EC 483/597 Midterm #1 Study Guide

The exam will consist of two essay style questions worth a total of 50% and ten short answer questions worth 50%. You may include graphs, intuition, and math as you see fit to answer each question. Note that it is not always the case that more writing is better. Be concise, be clear, and state your points. If you get to a question and feel that you don't know how to answer it, write whatever you know about the topic, which I am sure is more than you realize. Do not leave answers completely blank.

For the first half of the exam, the essay portion, I will provide you with four essay questions, and on exam day I will choose two of these four. The four essay questions are as follows:

1) In class we discussed some of the ideas within the field of welfare economics. Specifically, we talked about four schools of thought: the Utilitarian, or Benthamite, approach, the Rawlsian, or Maximin approach, the Nozickian approach, and finally commodity egalitarianism. Discuss some of the ideas of these schools of thought within welfare economics. Which school of thought do you find to be the most compelling and why? Be sure to include discussion of the equity-efficiency tradeoff and the idea of Arthur Okun's leaky bucket. What potential sources of leakage exist from the redistribution of wealth? How might we extend ideas of welfare economics to the market for medical care? Suppose an individual is unemployed, without food and shelter, and with no prospects for improvement. If you were a benevolent social planner and you had the opportunity to implement your preferred "school of thought," how would you handle cases such as this when individuals need emergency medical care, yet have no money with which to fund utilization? Would your system also provide preventative care to this individual?

2) The work of Michael Grossman in 1972 deepened our understanding of how we might theoretically model an individual's demand for health. Health is something that we can't simply go to the market and purchase, yet we all demand it. Discuss some of the ideas within Grossman's Model of Health Demand. What is the goal of the consumer within this model, and by what is the consumer constrained? Discuss the labor-leisure tradeoff within the model. One trait that makes the model unique is its treatment of health in that we don't only consume it, but also our actions today can heavily influence our health tomorrow. Hence, we all face a tradeoff between consumption today and investment in health for tomorrow. How might our health be influenced by other elements of life such as human capital and wages? Do you think good health leads to higher levels of human capital and higher wages, or do you think these latter factors lead to better health?

3) In class, we discussed empirical microeconomic techniques that are commonly used within health economics. Ideally, we would like to rely on controlled experiments to directly test all our hypotheses, however these types of experiments are typically not feasible. Hence, we often must rely on clever empirical techniques, aka identification strategies, in order to establish a causal relationship. In short, the goal of most identification strategies is to separate observations into two categories, one representing a treatment group and the other representing a control group. Though not ideal, this is often the best we can do to mimic a controlled experiment by using observational data. Briefly describe the experimental ideal, that is, the ideal controlled experiment that we would like to run in a perfect world. It may be helpful to use a real life example of a controlled experiment in your description. Why is a controlled experiment typically infeasible? Each of you are working with a specific paper for the end of semester course project. Briefly describe the empirical methodology/identification strategy employed in your group's paper. Does your paper attempt to establish a causal mechanism? Does it attempt to mimic a controlled experiment?

4) We all face the risk of adverse health outcomes. One way that we deal with uncertainty presented by this risk is by purchasing health insurance. There are several traits that characterize an ideal insurance system. Briefly discuss these traits and some of the issues such as moral hazard and adverse selection that make insurance such a complicated institution. Given all that we have discussed about insurance, and the many complications that arise both with and without its existence, do you believe that all members of society should be forced to have some form of health insurance? One of the provisions of our recent health care reform is an individual mandate requiring that everyone within the U.S. have health insurance. What was the purpose of this mandate, and do you think such a mandate is morally acceptable? Has the mandate been successful in achieving its intended task?

The remainder of the exam will consist of short answer/problem solving type questions. Here is non-exhaustive discussion of topics that we have covered:

Health Care Spending

In the U.S. we spend about twice as much per-capita on healthcare than the average of all other OECD countries. About one-third of our gross domestic product comes from health spending. This massive health care spending in the U.S. can be primarily attributed to spending in the later years of life. Despite this spending, our health outcomes are not necessarily better than that of the rest of the developed world. We actually have slightly lower life expectancy when compared to a number of other developed nations.

The data shows that life expectancy can vary both across gender and race. Females tend to outlive males, and blacks have the lowest life expectancy, on average, when compared to whites and Hispanics. Interestingly, Hispanics in the U.S. have a higher life expectancy than both blacks and whites. This is interesting, as Hispanics are considered to be a socio-economically disadvantaged group.

Welfare Economics and Arrow (1963)

Welfare economics is a branch of economics that use microeconomic techniques to evaluate fairness and well-being within society. When we use the term "welfare," we are referring to the aggregate well-being of society. Welfare economics is concerned with normative issues, as opposed to positive issues, in that it studies what "should be" as opposed to what "is." A key concept of welfare economics is the equity-efficiency tradeoff. When we refer to efficiency, we are referring to the idea of Pareto efficiency. A Pareto efficient outcome is one in which no one party can be made better off without harming another party. Note that a Pareto efficient outcome need not be unique, and need not be socially-acceptable. For example, I could have all the money and everyone else in the class has zero, this is a Pareto efficient outcome, though incredibly unequitable. The equity-efficiency tradeoff is a tradeoff that we face when determining the socially optimal outcome. Of course we want the most efficient outcome, i.e. the one that maximizes societal well-being, but we also are concerned with having an equitable outcome, or one that is fair. An analogy developed by Arthur Okun known as "Okun's leaky bucket" pretty well describes the equity-efficiency tradeoff. Consider a situation in which we seek to redistribute wealth by taking money from the wealthy, putting the money into a bucket, and then we deliver the bucket to the poor. According to Okun, the problem is that we have leaks in the bucket, and when we transfer a dollar to the poor, we lose some amount of that dollar through inefficiency. When you tax the income of the rich, they may have the incentives to work less, and this will result in a societal inefficiency due to decreased production. Moreover, when you give money to the poor through welfare programs, the poor may have the incentive to shirk at the labor market, further worsening productive inefficiency. Moreover, this potentially provides the poor with an incentive to remain in a worser economic state. These are the sources of leakage that Okun described in his analogy.

We discussed the idea behind the social welfare function. The social welfare function relates inputs (some measure of individual utility) to output (some aggregate level of utility). We might consider the role of a social planner (a benevolent leader of society) in attempting to achieve a maximum level of social welfare. We discussed a number of different "schools of thought" of social welfare, each with its own unique set of beliefs. For example, under a Utilitarian, or Benthamite, social welfare function, we might simply add up all of the societal members' levels of utility. This is a linear social welfare function. Alternatively, under the Maximin, or Rawlsian social welfare function, society is only as well off as its worse-off member. In this case, the functional form might be a min( ) function (recall the min function is L-shaped). A third system of beliefs is known as Nozickian social welfare. Under this belief system, it does not matter who has what in society, however there should be an equal distribution

of opportunities. In other words, I can have all the money and you can have zero, as long as we each had a fair opportunity in the beginning. Finally, we discussed commodity egalitarianism, a system of beliefs in which there should be some minimum or some floor of living standard. As long as the poorest of society is above this bare minimum standard, it doesn't matter how much inequality occurs beyond that. There may be significant levels of disagreement, however, in what comprises the bare minimum (should it include clothing, shelter, etc.). Finally, we discussed the idea presented by John Rawl's called the "veil of ignorance." In this analogy, Rawls claimed that the only way to redistribute assets in a fair way is to consider the idea that you are yet to be born. With a fifty percent probability you will be born into a very wealthy family, and with a fifty percent probability you will be born into a homeless family living on the streets. Would you be willing to take that gamble? Most would say no due to risk-aversion. Hence, according to Rawls, the only way to have a fair and just world is to have at least some level of wealth equality.

In class we discussed the extent of wealth and health care access inequality in the United States. According to the gini-coefficient, a measure of income inequality, the U.S. has more extreme income inequality than most other nations within the developed world. Additionally, according to a stream of health economics literature, the U.S. has extreme health care access inequality, with the wealthy class utilizing health care at much higher rates than the poor. Health and wealth inequality are two things that we should keep in mind when discussing ideas from welfare economics.

We discussed the Two Fundamental Theorems of Welfare Economics. The first fundamental theorem states that given an initial distribution of wealth, a perfectly competitive market will always result in a Pareto efficient outcome. The second fundamental theorem states that through a redistribution of initial wealth, and under a perfectly competitive market, we can achieve both a socially-acceptable and an efficient outcome. Kenneth Arrow in his 1963 paper discusses the market for health care within the context of welfare economics. According to Arrow (1963), the two fundamental theorems do not hold when it comes to health care because health care deviates substantially from a perfectly competitive market. The assumptions of perfect competition are that there are many firms, producing an identical product, in a market with zero barriers to entry. In the market for health care, there are not necessarily "many firms," there are actually few firms. This is particularly the case in the pharmaceutical industry (consider Epipen as an example). Firms do not sell identical products (consider the many slight variations of prescription drugs). There are very steep barriers to entry (medical professionals have to go to school for quite a long time, and to start a hospital one would need approval from various government bodies). Moreover, health care providers are not solely profit-maximizers in that they also presumably care about the well-being of their patients. There are three other factors that further complicate the market for health care: uncertainty, information asymmetry, and externalities. In health care, there is extreme uncertainty in many cases. For example, in diagnosing a complicated health condition there is uncertainty, as well as with the efficacy of treatment. Informational differences between the health care provider and the patient are extreme, however this knowledge gap has been lessened in recent years due to the internet and free-flow of valuable information (e.g. WebMD). Finally, externalities play a role in medicine (for example, when I smoke it negatively effects people through secondhand smoke, and when I

get a vaccination it positively affects others through protection from disease). These additional factors further complicate the market for health care and cause it to deviate from the standard framework of perfect competition. Hence, the two fundamental theorems of welfare economics may not hold in the case of medical care.

Health Production and the Demand for Health

In this section of the course we began to discuss how we can introduce some of the theoretical concepts from micreconomics and apply them to health. We begin our discussion by recalling the standard production function from intermediate microeconomics. The production function relates inputs (typically units of capital and labor) to output (some product produced by a firm). The most commonly used production function within economics is the Cobb-Douglas production function. Using the standard Cobb-Douglas functional form, we were able to calculate the marginal product of labor (i.e. the additional output from increasing labor inputs by one unit) and the marginal product of capital (i.e. the additional output from increasing capital inputs by one unit). Moreover, we were able to show whether a production function exhibits constant, increasing, or decreasing returns to scale. With constant returns to scale, doubling inputs leads our output to double. With increasing returns to scale, doubling inputs leads our output to more than double. Finally, with decreasing returns to scale, doubling inputs leads our output to less than double. We relate returns to scale with a concept known as functional homogeneity. In short, if a production function exhibits constant returns to scale, we say that it is homogeneous of degree 1. This is a standard assumption made in a lot of theoretical models. We discussed two other versions of production function: the linear production function used when inputs are perfectly substitutable, and the min production function used when inputs are perfect complements. Note that these functional forms also apply to utility functions and, hence indifference curves.

We discussed the standard assumptions of the production function. Assumption 1, inputs are finite and non-negative. In other words, we have a limited amount of input resources, and we cannot have negative inputs to production. Assumption 2, in order to get some output, we must use some positive amount of inputs. This is related to the idea that something cannot come from nothing, i.e. we cannot have a “free lunch.” Assumption 3, output is monotonically increasing with inputs. As we increase inputs, output must increase, or at least be non-decreasing. Assumption 4, the production function is continuous and twice-differentiable. Finally, Assumption 5, the production function is quasi-concave to the origin. This assumption is related to the idea of diminishing marginal productivity, i.e. our first installment of inputs leads to large increases in output, while later installments lead to lesser output gains.

Within the framework of health, we talked about how we might model health production. Our output is some measure of health status, for example healthy days, life expectancy, or mortality, while our inputs are medical care, lifestyle, genetics, environment, and perhaps luck. To simplify our problem, we assumed that the only input to health production was medical care, and we plotted a concave health production function with health status on the y-axis and medical

care consumption on the x-axis. Note that this production function does not start from the origin, and instead begins from some positive amount of health status. This is due to the fact that many people achieve a positive health status with zero consumed units of medical care. Our production function of health assumes that health status is strictly increasing in medical care, and it is increasing at a decreasing rate (it is concave). While looking at the health production function, we discussed the idea of "Flat-of-the-Curve" Medicine, i.e. the idea that we as a society are spending more and more on medical care, but we are not observing improvements in health status. This idea would imply that we are on the "flat-of-the-curve" of the health production function.

If we are at the flat-of-the-curve in terms of medical spending, then we should redirect resources away from medical care and perhaps into more productive inputs such as education. We discussed two pieces of literature, Schoder and Zweifel (2011) and Whaley (2016) that study both whether we are on the flat-of-the-curve, and if we are, if medical spending need be completely wasteful. Schoder and Zweifel (2011) study whether increased health spending leads to declines in the variance of death age, or the variation of age at death, within a society. The idea is that even if we do not improve the mean age of death through health spending, perhaps this spending lowers the amount of dispersion about the mean, or the variance, of the death age. According to this study, countries that spend more on health care tend to have a lower variation of age at death, so a larger proportion of members of society are dying near the mean death age. This would indicate that perhaps we are not yet at the flat-of-the-curve in terms of health care spending. A similar study by Whaley (2016) analyzes this question beyond simply the mean and variance by looking at the skewness of the health outcomes distribution. Whaley (2016) finds evidence that we are at the flat-of-the-curve in terms of spending, however additional spending may not necessarily be wasteful. For example, perhaps by going to the doctor more, people obtain some non-tangible source of utility in the form of "peace of mind." Moreover, perhaps by going to the doctor for preventative care, we are able to lower the probability of extremely bad health outcomes, and hence shrink the tail of the bad end of the health outcomes distribution.

Using the health production function, we introduced the Four Quadrant Model created by Wagstaff (1986) in Quiz #1. The four quadrant model allows us the visual framework to model consumer choice with the inclusion of a health production function. Consumers seek to maximize utility from the consumption of health and some vector of composite goods. Note that the consumer consumes health and does not gain utility from consumption of medical care. The consumer chooses levels of medical care and composite good subject to some budgetary constraint, and she uses medical care as an input in the production of the good health. In Quadrant 1, we display the consumer's indifference curve, in Quadrant 2, the health production function, in Quadrant 3, the consumer's budget constraint, and in Quadrant 4, we display a 45-degree line simply used as a tool to relate units of goods from quadrant 3 to units of good in quadrant 1. Beginning from the consumer's budget constraint in Quadrant 3, we are able to calculate all the feasible bundles of medical care and composite goods. Then, mapping horizontally over to the 45-degree line and then vertically to the indifference curve, while simultaneously mapping vertically to the health production function and horizontally over to the indifference curve, we are able to pinpoint which utility level the consumer will be at given a

certain combination of goods. Of course, whichever combination of goods puts the consumer on the most "far out" indifferent curve is the consumer's optimal choice.

We more formally introduced the demand for health by discussing the famous Grossman (1972) model. In this dynamic, finite horizon model, consumers gather utility from the consumption of health and some amount of home goods, i.e. all the other enjoyments in life. The consumer can produce both health and home good. Health is produced using medical care as an input as well as by using a time input, i.e. taking the time to go to the gym or to plan a nutritious diet. Home goods are produced by using market good inputs and additional time inputs. One example of a home good may be laundry, where we get utility from having clothes that are fresh and clean. To produce these clothes, we must use market good inputs, i.e. laundry detergent and fabric softener, and we must use time inputs, i.e. us actually taking the time to put the laundry in the washer, dryer, and finally ironing and folding the clothes. A second example of a home good may be watching Alabama football. Output is the utility that we gain from watching Bama win, while the inputs are a ticket to get into the game, and a time input capturing how long it takes for us to travel to the game.

Just like in any standard consumer optimization problem, in the Grossman (1972) model, the consumer seeks to maximize utility subject to some constraints. The consumer is constrained by a budget, where the consumer earns a wage rate for each hour spent working, and the consumer must exhaust all lifetime earnings by purchasing either medical care or market good inputs. Not only is the consumer constrained by budget, but they are also constrained by time. With each unit of time, they must decide whether to go to work at the labor market, take sick days at home away from work, or enjoy leisure activities through health and home good production. In each period, the consumer makes a labor/leisure decision and a consumption/investment decision. The summation of each period represents the consumer's lifetime decision path. All decisions are made at the beginning of time, i.e. this is a perfect foresight world.

After quite a lot of difficult math, the Grossman (1972) model comes to a few basic conclusions. First, if we assume that the rate of depreciation of health stock is increasing across the life cycle, i.e. as people get older their health deteriorates more quickly, then the shadow price of health rises with age. Here, the shadow price refers to a price for something that is not actually tangibly purchased at a market. The second result of the model is that the shadow price of health falls with education. In other words, more educated people are more efficient producers of health. The third result is that under certain conditions, an increase in the shadow price of health may simultaneously reduce the quantity of health demanded and increase the quantity of medical care demanded. This result can in a sense be related to what we are observing in the real world, with a majority of our health care spending coming from later-life spending.

Some criticism of the Grossman (1972) as outlined in Laporte (2015) includes the fact that current health behavior does not depend on past health behavior, the model does not predict that health is related to socioeconomic status (quite the contrary to all of the evidence supporting this phenomenon), and the model does not preclude an individual from living forever. There is an entire literature devoted to criticizing the Grossman (1972) model. Just like any economic

mode, whether theoretical or empirical, the Grossman (1972) model is far from perfect. Some of the insights gained from the model, however, have proven to be quite valuable over forty years after its creation.

Empirical Microeconomics

In this section of the course we deviated from economic theory and briefly outlined some of the popular empirical methods used in health economics. We began by discussing the difference between a theoretician and an empiricist. A theorist will claim: "Health production depends on medical care." In this claim, medical care is literally the only thing influencing health production in the model. In empirical economics, we introduce what is known as the disturbance term, or the error term. An empiricist might claim, "Health production depends on medical care and everything else that we are not controlling for, i.e. the disturbance term." By including this term, we admit that our model is not perfect, and we allow ourselves a bit of error within the estimation.

We discussed the goal of econometrics and the econometrician's concern in the breakdown of the assumptions underlying classical linear regression. Ordinary Least Squares (OLS) is the most commonly used estimator in economics, and it is the standard for estimating the classical linear model. There are a few standard assumptions of classical linear regression. First, the model is linear in its parameters and it is correctly specified. Second, the matrix of explanatory variables X must be of full rank, i.e. we must have at least as many observations as explanatory variables. Third, explanatory variables must be exogenous, i.e. the model must not suffer from endogeneity. Finally, the error term is assumed to be independently and identically distributed with a mean of zero and some constant variance.

The concern of endogeneity is the most important of all of the assumptions of classical linear regression. Endogeneity can occur through a number of different channels. First, if the explanatory variable is correlated with unobservables, i.e. if the explanatory variable is correlated with the error term. This is the textbook case of endogeneity. Second, perhaps we might have reverse causality, or simultaneity. In other words, maybe our outcome variable y is affecting our explanatory variable x instead of vice-versa. Often times, in order to solve problems of endogeneity, we have to resort to clever empirical techniques. These tools to solve endogeneity are known as identification strategies. In empirical microeconomics, fixing problems of endogeneity is the name of the game.

In class we discussed several regressions with ADHD diagnosis on the left hand side as a dependent variable. For example, in one model we wanted to control for everything that might affect ADHD diagnosis, so on the right hand side we included sex, age, race, total family income, etc. We discussed that if a p-value is below 0.05, the coefficient is considered significant at the 95% level. Moreover, if a p-value is below 0.01, then significance at the 99% level is achieved. That is, we can be 99% certain that the coefficient is statistically different than zero. We showed in a model that income and mother's education are negatively correlated with ADHD diagnosis of a child, indicating that as income goes up or the mother achieves education

at a higher level, the probability that he child is diagnosed with ADHD decreases. Alternatively, we showed that male children, older children, and white and black children are more likely to be diagnosed with the condition, as coefficients associated with these variables were positive and significant.

It is important to note that when we simply run OLS regressions like we did in the model discussed above, we are merely reporting correlations as opposed to inferring causality. To infer causality, we must use more sophisticated econometric techniques and we must formulate an identification strategy. We discussed that in an ideal world, we would love to be able to construct controlled experiments to test our hypotheses. In reality, however, controlled experiments are costly, take a long time and a lot of money to implement, and hence are usually not feasible. Thus, we rely on observational data and the formulation of identification strategies. Typically, we seek to find some form of a discontinuity, or we rely on a natural experiment in our best efforts to mimic a controlled experiment. Here, a natural experiment might be some form of exogenous policy change. We discussed some of the commonly used identification strategies in health economics including fixed effects, difference-in-differences, instrumental variables, and regression discontinuity. Ultimately, these techniques seek to separate one group of observations into the Treatment Group and another comparable group of observations into the Control Group. Hence, in most cases we are attempting to mimic a controlled experiment. One example of this that we discussed is with certain educational policies in a time period surrounding the No Child Left Behind Act of 2001 (NCLB). In these state-level policies, teachers or administrators were often provided with financial rewards if their students did well on standardized tests. Given that about half of ADHD diagnoses are first recommended by the teacher (Sax and Kautz, 2003), increases in ADHD diagnosis may have been an unintended consequence of these school accountability policies. For example, if I am a teacher that may be eligible to receive a reward for my students' performance, then I may have the incentive to recommend that a child has ADHD, particularly if this child is on the margin of ADHD diagnosis. Hence, we showed that we can separate observations into two groups: a Treatment group-i.e. children living in states that had rewards programs, and a Control Group-i.e. children living in states that did not offer such rewards. Then, interacting the policy variable with a time dummy, we can construct a difference-in-differences model showing the effect of a child living in a treated state on probability of the child receiving an ADHD diagnosis. Typically, we want to include a control for a pre-policy difference on the right hand side, as well as a control representing the post-policy treatment effect. Ideally, we would like to see that the pre-treatment dummy is insignificant, while the post-treatment dummy is significant.

Other identification strategies that we discussed are fixed effects estimation, instrumental variables estimation, and regression discontinuity design. Fixed effects estimation is used when one is concerned about unobserved heterogeneity. For example, if one is concerned with unobservables varying within states, one might include a state fixed effect. Alternatively, if the concern is with unobservables that vary across time, one might include a time fixed effect. In its simplest sense, state fixed effects are simply an added control for the state in which an individual resides. Similarly, a time fixed effect is simply an added control for the time period of an observation of data. For example, if we have a panel setup with ten years, we might include a

dummy variable on the right hand side for each of these years. This is a time fixed effect. Instrumental variables estimation (IV) is used when one is concerned that the explanatory variable of interest is correlated with the error term (i.e. endogenous). There are two requirements for a good IV. First, that it is highly correlated with the endogenous explanatory variable. Second, that it is uncorrelated with unobservables, i.e. it does not itself suffer from endogeneity. Finally, we briefly discussed regression discontinuity design, a method of analyzing observations that lie on either side of some defined cutoff. For example, we might look at school start dates compared to when a child turns five years old. Those born right before the cutoff must start kindergarten as the least mentally developed of their cohort. Those born right after the cutoff are not yet five years old when the school year starts, and hence must wait an entire calendar year to start school. Two articles, (Evans et al, 2010; Elder, 2010) use regression discontinuity design techniques and find that children born right before the school start date cutoff are significantly more likely to be diagnosed with ADHD than those born right after the school start date cutoff. This is potentially due to the idea that the younger children are less mentally developed when compared to their peers, and hence the teacher is most likely to point out behavioral problems within this subset of children.

Health Insurance

In this section of the course we began to discuss some of the theory underlying the supply of and demand for health insurance. We discussed the concept of risk-pooling. That is, we all face some risk of adverse health outcomes, so we might pool our money together and all take on the risk as one big group. That is the idea of insurance. There are four desirable qualities of insurance: the number of insureds is large and each person faces risk independently of one another, losses covered should be definite in time, place, and amount, the probability of a loss is measurable, and losses should be accidental. We discussed the idea of risk-aversion, that is, we would rather avoid risk if at all possible, even if it means paying a little bit more. Note that a risk averse utility function is one that is assumed to be concave to the origin. That is, we experience decreasing marginal utility of income. In Quiz #2 we showed an individual consumer's willingness to pay for insurance. We first calculated the individual's expected income and expected utility, composed of income and utility from the good state and income and utility from the bad state. Using the individual's expected utility, we were able to back out an amount of income representing the individual's certainty equivalent. The certainty equivalent is the amount of income that the individual would be willing to receive in order to completely avoid any risk. Despite the fact that the CE is less than expected income, the consumer would prefer the CE due to the well-being obtained from the avoidance of risk altogether. We were able to graphically depict the consumer's willingness to pay for insurance, which is a function of the expected loss and the risk premium. The risk premium is the additional amount that the insurance company can charge on top of the expected loss due to the consumer's preference to avoid risk. You might also consider the risk premium to contain loading fees or administrative costs associated with the insurance industry. Insurers must charge a little bit more than the consumer's expected loss to help the firm break even and continue to operate in the long run.

We formally discussed how to model the demand for and the supply of health insurance. In order to model a consumer's demand for health insurance, it is appropriate to gauge what might influence an individual's willingness to pay for insurance. In order to analyze how much insurance a consumer might want to purchase, we relied on the concepts of marginal costs and marginal benefits. Within the insurance framework, we defined marginal benefits as the expected marginal utility derived from the payout of an insurance policy. Similarly, we defined marginal costs as the expected marginal utility that the individual must give up in order to pay the insurance premium. In other words, the marginal cost here is a form of opportunity cost. By giving the insurance company a premium, the consumer cannot consume other goods using this money. It is important to note that with insurance, a person will only achieve benefits in the event of a bad state, and a person will only realize costs in the event of a good state.

We provided an intuitive background for why the marginal costs curve is upward sloping, and why the marginal benefits curve is downward sloping. In the case of marginal benefits, for each unit of insurance coverage, an individual derives some marginal benefit. For the first unit, for example, let's say the benefit is $500. For the second unit, let's also say the benefit is $500. It is the case, however, due to the law of diminishing marginal utility of income, that the consumer does not gain as much from the second unit as she did from the first. The same holds for each successive unit, and hence the marginal benefits curve is downward sloping. A similar argument can be made for the marginal costs curve. For the first unit of insurance, let's say the cost is $50. For the second unit, let's also say the cost is $50. With the second unit of insurance, however, the consumer is poorer than they were during the purchase of the first unit. Hence, the level of disutility, or the marginal cost, is higher in all successive purchases of insurance units, and the marginal costs curve is upward sloping. Again, the marginal costs of insurance are foregone utils that cannot be obtained through the consumption of other goods.

We showed graphically that the consumer will choose the amount of insurance at the point in which the marginal costs curve crosses the marginal benefits curve. We analyzed changes such as increased premiums, increases in expected losses, and improvements in wealth. If premiums increase, the marginal costs curve shifts upward and the marginal benefits curve shifts downward. This results in a lower optimal level of insurance coverage. Intuitively, this makes sense, as if insurance costs more money I'll purchase less of it. If expected losses increase, the marginal benefits curve shifts upward, and the consumer's optimal choice of insurance is higher. Intuitively, this makes sense as if my expected losses are higher, I prefer to purchase more insurance to shield myself from these potential losses. Finally, if my income increases, the marginal costs curve and the marginal benefits curve both shift downward. The size of the shift of each curve will determine whether I demand more or less insurance. To sum up the demand for insurance, both increases in the probability of loss and the size of loss leads us to demand higher levels of insurance. Moreover, since the marginal utility of income declines as income increases, the poor tend to demand larger amounts of insurance. Intuitively, if I make $750,000 in annual income, and I face a probability of incurring losses of $10,000, I am not too concerned about these losses and hence see little need to insure. Alternatively, if I make $20,000, then $10,000 is a substantial proportion of my income, and I have every incentive to purchase insurance. We completed our analysis of the demand for insurance by constructing a consumer

optimization model in which the consumer maximizes her expected utility by choosing some level of insurance benefits. Ultimately, the equilibrium concept of marginal costs equal to marginal benefits indicates that the consumer will choose the level of insurance benefits such that her MC=MB.

In order to model the supply of insurance, we assumed that the insurance industry is a perfectly competitive market. In perfect competition, long-run economic profits within the industry must be zero. If economic profit is greater than zero, then additional firms will enter the market, undercut the price, and this phenomenon will continue to occur until profits reach zero. Firms are willing to continue operating despite zero profits because these are not accounting profits, these are economic profits (profits after everyone working for our company receives a salary). Hence, it is beneficial for firms to continue to operate in the long run if they are making zero economic profits. We defined a firm's profit function within the insurance industry as premiums less overhead and payouts. Firms earn revenue through the form of premiums paid by insured individuals, firms offer payouts to these individuals in the case of bad outcomes, and firms pay for overhead/administrative costs to operate on a day-by-day basis. We showed that the premium charged by a firm can be tied directly to the probability of loss, and we used this relationship to show that in order for the equilibrium condition from the model of demand to hold, individuals must purchase full insurance. All of this is, of course, under the assumption of perfect competition, and under an additional assumption that administrative costs go to zero.

After modeling supply of and demand for insurance, we discussed two issues of imperfect information within health insurance: moral hazard and adverse selection. Moral hazard is the incentive for people to engage in undesirable behavior due to the fact that they are being imperfectly monitored. For example, if I have full health insurance coverage, maybe I am more likely to live carelessly and not take care of my health (as I know the insurance company will pay all of my medical bills). Though moral hazard is believed to be less of a problem with health insurance when compared to automobile or home insurance, it still can become a problem. Some potential solutions to problems of moral hazard are the use of deductibles or coinsurance. With these tools, the insurance company is able to deter people from utilizing health care in unnecessary situations. For example, if I have a high deductible, then I will only go to the doctor in completely necessary health situations. This might lower some of the "wasteful" health spending that we have become concerned with in the U.S. A second issue within health insurance markets that we discussed is adverse selection. It is common knowledge that sick people prefer to have insurance when compared to healthy people. Adverse selection is a problem that occurs when the seller of a good knows more about the good than the buyer. Within the context of health, the individual that is being insured certainly knows more about his or her health than the insurance company. This is the source of information asymmetry. The insurance company can only form beliefs about those that are insured, and using these beliefs it can charge a premium that reflects expected losses. This premium is likely higher than the average medical spending of a healthy individual, and hence healthy people may have the incentive to drop out of the market and not purchase insurance. This phenomenon is called adverse selection. Here, sick people are "adversely selecting" into health insurance markets. In some cases, when you go to purchase health insurance, you may be asked to fill out an extensive survey of health behaviors,

personal information, etc. This survey is a form of screening, and this is a tool commonly used by insurers to combat adverse selection. More recently, after the passage of the Affordable Care Act (ACA), the government mandates that everyone in the U.S. have some form of health insurance. This individual mandate was put into place in an effort to solve the adverse selection problem. The idea is that by forcing everyone to purchase health insurance, including healthy people that otherwise would not purchase it, we can get around the adverse selection problem. A question of morality, however, is whether it is acceptable for a government to force individuals to purchase a good that is typically considered to be a market good. Moreover, can all of these new insureds actually afford this insurance? These are questions that we will address later in the course when we formally discuss the ACA.

Finally, we concluded our discussion of health insurance theory by covering another topic related to adverse selection and incomplete information, the model of Rothschild and Stiglitz (1976). In this model, there are two types of consumers, high-risk consumers and low-risk consumers. Each derive utility solely through wealth, however they are each subject to some risk that they will lose a portion of wealth due to illness. The high-risk individuals face a larger risk of loss than those that are low-risk. To shield themselves from this risk, consumers purchase an insurance contract a = {a1, a2}, where a1 is the premium paid to the insurance company and a2 is the payout received in the event of a bad state. The consumer, as usual, seeks to maximize expected utility by choosing some level of insurance. We depicted this model graphically in a two-good world with wealth in the healthy state on the x-axis and wealth in the sick state on the y-axis. Using concepts analogous to the consumer's budget constraint and indifference curves, we were able to show a "tangency condition," i.e. an optimal choice of insurance of the consumer. We showed that is a consumer insures along the 45-degree line, this is a case of full insurance as wealth in the healthy state is equal to wealth in the sick state. Moreover, we constructed a contract line for the insurance company that is analogous to the consumer's budget constraint from microeconomic theory. Anywhere along this contract line, the insurance company is breaking even by offering an actuarially fair rate (assuming that loading fees are zero). By using a concept related to the tangency condition from intermediate micro, we were able to show that the optimal choice of insurance occurs at the point in which the consumer's indifference curve lies tangent to the insurance company's contract line, i.e. the marginal rate of substitution is equal to the price ratio (the price ratio we show to be equivalent to (1-p) / p, where p is the probability of sickness.

We looked at three different situations: a situation with only one type of consumer, a full information situation with both low- and high-cost consumers, and an incomplete information situation with both low- and high-cost consumers. With only one type of consumer, we showed that the consumer will purchase full insurance, i.e. the optimal point of tangency occurs directly at the 45-degree line. With complete information and two types of consumers, i.e. a situation in which the insurance company can directly observe which consumers are of which type, we once again showed that each type of consumer will fully insure, with each optimal point occurring directly at the 45-degree line (the low-risk individual of course purchasing full insurance at a cheaper price). Finally, we considered the situation in which there is incomplete information, i.e. the insurance company cannot distinguish between the two types (this is perhaps the most

realistic in the real world). We showed that with incomplete information, the high-risk individual will have the incentive to misrepresent her type and fully insure at the previous low-risk level. This is not surprising, as everyone would prefer a cheaper insurance contract. We showed one way that the insurance company may be able to deal with this situation. By offering two different contracts, the insurance company may be able to get both types of consumers to insure at appropriate levels. Contracts are structured such that everyone purchases some form of insurance, and neither type of consumer has any incentive to deviate from their own type's contract. In other words, the two types of contracts must be incentive compatible. To achieve a situation in which neither type has the incentive to deviate, the insurance company must first offer the same full-insurance type contract to the high-risk individual. Secondly, the company must offer some level of partial insurance to the low-risk individual, at a level such that this low-risk individual prefers it to being uninsured, and at a level such that the high-risk individual does not prefer to deviate away from his own full-insurance contract. Due to the shape of each type of consumer's indifference curve, the insurance company is able to offer a very particular contract that meets these requirements. Hence, the idea stemming from Rothschild and Stiglitz (1976) is that it may be ideal in the real world for high-risk individuals to obtain full insurance and low-risk individuals to obtain partial insurance. Moreover, a world in which insurance companies can only charge a single premium to all insureds may not be ideal.