue such a degree may be distorted under different conditions, especially in respect to large costs associated with the investment.

Bedard and Herman’s (2006) study is also pertinent because it employs the human capital investment model when examining a Bachelor of Science (B.Sc.) holder’s
decision to enroll in an advanced degree program. The authors assume that individuals evaluate the expected pecuniary and non-pecuniary returns associated with various educational paths and choose the option that maximizes their expected lifetime utility, which is in accordance with the preference framework as well. Specifically, they focus on an individual’s enrollment decision just after earning an undergraduate degree. In concurrence with the human capital model, they reiterate that the decision to enroll in a graduate/professional program depends on the expected advanced degree wage premium relative to the net educational cost incurred (Bedard & Herman, 2006).

To create a more stylized and individual-specific model, the authors incorporate differences in ability and educational background since wage opportunities and/or graduate stipend offers will also differ as a result of this (Bedard & Herman, 2006). The authors’ research is valuable in this sense because it considers factors beyond just earnings, allowing for a better understanding of how the individual choice process is affected by characteristics beyond simply the average expected return.

Since the response of enrollment to business cycle fluctuations is unclear, it must be addressed empirically (Bedard & Herman, 2006). In order to estimate the impact of the business cycle on the decision of individuals with different ability levels to enter an advanced degree program, Bedard and Herman (2006) maintain that detailed information is required that relates to family background, individual attributes, educational history, ability, and the economic conditions over many years. To acquire this information the authors use as their main data source the 1993-2001 National Survey of Recent College Graduates (NSRCG) provided by the Scientists and Engineers Statistical Data System (SESTAT). This contains detailed information on recently graduated science and
engineering bachelor and master’s students from 1990 to 2000 (Bedard & Herman, 2006).

The NSRCG is a series of five nationally representative samples of recent science and engineering B.Sc. and master’s graduates who earned a degree from a United States institution in the two academic years prior to the survey reference date (Bedard & Herman, 2006). The surveys primarily look at science and engineering degrees, classified into five categories: computer science and mathematics, life sciences, physical sciences, social sciences, and engineering (Bedard & Herman, 2006). To account for and measure business cycle fluctuations, the authors use state-level unemployment rates from the Local Area Unemployment Statistics program of the Bureau of Labor Statistics (BLS). Specifically, the annual state unemployment rate for the civilian non-institutional population age 20-24 is used.

Although the NSRCG covers a large portion of students, a fault of the data is that it fails to provide information on non-science and engineering students (business administration, arts, and humanities). However, an advantage of using these surveys is that academic history information can be linked to individual and family attributes as well as local unemployment rates over time, thus allowing for an examination of graduate school enrollment decisions over the business cycle (Bedard & Herman, 2006).

The dependent variable being measured is enrollment in graduate school. To measure ability, the authors use B.Sc. GPA as a measure of observable ability. In order to control for socioeconomic status and family background, the authors use data on age, number of months since B.Sc. graduation, ethnicity, state location of high school and B.Sc., an indicator for being married with an employed spouse, the number of children
residing in the household, and mother’s educational attainment, all of which are retrieved from the NSRCG (Bedard & Herman, 2006). Since the NSRCG contains only observable wage data for employed individuals during the reference period, data on expected wages, potential wage opportunities, or other wage offers are obviously not available for the entire sample, which prompted the use of state-level unemployment rates as an indicator of economic opportunities faced by recent graduates (Bedard & Herman, 2006).

The authors employ a probit model as their empirical approach since the dependent variable can only take two values (enrolled or not enrolled), therefore exhibiting a binary response (Bedard & Herman, 2006). The independent variable, denoted $G_{it}$, equals 1 if individual $i$ is enrolled in a particular type of graduate education program (denoted by $t$) in the survey period directly after obtaining a B.Sc. It follows that $G_{it} = 0$ otherwise. The choice problem is defined as below:

$$G_{it}^* = \beta UER_{it} + \text{GPA}_{it} \beta_2 + \text{MAJ}_{it} \beta_3 + T_t \theta + S_i \theta + X_{it} \gamma + \epsilon_{it}$$

($G_{it}$ equals 1 if $G_{it}^* > 0$ and $G_{it} = 0$ if $G_{it}^* \leq 0$)

Note that $G^*$ is the propensity to enter an advanced degree program directly after college, UER is the state-level unemployment rate, GPA is a set of three GPA indicator variables, MAJ is a set of dummy variables indicating the individual’s undergraduate major, $T$ is a vector of B.Sc. graduation year indicators, $S$ is a set of state of B.Sc. graduation indicators, $X$ represents observable characteristics, and $\epsilon$ is the error term (Bedard & Herman, 2006).

Distinguishing the effect of the business cycle on graduate enrollment is based on two assumptions set forth by the authors. First, Bedard and Herman (2006) assume that within state unemployment rate fluctuations are independent of the unobservable
characteristics that contribute to determining enrollment status, such as unobserved ability. Secondly, it is assumed that fluctuations appropriately capture the entry-level labor market economic conditions at B.Sc. graduation. Further, the regression equation is in reduced form because unemployment rate is meant to represent the combined response to potential wage and graduate stipend offers. That is, it reflects the net effect of the difference between the potential forgone wage and the potential graduate stipend offer (Bedard & Herman, 2006).

Applying the data to the choice problem defined by the equation, the authors obtain an unbiased estimate of the impact of the business cycle on graduate school enrollment. The results indicate that, not surprisingly, students with higher GPAs are more likely to pursue advanced degree programs (Bedard & Herman, 2006). In regards to major, physical science majors are the most likely to enroll in Ph.D. programs. The authors observe that male Ph.D. enrollment is counter-cyclical, male master’s degree enrollment is procyclical, female professional school enrollment is counter-cyclical and all other enrollment appears to be acyclical (Bedard & Herman, 2006).

Understanding why male master’s degree enrollment is procyclical is undoubtedly a major point of interest. The authors propose that the reason for this may be due to a substitution of master’s degrees for longer Ph.D. programs during economic expansions. In addition, there may be a reduction in the number of employers paying for part-time master’s programs if economic conditions are poor (Bedard & Herman, 2006).

In addition, the authors do admit that there is a substantial difference in B.Sc. majors across gender. There is an extreme concentration of women in social science, which makes it difficult to detect business cycle driven graduate school enrollment in
models due to the small cell sizes for many of the undergraduate major state-year cells (Bedard & Herman, 2006). Nonetheless, the authors conclude that male and female graduate school enrollment decisions are differently impacted by the business cycle, with men being much more responsive to labor market fluctuations (Bedard & Herman, 2006).

Although the article focuses on how decisions are affected by business cycle fluctuations, it is still relevant to the research question at hand because it looks at how individual decisions are affected by certain signals and information. It may be possible to use another variable, other than the unemployment rate, to empirically examine the decision making process of individuals and how the choices they make may result from distorted signals. In addition, the article was useful in incorporating individual-specific characteristics that may also affect the decisions being made and potentially distinguish how some individuals behave irrationally. Overall, the article demonstrated how individual decisions are affected by external information beyond a rational or neutral environment.

3.4: How have college decisions changed over time? An application of the conditional logistic choice model

Long (2003) notes that over the last 30 years, the market for higher education has changed dramatically, requiring an analysis of how the individual’s choice for college has changed. The author claims that as American higher education has grown from a local market to a market that is now integrated regionally and nationally, this has made the college enrollment decision increasingly complex. For this reason, the author insists that colleges have had to differentiate themselves, resulting in higher variation between
colleges and increasing options available to students, including tuition prices and quality (Long, 2003).

Although available subsidies offered to students has grown as well, Long (2003) claims that the list tuition price of colleges has nearly doubled in real terms over the last three decades. Nonetheless, the benefits associated with college have increased as well, leading people to respond by enrolling in greater numbers (Long, 2003). Although the author is concerned with simply how the changing dynamic of the education market has affected individual choices, it is possible to extend her analysis beyond this narrow scope and apply it to a herding framework to identify individuals who are making irrational decisions.

Since the author hypothesizes that individuals have altered the way in which they consider their college options, the paper attempts to examine how individuals from 1972, 1982, and 1992 chose whether to and where to attend college by estimating the importance of postsecondary costs and quality. This is applicable and important to the research question because it evaluates the decision of the individual. That is, by looking at person-specific factors, the paper improves on past research by using more individualistic characteristics rather than relying upon aggregated data (Long, 2003). In addition, the article claims to be the first to examine how the decision process has changed over time and how the role of price, quality, and distance has transformed based on the choices of individuals (Long, 2003).

The paper addresses two questions: how individuals decide between colleges, conditional on attendance, and how individuals decide whether to attend at all. The author notes that previous literature does provide estimates of the effect of price at
different times in the past, but points out that they are not comparable because each study involves different controls, different econometric models, and different data samples that have not been adjusted in order to provide comparable results (Long, 2003). The author stresses that it is important to study change over time because beyond influencing the enrollment decision, price may also affect which college an individual may choose. This is especially the case as it is becoming more important to not only attend college, but to choose a college that provides the best possible return (Long, 2003).

Long (2003) also argues that college quality has largely been ignored in previous literature. Although some studies have considered how it affects the return on the investment, it also has the ability to affect the ex ante decision of students (Long, 2003). The author reasons that this is increasingly important as the market is transitioning from a local market to a nationally integrated market, thus driving colleges to compete for students based on the quality of their institution and educational products rather than simply price (Long, 2003).

The author also insists that distance to college is an important factor to include in the decision model as well. A shift in the market to become more nationally competitive (paired with advances in transportation and communication), as well as the increased availability of college information, implies that distance to college is an important factor that must be assessed. Its inclusion is necessary to truly examine how the college enrollment decision has changed over time, allowing for a more complete insight into how students make decisions of whether, and where, to attend college (Long, 2003). This may also provide a better understanding of how individuals process information and yield
a stronger grasp on where the signal they receive can be distorted, hence generating irrational decisions.

The theoretical framework presented assumes that an individual has $J$ colleges to choose from. Any individual college, $j$, can be described by a vector of characteristics, $Y_j$, including measures such as prices, resources available for students, and location, for example (Long, 2003). Then, relating to both the human capital model and utility maximization framework based on preferences, Long (2003) indicates that the individual will infer how much value-added each school will produce for her human capital and the consumption goods the college offers. Further, the author introduces another vector, $X_i$, which contains the individual’s characteristics such as high-school performance and family income, which are likely to affect the individual’s demand for education and her opportunity set (Long, 2003).

Then, the value of the $j$th college to the $i$th decision maker may be given by $U(Y_j, X_i)$. The author clarifies that utility may have random elements so that all individuals are not assumed to have the same tastes. Individuals then compare the potential returns to attending different colleges, as well as the option to not enroll at all (implying they would enter the labor market immediately). This reflects how the decision to attend college is made up of two choices. That is, after determining the best college option, the individual must then decide whether to attend the college at all, ultimately deciding upon the choice that maximizes lifetime utility subject to her budget constraint (Long, 2003).

To empirically understand the how individuals make their college decisions, the author asserts that an empirical model should investigate the tradeoffs between the opportunity chosen and the available alternatives that are not selected. For this reason,
she uses the conditional logistical regression model, also referred to as McFadden’s choice model, since it exploits extensive detailed information on alternatives. That is, the variables of the conditional logistical regression model are choice-specific attributes rather than individual-specific characteristics. Specifically, this methodology allows for over 2,000 possible college alternatives and considers important match-specific information between the individual and the college (Long, 2003).

To estimate the conditional logit, the author organizes the data as pair-wise combinations of each student $i$ with each school $j$ with the observations stratified by individual into groups of $j$. Then, the conditional logit model is made up of $j$ equations, using these combinations, for each individual $i$, with each equation describing one of the alternatives (Long, 2003). This format is beneficial because it takes into account match-specific variables based on the interaction of individual $i$ with college $j$. The conditional logit model is then capable of determining the probability of enrollment at each of the colleges in the stratum relative to all other alternatives, with the dependent variable equal to one for the alternative that was chosen (Long, 2003).

Long (2003) uses three data sources that provide information on high-school graduates from 1972 to 1992. The NLS72, the HSB, and the NELS, 1988 deliver information on the backgrounds and college decisions of individuals who graduated from high school in 1972, 1982, and 1992, respectively. The data was all collected by the National Center for Education Statistics (NCES) with the explicit purpose of being made to be comparable to each other (Long, 2003).

To gather information on college costs, quality, and location, Long (2003) uses several sources. The Integrated Postsecondary Education Data System (IPEDS) and its
predecessor, the Higher Education General Information Survey (HEGIS), were used to furnish institutional characteristics and financial data. Supplementary data on median student-body scores was taken from Cass’ and Brinbaum’s *Comparative Guide to American Colleges* for the 1971-1972 school year and Barron’s *Profiles of American Colleges* for the 1981-1982 and 1991-1992 school years. Data on faculty quality was based on data from the College Board’s Annual Survey of Colleges.

Employing the conditional logistic model, the author matches each student with each possible college. College price, distance, and quality are defined as follows. Each student is assigned the in-state tuition price if they live in the designated state. To glean a better sense of net price, the amount of Federal Pell Grants that a student could expect to receive was subtracted from the list tuition price. Distance was calculated by comparing the zip codes of the college and high school attended by each person. Finally, quality was measured using the median SAT score of the college’s student body as well as instructional expenditures per student. To compute faculty quality, the author uses the college’s student-faculty ratio and the percentage of the faculty with a Ph.D. To control for supply constraints, she also includes admissions constraints to prevent the model from predicting that individuals with low test scores will attend highly selective institutions. In addition, the author includes undergraduate full-time equivalent (FTE) enrollment, which allows for the fact that larger schools have a higher level of visibility and are able to admit more students, thus giving these schools greater choice probability (Long, 2003).

Then, using the conditional logistic choice model, estimates are obtained of how the attributes of a college affect the probability of each individual choosing to attend that college, conditional on attending any postsecondary institution. Not surprisingly,
distance and tuition price negatively affect the likelihood of a student choosing a college, all else equal, whereas college quality positively affects the student’s choice (Long, 2003). However, a comparison across cohorts indicates that the significance of these factors has changed over time (Long, 2003).

Price has played a declining role and the role of distance has also weakened (Long, 2003). The author finds that college quality was not an important factor in the college decision for the 1972 and 1982 cohorts, but played a more important role for the class of 1992. Specifically, students from the 1992 cohort were more likely to choose schools that have higher scores than their own (Long, 2003). To help alleviate the concern that the differences may be related to the changing makeup of college-bound students rather than real changes in decision making patterns, the author standardizes later cohorts to reflect the 1972 population. This revealed that the negative influence of price on college as well as the role of distance weakened. The importance of quality, however, increased for more recent cohorts (Long, 2003).

The author also used a second method of reducing the differences between the cohorts by limiting the sample to students with specific characteristics. Similar results were found, but low-income students in later cohorts were still found to be sensitive to price. This implies that even after advancements in financial aid programs, price negatively impacted the choices of graduates in 1982 and 1992 in a manner similar to the way it influenced the class of 1972. However, high-ability students did not exhibit as much sensitivity to price (Long, 2003).

Although this model estimates the choice between colleges, the individual must also decide whether to attend college at all. The author maintains that it is acceptable to
treat the enrollment and choice decisions separately, contending that it is equivalent to determining the choices of students using backward induction, similar to how it is considered in game theory (Long, 2003). To estimate the decision to attend, she uses a logistic model, which controls for family background, student achievement, and the unemployment rate of the individual’s county (Long, 2003).

Similar to deciding between colleges, Long (2003) investigates how college characteristics affect the likelihood of an individual to enroll in college in general, estimating specifically the effect of price, distance, and quality on the decision to attend. The results show that the role of price and distance has fallen in terms of enrollment while quality did not play a role in determining whether an individual from any cohort chose to attend college (Long, 2003). Specifically, for the class of 1992, with the acceptance of distance, no college characteristic is estimated to have statistically significant effects on enrollment. The author points out that this may suggest that capital constraints are no longer a major deterrent in college decisions (Long, 2003).

From the analysis conducted on the college enrollment decision, it is possible to see a more competitive education market has emerged in terms of deciding between schools. The results suggest that quality has recently played a stronger role in determining which college to attend, conditional on enrollment (Long, 2003). It appears that earning a degree is no longer sufficient, but that it is also important to earn a degree of a certain prestige and caliber. If individuals are enrolling in larger numbers, they must compete with one another to earn a respected degree that signals to employers a higher worth in the labor force.
The results also illustrate that the role of college costs in the enrollment decision has been found to decrease over the time period, suggesting the factors beyond capital constraints play a role in the decision making process (Long, 2003). As an extension of the author’s model, it may be the case that students use other factors beyond price (such as herding) to aid in making their decisions. This may imply that decision makers are using irrelevant information to choose between colleges as well as whether to attend or not.

If individuals are receiving signals about factors beyond price alone, this suggests that the human capital model is not sufficient on its own in evaluating the college investment decision. In addition, as distance becomes less of an important factor, individuals have an increasing amount of colleges to choose from. This provides individuals with an abundance of choices, likely presenting the decision maker with an overwhelming amount of information and a daunting choice. Thus, in regards to the search process and information costs, it is probable that students are using some method, such as herding, to lessen the burden of sorting through information. This is especially the case since, as Long (2003) mentions, the market for education has shifted from a local market to a more nationally-integrated market.

3.5: “Overeducation” in the Labor Market

In regards to the returns on the investment in higher education, Sicherman (1991) declares that a potentially problematic observation is that a large number of workers report a discrepancy between their own level of schooling and the level of schooling that is required by their job. For this reason, he attempts to examine the causes of this
observed discrepancy and the resulting differences in returns to schooling. He argues that the synchronization between the education system and the labor market has been a major concern for several decades among educators, policymakers, and social scientists (Sicherman, 1991). Although this may reflect disequilibrium or inefficiency in the labor market and/or the schooling system, it also has the potential to reveal irrational decisions made by individuals if the return they realize does not equate to what was initially expected.

Sicherman (1991) explains that his article is motivated by two stylized facts. First, workers in occupations that require less schooling they what they have actually received are classified as ‘overeducated.’ Generally, they earn lower wages than workers with similar levels of schooling who hold jobs that require the level of schooling they have actually obtained. Nonetheless, these overeducated workers do earn more than their coworkers who have less education (but is the required amount for the job). The author insists that this implies inefficiency, as overeducated workers should be considered as underutilized workers (Sicherman, 1991).

As an extension, this may relate to how individuals are making their decisions to invest in higher education. If there exists a high amount of overeducated workers, it may be the case that herding is taking place if workers are forced to take lower-level positions simply because they cannot get jobs aligning with their actual level of schooling. Further, if overeducated workers are indeed receiving lower wages (relative to those with similar levels of education that hold jobs requiring this level of education), this could suggest that there is an oversupply of educated workers, leading to a decrease in realized earnings.
The second stylized fact presented is that undereducated workers, those who hold jobs that require more schooling than they have obtained, receive higher wages than workers with the same level of schooling who work in jobs that actually require their level of schooling. However, they do earn less than coworkers with a higher level of education (Sicherman, 1991).

Sicherman (1991) then aims to explore discrepancies between actual and required levels of schooling, and the resulting differences in the returns to schooling, by using a human capital mobility framework. The author’s goal is to provide a better understanding of relations among schooling, the patterns of wages, and workers’ mobility across firms and occupations. However, the article also has the potential to garner information on the effects of herding and an oversupply of educated workers in the labor force, suggesting a decision making process that may not be rational for some individuals.

The author first suggests that a possible explanation for such discrepancies involves the potential tradeoff between schooling and other components concerning the level of human capital. For example, he suggests that on-the-job training may makeup for a lack of education, which thus only indicates on the surface that workers are undereducated. In other words, workers may qualify for similar jobs even if they hold different levels of schooling because it is possible that they hold comparable levels of human capital (Sicherman, 1991). Hence, it is important to remember that schooling is not the only thing that increases human capital. Rather, it includes anything that enhances the workers skills and ability to perform in the labor market. The author also mentions that other factors, such as more or less market experience as well as more or
less innate, unmeasured ability are also examples of factors that contribute to human capital (Sicherman, 1991).

The author’s second hypothesis is that such discrepancies occur due to the mobility patterns of workers. That is, he believes that the phenomenon can at least in part be attributed to how individuals move between firms and occupations (Sicherman, 1991). Based on this hypothesis, Sicherman (1990) presents two theories that could explain a temporary discrepancy.

First, the matching theory suggests that a mismatch between the worker and the job exists, indicating simply a bad match in the short run. It is assumed that the employee will eventually change her job to one in which she is better qualified for to correct the problem (Sicherman, 1991). The second theory, career mobility, predicts that workers may temporarily work in mismatched jobs in order to advance their career in the future. In other words, the skills they obtain from the current job may be used at a later point to help with a higher-level job (Sicherman, 1991). Although the author presented several explanations for overeducation, the paper chooses to focus specifically on these two theories.

Sicherman (1991) uses the Panel Study of Income Dynamics (PSID) to obtain data. From this, he uses a sample of male heads of households, aged 18-60 to test his hypothesis. He then conducts a cross-sectional analysis with the 1976 and 1978 waves of the PSID, while the mobility models are estimated on the 1976-77 and 1978-79 waves. For the 1976 and 1978 waves, participants were asked how much formal education was required to get the job that they currently hold, with answers bracketed into seven classes: 0-5, 6-8, 9-11, 12, 13-15, 16, 17 (Sicherman, 1991).
Based on this information, individuals are defined as over or undereducated if their reported years of schooling lie outside the reported range of required years of schooling (Sicherman, 1991). The results indicate that about 40% of the workers in the sample were considered to be overeducated while 16% were undereducated. Further analysis reveals that overeducated workers on average did receive less on-the-job training, while undereducated workers reported more on-the-job training (Sicherman, 1991).

This supports the hypothesis that skills obtained by labor market experience may be substituted for insufficient schooling (Sicherman, 1991). Nonetheless, the prevalence of overeducation, as compared to undereducation, does possibly suggest an oversupply of educated workers if students are investing in education beyond what is actually required. It is not unreasonable to assume that this could potentially be affected by herding.

In regards to the matching theory, the author uses a regression model to estimate the likelihood that individuals with over/underschooling will change firms or occupations. Important variables include the number of years of schooling, experience in the career field, whether the individual is a member of a union, the individual’s race, city size, and whether the individual is married, disabled, and currently overeducated or undereducated. The results confirm that workers with years of schooling that deviate from those required on the job have higher probabilities of changing both their occupation and their firm, indicating a bad match (Sicherman, 1991). They author further hypothesizes that undereducated workers should experience a higher incidence of layoffs. However, this is rejected by the empirical results. No evidence supports the notion of a good or bad match in this case between the worker and the firm (Sicherman, 1991).
Sicherman (1991) next examines the theory of career mobility. He notes that an occupational change is defined to occur when the two-digit occupational category in the PSID reported by the worker in two successive surveys is different. In basic terms, it is assumed that an occupation change occurs if there is an apparent change in the tasks performed by the worker (Sicherman, 1991). Occupational mobility that results from career mobility is considered mobility in the sense that the individual moves to a higher-level occupation. This is measured by vertical distance and is calculated by measuring the difference in the mean levels of human capital needed to work in the occupation after required training is completed (Sicherman, 1991).

An example that the author gives to illustrate this phenomenon is of a police officer. Both a high school and college graduate may choose to enter their careers at this same level, that is, as a police officer. However, as opposed to the high school graduate, the college graduate is most likely using this entry-level position to obtain necessary skills to advance his career in the future. He may be promoted to sergeant or leave the institution to become a private detective (another higher-level occupation). Therefore, career mobility assumes that it is reasonable for workers to take on jobs in which the schooling requirements are lower than actual schooling levels (Sicherman, 1991).

Since this trend involves the case of overeducation (as opposed to undereducation), the author makes the decision to focus his study on this scenario only. He anticipates that workers who are overeducated are more likely to move to a higher-level occupation and higher wage level. Using the PSID and a regression model with the same key variables used for the matching theory, the empirical results suggest that
overeducated workers are indeed more likely to move to higher-level occupations than workers with the required level of schooling (Sicherman, 1991).

As a critique to his own work, Sicherman (1991) does mention that the discrepancy between actual and required schooling and observed wage differentials could be the effect of quality difference between different groups. In other words, overeducated workers may be classified as so if the quality of education they receive is lower or if they lack a general ability to perform based on individual characteristics. The opposite may also exist for undereducated workers (Sicherman, 1991). However, this argument is not tested nor included in Sicherman’s model.

Although the author’s models investigate why overeducation and discrepancies occur between the worker’s actual level of schooling and that which is required, the implications of overeducation go beyond what is directly analyzed, possibly providing a link to herding behavior. In support of this, Sicherman (1991) does note an increase in college attainment in the United States resulting from the baby boom generation reaching college age. This may have lead to the resulting reduction in the returns to schooling, reviving the notion of overeducation (Sicherman, 1991). Although the author perceives this as mostly a temporary situation, this does not mean that the consequences are not long lasting.

If the market for labor cannot adjust to such a situation, equilibrium will be distorted, suggesting the ex ante decisions made by individuals may differ from what is actually experienced. The paper attempts to provide an explanation for overeducation, supporting logical explanations for discrepancies. However, it does not necessarily consider how large numbers of people making the choice for higher education may
distort, for example, a tradeoff between schooling and other forms of human capital. Inefficiency must be viewed and judged in a broader context, as the author admits himself.

Although discrepancies may occur due to mismatches between the worker and the job at the beginning stages of her career, this may not account for all situations. It is also possible that individuals are taking lower-level positions because they have to. The empirical result that overeducated workers earn less than comparable individuals who hold jobs that require their same level of education suggests that the investment is less valuable to those who are not able to take a job that matches their actual level of schooling. Even if this only occurs at the start of a career, it may nonetheless still reduce wages over the individual’s lifetime. Thus, it is important to understand why overeducation is occurring and the implications it has on the investment decision and potential irrational choices being made by decision makers.

3.6: Conclusion

Each of the articles presented examines aspects related to the decision process involving the investment in higher education. Hussey and Swinton (2011) provide a framework for analyzing the return to education and stress that ex ante expected returns should be evaluated based on more than just the average return to earning a bachelor’s degree. Ability, personal characteristics, school selectivity, and probability of completion are all important factors to take into consideration. This provides a more holistic approach to the human capital model by incorporating additional predictors of the
expected earnings. In addition, this allows for a better analysis of how individuals are processing information as well as how the decision relates to individual preferences.

This may be taken to the next level in researching graduate school decisions, as explored by Bedard and Herman (2003). Further, as differential pricing is more common at graduate institutions, analyzing this decision is likely to reduce the effects of subsidizing, as illustrated by Stange (2013). If individual-specific characteristics are incorporated, such as undergraduate field and ability, it is easier to identify how individuals are affected by certain signals and information that extend beyond the rational environment.

In addition, by researching the choice to enroll in graduate school, this may reduce extra ‘noise’ associated with the decision making process, possibly making it easier to isolate irrational decisions. As Long (2003) points out, this may be especially important to consider in a more nationally integrated market for education. By looking at what information individuals do rely on to make decisions, it may be possible to determine how decision makers are processing signals and if herding is occurring to aid in the search process by reducing information costs.

This may potentially be measured by evaluating what types of occupations individuals are entering. If workers are overeducated, as addressed by Sicherman (1991), it may be the case that there is an oversupply of educated workers who are receiving returns that are lower than expected, based on the ex ante decision. This has the potential to support a herding model and illustrate how some individuals may be making irrational decisions to enroll in college according to how they process information when engaging in the decision process.
IV. Model Specification

Just as enrollment in undergraduate education is trending upward, so too is the enrollment in graduate school (“Fast Facts,” 2012). This phenomenon presents an opportunity to proxy for herding within the context of higher education. The following chapter outlines an empirical approach to investigating the effects of earning a higher degree in the face of potential herding behavior and an increasing labor supply. The method for analyzing the investment decision is based upon the human capital model. This implies that the decision to invest is deemed rational if the payoff of attending proves to be greater than alternatives the individual could have realized if she had gathered information on her own.

Although herding behavior has primarily been discussed in the context of undergraduate education, the empirical model defined discusses the payoff of continuing education beyond an undergraduate degree and the implications associated with this decision. Since an individual’s earnings is the basis of analysis in determining if the investment is worthwhile, the model attempts to examine the difference in earnings between those who have obtained bachelor’s degrees relative to higher degrees, hypothesizing that graduate school will not result in as high of a payoff compared to earning an undergraduate degree.

The Oaxaca-Blinder decomposition procedure is introduced in terms of its original use in measuring discrimination, and is then extended to investigate the earnings difference between degree types by decomposing this difference. This will allow for a better understanding of the payoff associated with earning a higher degree in the presence of increasing enrollment. An earnings function is also specified in order to define
regression equations necessary for the decomposition procedure. Lastly, The National Survey of College Graduates is presented and discussed, as it will serve as the dataset used for the empirical testing.

4.1: The Graduate School Decision

The empirical model used to investigate the prevalence of herding in the decision to pursue higher education is based on the Oaxaca-Blinder decomposition method. Before applying this method, however, it is first essential to describe how a rational decision will be defined according to this paper. When evaluating an individual’s choice, the human capital model will be employed to follow the traditional framework used to analyze an individual’s decision to invest in college. This model is appropriate to use based on its prevalence within economic theory and because of its simplicity, allowing the decision to be evaluated based on pecuniary benefits. If an individual makes a choice to invest in higher education, it is assumed that she does so because she is enhancing her ability to perform or produce labor at a higher quality and economic value. Thus, she should earn a higher return if she chooses to invest, relative to not investing.

Although herding behavior has been discussed in terms of undergraduate degrees, it nonetheless pertains to the decision to obtain degrees beyond a bachelor’s degree as well. If herding is indeed taking place, more and more people should be attending school today. Based on The National Center for Education Statistics, the primary federal entity for collecting and analyzing data related to education, enrollment has increased in recent years:

Enrollment in degree-granting institutions increased by 11 percent between 1990 and 2000. Between 2000 and 2010, enrollment increased 37 percent, from 15.3
The theory of herding suggests that as more students are currently attending college, many may feel as if they need to attend graduate school to set themselves apart and distinguish themselves in the labor market.

In this sense, herding may cause graduate school to be viewed today as more of a necessity. With an increase in demand for undergraduate education, more students may feel as if pursuing an even higher degree has become less of a choice and more of a requirement. Those who choose to attend may base their decisions upon what they observe their peers doing, as well as those individuals who made similar decisions in prior years.

In the presence of herding, it is expected that the payoff of attending graduate school will not be as high as initially expected. In the past, there may not have been as much pressure to pursue higher education, so only those students who were intrinsically interested in pursuing a degree for personal preferences did so. This would, therefore, result in a higher payoff for attending graduate school in the past based on a lower supply of highly educated students entering the labor market.

This is not to say that attending graduate school will generate no payoff for the individual today. Rather, it may be the case that the payoff is less than initially expected. For this reason, the decision to invest may be irrational in terms of the human capital investment model. If an individual is expecting certain returns and instead receives a return lower than anticipated, this may signal an irrational decision. This is based on the
argument that too many people are following the choices made by those who have gone before them. This does not suggest that every individual making the decision to attend graduate school is engaging in an irrational decision. Instead, this suggests that some of these people are choosing an option that is less beneficial than an alternative, based on the pecuniary analysis presented by the human capital theory.

In other words, herding has the potential to induce irrational decisions if individuals are realizing returns that are lower than an alternative they could have realized if they had acquired information on their own. These alternatives may include entering the labor force immediately, or perhaps even just choosing a different area of study. Therefore, for some individuals, the outcome of their choice is an irrational decision directly related to herding. If they would have made their decision based on other information, rather than just succumbing to the signal they gleaned from herding, they could have possibly made a better decision that led to a better outcome.

If it is true that a bachelor’s degree is becoming a basic requirement for obtaining a decent job, then it may be reasonable to assume that the returns to a bachelor’s degree will most likely offer a higher return on the investment, relative to a graduate degree. This does not mean that those with higher degrees will not earn higher salaries. Instead, this suggests that the return, relative to the investment, may not be as high for a graduate degree. That is, each degree earned is expected to have decreasing returns in regards to earnings, especially in the face of herding. Obtaining a bachelor’s degree may allow a person to jump from a minimum wage job to a career with a salary. However, a graduate degree is not as likely to offer such a significant return relative to an undergraduate degree, especially in the present as compared to the past.
If individuals make their decision to attend school based on what they witness their peers doing, or based on the decisions made by those in the past, they may encounter a return that is less than originally perceived. Although those with higher degrees do typically earn more, the herding model suggests that they are merely not earning as much as expected, or as much as in previous years, especially relative to a bachelor’s degree.

4.2: The Oaxaca-Blinder Decomposition Procedure

The Oaxaca-Blinder decomposition procedure can be used to analyze the effects of herding by applying this model to an earnings function. This procedure is most commonly used to measure the effects of discrimination (Hoffman & Averett, 2010). Typically, the model is applied to understand the proportion of the gender gap in earnings that is the result of ‘other factors’ (mainly discrimination). The rest of the percentage is what can be attributed to an actual difference in skills (Hoffman & Averett, 2010).

Although a more complex earnings function will be defined latter, for simplicity, assume that a single variable, $X$, determines an individual’s earnings and accounts for his or her skill level. In other words, $X$ accounts for the individual’s total amount of human capital. It follows that:

$$Y_i = \beta X_i + \epsilon_i$$

In this equation, $Y$ accounts for earnings. Further, if $\bar{Y}$ stands for average earnings and $\bar{X}$ stands for average level of skills, it must be that:

$$\bar{Y} = \beta \bar{X}$$
This simple equation is telling because it indicates that only two things may account for differences in earnings. It must be the case that either $\bar{X}$ is different between the two groups or $\beta$ is different. If $\bar{X}$ is different, then the difference in earnings can be accounted for by differences in human capital. However, if $\beta$ is the variable that differs, the difference is likely due to discrimination (Hoffman & Averett, 2010).

To determine how much of the difference is accounted for by skills and how much by discrimination (or other factors), regression analysis can be used, supplemented by the application of the Oaxaca-Blinder decomposition procedure. It is first necessary to determine regression equations for each group, estimated separately (Hoffman & Averett, 2010). Although the analysis above is primarily used to interpret discrimination that females most often face, it can also be applied to understand possible effects of herding. Instead of measuring how discrimination affects earnings, it would be more useful to think of discrimination, in this scenario, as the effects of herding.

In terms of gender, part of the reason that men make more is because some of them do have higher skill levels. The rest is attributed to ‘other factors.’ Similarly, in the scenario outlined by this paper, those with higher degrees will earn more in part because they have a higher skill level. However, if those earning a graduate degree do not receive the same level of reward as seen in earning a bachelor’s degree, this can, likewise, be attributed to ‘other factors.’ Herding indicates that the payoff is not just a result of differing levels of human capital, but also the effect of other factors that may induce irrational decisions.

Separating our sample into two subgroups such that one group contains individuals who have earned a bachelor’s degree with the other group consisting of those
individuals with higher degrees (any degree above a bachelor’s degree), two regression
equations can be determined. They are written as follows:

\[
\bar{Y}_{HD} = \beta_{HD} \bar{X}_{HD} \quad \text{for higher degrees and} \quad \bar{Y}_B = \beta_B \bar{X}_B \quad \text{for bachelor’s degrees}
\]

Then, the difference in average earnings based on degree type can be written as:

\[
\bar{Y}_{HD} - \bar{Y}_B = \beta_{HD} \bar{X}_{HD} - \beta_B \bar{X}_B
\]

The Oaxaca-Blinder decomposition procedure, written below, then allows for the
difference to be separated and explained in terms of ‘difference due to qualifications’ as
well as an ‘unexplained’ difference. The ‘difference due to qualifications’ is the first
term on the right-hand side of the equation and the ‘unexplained’ difference is the second
term on the right-hand side:

\[
\bar{Y}_{HD} - \bar{Y}_B = (\bar{Y}_{HD} - \bar{Y}_B^*) + (\bar{Y}_B^* - \bar{Y}_B)
\]

(Blau, Ferber, & Winkler, 2014)

Then, these components can be broken down, such that:

\[
\bar{Y}_{HD} = \beta_{0HD} + \beta_{1HD} \times \bar{X}_{HD}
\]
\[
\bar{Y}_B = \beta_{0B} + \beta_{1B} \times \bar{X}_B
\]
\[
\bar{Y}_B^* = \beta_{0HD} + \beta_{1HD} \times \bar{X}_B
\]

(Blau et al., 2014)

Substitution of these terms into the original equation then allows for an easier analysis of
what exactly the ‘explained’ and ‘unexplained’ difference accounts for in terms of the
theory outlined in this paper.
Substitution yields:

\[
Y_{HD} - Y_B = (Y_{HD} - Y_B^*) + (Y_B^* - Y_B)
\]

\[
Y_{HD} - Y_B = [(\beta_{0HD} + \beta_{1HD} \times X_{HD}) - (\beta_{0HD} + \beta_{1HD} \times X_B)] + [(\beta_{0HD} + \beta_{1HD} \times X_B) - (\beta_{0B} + \beta_{1B} \times X_B)]
\]

\[
Y_{HD} - Y_B = [\beta_{1HD} \times (X_{HD} - X_B)] + [(\beta_{0HD} - \beta_{0B}) + (\beta_{1HD} - \beta_{1B}) \times X_B]
\]

(Hoffman & Averett, 2010)

The first term on the right-hand side of the equation is the difference in skill level displayed by varying degree types multiplied by \(\beta_{1HD}\), the value of a unit of \(X\) for those with higher degrees. In other words, this is obtained by taking the difference in the average level of human capital for those with higher degrees and bachelor’s degrees and multiplying it by the higher degree regression coefficients. This represents the dollar amount of the earnings difference that can be attributed to actual differences in skill. This is typically referred to as the ‘explained’ portion of the earnings difference (Hoffman & Averett, 2010).

The second term on the right-hand side of the equation accounts for the ‘unexplained’ portion of the difference in earnings between the two groups. This term measures the difference in the market value of skills for each degree type multiplied by the average skill level of those with bachelor’s degrees. This, therefore, measures the dollar amount of the earnings gap that is the result of differences in monetary return or reward for skills (Hoffman & Averett, 2010). It can be thought of as how much those with bachelor’s degrees would earn if they were rewarded in the same way as those who earn a higher degree are (retaining their current skill level). Basically, this portion is
determined by the difference between what those with bachelor’s degrees earnings are when their average level of human capital is evaluated using the higher degree regression versus when they are evaluated using the bachelor’s degree regression (Hoffman & Averett, 2010).

Then, dividing each of these terms by the total earnings difference (the left-hand side of the equation) yields the explained and unexplained portions in percentage terms, allowing researchers to state what percent of the earnings gap is due to unexplained factors (Hoffman & Averett, 2010).

4.3: Explained and Unexplained Portions in Terms of the Hypothesis

In terms of the hypothesis that graduate school does not generate as high of a payoff as undergraduate school (relative to the investment), this indicates that the unexplained portion of the Oaxaca-Blinder decomposition should be rather small since this portion would be linked to the effects of irrational decisions made by individuals in terms of receiving a payoff that is not as large as originally presumed. Looking at what this portion represents mathematically \((\hat{\beta}_{HD} - \hat{\beta}_{1B}) \times X_B\), the main determinant of this figure is the difference in regression coefficients between those who have earned higher degrees and those who have earned bachelor’s degrees.

The hypothesis argues that earning a higher degree will not result in a huge increase in payoff, relative to a bachelor’s degree. That is, the regression coefficients should not be all that different. Since the difference between them should be small, it follows that the unexplained portion should also be small within this framework. Appealing to common sense, this makes sense because it is the opposite of the gender gap.
model, which asserts that men are making much more than women due to ‘other factors.’ On the contrary, this model suggests that those with higher degrees are not making that much more than those with bachelor’s degrees, implying that this portion should thus be small.

This directly relates to the effects of herding and the premise that earning a graduate degree in the face of an increasing labor supply should result in a lower return, relative to a bachelor’s degree. This is not to say that those with higher degrees will earn less than those with bachelor’s degrees. Rather, it is saying that the return on the investment of earning a higher degree is not as great as the return for a bachelor’s degree. For this reason, the unexplained portion is of particular importance to the hypothesis in understanding and explaining the effects of herding within the context of higher education.

The explained portion, on the other hand, should assess differences in earnings that are simply a result of higher skill levels. This is expected to be larger than the unexplained portion if differences in earnings are linked to more than just obtaining a higher degree, but to other factors that contribute to human capital as well, such as experience or training. Clearly, a majority of the decomposition must be attributed to this or else human capital would play no role at all. Based on the mathematical calculation of this portion \((\beta_{1_{HD}} \times (\bar{X}_{HD} - \bar{X}_{B}))\), if the difference between the levels of human capital between each subsample was zero, then the amount of human capital that an individual accrues plays no role in their earnings. This is certainly not expected since the human capital theory indicates that an individual obtains more human capital to earn more in the long run. It may be other forms of human capital, however, (as opposed to schooling)
that play a more significant role in regards to earnings, especially if herding results in a decrease in the value of a higher degree.

Further, as discussed previously, those who earn a higher degree are anticipated to earn more. Some occupations require graduate degrees and are therefore expected to pay a higher salary. The decision to obtain a higher degree only becomes irrational when the individual makes choices based on ‘unreliable’ signals, such as herding, that result in a payoff that is less than what is expected and less than the outcome of an alternative choice. Those with higher degrees, and therefore higher levels of human capital, should, in theory, earn more. However, the payoff may indeed fall short of initial expectations.

Nonetheless, it is not so easy to describe an earnings function based on one variable, \( X \), that accounts for an individual’s entire stock of human capital. Instead, an individual’s earnings model is likely to be based on a variety of factors.

4.4: The Earnings Function

First and foremost, earnings functions are usually written in log-linear form. This is because the independent variables are expected to have a non-constant impact on earnings. The log-linear form dictates that a unit change in an \( X \) variable will result in a \( \beta \times 100 \) percent change in \( Y \). This is useful in describing the expected concave appearance of an earnings graph, depicted in the human capital model. For the empirical model used in this paper, earnings are based on total earned annual income before deductions.

Aside from investment in human capital being the major contributor to an individual’s earnings function, demographic characteristics are no less important
variables to consider. An individual’s age is a reasonable variable to consider when
determining the amount an employee will earn. Typically, the older a person is, the more
money he or she will make. Since this paper would like to test the effects of acquiring
higher education on earnings, and essentially, whether the investment is worthwhile, it is
sensible to look at recent graduates. This is because their earnings are more likely to
reflect and isolate the effects of the degree earned. For this reason, the model will
address individuals in a five-year age range, from 28 to 33. This is meant to control for
other unknown factors that would be affecting an individual’s salary beyond their
educational attainment. By using recent graduates, the hope is that their earnings will
show evidence of the effects of herding before other forms of human capital influence the
individual’s salary more heavily.

As pointed out by the use of the typical Oaxaca-Blinder decomposition procedure,
women are likely to earn less than men. Although some of this can be accounted for
based on differences in skill level and human capital, discrimination as well as ‘other
factors,’ (such a cultural and societal influences, for example) are also likely to attribute
to a common trend of women earning less than men. For this reason, sex must be
incorporated in the model to account for differences in earnings between men and
women. The prediction is that being female will negatively affect earnings, despite the
level of degree earned.

Race, as well, plays a large role in determining the earnings of an individual. As
with sex, this may be due to actual differences in human capital attainment, a bias or
discrimination on the demand-side of labor, or because of larger societal issues. For
example, African Americans are often cited as earning less than their white counterparts.
This may be due to inherent discrimination or perhaps a reflection of some other societal factor. Along similar lines, those of Asian descent are likely to earn more because many Asian cultures place a heavy emphasis on achieving academic success, focusing on studying, and earning high grades. This is likely to spill over in the labor force, with the prediction that Asians are pressured to work harder, contributing to higher earnings.

Thus, race must be included to capture characteristics that are not included elsewhere but nonetheless affect earnings. The earnings function outlined in this paper will include two dummy variables to account for different ethnicities. These dummy variables represent African American and Asians. The third, implicit category that these variables are being compared against is White/Caucasian.

Beyond demographic variables, however, the human capital model draws mainly upon the variables of educational attainment and experience. Clearly, education is accounted for in each regression since two equations are run separately for the two subgroups: those who have earned a bachelor’s degree and those who have earned higher degrees. Experience, however, must be included in the equation to account for factors such as on-the-job training. This no doubt contributes to earnings and accounts for a significant portion of an individual’s human capital. Although the analysis is attempting to look at recent college graduates (who thus have less experience) to capture the initial effects of each degree type on earnings, experience still must be included to account for different qualifications that each individual possesses.

Nonetheless, there is expected to be a non-linear relationship that exists between experience and earnings. Typically, experience has a positive effect on earnings, but at a decreasing rate. Thus, a quadratic must be employed within the earnings function to
reflect the inverted U-shape of the relationship. This implies that earnings will increase (at a decreasing rate) up until a turning point, where it will then decrease with increasing experience. Thus, the experience variable must also be included as a squared variable (in addition to the original experience variable) to account for the effects of experience over time. This is standard approach and again is compatible with the human capital model, which depicts the earnings graph to be increasing at a decreasing rate.

From the above description of the earnings function, it follows that the regression equation should look as follows, where the variable $Y$ denotes earnings (total earned annual income before deductions):

$$\ln(Y) = \beta_0 + \beta_1 \text{Age} + \beta_2 \text{Female} + \beta_3 \text{Asian} + \beta_4 \text{Black} + \beta_5 \text{Experience} + \beta_6 \text{Experience}^2$$

Note, again, that the variable for education is not included since it will be accounted for using the Oaxaca-Blinder decomposition procedure when running the regressions for the two subgroups.

### 4.5: Data

The source used to provide data for the empirical analysis is The National Survey of College Graduates (NSCG). This survey is a longitudinal survey conducted every other year. The NSCG has taken place since the 1970s to provide data on the nation’s college graduates, with particular focus on those in the science and engineering workforce (National Science Foundation, 2010). This is useful in that the data source itself controls partially for differences in degrees and occupations. Those who have been surveyed are mostly those individuals who have earned similar degrees and pursue comparable jobs. Thus, the data partially eliminates differences in earnings that should occur based on differing career fields.
In addition, STEM fields (science, technology, engineering, and math) have been recently been noted as fields that are linked to fast-growing, high-earning occupations, and as vital to what the nation needs in terms of growth and competition with other countries. This should, therefore, reflect herding based on the idea that educators are encouraging students to enter such fields. Clearly, students are receiving signals from recent graduates, peers, and teachers that these are reliable majors to strive for.

Further, the survey requires individuals to be living in the United States during the survey week and have at least a bachelor’s degree. This survey is also a unique source in that it provides demographic information, educational information, and occupational information, which therefore allows specifically for analyzing the relationship between such characteristics.

Specifically, this paper uses the most recent cycle of the NSCG, which was conducted in 2010. The NSCG uses a stratification sampling design to select its sample from the eligible sampling frame. Within the sampling strata, the NSCG used probability proportional to size or systematic random sampling techniques to select the NSCG sample. The U.S. Census Bureau conducted the NSCG for the National Center for Science and Engineering Statistics (NCSES). Data was collecting using both a self-administered mail survey as well as a self-administered web survey (National Science Foundation, 2010).

The demographic variables listed in the earnings function outlined in this paper are all available through the NSCG. (These include age, gender, and race.) Those individuals who were not in the outlined age range (28-33) were dropped from the data set. The experience variable then had to be created using a proxy. Subtracting the year
the individual earned his or her degree from 2010 (the year the survey was administered) was the method used to generate this variable. This is a widely accepted method for proxying for experience by assuming that these years account for on-the-job experience and training.

The variable earnings can be described, based on the survey, as total earned income before deductions in the previous year. Further, it must be noted that the data was paired down to only those respondents who worked full-time, which is defined in this paper as those individuals who work 48 hours per week or more.

In order to stratify the sample based on degree type, the individuals were divided based on the highest degree they had earned. Respondents were able to choose one of the following four choices to describe their degree: bachelor’s, master’s, doctorate, or professional. Those who chose the option ‘bachelor’s degree’ as their highest degree earned were put into a group containing only individuals who had obtained bachelor’s degrees. Those who had chosen one of the three later degrees were placed in a second group, accounting for ‘higher degrees.’

Of the 2,599 individuals included in the final data sample fitting the criteria specified, 1,260 observations accounted for those who had earned a bachelor’s degree and 1,339 observations represented those who had earned higher degrees. Thus, the data was split fairly evenly between the two subsamples. Table 4.4.1 contains a summary of variable names, definitions, and expected signs.
<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Description</th>
<th>Expected Sign</th>
</tr>
</thead>
<tbody>
<tr>
<td>logEarnings</td>
<td>The log of total earned income before deductions in the previous year</td>
<td>N/A</td>
</tr>
<tr>
<td>HigherDegree</td>
<td>Dummy variable used to account for highest degree type that an individual holds; Consists of two categories: Bachelor’s Degree and Higher Degree (which is made up of individual’s who reported earning a master’s, doctorate, or professional degree)</td>
<td>Positive</td>
</tr>
<tr>
<td>Age</td>
<td>Age of the individual</td>
<td>Positive</td>
</tr>
<tr>
<td>Female</td>
<td>Dummy variable used to account for the number of individuals who self-identified as female, relative to male</td>
<td>Negative</td>
</tr>
<tr>
<td>Asian</td>
<td>Dummy variable used to account for the number of individuals in the sample who identified themselves as Asian only</td>
<td>Positive</td>
</tr>
<tr>
<td>Black</td>
<td>Dummy variable used to account for the number of individuals in the sample who identified themselves as black only</td>
<td>Negative</td>
</tr>
<tr>
<td>Experience</td>
<td>Proxy Variable created to account for the years of experience an individual has; Generated by subtracting the year of award of the highest degree earned from 2010 (the year of the survey)</td>
<td>Positive</td>
</tr>
<tr>
<td>ExperienceSquared</td>
<td>The number of years of experience squared; Created to account for the non-linear relationship between experience and earnings, which is expected to be increasing at a decreasing rate</td>
<td>Negative</td>
</tr>
</tbody>
</table>
4.6: Conclusion

Herding is not confined to only the decision to acquire an undergraduate degree. This chapter has extended the herding framework to the decision to attend graduate school, revealing how many may feel as if it is less of a choice and more of a necessity. Nonetheless, herding still has implications for the payoff on a higher degree, especially if the return on earning a higher degree is expected to be smaller than the return realized by earning a bachelor’s degree.

The Oaxaca-Blinder decomposition procedure can be used to better understand the difference in earnings between the two degree types based on the hypothesis of a large explained portion and a small unexplained portion. This suggests that the two subsamples will not exhibit significantly different regression coefficients, indicating a small payoff in regards to attaining a higher degree. The data obtained from the National Survey of College Graduates will be used in Chapter V to run regressions according to the earnings function outlined, which are then examined using the decomposition procedure.
V. Results and Analysis

The following chapter is an extension of the model specification, applying the National Survey of College Graduates dataset to the Oaxaca-Blinder decomposition procedure. This process is employed to gain an insight into the effects of herding by decomposing the earnings difference between those who have earned a bachelor’s degree and those who have earned a higher degree.

Before the Oaxaca-Blinder decomposition procedure is applied, a basic regression is run using the variable HigherDegree as a dummy variable. This allows for an analysis of the return on earnings by obtaining a higher degree before the Oaxaca-Blinder decomposition is applied.

Next, a Chow test is employed to determine if the regression coefficients for the two subsamples are equivalent. This preliminary test dictates that it is appropriate to exercise the Oaxaca-Blinder decomposition procedure. After the decomposition is performed, the results are discussed and it is discovered that the model may be more useful when controlling for the variable Experience, as it differs between the two groups.

After adjusting the age range for each group to control for experience, the Oaxaca-Blinder decomposition procedure is employed again for the varying degree types. The results, however, are not consistent with the hypothesis of a large explained portion and a small unexplained portion. This suggests that there still exists a significant return on earning a higher degree.
5.1: The Basic Regression

In order to understand the earnings function outlined, it is first advantageous to run a regression for the equation including HigherDegree as a dummy variable within the model. This will allow for an analysis of the return on earnings by obtaining a higher degree using a basic approach. Running a log-linear regression yields coefficients presented in Table 5.1.1, with coefficients followed by t-scores.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Regression Output</th>
<th>t statistics in parentheses</th>
</tr>
</thead>
<tbody>
<tr>
<td>HigherDegree</td>
<td>0.318***</td>
<td>(8.69)</td>
</tr>
<tr>
<td>Age</td>
<td>0.0271**</td>
<td>(2.97)</td>
</tr>
<tr>
<td>Female</td>
<td>-0.182***</td>
<td>(-6.68)</td>
</tr>
<tr>
<td>Asian</td>
<td>0.138***</td>
<td>(4.01)</td>
</tr>
<tr>
<td>Black</td>
<td>-0.000905</td>
<td>(-0.02)</td>
</tr>
<tr>
<td>Experience</td>
<td>0.189***</td>
<td>(12.30)</td>
</tr>
<tr>
<td>ExperienceSquared</td>
<td>-0.00939***</td>
<td>(-6.85)</td>
</tr>
<tr>
<td>Constant</td>
<td>9.357***</td>
<td>(35.00)</td>
</tr>
</tbody>
</table>

* p < 0.05, ** p < 0.01, *** p < 0.001

The regression coefficients each exhibit the anticipated sign. A single asterisk denotes that the variable is significant at the 5% level of significance; two asterisks indicate that the variable is significant at the 1% level; and lastly, three asterisks mean that the variable is significant at the 0.1% level. Nonetheless, the results and significance
of each variable cannot be taken at face value. Before the significance of each can be accurately determined, it is necessary to also explore the correlation between the variables.

Using the `pwcorr` command in Stata, the highest correlation (of 0.9546) exists between `Experience` and `ExperienceSquared`. This result can be found in Appendix A. Nonetheless, this is to be expected since the variable `ExperienceSquared` is generated from the `Experience` term. No other variables exhibit unsuspected correlations that reveal potentially problematic issues.

Further, it is also important to determine if there is evidence of multicollinearity. Multicollinearity occurs when two or more predictor variables are highly correlated such that one can be linearly predicted from the others with little inaccuracy. This may result in violations of the Classical Assumptions where the variances and standard errors are inflated, thus yielding estimated coefficients that are fragile and unstable to small changes in the sample or model specification. In addition, the $t$-scores will fall, making it more likely for the researcher to commit a Type II error (the failure to reject a false null hypothesis).

Although there is no specific method that can detect the presence of multicollinearity, a common approach to assessing whether multicollinearity exists within a model is to look at the variance inflation factor, or VIF. A researcher may suspect the presence of multicollinearity if the $R^2$ of the model is relatively high in the presence of many insignificant variables. Since $VIF = \frac{1}{1-R^2}$, a general rule of thumb is that a VIF of five or greater may imply multicollinearity because this equates to an $R^2$ value of 0.8.
However, the researcher must evaluate the VIF relative to the significance of the variables and other indicators of a well or poorly specified model.

Using the vif command in Stata for this basic regression, the generated mean VIF is 4.40, as shown in Appendix A. Since this number is less than five and only one variable in the model appears to be insignificant, this leads to the likely conclusion that multicollinearity is not present within the model.

The next step is to test for heteroskedasticity within the model. This violates that Classical Assumption that the population error term has a constant variance. Although the presence of heteroskedasticity does not bias the regression coefficients, it does typically lead to inefficient estimators. This is because the ordinary least squares regression is no longer minimum variance, resulting in an underestimate of the standard errors. This, in turn, results in an overestimation of $t$-scores and also invalidates the $F$-test. Thus, in the presence of heteroskedasticity, a Type I error (the incorrect rejection of a true null hypothesis) is much more likely.

The White test is used to test the null hypothesis that homoscedasticity exists. Using Stata, the calculated chi-square is 40.24, compared to the critical value of 42.557, for 29 degrees of freedoms at the 5% level of significance. This can again be seen in Appendix A. Since the calculated value is less than the critical value (40.24 < 42.557), the null hypothesis fails to be rejected, implying homoscedasticity. Thus, the model does not need to be corrected for heteroskedasticity.

After testing for multicollinearity and heteroskedasticity, it can be determined that it is appropriate to use the original results to run an $F$-test, as well as consider the significance of each independent variable. The $F$-test can be employed to test the null
hypothesis that all the slope coefficients in the equation equal zero simultaneously

\( H_0: \beta_1 = \beta_2 = \cdots = \beta_K = 0 \). Obviously, it follows that the alternative hypothesis is that \( H_0 \) is not true.

As shown in Appendix A, Stata declares a calculated \( F \)-statistic of 59.11, which must then be compared to the critical \( F \)-statistic. It follows that the critical value is 2.01 at the 5% level of significance, based on 7 numerator degrees of freedom and 2,583 denominator degrees of freedom. Since the calculated value is greater than the critical value (59.11 > 2.01), the null hypothesis can be rejected. This conveys that the model as a whole has explanatory power and that the \( R^2 \) value is significant. Thus, having an \( R^2 \) of 0.1381 means that 13.81% of variation in the dependent variable is explained by the model.

Since the model as a whole has been shown to have explanatory power, it is necessary to further investigate the significance of each independent variable separately since multicollinearity and heteroskedasticity have been ruled out. In order to do so, the \( p \)-value can be used as a basis for testing. The \( p \)-value for a \( t \)-score is the probability of observing a \( t \)-score the size of the one actually observed or larger (in absolute value) if the null hypothesis were true.

Since the \( p \)-value is being used to test the significance of each variable, the null hypothesis (using a two-tailed test) is that the coefficient is equal to zero. If there is a significant linear relationship between the independent variable and dependent variable, however, the coefficient will not equal zero. Thus, the alternative hypothesis is that the coefficient is different from zero.
Typically, results are tested at the 1%, 5%, or 10% level of significance. Since a significance level of 5% is usually the norm and is widely accepted in academic research, this paper chooses to focus on this level. Although this is not as rigorous as the 1% level (or even the 0.1% level), it still provides confidence in the results, especially relative to the 10% level of significance. The 5% level of significance will be discussed primarily in this paper, but the variables can also be examined at the 1% and 0.1% levels provided by the tables illustrating regression outputs.

Using the 5% level of significance, the \( p \)-value less than 0.05 rule can be used to determine the significance of the individual variables. That is, if the \( p \)-value is less than 0.05, the null hypothesis can be rejected at the 5% level of significance, suggesting that the variable is significant. As shown in Appendix A, The \( p \)-value for each independent variable in the regression is less than 0.05 except for the variable Black. Thus, at the 5% level of significance, it can be assumed that every variable except Black is significant within the model. Although Black is insignificant based on its \( p \)-value, it should still be included based on its theoretical importance.

Now that the model has been evaluated in terms of the regression equation as a whole as well as individual variables, it can be determined that earning a higher degree results in a 31.8% increase in \( \log \text{Earnings} \), ceteris paribus. This indicates that earning a higher degree does result in higher earnings and a positive return on the investment.

5.2: The Chow Test

To apply the Oaxaca-Blinder decomposition method, it is necessary to remove the independent dummy variable \( \text{HigherDegree} \) from the equation and stratify the sample
into two groups: those who have earned only a bachelor’s degree and those who have earned higher degrees. Before this is done, however, it would be helpful to run a Chow test to determine if the two sets of data contain significantly different regression coefficients for the same theoretical equation. If the regression coefficients are indeed significantly different, it follows that a structural difference exists between the two groups. The null hypothesis of this test is that the regression/slope coefficients are the same in the two subsamples. If this were the case, the coefficients would act no differently for the two groups, implying that graduate school does not pay off at all. The alternative hypothesis is that the coefficients do differ between the two groups.

The Chow test is an application of the $F$-test and requires three regressions to be run. The first regression is the earnings function run for the entire sample without the HigherDegree dummy variable. The next two regressions are identical in form, except that the first includes only those individuals who have earned a bachelor’s degree while the second consists of only those who have earned a higher degree. These regressions, just like the original regression, should undergo basic analysis. This includes an examination for multicollinearity, testing for heteroskedasticity, and testing the significance of the equation as well as the significance of the independent variables. All of the results for the regressions and tests discussed in this section can be found in Appendix B.

The correlations between variables for the first regression (before the sample is stratified) are the same as in the model using HigherDegree as a dummy variable. Again, none of the correlations present a concern. Only Experience and ExperienceSquared exhibit a high correlation, which is to be expected as discussed previously. Then, using
Stata to determine the mean VIF, the result is 4.64. This again suggests that multicollinearity is not a problem, especially in the face of high t-scores.

Testing for heteroskedasticity, the White test is again employed. The calculated chi-square value using Stata is 35.66. For 22 degrees of freedom, the critical value is 33.924 at the 5% level of significance. In this scenario, the calculated value is greater than the critical value (35.66 > 33.924), so the null hypothesis of homoscedasticity is rejected, suggesting the model is heteroskedastic. To correct for this, the robust command in Stata is used.

Using an F-test to test the validity of the model and the significance of the $R^2$ value, the calculated F-value of 49.45 must be compared to the critical value. With 6 numerator degrees of freedom and 2,584 denominator degrees of freedom, it follows that the critical value is 2.10 at the 5% level of significance. Since the calculated value is greater than the critical value (49.45 > 2.10), the null hypothesis that all slope coefficients are equal to zero can be rejected. This, again, conveys that the $R^2$ value is significant. That is, the model explains 11.29% of variation in the dependent variable. In regards to the independent variables, each displays the expected sign. Looking at the $p$-values, only the $p$-value for Black is greater than 0.05. Hence, Black is once again the only variable that is not significant at the 5% level of significance.

The next regression uses the same form, but is only run for those who have earned a bachelor’s degree. Running a pairwise correlation in Stata reveals that only Experience and ExperienceSquared display a very high correlation (0.9789). However, as before, this is to be expected. Next, the mean VIF generated by Stata is 9.61, suggesting that multicollinearity may be present within the model. However, based on theory,
appropriate correlation coefficients, and generally high $t$-scores, this high mean VIF is most likely the result of the high correlation between $Experience$ and $ExperienceSquared$.

Using the White test to test for heteroskedasticity, the resulting calculated chi-square value is 19.79. The critical value at the 5% level of significance using 22 degrees of freedom is once again 33.924. Since $19.79 < 33.924$, the null hypothesis of homoscedasticity fails to be rejected, implying that heteroskedasticity is not present within the model.

Following, the calculated $F$-value is 19.47. Still using 6 numerator degrees of freedom, and 1,250 denominator degrees of freedom, the critical $F$-value is 2.10 at the 5% level of significance. Since $19.47 > 2.10$, the null hypothesis that all slope coefficients are equal to zero can be rejected. Thus, the significant $R^2$ value conveys that the model explains 8.55% of variation in the dependent variable.

Looking at the $p$-values for each independent variable, however, three variables now exhibit a $p$-value that is greater than 0.05, suggesting they are insignificant at the 5% level of significance. These variables are $Age$, $Black$, and $ExperienceSquared$. Nonetheless, the remaining variables are significant at the 5% level. In terms of the signs for each regression coefficient, only $Age$ has a sign that is unexpected.

Finally, the same procedure is followed for the third regression, which is run for only those individuals with higher degrees. Again, only $Experience$ and $ExperienceSquared$ show a high correlation of 0.9464. Further, the mean VIF of 3.93 suggests that multicollinearity is not present within the model.

The White test, employed using Stata, yields a calculated chi-square value of 43.25. Again, with 22 degrees of freedom, the critical chi-square value at the 5% level of
significance is 33.924. Since 43.25 > 33.925, the null hypothesis of homoscedasticity is rejected, indicating that the regression needs to be corrected for heteroskedasticity using the robust command in Stata.

This regression can then be analyzed using the $F$-test. The calculated $F$-value equates to 46.67. With 6 numerator degrees of freedom and 1,327 denominator degrees of freedom, the critical $F$-value is 2.10 at the 5% level of significance. Since the calculated value exceeds the critical value (51.54 > 2.10), the null hypothesis that all slope coefficients are equal to zero can again be rejected. Thus, the significant $R^2$ value conveys that the model explains 18.90% of variation in the dependent variable.

Looking at the $p$-values for each independent variable, once again, only Black shows a $p$-value that is greater than 0.05, suggesting that it is insignificant at the 5% level. Thus, the rest of the independent variables are significant within the model run for higher degrees. Black is also the only variable that possessed a sign that is opposite of what was expected. Important regression statistics are listed in Table 5.2.1.
Now that these three regressions have been run, the Chow test can be executed.

The calculated $F$-statistic for the Chow test can be determined using the following formula:

$$F = \frac{(RSS_T - RSS_1 - RSS_2)/(K + 1)}{(RSS_1 + RSS_2)/(N_1 + N_2 - 2K - 2)}$$

Where:

$K = $ The Number of Independent Variables

$N_1 = $ The Number of Observations in Sample 1

$N_2 = $ The Number of Observations in Sample 2
Thus, substitution yields:

$$F = \frac{(RSS_T - RSS_1 - RSS_2)/(K + 1)}{(RSS_1 + RSS_2)/(N_1 + N_2 - 2K - 2)}$$

$$F = \frac{(1173.24175 - 535.150324 - 597.824148)/(6 + 1)}{(535.150324 + 597.824148)/(1257 + 1334 - 2(6) - 2)}$$

$$F = \frac{(40.267278)/(7)}{(1132.97447)/(2577)}$$

$$F = \frac{5.75246829}{0.43964861} = 13.0842408$$

Comparatively, the critical $F$-value can be found for $(K + 1)$ numerator degrees of freedom (equal to 7) and $(N_1 + N_2 - 2K - 2)$ denominator degrees of freedom (equal to 2,577). Hence, the critical $F$-value is 2.01. Since $13.08 > 2.01$, the null hypothesis that the slope coefficients in the two samples are the same can be rejected. This clarifies that the regression coefficients do indeed act differently for the two groups. Thus, this indicates that earning a higher degree does equate to higher earnings and that the model should be stratified.

It follows that the coefficient values for the two subgroups are only relevant to the particular group they define. Therefore, a comparison of the coefficients can be done, allowing for an examination of how each variable affects earnings differently for the two samples. Table 5.2.2 compares the regression coefficients for each degree type.
Table 5.2.2: A Comparison of Regression Coefficients Based on the Chow Test

<table>
<thead>
<tr>
<th>Variable</th>
<th>Bachelor’s Degree</th>
<th>Higher Degree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>-0.00627</td>
<td>0.0461***</td>
</tr>
<tr>
<td>Female</td>
<td>-0.191*</td>
<td>-0.176***</td>
</tr>
<tr>
<td>Asian</td>
<td>0.159**</td>
<td>0.130**</td>
</tr>
<tr>
<td>Black</td>
<td>-0.0230</td>
<td>0.0416</td>
</tr>
<tr>
<td>Experience</td>
<td>0.141***</td>
<td>0.234***</td>
</tr>
<tr>
<td>ExperienceSquared</td>
<td>-0.00439</td>
<td>-0.0163***</td>
</tr>
</tbody>
</table>

*p < 0.05, ** p < 0.01, *** p < 0.001

From this comparison, it can be seen that the variable Age not only has different coefficients for the two groups, but different signs as well. The variable is not significant in the bachelor’s degree subsample, and displays a sign opposite of expected, suggesting that a one-year increase in age results in 0.627% decrease in logEarnings, ceteris paribus. For those with higher degrees, by contrast, an increase in age by one year should result in a 4.61% increase in logEarnings, ceteris paribus. This likely captures characteristics of the individual that are not captured elsewhere in the model. For example, the older an individual is, the more likely it is that he or she will have greater experience gained from other areas outside of the workforce (which is in turn accounted for by the Experience variable). For instance, an individual may be able to gain experience doing volunteer work, which a younger individual may not have.

The variable Female indicates a negative effect on earnings for both groups. However, the regression results show that a higher penalty exists for those females who hold bachelor’s degrees. That is, the coefficient implies that being female at the bachelor’s degree level equates to a 19.1% decrease in logEarnings, ceteris paribus. This can be compared to a 17.6% decrease in logEarnings for females with a higher degree, ceteris paribus.
As expected, the regression coefficients for *Asian* are positive for both subgroups, suggesting that individuals who are Asian will earn more relative to other races. The results indicate that *Asian* has a greater effect on earnings for the bachelor’s degree subsample relative to the higher degree subsample. That is, an individual who is Asian can expect a 15.9% increase in *logEarnings*, ceteris paribus, if they hold a bachelor’s degree, but only a 13.0% increase in *logEarnings*, ceteris paribus, if they hold a higher degree.

Contrary to *Asian* the variable *Black* is expected to be negative. This is only the case for the subsample containing individuals with a bachelor’s degree. Further, the results are not significant for either subgroup. Disregarding significance, however, the results would signal that individuals who identified themselves as black would experience a decrease in *logEarnings* by 2.3%, ceteris paribus, if they held a bachelor’s degree. However, if they held a higher degree, they would actually experience an increase in *logEarnings* by 4.16%, ceteris paribus.

Lastly, the variables representing experience must be analyzed. As discussed in the model specification, a non-linear relationship exists between experience and earnings. Thus, a quadratic was employed to represent the inverted U-function. This illustrates that as experience increases, earnings increase at a decreasing rate, eventually reaching a turning point where earnings actually decrease with greater experience.

For this reason, *Experience* and *ExperienceSquared* must be interpreted together. Although the previous variables have been interpreted independently, it is not permissible to follow that process in this scenario. This is because it is not possible for *Experience* to change without also changing *ExperienceSquared* (and vice versa). Therefore, to assess
the change in logEarnings associated with the change in the individual’s level of experience, calculus must be used. Taking the partial derivative of earnings with respect to experience (such that $Y$ denotes $\logEarnings$ and $Exp$ denotes $Experience$ in this case), it follows that:

$$Y = \beta_0 + \cdots + \beta_{Exp}Exp + \beta_{Exp^2}Exp^2$$

$$\frac{\partial Y}{\partial Exp} = \beta_{Exp} + 2\beta_{Exp^2}Exp$$

This equation, as a whole, therefore represents the expected change in $\logEarnings$ as $Experience$ changes, all else being equal. Thus, the marginal effects are not constant and vary with each point on the curve.

Now, looking at the two subsamples, for those with bachelor’s degrees, the regression coefficient for the variable $Experience$ is positive, accompanied by a negative coefficient for $ExperienceSquared$. This is as expected and suggests that the function is increasing at a decreasing rate. More specifically, the regression coefficient for $Experience$ is 0.141 while the coefficient for $ExperienceSquared$ is -0.00459 (although insignificant). Thus, using the model above, the derivative of earnings with respect to experience for bachelor’s degrees is:

$$\frac{\partial Y}{\partial Exp} = 0.141 - 0.00918(Exp)$$

Hence, it follows that the change in $\logEarnings$ for those with bachelor’s degrees is:

$$\Delta Y = (0.141 - 0.00918(Exp))(\Delta Exp)$$

This equation can then be used to estimate the marginal effects of a change in experience at each point on the curve.
The same method can be used to analyze the subgroup of individuals holding higher degrees. Again, the regression coefficient for the variable *Experience* is positive while the coefficient for the variable *ExperienceSquared* is negative, as anticipated. Specifically, the regression coefficient for *Experience* is 0.234 and the coefficient for *ExperienceSquared* is -0.0163. Therefore, for higher degrees, the derivative of earnings with respect to experience is:

$$\frac{\partial \hat{Y}}{\partial Exp} = 0.234 - 0.0326(Exp)$$

It follows that the change in *LogEarnings* for higher degrees is thus:

$$\Delta Y = (0.234 - 0.0326(Exp))(\Delta Exp)$$

Again, this equation can be used at each point along the curve to estimate the marginal effects on earnings as experience changes.

Overall, the above comparisons illustrate how the models differ in terms of their regression coefficients. Seeing that the two sets of data cannot be treated in the same way due to different regression coefficients, it follows that the two subgroups are structurally different. Therefore, it is reasonable to use the Oaxaca-Blinder decomposition procedure to further investigate the difference in earnings between the samples and determine how much of it is due to explained and unexplained factors.

### 5.3: An Application of the Oaxaca-Blinder Decomposition Procedure

As a reminder, to examine the difference in earnings between those with bachelor’s degrees and those with higher degrees, the Oaxaca-Blinder decomposition procedure can be applied using the following formula:
\[
\bar{Y}_{HD} - \bar{Y}_B = (\bar{Y}_{HD} - \bar{Y}_B^*) + (\bar{Y}_B^* - \bar{Y}_B)
\]

(Blau et al., 2014)

Again, each component can be represented as follows:

\[
\bar{Y}_{HD} = \beta_{0_{HD}} + \beta_{1_{HD}} \times \bar{X}_{HD}
\]

\[
\bar{Y}_B = \beta_{0_{B}} + \beta_{1_{B}} \times \bar{X}_B
\]

\[
\bar{Y}_B^* = \beta_{0_{HD}} + \beta_{1_{HD}} \times \bar{X}_B
\]

(Blau et al., 2014)

Using the separate regressions employed for each subsample in the Chow test, the calculations for each component of decomposition procedure are shown in Table 5.3.1, 5.3.2, and 5.3.3. As a note, all statistics used and discussed in this section can be found in Appendix C

Table 5.3.1: \(\bar{Y}_{HD}\)

| \(\bar{Y}_{HD}\) |
|---|---|---|---|---|
| **X Variable Name** | \(\beta_{0_{HD}}\) | \(\bar{X}_{HD}\) | \(\beta_{0_{HD}} \times \bar{X}_{HD}\) | Antilog(\(\bar{Y}_{HD}\)) |
| Age | 0.0461481 | 30.57132 | 1.410808332 | --- |
| Female | -0.1761489 | 0.4660194 | -0.082088605 | --- |
| Asian | 0.1302865 | 0.2337565 | 0.030455316 | --- |
| Black | 0.0416191 | 0.091128 | 0.009392033 | --- |
| Experience | 0.23386 | 3.536221 | 0.826980643 | --- |
| ExperienceSquared | -0.042666 | 18.2815 | -0.297372155 | --- |
| Constant (\(\beta_{0_{HD}}\)) | --- | --- | 9.064363 | --- |
| Total (\(\bar{Y}_{HD}\)) | --- | --- | 10.9469384 | $56,779.9395$ |
Table 5.3.2: $Y_B$

<table>
<thead>
<tr>
<th>$X$ Variable Name</th>
<th>$\beta_B$</th>
<th>$X_B$</th>
<th>$\hat{\beta}_B \times X_B$</th>
<th>Antilog($Y_B$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>-0.0662673</td>
<td>30.38571</td>
<td>-0.19043636</td>
<td>---</td>
</tr>
<tr>
<td>Female</td>
<td>-0.1913773</td>
<td>0.315873</td>
<td>-0.060450922</td>
<td>---</td>
</tr>
<tr>
<td>Asian</td>
<td>0.1593942</td>
<td>0.131746</td>
<td>0.020999548</td>
<td>---</td>
</tr>
<tr>
<td>Black</td>
<td>-0.0229668</td>
<td>0.111111</td>
<td>-0.002551866</td>
<td>---</td>
</tr>
<tr>
<td>Experience</td>
<td>0.1413166</td>
<td>7.18254</td>
<td>1.015012132</td>
<td>---</td>
</tr>
<tr>
<td>ExperienceSquared</td>
<td>-0.0045916</td>
<td>56.6619</td>
<td>-0.26016878</td>
<td>---</td>
</tr>
<tr>
<td>Constant ($\beta_{0B}$)</td>
<td>---</td>
<td>---</td>
<td>10.44329</td>
<td>---</td>
</tr>
<tr>
<td>Total ($Y_B$)</td>
<td>---</td>
<td>---</td>
<td>10.96569375</td>
<td>$57,854.91857$</td>
</tr>
</tbody>
</table>

Table 5.3.3: $Y^*_B$

<table>
<thead>
<tr>
<th>$X$ Variable Name</th>
<th>$\beta_{HD}$</th>
<th>$X_B$</th>
<th>$\hat{\beta}_{HD} \times X_B$</th>
<th>Antilog($Y^*_B$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.0461481</td>
<td>30.38571</td>
<td>1.402242784</td>
<td>---</td>
</tr>
<tr>
<td>Female</td>
<td>-0.1761489</td>
<td>0.315873</td>
<td>-0.555640681</td>
<td>---</td>
</tr>
<tr>
<td>Asian</td>
<td>0.1302865</td>
<td>0.131746</td>
<td>0.017164725</td>
<td>---</td>
</tr>
<tr>
<td>Black</td>
<td>0.0416191</td>
<td>0.111111</td>
<td>0.004624344</td>
<td>---</td>
</tr>
<tr>
<td>Experience</td>
<td>0.23386</td>
<td>7.18254</td>
<td>1.679708804</td>
<td>---</td>
</tr>
<tr>
<td>ExperienceSquared</td>
<td>-0.0162666</td>
<td>56.6619</td>
<td>-0.921696463</td>
<td>---</td>
</tr>
<tr>
<td>Constant ($\beta_{0HD}$)</td>
<td>---</td>
<td>---</td>
<td>9.054363</td>
<td>---</td>
</tr>
<tr>
<td>Total ($Y^*_B$)</td>
<td>---</td>
<td>---</td>
<td>11.18076651</td>
<td>$71,737.32861$</td>
</tr>
</tbody>
</table>
Using these calculations, the Oaxaca-Blinder Decomposition procedure yields:

\[ \text{Difference} = \text{Explained} + \text{Unexplained} \]

\[ \bar{Y}_{HD} - \bar{Y}_B = (\bar{Y}_{HD} - \bar{Y}_B^*) + (\bar{Y}_B^* - \bar{Y}_B) \]

\[-1,074.979028 = -14,957.389 + 13,882.41 \]

The results show that the explained portion is -$14,957.39 while the unexplained portion is $13,882.41. These two values can now be interpreted to better understand the difference in earnings between those with bachelor’s degrees and those with higher degrees.

As a reminder, the explained portion is the difference in average skill level for the two different degree types, evaluated using the higher degree regression coefficients. More simply, this is the dollar amount of the earnings difference that can be attributed to actual differences in ability. It was suspected that this number would be positive, according to the premise that those with higher degrees should earn more. The negative number produced, however, implies that those with bachelor’s degrees are actually earning more. At first glance, this is surprising. Although herding suggested that earning a higher degree might not provide as significant of a return relative to the investment when compared to a bachelor’s degree, it is still expected to yield higher earnings on average. However, it is important to address the age group that the model is analyzing.

Clearly, since earning a higher degree takes time, those who further their education are delaying their entry into the labor force. This, again, is represented by the human capital model, which suggests that individuals take on costs in the short run
(explicit and implicit) to achieve a higher payoff in the long run. Therefore, those who make the decision to pursue a higher degree (and thus enter the labor market later than individuals who earn a bachelor’s degree) are sacrificing not only earnings but also experience. As a result, the variable *Experience* is what is most likely causing those with bachelor’s degrees to earn more. Since the model addresses individuals early on in their career, this age group depicts those with higher degrees as doing worse because they have just recently left school and lack the experience that those with bachelor’s degrees have.

Comparing the averages for the variable *Experience*, those with bachelor’s degrees have a mean of 7.18 years of experience while those who hold higher degrees have an average of 3.54 years of experience. Therefore, the negative difference in earnings can most likely be attributed to the young age group used.

It is possible that the payoff on higher degrees will be realized later in life. However, the results suggest that there may be no immediate payoff for graduate school. With the rising cost of tuition, this signifies the importance of weighing the costs against the benefits, especially in the face of large loans. If the cost of attending is much more than the resulting increase in earnings, the investment may not be worth it. In fields with less job opportunity, this is especially important to consider. Even if a higher salary is guaranteed, the extra cost of education and accumulated debt might be a heavy burden when the return on the investment is slow. Assessing how quickly an increase in earnings will offset tuition loans and costs undertaken by delaying entry into the workforce is an important measure to take.

Looking at the unexplained portion ($13,882.41), this value measures the difference in the market value of skills for each degree type multiplied by the average
skill level of those with bachelor degrees. It is how much more those with bachelor’s degrees would earn if they were rewarded in the same way as those who earn a higher degree. In essence, it reflects the difference in monetary return or reward for skills. Since the main determinant of this value is the difference in regression coefficients between those who have earned a higher degree and those who have earned a bachelor’s degree, the term should be small if earning a higher degree does not result in a huge payoff.

This value must also be analyzed carefully considering the age group at hand. Again, because the term uses the characteristics of those who have earned a bachelor’s degree, this number is most likely large and positive because these individuals have more experience. Thus, the term may be influenced heavily by the Experience variable once again, resulting in the positive value. Thus, it may be that Experience is overriding most of the human capital characteristics in regards to education.

Although the positive unexplained value is not inconsistent with the theory because it is predicted that a higher degree should yield higher earnings, it is difficult to assess whether the payoff of pursuing more education is large or small due to the opposite signs of the explained and unexplained portions. The variable Experience makes it difficult to isolate the effects of herding and analyze the return on the investment to determine if earning a higher degree pays off. For this reason, it may be helpful to extend the model further by attempting to control at least partially for experience.
5.4: Controlling for Experience

Although it would still be beneficial to use recent graduates to isolate the effects of the degree, the idea of a ‘recent graduate’ means different things for the varying degree types. Assuming that most individuals choose to go to graduate school right after earning their undergraduate degree, it is possible to adjust the age ranges for each group to better reflect each category. Using ages 28 to 32 for those with bachelor’s degrees assumes that these individuals are early on in their careers, but nonetheless still established. For those with higher degrees, however, we can use the age range directly following this, 33 to 37 years. This assumes that these individuals have been given time to complete their degree and are more on par with the experience level of those who hold bachelor’s degrees given that they entered the labor force at later points in their lifetimes.

Once again, two regressions must be run separately before applying the Oaxaca-Blinder decomposition procedure. All results discussed in this section can be found in Appendix D. Table 5.4.1 outlines regression coefficients and t-statistics.
Table 5.4.1: Regressions Controlling for Experience

<table>
<thead>
<tr>
<th>Variable</th>
<th>Bachelor’s Degree</th>
<th>Higher Degree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.00515</td>
<td>0.0358*</td>
</tr>
<tr>
<td>Female</td>
<td>-0.194***</td>
<td>-0.186***</td>
</tr>
<tr>
<td>Asian</td>
<td>0.130*</td>
<td>0.132**</td>
</tr>
<tr>
<td>Black</td>
<td>-0.0302</td>
<td>-0.0439</td>
</tr>
<tr>
<td>Experience</td>
<td>0.184***</td>
<td>0.169***</td>
</tr>
<tr>
<td>ExperienceSquared</td>
<td>-0.00744</td>
<td>-0.00829***</td>
</tr>
<tr>
<td>Constant</td>
<td>9.969***</td>
<td>9.423***</td>
</tr>
</tbody>
</table>

*t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The first regression is run using the established earnings function including only those defined as having an age of 28 to 32 years who hold a bachelor’s degree. Using Stata to look at the correlations between variables, once again, it is no surprise that *Experience* and *ExperienceSquared* are highly correlated (0.9776). No other variables are correlated in an unexpected way. The mean VIF is 9.10, suggesting there may be multicollinearity, but this is most likely attributed to the high correlation mentioned since the *t*-scores are high and theory dictates that these variables are appropriate.

Testing for heteroskedasticity, the White test is once again employed. The calculated chi-square value is 33.52. Using 22 degrees of freedom, the critical chi-square value at the 5% level is 33.924. Since the calculated value is less than the critical value (33.52 < 33.924), the null hypothesis of homoscedasticity fails to be rejected, suggesting the model is not heteroskedastic.
The regression can then be analyzed using the $F$-test. The calculated $F$-value is equal to 16.80. Further, the regression contains 6 numerator degrees of freedom and 1,056 denominator degrees of freedom. This equates to a critical $F$-value of 2.10 at the 5% level of significance. Since $16.80 > 2.10$, the null hypothesis that all regression coefficients equal zero can be rejected. Therefore, the $R^2$ value of 0.0871 indicates that the model explains 8.71% of variation in the dependent variable.

Next, when evaluating each independent variable, each coefficient exhibits the expected sign. Using $p$-values to consider the significance of each variable, *Female*, *Asian*, and *Experience* all have a $p$-value less than 0.05, suggesting they are indeed significant at the 5% level of significance. However, all other variables (*Age*, *Black*, and *ExperienceSquared*) exhibit $p$-values greater than 0.05, indicating insignificance at the 5% level.

A second regression is then run with an identical earnings function, but this time for the group defined by those who hold a higher degree in the age range of 33 to 37 years of age. Looking at pairwise correlations, *Experience* and *ExperienceSquared* exhibit a high correlation of 0.9581. This is not unsuspected and is the only correlation that is high enough to generate a second look. The mean VIF for this last regression is 4.86, implying that multicollinearity most likely does not exist within the model.

This model, as well, must be tested for heteroskedasticity using the White test. Stata produces a calculated chi-square value of 18.69. The critical value, again for 22 degrees of freedom, is 33.924 at the 5% level. It follows that the null hypothesis of homoscedasticity fails to be rejected at the 5% level since the critical value exceeds the
calculated value \( (18.69 < 33.924) \). Therefore, it is not suspected that the regression is heteroskedastic.

Using the \( F \)-test to test the explanatory model of the power, the calculated \( F \)-value of 31.63 can be compared to the critical \( F \)-value of 2.10 based on 6 numerator degrees of freedom and 1,166 denominator degrees of freedom at the 5% level of significance. Following, the calculated value is greater than the critical value \( (31.63 > 2.10) \), verifying a significant \( R^2 \) term at the 5% level. The \( R^2 \) value of .1400 means that the model explains 14.00% of variation in the dependent variable.

In terms of the independent variables, all exhibit the expected sign. Regarding \( p \)-values, only the variable \textit{Black} has a \( p \)-value that is greater that 0.05, suggesting insignificance at the 5% level. All other variables are significant at the 5% level, boasting \( p \)-values less than 0.05.

Next, comparing the summary statistics of each regression, those in the bachelor’s degree subsample average 6.81 years of experience compared to 6.58 years of experience for those in the higher degree category. Thus, since the average experience for each group is relatively equal, applying the Oaxaca-Blinder decomposition procedure in this scenario may produce more helpful results. The calculations for each component of the decomposition are shown in tables 5.4.2, 5.4.3, and 5.4.4.
### Table 5.4.2: $\bar{Y}_{HD}$

<table>
<thead>
<tr>
<th>$X$ Variable Name</th>
<th>$\beta_{HD}$</th>
<th>$\bar{X}_{HD}$</th>
<th>$\bar{Y}<em>{HD} \times \bar{X}</em>{HD}$</th>
<th>Antilog($\bar{Y}_{HD}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.0358405</td>
<td>35.05952</td>
<td>1.256550727</td>
<td>---</td>
</tr>
<tr>
<td>Female</td>
<td>-0.1863469</td>
<td>0.3860544</td>
<td>-0.07994041</td>
<td>---</td>
</tr>
<tr>
<td>Asian</td>
<td>0.1317421</td>
<td>0.2406463</td>
<td>0.03170249</td>
<td>---</td>
</tr>
<tr>
<td>Black</td>
<td>-0.0438799</td>
<td>0.0960884</td>
<td>-0.004216349</td>
<td>---</td>
</tr>
<tr>
<td>Experience</td>
<td>0.1694012</td>
<td>6.582483</td>
<td>1.15086519</td>
<td>---</td>
</tr>
<tr>
<td>ExperienceSquared</td>
<td>-0.0082886</td>
<td>55.93112</td>
<td>-0.0463500681</td>
<td>---</td>
</tr>
<tr>
<td>Constant ($\beta_0$)</td>
<td>---</td>
<td>---</td>
<td>9.422825</td>
<td>---</td>
</tr>
<tr>
<td>Total ($\bar{Y}_{HD}$)</td>
<td>---</td>
<td>---</td>
<td>11.28641242</td>
<td>$79,730.89441$</td>
</tr>
</tbody>
</table>

### Table 5.4.3: $\bar{Y}_{B}$

<table>
<thead>
<tr>
<th>$X$ Variable Name</th>
<th>$\beta_{B}$</th>
<th>$\bar{X}_{B}$</th>
<th>$\hat{Y}<em>{B} \times \bar{X}</em>{B}$</th>
<th>Antilog($\bar{Y}_{B}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.0051468</td>
<td>29.90704</td>
<td>0.15392553</td>
<td>---</td>
</tr>
<tr>
<td>Female</td>
<td>-0.1939622</td>
<td>0.313615</td>
<td>-0.060829455</td>
<td>---</td>
</tr>
<tr>
<td>Asian</td>
<td>0.1302177</td>
<td>0.1286385</td>
<td>0.01675101</td>
<td>---</td>
</tr>
<tr>
<td>Black</td>
<td>-0.0303762</td>
<td>0.1079812</td>
<td>-0.003262782</td>
<td>---</td>
</tr>
<tr>
<td>Experience</td>
<td>0.1838526</td>
<td>6.807512</td>
<td>1.259758781</td>
<td>---</td>
</tr>
<tr>
<td>ExperienceSquared</td>
<td>-0.007423</td>
<td>50.29108</td>
<td>-0.374281305</td>
<td>---</td>
</tr>
<tr>
<td>Constant ($\beta_0$)</td>
<td>---</td>
<td>---</td>
<td>9.968688</td>
<td>---</td>
</tr>
<tr>
<td>Total ($\bar{Y}_{B}$)</td>
<td>---</td>
<td>---</td>
<td>10.9525698</td>
<td>$57,100.59422$</td>
</tr>
</tbody>
</table>
Then, the Oaxaca-Blinder Decomposition procedure yields:

\[
\text{Difference} = \text{Explained} + \text{Unexplained}
\]

\[
\bar{Y}_H - \bar{Y}_B = (\bar{Y}_H - \bar{Y}_B^*) + (\bar{Y}_B^* - \bar{Y}_B)
\]

\[
$22,630.30019 = $7,701.1779 + $14,929.12229
\]

The explained portion ($7,701.18) represents the difference in earnings that can be attributed to actual differences in skill. It was expected, based on the hypothesis, that this portion would be larger than the unexplained term ($14,929.12), suggesting that earnings are linked to more than just holding a higher degree. However, contrary to what was presumed, the explained term is smaller than the unexplained term, suggesting that there is a significant payoff associated with earning a higher degree.
Looking at this using a percentage, the explained portion accounts for 34.03% of the earnings difference. That is, about a third of the difference can be attributed to this term. Similarly, based on the hypothesis, the explained percentage was expected to be larger than the unexplained percentage since it was suspected that the difference in earnings was linked to other factors that contribute to earnings as well. Since the explained term is a smaller percentage, however, the results suggest a high return connected to earning a higher degree.

Based on a pecuniary analysis, in terms of the human capital investment model, if the costs associated with attending school are much higher than the payoff, the individual should choose not to invest. Since the results, however, do show that earning a higher degree offers a fairly significant return; this implies that for most individuals it may be beneficial to invest in a higher degree.

Beyond the explained term, the unexplained portion of the earnings difference must be assessed as well. Since the main determinant of this figure is the difference in regression coefficients between those who have earned higher degrees and those who have earned bachelor’s degrees, the unexplained term was expected to be small based on the assumption that the regression coefficients would not be all that different between the two groups. This was suspected based on the hypothesis that earning a higher degree would not result in a huge increase in payoff, relative to a bachelor’s degree.

The results, again, do not support the hypothesis since $14,929.12 is larger than $7,701.17. This demonstrates that the coefficients between the two regressions were indeed fairly different, attributing to the differing returns. In percentage terms, the unexplained portion accounts for 65.97% of the earnings difference. Once more, this
suggests that earning a higher degree does result in a higher return, relative to a bachelor’s degree.

Although the findings do not support the stated hypothesis, the results of the Oaxaca-Blinder decomposition do, however, reinforce the results found in the preliminary regression using \textit{HigherDegree} as a dummy variable. This regression indicated that earning a higher degree yields a significant return to earnings. Specifically, earning a higher degree can be equated with a 31.8\% increase in $\log\text{Earnings}$, ceteris paribus.

\textbf{A Further Examination of the Results}

Now that the findings have been presented, however, it may still be possible to look at the investment in terms of herding. Although the results should be taken at face value first, it may be possible to hypothesize why the findings appeared as they did. While the investment appears to provide a significant return for the average investor, the decision to invest might still be irrational for some if the costs are greater than the benefits.

Since the explained term does account for about a third of the difference, a sizable amount, it appears that other forms of human capital besides education do still play a role in regards to earnings, even if the part they play is less than the magnitude exhibited by earning a higher degree. Therefore, this may suggest that individuals should still carefully analyze the investment decision on a personal basis. Even though the explained term is less than the unexplained term, it nonetheless represents a significant portion of
the difference. This suggests that the explained portion should not be readily dismissed, despite the smaller value.

If the percentage were much smaller than 34.03%, it would be easier to come to the conclusion that earning a higher degree pays off for mostly everyone. Further, even if the investment is likely to pay off for most, the decision to obtain a higher degree may still be irrational for some if the decision maker chooses her option based on ‘unreliable’ signals, such as herding, that result in a payoff that is less than what is expected and less than the outcome of an alternative choice. In other words, although the payoff of a higher degree is significant, other alternatives may still offer a more significant reward.

Even if the choice is irrational for only a minority of individuals, it is no less important that they gather information on their own to make an informed decision resulting in the best possible solution. This is especially the case if herding is becoming more prevalent, based on increasing enrollment. Although the data used suggests a higher payoff associated with a higher degree today, it is still possible that herding will cause a decrease in returns in the future due to an increasing supply of highly educated workers entering the labor force.

In regards to the unexplained portion, the results are not entirely inconsistent with the hypothesis since it was expected that those with higher degrees would indeed earn more. Although the results did indicate that these earnings were higher than anticipated, it must again be remembered that this is a reflection of recent data. Thus, it is still important to be wary of the consequences of herding in the coming years if this phenomenon is indeed occurring.
If tuition continues to rise, even if those with higher degrees do earn more, higher earnings may not be enough for some individuals to make up for the increase in costs. This is especially true if other alternatives are more attractive in the sense that they yield higher returns. Despite results inconsistent with the hypothesis, it is still important that individuals consider the investment based on private information, as a more informed decision is likely to result in the best outcome.

Although the findings suggest that earning a higher degree does equate to higher earnings, individuals should not take this as a given. Despite results that indicate that most should profit from the higher return, this does not mean that everyone will experience the same fate. Individuals should still consider the investment carefully in terms of individual costs and benefits.

5.5: Conclusion

The empirical results do not support the hypothesis, suggesting that the payoff of earning a higher degree is significant relative to a bachelor’s degree. Running a regression with HigherDegree as a dummy variable, the results indicate that acquiring a higher degree should result in a 31.8% increase in logEarnings, ceteris paribus. Then, removing this dummy variable from the equation allowed for a Chow test to be run. This was used to determine if the two subsamples (bachelor’s degrees versus higher degrees) exhibited equivalent regression coefficients. The results indicated that the two subgroups were indeed structurally different, indicating that it was reasonable to apply the Oaxaca-Blinder decomposition procedure.
The results of the decomposition suggested that those with higher degrees were actually earning less, based on a difference in earnings of -$1,074.98. The explained and unexplained portions displayed opposite signs, with the former equal to -$14,975.39, and the latter equal to $13,882.41. After careful consideration of the data at hand, it became obvious that the difference in experience between the two groups was what likely led to this outcome.

In order to then control for experience, two subsamples with varying age ranges were created to establish two groups that displayed more similar characteristics. The Oaxaca-Blinder decomposition procedure was then applied again, resulting in an earnings difference between the two groups of $22,630.30, such that those with higher degrees were now earning more. Inconsistent with the hypothesis, the explained portion of the difference ($7,701.18) was smaller than the unexplained term ($14,929.12) and accounted for 34.03% of the earnings difference. The unexplained term accounted for 65.97% of the earnings difference.

This does not provide evidence in support of the hypothesis that earning a higher degree may not provide a payoff as great as initially expected. However, although most individuals should receive a significant return if they obtain a higher degree, this does not mean that the few who do not should be discounted. Individuals are engaging in irrational decisions if it is the case that they could have been better off by choosing an alternative. Specifically, an alternative that they could have realized had they acquired information on their own instead of engaging in herding behavior. This alternative may be as simple as choosing a different field of study. Even if this outcome is only
representative of the minority, it is still important to address the phenomenon and its implications.

In extension, the unexplained portion, despite being smaller than the explained portion, accounted for about a third of the earnings difference. This was significant enough to suggest that earning a higher degree was not the only component that explained differences in earnings between those with bachelor’s degrees and those with higher degrees. Thus, the Oaxaca-Blinder decomposition presented here still stresses the importance of acquiring information autonomously. It is still important to base the choice to invest in higher education on factors beyond simply the decisions of others, in the present and past. Individuals must decipher the signals that they receive carefully when making their investment decision.

Further, if herding is becoming more prevalent, this may suggest a different outcome regarding the decomposition of earnings in the future. Since the decision to acquire a higher degree is associated with explicit and implicit costs, not only must an individual pay for tuition itself, but she must also consider the experience and earnings she foregoes by delaying entry into the workforce. This is especially important to consider if tuition rates continues to rise. The direct costs may contribute to a large amount of debt accompanied by a slow return on the investment (as well as no actual guarantee on earning a higher salary).

Especially during economic downturns, having an advanced degree may actually be a liability, considering an individual may be too qualified and thus unaffordable within the labor market. Nonetheless, for others, a higher degree may be required for work they wish to pursue and result in a payoff that validates the investment. Clearly, each
individual must evaluate the decision to pursue higher education carefully, based on individual circumstances and goals.

The Oaxaca-Blinder decomposition procedure presented in this paper has thus provided insight into the returns on earning a higher degree, suggesting that the investment must be considered carefully in terms of the context of the individual at hand. Especially in the future, if the payoff associated with a higher degree declines as a result of herding behavior, it is not wise for decision makers to simply follow the choices made by others before them. Instead, individuals should make their decision by assessing alternatives and gaining a better understanding of the returns.

Nonetheless, the results of this analysis do show that earning a higher degree is still associated with a significant return and higher earnings for the time being. It is unclear whether this will remain the case, but in the meantime, it appears as if the majority of individuals should benefit from pursuing a higher degree.

The model presented, however, is not without limitations. The next chapter discusses future research that can be done to further explore the effects of herding and provide a deeper analysis that complements the research provided and discussed in this chapter.
VI. Conclusion

The decision to pursue higher education is a choice faced by many individuals. The human capital model suggests that individuals engage in the decision to invest in a degree only if the benefits outweigh the costs. Nonetheless, this method is employed using an ex ante procedure, requiring an individual to make predictions about expected returns. Since acquiring information has costs as well, it is likely that many individuals look to see what others have done before them. This behavior is known as herding, which occurs when decision makers abandon their own private signals, and instead make their choice based on the decisions of others who have gone through the process previously.

Although pursuing higher education has resulted in a significant payoff in the past, this does not necessarily imply that the payoff will be the same today. With an increasing supply of educated workers entering the labor force, it is likely that individuals will experience decreasing returns. That is, if many individuals hold the same credentials, they are no longer able to signal that they are more talented than any of their fellow herd members. They are essentially lost in the crowd, causing a decrease in returns based on an increasing proportion of consensuals.

The implications of this are reflected in the fact that students make ex ante decisions based on the human capital model. Therefore, based on herding, decision makers may hold false expectations of the present value of earnings, exaggerating the true outlook. In other words, the benefits actually realized might be less than the benefits that are expected, suggesting that the decision should be evaluated entirely differently.
This is not to argue that earning a degree does not result in higher earnings, but rather, that the earnings may be less than initially anticipated. Although some individuals will still gain from the investment, the decision may be irrational for others if they could have been better off by choosing an alternative, whether this be entering the labor force immediately, or simply choosing a different major.

Current literature has not yet examined the theory of herding when applied to the decision to invest in higher education. Although previous literature has attempted to understand the decision making process and the returns associated with earning a degree, none have gone so far as to examine the role of other people’s decisions on the choice of the individual at hand. The empirical work of this paper attempts to fill this gap by examining the returns to earning a degree beyond the undergraduate level in the presence of an increasing number of students choosing to enroll in higher education.

The herding theory suggests that as more students are currently earning undergraduate degrees, many may feel as if they need to attend graduate school to set themselves apart and distinguish themselves in the labor market. Graduate school, however, is not immune to herding since those who choose to attend may still base their decisions upon what they observe their peers doing, as well as those individuals who made similar decisions in prior years.

The choice to pursue education beyond the bachelor’s level is also useful for analysis because the decision might be more isolated from external pressures, thus reducing possible herding forces. Those who are considering enrollment are most likely less influenced by parental pressures and know more about their own preferences and what they want to do in the future. Having already experienced some college, they may
also be more familiar with the costs associated with earning a degree. Further, they may hold more information that is necessary for making an informed decision, such as information regarding financial and credit constraints that they face, for example. Having already survived the four years it took to earn a bachelor’s degree, it is also likely that the individual is aware of what goes into earning a degree, and is thus much more likely to complete their education as opposed to dropping out.

Assuming that a bachelor’s degree has become a basic requirement for obtaining a decent job, it was hypothesized that the returns to a bachelor’s degree were more significant than the returns to a higher degree, relative to the investment. Even if obtaining a higher degree does allow an individual to earn more overall, it may not be as much as expected, or as much in previous years.

To decompose the difference in earnings between the degree types, the Oaxaca-Blinder decomposition procedure was used to compare the two subgroups. The explained portion of the decomposition was hypothesized to be large since it reflected actual differences in skill level. The unexplained portion, on the other hand, was expected to be small since this term reflects the difference in the market value of skills for each degree type, multiplied by the average skill level of those with a bachelor’s degree. Since it was anticipated that the return on a higher degree would not be that much greater than the payoff of a bachelor’s degree, the unexplained term was expected to be smaller, as this would imply that the regression coefficients between the two groups do not vary by much.

Applying the decomposition procedure to the data obtained from the National Survey of College Graduates, the results indicated that those with bachelor’s degrees
actually earned more than those with higher degrees. Nonetheless, careful examination indicated that this was likely a reflection of the higher experience level held by those with bachelor’s degrees. This can be attributed to the fact that they had been in the workforce for a longer period of time. Then, an attempt to control for experience required using different age ranges for the two subgroups. The decomposition procedure was then applied a second time, controlling for the difference between the two subsamples. The results did not support the hypothesis, however, exhibiting a smaller explained term compared to a larger unexplained term. The explained portion accounted for 34.03% of the difference while the unexplained portion accounted for 65.97%.

Despite the fact that the results do not provide evidence in favor of the hypothesis, the explained portion was still high enough to insinuate that although the majority of individuals will benefit from pursuing a higher degree, others must still evaluate the investment decision carefully. Even if a significant return can be expected for most, this return is not necessarily guaranteed for everyone if other factors do still play a role in the earnings difference. In addition, although the payoff of earning a higher degree does appear to be significant, this does not imply that other alternatives fail to provide a more significant return, thus making the decision to invest in higher education irrational based on a pecuniary analysis. Further, even though the findings did not support a theory of herding behavior and decreasing returns, this does not insinuate that such behaviors will not take place in the future.

The model used in the paper, however, is limited in that it attempted to isolate the effects of the degree by considering the payoff in terms of recent graduates who just entered the workforce. A more complete story could be told, however, if the analysis was
extended to later years in the individual’s life, so long as the effects of earning the degree could still be isolated in some way.

Further, future analysis could extend upon the results of this paper by looking at experience in a more controlled way. Although the proxy variable used in this model is widely accepted as an appropriate means for accounting for experience, it is not fully accurate, as some individuals may have chosen to take time off of their career and reenter the labor force later in life. In addition, individuals may have more experience if they gained skills through activities outside of the workforce, such as volunteering, for example. This leads to the question of the quality of experience, as well. The model used assumes that each year of experience for an individual is of the same quality. However, different individuals no doubt acquire different levels and qualities of experience over the time frame.

Another drawback of the model employed is that it only considers the pecuniary benefits of the degree earned. Although this is consistent with the human capital model, theory also suggests that individuals acquire degrees for non-pecuniary reasons as well. For example, an individual may choose to earn a higher degree simply because they have a love for knowledge and enjoy learning. Others may hold preferences for certain careers that require a higher degree. For example, an individual who wants to be a veterinarian has no choice but to continue with her education. Thus, incorporating preferences would greatly add to the analysis presented.

The results of this paper would also benefit from a model that attempted to understand what factors played a role in the decision of the individual, whether this be innate preferences, parental pressures, signals gleaned from the market, or signals
gleaned from observing how others tackled the same decision (i.e. herding signals). This complementary approach could better assess why individuals made the decisions they did, also accounting for preferences and non-pecuniary benefits.

In addition, although the model focused on STEM majors and career paths, looking at other majors and distinguishing between them could improve the research. This would allow for a more complete understanding of the rewards associated with earning a higher degree, by stratifying the sample further. It may be the case that the payoff of earning a higher degree is much different for majors not in STEM fields.

Lastly, a better means for accounting for herding behavior could also greatly add to the research presented in this paper. One way of accounting for the effects of herding is to compare the returns of a higher degree today to the returns realized in the past. That is, the same survey from a few generations ago could have been employed using the same Oaxaca-Blinder decomposition procedure. If the returns were higher in the past, it is likely that herding and an increase in the supply of educated workers has contributed to this decrease in payoff.

Another way to account for herding is to create a ‘social interaction’ variable when looking at an individual’s decision to pursue a higher degree. For example, a variable could be introduced that accounts for the number of individuals who made the decision to pursue a higher degree in the last few years. This could proxy for herding by factoring in the decisions of others into the choice faced by the individual at hand.

Although the research presented in this paper provides a basic analysis of herding and its effects on the return to higher education, an extension of such research would greatly add to the understanding of such a phenomenon. Although, currently, pursuing
higher education appears to result in a significant return for most individuals, the effects of herding behavior may have not yet been seen. If the power of custom continues to prevail, individuals may be forced to reevaluate the decision to invest in higher education. Further, if individuals are finding themselves in situations in which they are worse off than if they had chosen an alternative path, the choice to pursue a higher degree may have been an irrational decision.

This does not mean that America’s education system is becoming worthless. Rather, this theory suggests that individuals would benefit from gathering information on their own rather than following the herd, hopefully resulting in a better outcome than if the individual follows the signals of others. Although the investment appears to be worthwhile for most, for others, it may still pay off to choose a different path.

This also does not imply that the country should stop focusing on education. The positive externalities of education reach far and wide and are necessary for the advancement and competitiveness of the nation. Instead, education may simply need some reform. With an abundance of intelligent individuals, willing teachers, and creative resources, it may just take a better way to put these components together to result in a worthy payout for the individual making the choice. The more an individual is informed about her options, the more likely she is to make a choice that is best for her, regardless of what others are doing.
VII. Appendices

Appendix A: The Basic Regression

Statistics of Variables used in the Data Set:

```
. describe Earnings logEarnings HigherDegree Age Female Asian Black Experience ExperienceSquared HRSWK
storage   display    value
variable name   type    format     label      variable label
------------------------------------------------------------
---------------
---------
Earnings        long    %8.0g
logEarnings     float   %9.0g
HigherDegree    float   %9.0g
Age             byte    %8.0g
Female          float   %9.0g
Asian           float   %9.0g
Black           float   %9.0g
Experience      float   %9.0g
ExperienceSquared float   %9.0g
HRSWK           byte    %8.0g
```

```
. summarize Earnings logEarnings HigherDegree Age Female Asian Black Experience ExperienceSquared HRSWK
Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------
Earnings |      2599  72696.31     68007.8          0     999996
logEarnings |      2591   10.95519    .7145836   4.174387   13.81551
HigherDegree |      2599   .5151982    .4998651          0          1
Age |      2599  30.48134    1.701881         28         33
Female |      2599   .3932282    .4885608          0          1
Asian |      2599   .1843017    .3878046          0          1
Black |      2599   .100808    .3011323          0          1
Experience |      2599   5.303963     2.95969          0         15
ExperienceSquared |      2599  36.88842    33.16927          0        225
HRSWK |      2599  55.90958   8.617654         48         96
```

Regression run with HigherDegree as a Dummy Variable:

```
. regress logEarnings HigherDegree Age Female Asian Black Experience ExperienceSquared
Source |       SS       df       MS
-------------
F(  7,  2583) =   59.11
Model |  182.613089     7  26.0875842           Prob > F      =  0.0000
Residual |  1139.91776  2583   .44131543           R-squared     =  0.1381
Total |  1322.53085  2590   .51062967           Root MSE      =  .66432

logEarnings |      Coef.    Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------
HigherDegree |    .318277    .036627     8.69   0.000     .2464557    .3900982
Age |   .0270612   .0091268     2.97   0.003     .0091647    .0449578
Female | -.1815939   .027195    -6.68   0.000    -.23492    -.1282677
Asian |   .1384522   .034564     4.01   0.000     .0706761    .2062282
Black | -.0009048   .0440587    -0.02   0.984    -.0872987    .0854892
Experience |   .189099   .0153748    12.30   0.000     .1589509    .2192471
ExperienceSquared | -.0093928   .0013713    -6.85   0.000    -.0120817    -.0067039
_cons |   9.356601   .2673217    35.00   0.000     8.832415    9.880788
```
Pairwise Correction for Regression using HigherDegree as a Dummy Variable:

```plaintext
. pwcorr logEarnings HigherDegree Age Female Asian Black Experience ExperienceSquared

<table>
<thead>
<tr>
<th></th>
<th>logEarnings</th>
<th>HigherDegree</th>
<th>Age</th>
<th>Female</th>
<th>Asian</th>
<th>Black</th>
<th>Experience</th>
<th>ExperienceSquared</th>
</tr>
</thead>
<tbody>
<tr>
<td>logEarnings</td>
<td>1.0000</td>
<td>-0.0136</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HigherDegree</td>
<td>-0.0136</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>0.1754</td>
<td>0.0545</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>-0.1136</td>
<td>0.1536</td>
<td>-0.0240</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>0.0943</td>
<td>0.1315</td>
<td>-0.0137</td>
<td>-0.0129</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>-0.0222</td>
<td>-0.0332</td>
<td>-0.0144</td>
<td>0.0627</td>
<td>-0.1592</td>
<td>1.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experience</td>
<td>0.2583</td>
<td>-0.6158</td>
<td>0.3488</td>
<td>-0.0601</td>
<td>-0.0250</td>
<td>0.0015</td>
<td>1.0000</td>
<td></td>
</tr>
<tr>
<td>ExperienceSquared</td>
<td>0.2150</td>
<td>-0.5784</td>
<td>0.4037</td>
<td>-0.0584</td>
<td>-0.0206</td>
<td>-0.0114</td>
<td>0.9546</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>logEarnings</th>
<th>HigherDegree</th>
<th>Age</th>
<th>Female</th>
<th>Asian</th>
<th>Black</th>
<th>Experience</th>
<th>ExperienceSquared</th>
</tr>
</thead>
<tbody>
<tr>
<td>logEarnings</td>
<td>1.0000</td>
<td>-0.0136</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HigherDegree</td>
<td>-0.0136</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>0.1754</td>
<td>0.0545</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>-0.1136</td>
<td>0.1536</td>
<td>-0.0240</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>0.0943</td>
<td>0.1315</td>
<td>-0.0137</td>
<td>-0.0129</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>-0.0222</td>
<td>-0.0332</td>
<td>-0.0144</td>
<td>0.0627</td>
<td>-0.1592</td>
<td>1.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experience</td>
<td>0.2583</td>
<td>-0.6158</td>
<td>0.3488</td>
<td>-0.0601</td>
<td>-0.0250</td>
<td>0.0015</td>
<td>1.0000</td>
<td></td>
</tr>
<tr>
<td>ExperienceSquared</td>
<td>0.2150</td>
<td>-0.5784</td>
<td>0.4037</td>
<td>-0.0584</td>
<td>-0.0206</td>
<td>-0.0114</td>
<td>0.9546</td>
<td></td>
</tr>
</tbody>
</table>

Mean VIF:

<table>
<thead>
<tr>
<th>Variable</th>
<th>VIF</th>
<th>1/VIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experience</td>
<td>12.15</td>
<td>0.082313</td>
</tr>
<tr>
<td>ExperienceSquared</td>
<td>12.11</td>
<td>0.082570</td>
</tr>
<tr>
<td>HigherDegree</td>
<td>1.97</td>
<td>0.508303</td>
</tr>
<tr>
<td>Age</td>
<td>1.42</td>
<td>0.705995</td>
</tr>
<tr>
<td>Asian</td>
<td>1.05</td>
<td>0.949170</td>
</tr>
<tr>
<td>Female</td>
<td>1.04</td>
<td>0.964858</td>
</tr>
<tr>
<td>Black</td>
<td>1.03</td>
<td>0.968627</td>
</tr>
</tbody>
</table>

Mean VIF | 4.40

Test for Heteroskedasticity:

```plaintext
. imtest, white

White's test for H0: homoskedasticity
against Ha: unrestricted heteroskedasticity

<table>
<thead>
<tr>
<th>Source</th>
<th>chi2</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heteroskedasticity</td>
<td>40.24</td>
<td>29</td>
<td>0.0801</td>
</tr>
<tr>
<td>Skewness</td>
<td>9.20</td>
<td>7</td>
<td>0.2390</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>3.55</td>
<td>1</td>
<td>0.0596</td>
</tr>
<tr>
<td>Total</td>
<td>52.98</td>
<td>37</td>
<td>0.0429</td>
</tr>
</tbody>
</table>
```
Appendix B: The Chow Test

Statistics of Variables used in the Data Set:

```plaintext
. describe Earnings logEarnings HigherDegree Age Female Asian Black Experience ExperienceSquared HRSWK

<table>
<thead>
<tr>
<th>variable name</th>
<th>storage</th>
<th>display</th>
<th>value</th>
<th>label</th>
<th>variable label</th>
</tr>
</thead>
<tbody>
<tr>
<td>Earnings</td>
<td>long</td>
<td>%8.0g</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>logEarnings</td>
<td>float</td>
<td>%9.0g</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HigherDegree</td>
<td>float</td>
<td>%9.0g</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>byte</td>
<td>%8.0g</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>float</td>
<td>%9.0g</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>float</td>
<td>%9.0g</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>float</td>
<td>%9.0g</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experience</td>
<td>float</td>
<td>%9.0g</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ExperienceSquared</td>
<td>float</td>
<td>%9.0g</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HRSWK</td>
<td>byte</td>
<td>%8.0g</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```

```plaintext
. summarize Earnings logEarnings HigherDegree Age Female Asian Black Experience ExperienceSquared HRSWK

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Earnings</td>
<td>2599</td>
<td>72696.31</td>
<td>68007.8</td>
<td>0</td>
<td>999996</td>
</tr>
<tr>
<td>logEarnings</td>
<td>2591</td>
<td>10.95519</td>
<td>.7145836</td>
<td>4.174387</td>
<td>13.81551</td>
</tr>
<tr>
<td>HigherDegree</td>
<td>2599</td>
<td>.5151982</td>
<td>.4998651</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Age</td>
<td>2599</td>
<td>30.48134</td>
<td>1.701881</td>
<td>28</td>
<td>33</td>
</tr>
<tr>
<td>Female</td>
<td>2599</td>
<td>.3932282</td>
<td>.4885608</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Asian</td>
<td>2599</td>
<td>.1843017</td>
<td>.3878046</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Black</td>
<td>2599</td>
<td>.100808</td>
<td>.3011323</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Experience</td>
<td>2599</td>
<td>5.303963</td>
<td>2.959690</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td>ExperienceSquared</td>
<td>2599</td>
<td>36.88842</td>
<td>33.16927</td>
<td>0</td>
<td>225</td>
</tr>
<tr>
<td>HRSWK</td>
<td>2599</td>
<td>55.90958</td>
<td>8.617654</td>
<td>48</td>
<td>96</td>
</tr>
</tbody>
</table>
```

Regression Run without HigherDegree:

```plaintext
. regress logEarnings Age Female Asian Black Experience ExperienceSquared HRSWK

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>149.2891</td>
<td>6</td>
<td>24.8815167</td>
</tr>
<tr>
<td>Residual</td>
<td>1173.24175</td>
<td>2584</td>
<td>.454040923</td>
</tr>
<tr>
<td>Total</td>
<td>1322.53085</td>
<td>2590</td>
<td>.51062967</td>
</tr>
</tbody>
</table>

| logEarnings | Coef. | Std. Err. | t | P>|t| | [95% Conf. Interval] |
|-------------|-------|-----------|---|-----|---------------------|
| Age         | .0568379 | .0085802 | 6.62 | 0.000 | .0400132 .0736627 |
| Female      | -.142055 | .0271954 | -5.22 | 0.000 | -.1953821 -.088728 |
| Asian       | .1863462 | .0346102 | 5.38 | 0.000 | .1184796 .2542128 |
| Black       | -.0104684 | .0046755 | -.23 | .815 | -.0980718 .0771349 |
| Experience  | .1582335 | .015173 | 10.43 | 0.000 | .1284811 .187986 |
| ExperienceSquared | -.0010179 | .0013883 | -7.28 | 0.000 | -.0128303 -.0073854 |
| _cons       | 8.779452 | .2624661 | 33.43 | 0.000 | 8.264434 9.29447 |
```
Pairwise Correlation for Regression without *HigherDegree*:
```
pwcorr logEarnings Age Female Asian Black Experience ExperienceSquared
```

<table>
<thead>
<tr>
<th></th>
<th>logEarnings</th>
<th>Age</th>
<th>Female</th>
<th>Asian</th>
<th>Black</th>
<th>Experience</th>
<th>ExperienceSquared</th>
</tr>
</thead>
<tbody>
<tr>
<td>logEarnings</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>0.1754</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>-0.1136</td>
<td>-0.0240</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>0.0943</td>
<td>-0.0137</td>
<td>-0.0129</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>-0.0222</td>
<td>-0.0144</td>
<td>0.0627</td>
<td>-0.1592</td>
<td>1.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experience</td>
<td>0.2583</td>
<td>0.3488</td>
<td>-0.0601</td>
<td>-0.0250</td>
<td>0.0015</td>
<td>1.0000</td>
<td></td>
</tr>
<tr>
<td>ExperienceSquared</td>
<td>0.2150</td>
<td>0.4037</td>
<td>-0.0584</td>
<td>-0.0206</td>
<td>-0.0114</td>
<td>0.9546</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

Mean VIF:
```
vif
```

<table>
<thead>
<tr>
<th>Variable</th>
<th>VIF</th>
<th>1/VIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experience-d</td>
<td>12.07</td>
<td>0.082869</td>
</tr>
<tr>
<td>Experience</td>
<td>11.50</td>
<td>0.086954</td>
</tr>
<tr>
<td>Age</td>
<td>1.22</td>
<td>0.821846</td>
</tr>
<tr>
<td>Black</td>
<td>1.03</td>
<td>0.969232</td>
</tr>
<tr>
<td>Asian</td>
<td>1.03</td>
<td>0.973935</td>
</tr>
<tr>
<td>Female</td>
<td>1.01</td>
<td>0.992646</td>
</tr>
<tr>
<td>Mean VIF</td>
<td>4.64</td>
<td></td>
</tr>
</tbody>
</table>

Test for Heteroskedasticity:
```
imtest, white
```

White's test for Ho: homoskedasticity
against Ha: unrestricted heteroskedasticity

```
chisq(22) = 35.66
Prob > chisq = 0.0330
```

Cameron & Trivedi's decomposition of IM-test
```
<table>
<thead>
<tr>
<th>Source</th>
<th>chisq</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heteroskedasticity</td>
<td>35.66</td>
<td>22</td>
<td>0.0330</td>
</tr>
<tr>
<td>Skewness</td>
<td>7.65</td>
<td>6</td>
<td>0.2646</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>3.45</td>
<td>1</td>
<td>0.0633</td>
</tr>
<tr>
<td>Total</td>
<td>46.77</td>
<td>29</td>
<td>0.0197</td>
</tr>
</tbody>
</table>
```

Correction for Heteroskedasticity:
```
regress logEarnings Age Female Asian Black Experience ExperienceSquared, robust
```

Linear regression
```
Number of obs = 2591
F(  6,  2584) = 49.45
Prob > F = 0.0000
R-squared = 0.1129
Root MSE = 0.67383
```

| logEarnings | Robust Coef. | Std. Err. | t   | P>|t| | [95% Conf. Interval] |
|-------------|--------------|------------|-----|-----|----------------------|
| Age         | 0.0568379    | 0.0087026  | 6.53| 0.000| 0.0397733  0.0739026 |
| Female      | -0.142055    | 0.0267398  | -5.31| 0.000| -0.1944886 -0.0896215 |
| Asian       | 0.1863462    | 0.0366986  | 5.08| 0.000| 0.1143845  0.2583079 |
| Black       | -0.0104684   | 0.0472949  | -0.22| 0.825| -0.1032082 0.0822714 |
| Experience  | 0.1582335    | 0.0167446  | 9.45| 0.000| 0.1253993  0.1910678 |
| ExperienceSquared | -0.0101078 | 0.0015191  | -6.65| 0.000| -0.0130865 -0.0071291 |
| _cons       | 8.779452     | 0.2681293  | 32.74| 0.000| 8.253682  9.305222  |
Regression Run for Bachelor’s Degrees Only:

```
. regress logEarnings Age Female Asian Black Experience ExperienceSquared if HigherDegree == 0
```

```
Source | SS           df | MS    Number of obs = 1257
---------|---------------|-------|-------------------|
Model | 50.0083153 | 6 | 8.33471922 | F( 6, 1250) = 19.47
Residual | 535.150324 | 1250 | .428120259 | Prob > F = 0.0000
---------|------------|-----|-----------|-------------------|
Total | 585.158639 | 1256 | .465890636 | Adj R-squared = 0.0811
```

```
+----------------------------------------+-----------------+---------|
| logEarnings | Coef. | Std. Err. | t  | P>|t| | [95% Conf. Interval] |
|-------------|-------|-----------|----|-------|-------------------------|
| Age | -.0062673 | .0146614 | -0.43 | 0.669 | -.035031 | .024964 |
| Female | -1.1913773 | .0398765 | -4.80 | 0.000 | -.2696096 | -.113145 |
| Asian | .1593942 | .0553813 | 2.88 | 0.004 | .0507437 | .2680448 |
| Black | -.0229668 | .0597445 | -0.38 | 0.701 | -.1401774 | .0942438 |
| Experience | .1413166 | .0410553 | 3.44 | 0.001 | .0607717 | .2218616 |
| ExperienceSquared | -0.0045916 | .0029661 | -1.55 | 0.122 | -.0104107 | .0012275 |
| _cons | 10.44329 | .4696612 | 22.24 | 0.000 | 9.521876 | 11.3647 |
```

Summary Statistics for those with a Bachelor’s Degree:

```
. summarize logEarnings Age Female Asian Black Experience ExperienceSquared if HigherDegree == 0

Variable | Obs | Mean | Std. Dev. | Min | Max
---------|-----|------|-----------|-----|------
logEarnings | 1257 | 10.96521 | .6825618 | 4.174387 | 13.81551 |
Age | 1260 | 30.38571 | 1.706071 | 28 | 33 |
Female | 1260 | .315873 | .4650472 | 0 | 1 |
Asian | 1260 | .131746 | .3383487 | 0 | 1 |
Black | 1260 | .111111 | .3143945 | 0 | 1 |
Experience | 1260 | 7.18254 | 2.253233 | 1 | 15 |
Experience-d | 1260 | 56.6619 | 32.82875 | 1 | 225 |
```

Pairwise Correlation for Regression Run for Bachelor’s Degrees:

```
. pwcorr logEarnings Age Female Asian Black Experience ExperienceSquared if HigherDegree == 0
```

```
<table>
<thead>
<tr>
<th>logEarnings</th>
<th>Age</th>
<th>Female</th>
<th>Asian</th>
<th>Black</th>
<th>Experience</th>
<th>ExperienceSquared</th>
</tr>
</thead>
<tbody>
<tr>
<td>logEarnings</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>0.1244</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>-0.1243</td>
<td>-0.0235</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>0.0991</td>
<td>-0.0028</td>
<td>-0.0173</td>
<td>1.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>-0.0408</td>
<td>0.0104</td>
<td>0.0586</td>
<td>-0.1377</td>
<td>1.0000</td>
<td></td>
</tr>
<tr>
<td>Experience</td>
<td>0.2428</td>
<td>0.6028</td>
<td>0.0321</td>
<td>0.0716</td>
<td>-0.0522</td>
<td>1.0000</td>
</tr>
<tr>
<td>Experience-d</td>
<td>0.2276</td>
<td>0.6497</td>
<td>0.0331</td>
<td>0.0736</td>
<td>-0.0588</td>
<td>0.9789</td>
</tr>
</tbody>
</table>
```

Mean VIF:

```
. vif

<table>
<thead>
<tr>
<th>Variable</th>
<th>VIF</th>
<th>1/VIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experience-d</td>
<td>27.71</td>
<td>0.036086</td>
</tr>
<tr>
<td>Experience</td>
<td>25.05</td>
<td>0.039923</td>
</tr>
<tr>
<td>Age</td>
<td>1.83</td>
<td>0.545290</td>
</tr>
<tr>
<td>Black</td>
<td>1.03</td>
<td>0.970171</td>
</tr>
<tr>
<td>Asian</td>
<td>1.03</td>
<td>0.973797</td>
</tr>
<tr>
<td>Female</td>
<td>1.01</td>
<td>0.991237</td>
</tr>
</tbody>
</table>
```

Mean VIF | 9.61
Test for Heteroskedasticity:
.imtest, white

White's test for Ho: homoskedasticity
against Ha: unrestricted heteroskedasticity

\[ \chi^2(22) = 19.79 \]
Prob > \chi^2 = 0.5962

Cameron & Trivedi's decomposition of IM-test

<table>
<thead>
<tr>
<th>Source</th>
<th>\chi^2</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heteroskedasticity</td>
<td>19.79</td>
<td>22</td>
<td>0.5962</td>
</tr>
<tr>
<td>Skewness</td>
<td>5.52</td>
<td>6</td>
<td>0.4792</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>2.51</td>
<td>1</td>
<td>0.1134</td>
</tr>
<tr>
<td>Total</td>
<td>27.81</td>
<td>29</td>
<td>0.5278</td>
</tr>
</tbody>
</table>

Regression Run for Higher Degree Only:
.regress logEarnings Age Female Asian Black Experience ExperienceSquared if HigherDegree ==1

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>Number of obs = 1334</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>139.302945</td>
<td>6</td>
<td>23.2171574</td>
<td>F(6, 1327) = 51.54</td>
</tr>
<tr>
<td>Residual</td>
<td>597.824148</td>
<td>1327</td>
<td>.450508024</td>
<td>Prob &gt; F = 0.0000</td>
</tr>
<tr>
<td>Total</td>
<td>737.127093</td>
<td>1333</td>
<td>.552983566</td>
<td>Adj R-squared = 0.1853</td>
</tr>
</tbody>
</table>

| logEarnings   | Coef.  | Std. Err. | t    | P>|t| | [95% Conf. Interval] |
|---------------|--------|-----------|------|-----|----------------------|
| Age           | .0461481 | .0118743  | 3.89 | 0.000 | .0228537   .0694425 |
| Female        | -.1761489 | .0371138  | -4.75 | 0.000 | -.2489569  -.1033408 |
| Asian         | .1302865  | .0443845  | 2.94 | 0.003 | .0432151   .2173579 |
| Black         | .0416191  | .0650073  | 0.64 | 0.522 | -.0859091  .1691474 |
| Experience    | .23386    | .0236812  | 9.87 | 0.000 | .1874001   .2803199 |
| ExperienceSquared | -.0152666 | .0028327  | -5.74 | 0.000 | -.0218236  -.0107096 |
| _cons         | 9.054363  | .359649   | 25.18 | 0.000 | 8.348821   9.759906 |

Summary Statistics for those with a Higher Degree:
.summarize logEarnings Age Female Asian Black Experience ExperienceSquared if HigherDegree ==1

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>logEarnings</td>
<td>1334</td>
<td>10.94575</td>
<td>.7436286</td>
<td>6.907755</td>
<td>13.81551</td>
</tr>
<tr>
<td>Age</td>
<td>1339</td>
<td>30.57132</td>
<td>1.693369</td>
<td>28</td>
<td>33</td>
</tr>
<tr>
<td>Female</td>
<td>1339</td>
<td>.4660194</td>
<td>.4990304</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Asian</td>
<td>1339</td>
<td>.2337565</td>
<td>.4237727</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Black</td>
<td>1339</td>
<td>.0911128</td>
<td>.2878769</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Experience</td>
<td>1339</td>
<td>3.536221</td>
<td>2.404373</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>Experience-d</td>
<td>1339</td>
<td>18.28155</td>
<td>20.16859</td>
<td>0</td>
<td>121</td>
</tr>
</tbody>
</table>
Pairwise Correlation for Regression Run for Higher Degrees:

. pwcorr logEarnings Age Female Asian Black Experience ExperienceSquared if HigherDegree ==1

<table>
<thead>
<tr>
<th>logEarnings</th>
<th>Age</th>
<th>Female</th>
<th>Asian</th>
<th>Black</th>
<th>Experience</th>
<th>ExperienceSquared</th>
</tr>
</thead>
<tbody>
<tr>
<td>logEarnings</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>0.2221</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>-0.1037</td>
<td>-0.0411</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>0.0958</td>
<td>-0.0353</td>
<td>-0.0455</td>
<td>1.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>-0.0057</td>
<td>-0.0363</td>
<td>0.0788</td>
<td>-0.1749</td>
<td>1.0000</td>
<td></td>
</tr>
<tr>
<td>Experience</td>
<td>0.3773</td>
<td>0.3830</td>
<td>0.0544</td>
<td>0.0721</td>
<td>0.0028</td>
<td>1.0000</td>
</tr>
<tr>
<td>ExperienceSquared</td>
<td>0.3149</td>
<td>0.3938</td>
<td>0.0480</td>
<td>0.0720</td>
<td>-0.0039</td>
<td>0.9464</td>
</tr>
</tbody>
</table>

Mean VIF:

. vif

<table>
<thead>
<tr>
<th>Variable</th>
<th>VIF</th>
<th>1/VIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experience</td>
<td>9.70</td>
<td>0.103104</td>
</tr>
<tr>
<td>Experience</td>
<td>9.61</td>
<td>0.104041</td>
</tr>
<tr>
<td>Age</td>
<td>1.20</td>
<td>0.834333</td>
</tr>
<tr>
<td>Asian</td>
<td>1.05</td>
<td>0.956733</td>
</tr>
<tr>
<td>Black</td>
<td>1.04</td>
<td>0.961772</td>
</tr>
<tr>
<td>Female</td>
<td>1.02</td>
<td>0.984988</td>
</tr>
</tbody>
</table>

Mean VIF | 3.93

Test for Heteroskedasticity:

. imtest, white

White's test for Ho: homoskedasticity
against Ha: unrestricted heteroskedasticity

<table>
<thead>
<tr>
<th>Source</th>
<th>chi2</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heteroskedasticity</td>
<td>43.25</td>
<td>22</td>
<td>0.0044</td>
</tr>
<tr>
<td>Skewness</td>
<td>14.13</td>
<td>6</td>
<td>0.0012</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>10.44</td>
<td>1</td>
<td>0.0012</td>
</tr>
</tbody>
</table>

Total | 67.82 | 29 | 0.0001 |

Correction for Heteroskedasticity:

. regress logEarnings Age Female Asian Black Experience ExperienceSquared if HigherDegree ==1, robust

Linear regression

Number of obs = 1334
F( 6, 1327) = 46.67
Prob > F = 0.0000
R-squared = 0.1890
Root MSE = .6712

| logEarnings | Coef. | Std. Err. | t | P>|t| | [95% Conf. Interval] |
|-------------|-------|-----------|---|-----|-----------------|
| Age | 0.0461481 | 0.0123605 | 3.73 | 0.000 | 0.0218999 | 0.0703962 |
| Female | -0.1761489 | 0.0369727 | -4.76 | 0.000 | -0.2486801 | -0.1036176 |
| Asian | 0.1302865 | 0.0470437 | 2.77 | 0.006 | 0.0379983 | 0.2225748 |
| Black | 0.0416191 | 0.0678912 | 0.61 | 0.540 | -0.0915667 | 0.174805 |
| Experience | 0.23386 | 0.0546994 | 4.33 | 0.000 | 0.1818746 | 0.2858453 |
| ExperienceSquared | -0.0162666 | 0.0031128 | -5.23 | 0.000 | -0.0223731 | -0.01016 |
| _cons | 9.054363 | 3.75019 | 24.14 | 0.000 | 8.318669 | 9.790058 |

152
Appendix C: The Initial Oaxaca-Blinder Decomposition Procedure

Statistics of Variables used in the Data Set:

<table>
<thead>
<tr>
<th>variable name</th>
<th>storage</th>
<th>display</th>
<th>value format</th>
<th>value label</th>
</tr>
</thead>
<tbody>
<tr>
<td>Earnings</td>
<td>long</td>
<td>%8.0g</td>
<td></td>
<td></td>
</tr>
<tr>
<td>logEarnings</td>
<td>float</td>
<td>%9.0g</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HigherDegree</td>
<td>float</td>
<td>%9.0g</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>byte</td>
<td>%8.0g</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>float</td>
<td>%9.0g</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>float</td>
<td>%9.0g</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>float</td>
<td>%9.0g</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experience</td>
<td>float</td>
<td>%9.0g</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ExperienceSquared</td>
<td>float</td>
<td>%9.0g</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HRSWK</td>
<td>byte</td>
<td>%8.0g</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Regression Run for Bachelor’s Degrees Only:

. regress logEarnings Age Female Asian Black Experience ExperienceSquared if HigherDegree == 0

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>Number of obs = 1257</th>
</tr>
</thead>
<tbody>
<tr>
<td>F(  6,  1250)</td>
<td>19.47</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model</td>
<td>50.0083153</td>
<td>6</td>
<td>8.33471922</td>
<td>Prob &gt; F = 0.0000</td>
</tr>
<tr>
<td>Residual</td>
<td>535.150324</td>
<td>1250</td>
<td>.428120259</td>
<td>R-squared = 0.0855</td>
</tr>
<tr>
<td>Total</td>
<td>585.158639</td>
<td>1256</td>
<td>.46589636</td>
<td>Adj R-squared = 0.0811</td>
</tr>
<tr>
<td>Root MSE</td>
<td>.65431</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| logEarnings | Coef.  | Std. Err. | t   | P>|t| | [95% Conf. Interval] |
|-------------|--------|-----------|-----|-----|---------------------|
| Age         | -.0062673 | .0146614 | -0.43 | 0.669 | -.035031 to .0224964 |
| Female      | -.1913773 | .0398765 | -4.80 | 0.000 | -.2696096 to -.113145 |
| Asian       | .1593942 | .0553813 | 2.88 | 0.004 | .0507437 to .2680448 |
| Black       | -.0229668 | .0597445 | -0.38 | 0.701 | -.1401774 to .0942438 |
| Experience  | .1413166 | .0410553 | 3.44 | 0.001 | .0607717 to .2218616 |
| ExperienceSquared | -.0045916 | .0029661 | -1.55 | 0.122 | -.0104107 to .0012275 |
| _cons        | 10.44329 | .4696612 | 22.24 | 0.000 | 9.521876 to 11.3647 |
Summary Statistics for those with a Bachelor’s Degree:

```
 summarize logEarnings Age Female Asian Black Experience ExperienceSquared if HigherDegree ==0
```

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>logEarnings</td>
<td>1257</td>
<td>10.96521</td>
<td>4.174387</td>
<td>0.6825618</td>
<td>13.81551</td>
</tr>
<tr>
<td>Age</td>
<td>1260</td>
<td>30.38571</td>
<td>1.706071</td>
<td>28</td>
<td>33</td>
</tr>
<tr>
<td>Female</td>
<td>1260</td>
<td>.315873</td>
<td>.4650472</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Asian</td>
<td>1260</td>
<td>.131746</td>
<td>.3383487</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Black</td>
<td>1260</td>
<td>.1111111</td>
<td>.3143945</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Experience</td>
<td>1260</td>
<td>7.18254</td>
<td>2.253233</td>
<td>1</td>
<td>15</td>
</tr>
<tr>
<td>Experience~d</td>
<td>1260</td>
<td>56.6619</td>
<td>32.82875</td>
<td>1</td>
<td>225</td>
</tr>
</tbody>
</table>

Pairwise Correlation for Regression Run for Bachelor’s Degrees:

```
pwcorr logEarnings Age Female Asian Black Experience ExperienceSquared if HigherDegree ==0
```

<table>
<thead>
<tr>
<th></th>
<th>logEarnings</th>
<th>Age</th>
<th>Female</th>
<th>Asian</th>
<th>Black</th>
<th>Experience</th>
<th>Experience~d</th>
</tr>
</thead>
<tbody>
<tr>
<td>logEarnings</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>0.1244</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>-0.1243</td>
<td>-0.0235</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>0.0991</td>
<td>-0.0028</td>
<td>-0.0173</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>-0.0408</td>
<td>0.0104</td>
<td>0.0586</td>
<td>-0.1377</td>
<td>1.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experience</td>
<td>0.2428</td>
<td>0.6028</td>
<td>0.0321</td>
<td>0.0716</td>
<td>-0.0522</td>
<td>1.0000</td>
<td></td>
</tr>
<tr>
<td>Experience~d</td>
<td>0.2276</td>
<td>0.6497</td>
<td>0.0331</td>
<td>0.0736</td>
<td>-0.0588</td>
<td>0.9789</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

Mean VIF:

```
vif
```

<table>
<thead>
<tr>
<th>Variable</th>
<th>VIF</th>
<th>1/VIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experience~d</td>
<td>27.71</td>
<td>0.036086</td>
</tr>
<tr>
<td>Experience</td>
<td>25.05</td>
<td>0.039923</td>
</tr>
<tr>
<td>Age</td>
<td>1.83</td>
<td>0.545290</td>
</tr>
<tr>
<td>Black</td>
<td>1.03</td>
<td>0.970171</td>
</tr>
<tr>
<td>Asian</td>
<td>1.03</td>
<td>0.973797</td>
</tr>
<tr>
<td>Female</td>
<td>1.01</td>
<td>0.991237</td>
</tr>
</tbody>
</table>

Mean VIF | 9.61

Test for Heteroskedasticity:

```
imtest, white
```

White’s test for Ho: homoskedasticity
against Ha: unrestricted heteroskedasticity

```
chi2(22) = 19.79
Prob > chi2 = 0.5962
```

Cameron & Trivedi’s decomposition of IM-test

<table>
<thead>
<tr>
<th>Source</th>
<th>chi2</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heteroskedasticity</td>
<td>19.79</td>
<td>22</td>
<td>0.5962</td>
</tr>
<tr>
<td>Skewness</td>
<td>5.52</td>
<td>6</td>
<td>0.4792</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>2.51</td>
<td>1</td>
<td>0.1134</td>
</tr>
<tr>
<td>Total</td>
<td>27.81</td>
<td>29</td>
<td>0.5278</td>
</tr>
</tbody>
</table>
Regression Run for Higher Degree Only:
.regress logEarnings Age Female Asian Black Experience ExperienceSquared if HigherDegree ==1

Source | SS    df   MS   Number of obs = 1334
-----------------------------
Model  | 139.302945  6   23.2171574  F(  6,  1327) = 51.54
Residual | 597.824148  1327 .450508024  Prob > F = 0.0000
-----------------------------
Total  | 737.127093  1333 .552983566  R-squared = 0.1890
Adj R-squared = 0.1853

logEarnings | Coef.  Std. Err.  t   P>|t|  [95% Conf. Interval]
-----------------------------
Age | .0461481  .0118743  3.89   0.000   .0228537    .0694425
Female | -.1761489  .0371138 -4.75  0.000  -.2489569  -.1033408
Asian | .1302065  .0443845  2.94  0.003   .0432151    .2173579
Black | .0416191  .0650073  0.64  0.522  -.0859091    .1691474
Experience | .23386   .0236828  9.87  0.000   .1874001    .2803199
ExperienceSquared | -.0162666  .0028327 -5.74  0.000   -.0218236    -.0107096
_cons | 9.054363    .359649  25.18  0.000   8.348821    9.759906

Summary Statistics for those with a Higher Degree:
.summarize logEarnings Age Female Asian Black Experience ExperienceSquared if HigherDegree ==1

Variable |       Obs        Mean  Std. Dev.   Min       Max
------------+------------------------------------------
logEarnings |  1334    10.94575   .7436286   6.907755   13.81551
Age |  1339    30.57132   1.693639     28        33
Female |  1339    .4660194   .4990304          0        1
Asian |  1339    .2337565   .4233772          0        1
Black |  1339    .0911128   .2878769          0        1
Experience |  1339    3.536221 2.404373      0        11
Experience-d |  1339   18.28155 20.18659      0        121

Pairwise Correlation for Regression Run for Higher Degrees:
.pwcorr logEarnings Age Female Asian Black Experience ExperienceSquared if HigherDegree ==1

<table>
<thead>
<tr>
<th>logEarnings</th>
<th>Age</th>
<th>Female</th>
<th>Asian</th>
<th>Black</th>
<th>Experience</th>
<th>Experience-d</th>
</tr>
</thead>
<tbody>
<tr>
<td>logEarnings</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>.2221</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>-.01037</td>
<td>-.0411</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>.0958</td>
<td>-.0353</td>
<td>-.0455</td>
<td>1.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>-.0057</td>
<td>-.0363</td>
<td>.0788</td>
<td>-.1749</td>
<td>1.0000</td>
<td></td>
</tr>
<tr>
<td>Experience</td>
<td>.3773</td>
<td>.3830</td>
<td>.0544</td>
<td>.0721</td>
<td>.0028</td>
<td>1.0000</td>
</tr>
<tr>
<td>Experience-d</td>
<td>.3149</td>
<td>.3938</td>
<td>.0480</td>
<td>.0720</td>
<td>-.0039</td>
<td>.9464</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Mean VIF:
.vif

<table>
<thead>
<tr>
<th>Variable</th>
<th>VIF</th>
<th>1/VIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experience-d</td>
<td>9.70</td>
<td>0.103104</td>
</tr>
<tr>
<td>Experience</td>
<td>9.61</td>
<td>0.104041</td>
</tr>
<tr>
<td>Age</td>
<td>1.20</td>
<td>0.834333</td>
</tr>
<tr>
<td>Asian</td>
<td>1.05</td>
<td>0.956333</td>
</tr>
<tr>
<td>Black</td>
<td>1.04</td>
<td>0.961772</td>
</tr>
<tr>
<td>Female</td>
<td>1.02</td>
<td>0.984988</td>
</tr>
</tbody>
</table>

Mean VIF | 3.93
Test for Heteroskedasticity:
. `imtest, white`

White's test for Ho: homoskedasticity against Ha: unrestricted heteroskedasticity

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>chi2(22)</td>
<td>43.25</td>
<td>22</td>
<td>0.0044</td>
</tr>
<tr>
<td>Prob &gt; chi2</td>
<td>0.0044</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Cameron & Trivedi's decomposition of IM-test

<table>
<thead>
<tr>
<th>Source</th>
<th>chi2</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heteroskedasticity</td>
<td>43.25</td>
<td>22</td>
<td>0.0044</td>
</tr>
<tr>
<td>Skewness</td>
<td>14.13</td>
<td>6</td>
<td>0.0282</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>10.44</td>
<td>1</td>
<td>0.0012</td>
</tr>
<tr>
<td>Total</td>
<td>67.82</td>
<td>29</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

Correction for Heteroskedasticity:
. `regress logEarnings Age Female Asian Black Experience ExperienceSquared if HigherDegree ==1, robust`

Linear regression

Number of obs = 1334
F(  6, 1327) = 46.67
Prob > F      = 0.0000
R-squared     = 0.1890
Root MSE      = .6712

| logEarnings                  | Coef.   | Std. Err. | t     | P>|t| | [95% Conf. Interval] |
|------------------------------|---------|-----------|-------|-----|----------------------|
| Age                          | .0461481| .0123605  | 3.73  | 0.000 | .0218999 - .0703962  |
| Female                       | -.1761489| .0369727  | -4.76 | 0.000 | -.2486801 - -.1036176|
| Asian                        | .1302865| .0470437  | 2.77  | 0.006 | .0379983 - .2225748  |
| Black                        | .0416191| .0678912  | 0.61  | 0.540 | -.0915667 - .174805  |
| Experience                   | .23386  | .0264937  | 8.83  | 0.000 | .1818746 - .2858453  |
| ExperienceSquared            | -.0162666| .0031128  | -5.23 | 0.000 | -.0223731 - -.01016  |
| _cons                        | 9.054363| .375019   | 24.14 | 0.000 | 8.318669 - 9.790058  |
Appendix D: The Oaxaca-Blinder Decomposition Procedure Controlling for Experience

Statistics of Variables used in the Data Set:
```
describe Earnings logEarnings HigherDegree Age Female Asian Black Experience ExperienceSquared
storage   display    value
variable name   type    format     label      variable label
--------------------------------------------------------------
Earnings        long    %8.0g
logEarnings     float   %9.0g
HigherDegree    float   %9.0g
Age             byte    %8.0g
Female          float   %9.0g
Asian           float   %9.0g
Black           float   %9.0g
Experience      float   %9.0g
ExperienceSquared float   %9.0g
HRSWK           byte    %8.0g
```

```
summarize Earnings logEarnings HigherDegree Age Female Asian Black Experience ExperienceSquared HRSWK
Variable |       Obs       Mean    Std. Dev.     Min        Max
-------------+-------------------------------------
Earnings  |      4328      83631.72    78080.34          0     999996
logEarnings |      4316    11.08303    .7445572    3.78419   13.81551
HigherDegree |      4328    .5281885    .4992625          0          1
Age       |      4328    32.50416    2.896326         28         37
Female    |      4328     .370841    .4830858          0          1
Asian     |      4328    .1885397    .3911878          0          1
Black     |      4328    .0975046     .296678          0          1
Experience |      4328     6.78073    3.888762          0         20
ExperienceSquared |      4328    61.09727    58.54003          0        400
HRSWK     |      4328    55.59219    8.279316         48         96
-------------+-------------------------------------
```

Regression Run for Bachelor’s Degrees (Age 28-32):
```
regress logEarnings Age Female Asian Black Experience ExperienceSquared if HigherDegree ==0 & Age >27 & Age <33
```

```
Source |       SS       df       MS              Number of obs =    1063
-------------+-------------------------------------
Model |  43.844655     6    7.3074425           Prob > F      =  0.0000
Residual |  459.36352  1056  .435003333           R-squared     =  0.0871
Total |  503.208175  1062  .473830673           Root MSE      =  .65955
-------------+-------------------------------------
logEarnings |      Coef.   Std. Err.      t    P>|t|      [95% Conf. Interval]
-------------+-------------------------------------
Age         |  .0051468   .0189362     0.27   0.786    -.03201    .0423036
Female      | -.1939622   .0439331    -4.41   0.000    -.2801683    -.1077561
Asian       |  .1302177   .0612894     2.12   0.034     .0099549    .2504806
Black       | -.0302162   .0660595    -0.46   0.647    -.1598391    .0994066
Experience  |  .1838526   .0496704     3.70   0.000     .0863886    .2813166
ExperienceSquared | -.0074423   .0383292    -1.94   0.052    -.1049632    .090786
_cons       |   9.968688   .5935599    16.81   0.000     8.804978    11.1324
```
Summary Statistics for Regression Run for Bachelor’s Degrees (Age 28-32):

```
. summarize logEarnings Age Female Asian Black Experience ExperienceSquared if HigherDegree ==0 & Age >27 & Age <33

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>logEarnings</td>
<td>1063</td>
<td>10.952</td>
<td>.6883536</td>
<td>4.174387</td>
<td>13.81551</td>
</tr>
<tr>
<td>Age</td>
<td>1065</td>
<td>29.90704</td>
<td>1.400791</td>
<td>28</td>
<td>32</td>
</tr>
<tr>
<td>Female</td>
<td>1065</td>
<td>.313615</td>
<td>.4641799</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Asian</td>
<td>1065</td>
<td>.1286385</td>
<td>.3349567</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Black</td>
<td>1065</td>
<td>.1079812</td>
<td>.3105025</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Experience</td>
<td>1065</td>
<td>6.807512</td>
<td>1.988108</td>
<td>1</td>
<td>15</td>
</tr>
<tr>
<td>Experience~d</td>
<td>1065</td>
<td>50.29108</td>
<td>27.01284</td>
<td>1</td>
<td>225</td>
</tr>
</tbody>
</table>
```

Pairwise Correlation for Regression Run for Bachelor’s Degrees (Age 28-32):

```
. pwcorr logEarnings Age Female Asian Black Experience ExperienceSquared if HigherDegree ==0 & Age >27 & Age <33

<table>
<thead>
<tr>
<th></th>
<th>logEarnings</th>
<th>Age</th>
<th>Female</th>
<th>Asian</th>
<th>Black</th>
<th>Experience</th>
<th>Experience~d</th>
</tr>
</thead>
<tbody>
<tr>
<td>logEarnings</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>0.1353</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>-0.1200</td>
<td>-0.0447</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>0.0796</td>
<td>-0.0246</td>
<td>-0.0179</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>-0.0362</td>
<td>-0.0072</td>
<td>0.0517</td>
<td>-0.1337</td>
<td>1.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experience</td>
<td>0.2477</td>
<td>0.5676</td>
<td>0.0522</td>
<td>0.0485</td>
<td>-0.0455</td>
<td>1.0000</td>
<td></td>
</tr>
<tr>
<td>Experience~d</td>
<td>0.2295</td>
<td>0.6167</td>
<td>0.0514</td>
<td>0.0421</td>
<td>-0.0570</td>
<td>0.9776</td>
<td>1.0000</td>
</tr>
</tbody>
</table>
```

Mean VIF:

```
. vif

<table>
<thead>
<tr>
<th>Variable</th>
<th>VIF</th>
<th>1/VIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experience~d</td>
<td>26.04</td>
<td>0.038396</td>
</tr>
<tr>
<td>Experience</td>
<td>23.74</td>
<td>0.042116</td>
</tr>
<tr>
<td>Age</td>
<td>1.71</td>
<td>0.583511</td>
</tr>
<tr>
<td>Black</td>
<td>1.03</td>
<td>0.971961</td>
</tr>
<tr>
<td>Asian</td>
<td>1.02</td>
<td>0.976420</td>
</tr>
<tr>
<td>Female</td>
<td>1.02</td>
<td>0.983939</td>
</tr>
</tbody>
</table>

Mean VIF | 9.10
```

Test for Heteroskedasticity:

```
. imtest, white

White’s test for Ho: homoskedasticity against Ha: unrestricted heteroskedasticity

chi2(22) = 33.52
Prob > chi2 = 0.0549

Cameron & Trivedi’s decomposition of IM-test

<table>
<thead>
<tr>
<th>Source</th>
<th>chi2</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heteroskedasticity</td>
<td>33.52</td>
<td>22</td>
<td>0.0549</td>
</tr>
<tr>
<td>Skewness</td>
<td>5.35</td>
<td>6</td>
<td>0.4996</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>2.22</td>
<td>1</td>
<td>0.1359</td>
</tr>
<tr>
<td>Total</td>
<td>41.10</td>
<td>29</td>
<td>0.0675</td>
</tr>
</tbody>
</table>
```

158
Regression Run for Higher Degrees (Age 33-37):

```
. regress logEarnings Age Female Asian Black Experience ExperienceSquared if HigherDegree ==1 & Age >32 & Age <38
```

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>Number of obs = 1173</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model</td>
<td>90.7705891</td>
<td>6</td>
<td>15.1284315</td>
<td>F(  6,  1166) = 31.63</td>
</tr>
<tr>
<td>Residual</td>
<td>557.633024</td>
<td>1166</td>
<td>.478244446</td>
<td>Prob &gt; F = 0.0000</td>
</tr>
<tr>
<td>Total</td>
<td>648.403613</td>
<td>1172</td>
<td>.553245404</td>
<td>R-squared = 0.1400</td>
</tr>
</tbody>
</table>

```

Source | SS   | df  | MS       | Number of obs = 1173
-------|------|-----|----------|----------------------|
Model   | 90.7705891 |  6 | 15.1284315 | F(  6,  1166) = 31.63
Residual | 557.633024 | 1166 | .478244446 | Prob > F = 0.0000
Total   | 648.403613 | 1172 | .553245404 | R-squared = 0.1400

Adj R-squared = 0.1356
Root MSE = .69155

```

| logEarnings | Coef.  | Std. Err. | t    | P>|t| | [95% Conf. Interval] |
|-------------|--------|-----------|------|-----|----------------------|
| Age         | .0358405 | .0153297 | 2.34 | 0.020 | .0057636 -.0659174 |
| Female      | -.1863469 | .0419228 | -.45 | 0.658 | -.2685995 .1040944 |
| Asian       | .1317421 | .0482479 | 2.73 | 0.006 | .0370797 .2264046 |
| Black       | -.0438799 | .0698892 | -.63 | 0.530 | -.1810026 .0932428 |
| Experience  | .1694012 | .0199466 | 8.49 | 0.000 | .1302659 .2085365 |
| ExperienceSquared | -.0082886 | .0014512 | -5.71 | 0.000 | -.0111359 -.0054413 |
| _cons       | 9.422825 | .5371054 | 17.54 | 0.000 | 8.369024 .10.47663 |

```

Summary Statistics for Regression Run for Higher Degrees (Age 33-37):

```
. summarize logEarnings Age Female Asian Black Experience ExperienceSquared if HigherDegree ==1 & Age >32 & Age <38
```

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>logEarnings</td>
<td>1173</td>
<td>11.28614</td>
<td>.7438047</td>
<td>3.78419</td>
<td>13.65681</td>
</tr>
<tr>
<td>Age</td>
<td>1176</td>
<td>35.05952</td>
<td>1.417169</td>
<td>33</td>
<td>37</td>
</tr>
<tr>
<td>Female</td>
<td>1176</td>
<td>.3860544</td>
<td>.4870504</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Asian</td>
<td>1176</td>
<td>.2406463</td>
<td>.4276578</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Black</td>
<td>1176</td>
<td>.0960884</td>
<td>.2948379</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Experience</td>
<td>1176</td>
<td>6.582483</td>
<td>3.551446</td>
<td>0</td>
<td>17</td>
</tr>
<tr>
<td>ExperienceSquared</td>
<td>1176</td>
<td>55.93112</td>
<td>49.47012</td>
<td>0</td>
<td>289</td>
</tr>
</tbody>
</table>

Pairwise Correlation for Regression Run for Higher Degrees (Age 33-37):

```
. pwcorr logEarnings Age Female Asian Black Experience ExperienceSquared if HigherDegree ==1 & Age >32 & Age <38
```

<table>
<thead>
<tr>
<th></th>
<th>logEarnings</th>
<th>Age</th>
<th>Female</th>
<th>Asian</th>
<th>Black</th>
<th>Experience</th>
<th>ExperienceSquared</th>
</tr>
</thead>
<tbody>
<tr>
<td>logEarnings</td>
<td>1.0000</td>
<td>0.1428</td>
<td>-0.1280</td>
<td>0.0859</td>
<td>-0.0486</td>
<td>0.3055</td>
<td>0.2538</td>
</tr>
<tr>
<td>Age</td>
<td>0.1428</td>
<td>1.0000</td>
<td>-0.0765</td>
<td>0.0030</td>
<td>0.0270</td>
<td>0.3239</td>
<td>0.3565</td>
</tr>
<tr>
<td>Female</td>
<td>-0.1280</td>
<td>-0.0765</td>
<td>1.0000</td>
<td>-0.0968</td>
<td>0.0793</td>
<td>-0.0066</td>
<td>-0.0232</td>
</tr>
<tr>
<td>Asian</td>
<td>0.0859</td>
<td>0.0030</td>
<td>-0.0968</td>
<td>1.0000</td>
<td>0.0135</td>
<td>-0.0381</td>
<td>-0.0268</td>
</tr>
<tr>
<td>Black</td>
<td>-0.0486</td>
<td>0.0270</td>
<td>0.0793</td>
<td>-0.0381</td>
<td>1.0000</td>
<td>0.0135</td>
<td>-0.0385</td>
</tr>
<tr>
<td>Experience</td>
<td>0.3055</td>
<td>0.3239</td>
<td>-0.0066</td>
<td>0.0135</td>
<td>0.0135</td>
<td>1.0000</td>
<td>0.9581</td>
</tr>
<tr>
<td>ExperienceSquared</td>
<td>0.2538</td>
<td>0.3565</td>
<td>-0.0232</td>
<td>0.0268</td>
<td>-0.0385</td>
<td>0.9581</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

Mean VIF:

```
. vif
```

<table>
<thead>
<tr>
<th>Variable</th>
<th>VIF</th>
<th>1/VIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experience</td>
<td>12.60</td>
<td>0.079386</td>
</tr>
<tr>
<td>Experience</td>
<td>12.28</td>
<td>0.081401</td>
</tr>
<tr>
<td>Female</td>
<td>1.16</td>
<td>0.863119</td>
</tr>
<tr>
<td>Asian</td>
<td>1.04</td>
<td>0.958833</td>
</tr>
<tr>
<td>Black</td>
<td>1.04</td>
<td>0.959100</td>
</tr>
<tr>
<td>Female</td>
<td>1.02</td>
<td>0.979431</td>
</tr>
</tbody>
</table>

Mean VIF = 4.86
Test for Heteroskedasticity:
  . imtest, white

White's test for Ho: homoskedasticity
  against Ha: unrestricted heteroskedasticity

  chi2(22) = 18.69
  Prob > chi2 = 0.6646

Cameron & Trivedi's decomposition of IM-test

<table>
<thead>
<tr>
<th>Source</th>
<th>chi2</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heteroskedasticity</td>
<td>18.69</td>
<td>22</td>
<td>0.6646</td>
</tr>
<tr>
<td>Skewness</td>
<td>4.79</td>
<td>6</td>
<td>0.5704</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>1.65</td>
<td>1</td>
<td>0.1989</td>
</tr>
<tr>
<td>Total</td>
<td>25.13</td>
<td>29</td>
<td>0.6714</td>
</tr>
</tbody>
</table>
VIII. Bibliography


