[SQUEAKING]

[RUSTLING]

[CLICKING]

JON GRUBER: Today we're going to continue our tax discussion talking about chapter 20. Remember, the first three chapters of tax are really about tax theory. Chapter 18 was about thinking about how we define the tax base. Chapter 19, which we went through rather quickly, was about tax fairness and the topic of what economists call tax incidence. Now, we turn to the topic of tax efficiency.

Tax efficiency is in some sense, the more fun topic because it really is the ultimate place we see the difference between intended and unintended effects that we discussed in chapter 1. And to see that, there's no better place to start than about halfway through your packet, figure 20-3. Let's talk about the window tax. What is the window tax? So in 1696, English King William III needed to finance a war with France. He wanted to tax people who had the ability to pay taxes in order to raise money.

So how do you measure that in 1696? We don't have W-2 forms in 1696. Well, King William decided that a good sign of who was wealthy was how many windows they had in their house. That the more windows you had in your house, the wealthier you were. Glass was actually really expensive back then. If you go tour ancient Europe, you'll see windows were often made at the bottom of wine bottles because it was so hard to make glass, to get glass. So if you windows, that meant you're wealthy.

So basically he said, we're going to tax people based on how many windows they have in their house. The problem that is an easy tax to avoid by either boarding up your windows or building houses with no windows. So it was an easy-- when you tax people based on how many windows they had, he set up a tax that was rather easy to avoid. Moreover, the problem is when you avoid that tax by boarding up your windows, in 1696, it's incredibly unhealthy because basically you don't have fresh air coming into your house.

Indeed, Charles Dickens wrote about this. There was a typhus epidemic in 1781, in Carlisle, which was traced to a house inhabited by six poor families. In order to reduce the window tax, every window was built up and the source of ventilation was thus removed. The smell in this house was overwhelming and offensive to an unbearable extent. And 52 people died from typhus starting in this house.

So we have here is an example of what we call the inefficiency of tax avoidance of people trying to avoid taxes based by undertaking other activities that could be inefficient. Now, how do we know people responded? Well, this is a very clever empirical study. There's a study where these economists actually collected data on the number of windows in every house in a large section of England in the 1700s.

And they looked at the data on how many windows houses had. And they noticed that the way the window tax worked was if there was fewer than 10 windows in the house, there was a tax of 6 pence per window. But once you got to the 10th window, the tax went to 9 pence per window, including the first nine. So the ninth window was taxed at 6 pence. The 10th window was taxed at 100 pence, because you paid-- or no, 46 pence, because you paid an extra $0.04 on all the previous nine windows, plus the $0.10 for the 10th window.

And if you look at the data in figure 20-3, what do you see? A huge spike in the number of windows and houses at nine windows. Indeed, there was another notch at the 15th window. The tax went up higher. What do you see? Another spike in the number of windows at 14 windows. So basically, this is evidence of people avoiding the tax. Indeed, what's really cool is they then changed the tax where later, the big jump in tax start at the eighth window.

And so flip to figure 20-4, there became a huge spike at seven windows. This is not unusual. There is a history of people doing things to avoid taxes. Any have you ever heard of Tuscan bread? Tuscan bread is saltless bread. It's a term for saltless bread. Why is that called Tuscan bread? It's called Tuscan bread because in-- I got to get the year right --in the papal state of Tuscany, I don't have the year, the pope imposed a tax on salt. So they started making bread without salt.

If you visit many developing countries, you will see what looks like an unreasonable number of unfinished buildings. If you've been to other countries, you'll notice buildings often look like the top floor is unfinished. That's because many countries have taxes on finished buildings. So they do this. They just build buildings with an extra floor that's unfinished and avoid the tax on finished buildings.

And one of my favorite examples is there's a tax on shoes in the US that's 40%. The tax on slippers is 3%. What is interesting, shoes and slippers, it's basically how the bottom, what the bottom looks like. So actually, Chuck Taylor Sneakers by Converse have a thin layer of felt on the bottom that wear away after you use them a few times, but because of that, they can be classified as slippers, not sneakers, and pay lower tax.

The point of these fun examples is that basically, people can do lots of things to avoid taxation. And the fundamental point of this chapter is that those efforts they take to avoid taxation impose efficiency costs on society. So to see this, let's go back in figure 20-1 about halfway through your packet. Let's go back to Figure 20-1 to our tax on gasoline we had last time we talked about tax fairness.

Remember, we were in initial equilibrium with 100 billion gallons of gas being sold at $1.50 a gallon. The government comes in and taxes the suppliers of gas $0.50 a gallon. What did that do? That shifted the supply curve up by $0.50, moving our new equilibrium from point A to point B, with a new price of $1.80, which meant that producers avoided much of the incidence of the tax. And that's what we focused on last time, but also new quantity of $90 billion.

Now, we're going to turn our focus to quantity. Tax fairness is about prices. Tax efficiency is about quantities. What we see here is what we learned about in 14-01, that this tax causes a deadweight loss. The deadweight loss is the units that were efficiently traded before the tax that are no longer traded.

So before the tax, the 99 billionth gallon of gas was something where the consumer surplus from consuming that gallon exceeded the consumer value of consuming that unit, exceeded the marginal cost of producing that unit. As a result, producing that unit delivered both consumer surplus and producer surplus. When we tax, and therefore reduce the number of gallons sold to $90 billion, we cause deadweight loss. We destroy productive trades.

Trades that are being made that were delivering surplus to society are now destroyed. That creates our classic deadweight loss triangle. Remember, the deadweight loss triangle, we talked about this in chapter 6, it points to the efficient outcome. So the efficient outcome is point A, the deadweight loss triangle points there and then spreads out from there. So the further you get from the efficient outcome, the bigger the deadweight loss. Yeah, Alec.

AUDIENCE: If the demand curve was completely vertical, would there be--

JON GRUBER: Hold on. Let me-- hold on, hold on, hold on there. We'll get there. Come on, you got to leave me a little mystery here. OK. So basically, this is our fundamental inefficiency caused by taxation. OK. Question about the basic analysis? This should be pretty familiar. You saw this in 14-01. We know about deadweight loss, but the key thing is to remember the intuition of deadweight loss. It is trades that deliver social surplus that are destroyed by these new taxation.

And this is our first hint about why we can't just spend all that money in the first 2/3 of the book without thinking about how you raise the money. Because when you raise the money, you create an efficiency, and that's going to matter. This is the other leak in Okun's bucket. Remember, we talked about Okun's leaky bucket, and we talked about three sources of leakage. One was the person who has to carry the bucket. One was distorted labor supply among the people getting the transfers. This is the third source of leakage, a distortion in economic activity by raising the money.

Now, as Alec's question pointed out, let's go to the next figure. You can see that deadweight loss is fundamentally determined by elasticities. Remember, incidence was determined by elasticities. So is deadweight loss. When demand is very inelastic, as Alec pointed out, there's a very small deadweight loss. Why? Alec, why? Why is deadweight loss so small when the demand is very inelastic?

AUDIENCE: Because less trades get destroyed.

JON GRUBER: Yeah, basically it's perfectly inelastic. There's no inefficiency because nothings happened. When you tax insulin, there's no inefficiency. Everybody just keeps buying insulin. But when you tax McDonald's but not Burger King, there's massive inefficiency because everyone just switches over. It's perfectly elastic. Everyone switches over. Yeah.

AUDIENCE: Isn't consumer surplus still lost though, due to the higher prices?

JON GRUBER: Consumer surplus is lost through the higher prices, absolutely. Well, the thing is what's tricky about that is with perfectly inelastic demand, consumer surplus is infinite technically. So infinity minus x is still infinity. So yes, it's small-- it's not really smaller because in its completely made up world of perfectly inelastic demand, there's infinite consumer surplus. But it is true in general, if demand is very, very inelastic, you won't get a lot of distortion but you'll get a big loss in consumer surplus.

AUDIENCE: Does the consumer surplus count as deadweight loss?

JON GRUBER: You will get-- let me back up a second. No, I'm sorry, you won't. You're right. You're absolutely right. No, you're right. Let me retract that. You will not get a change in consumer surplus because consumer surplus, remember, is the area, basically, if the demand curve is vertical, then there's always infinite consumer surplus. It's a weird case. So it can't fall.

So basically, at the end of the day, you get some-- so I guess the point is that if you look at figure 20-2 a, to come to Alec's question, there's a tiny deadweight loss. Most of it is lost consumer surplus. This is, I guess, the better answer to your question, Alec. With varying elastic but not perfectly inelastic demand, most of the loss and surplus is consumer surplus. But it's not perfectly consumer surplus and it's not a lot because you didn't change your behavior that much.

When demand is very inelastic, very elastic, you see a much larger loss in surplus. And once again, if demand is perfectly elastic, you would see a loss in surplus, but it'll all be on producers. So it would be a measurable loss of surplus. You'd see a loss in producer surplus. Yeah.

AUDIENCE: I think what I meant was like, if the consumers were consuming past point B, or consuming either way, they lose surplus due to the higher prices though.

JON GRUBER: No, but they would-- no, no, no. No, because the point is that technically, consumer surplus is the area above the demand curve. Above the price, but below the demand curve. But there's no below the demand curve if it's perfectly elastic-- if perfectly inelastic.

AUDIENCE: I mean, when it's not perfectly elastic.

JON GRUBER: When it's not perfectly elastic, yes, they lose some. That's the top gray part.

AUDIENCE: Right, but what about past the-- for the people like from past point B. ?

JON GRUBER: Past which way?

AUDIENCE: To the left. Down to the original price.

JON GRUBER: Right, that's that upper gray triangle.

AUDIENCE: But now, isn't it only down to the second price?

JON GRUBER: I'm sorry. So basically, the price has gone up from P1 to P2.

AUDIENCE: Right.

JON GRUBER: Consumer surplus that has been lost is the area below the demand curve above the price. So it's the gray triangle above the P1 line. We can talk about it afterwards. So basically, essentially, elasticity is determinant. In fact, more generally, we can write the formula for deadweight loss. This is derived-- let me try this. Try to just wander over here.

This is derived in the appendix. We can write the formula for deadweight loss as deadweight loss-- at least we know my handwriting can't get any worse --is minus eta s eta d, over 2 times eta s minus eta d, times tau squared times Q over P. A very, very important formula you need to know cold. And let me explain why you need to know this cold. Because this formula is not only mathematically relevant but intuitively relevant. OK.

This formula shows that the deadweight loss of taxation depends on two things. Incidence, remember, depend on elasticities. They also depend on two things. First of all, depends on elasticities. Because this is multiplicative and this is additive, if either elasticity gets bigger, deadweight loss gets bigger. So basically, deadweight loss is larger, the larger either supply or demand elasticities. And that's intuitive from the graph.

But there's a second feature of this formula, which is very, very important, which is the tau squared term. Here's another way to think about this. Here's another way we could write this that might be helpful. Differentiate this with respect to tax. We get that the marginal deadweight loss from an increase in tax is minus eta s eta d over 2 times eta s minus eta d-- actually, this goes away --times tau times q over p.

The marginal deadweight loss depends on the tax. The higher is the initial level of tax, the more is the deadweight loss from additionally increasing tax. And that's why we call this the tau squared rule. The tau squared rule is taxes cause more inefficiency, A, if things are more elastic, or B, the further they take you from the efficient outcome. To see this, go to figure 20-5, and this will show you graphically. That's after the window tax figures. This will show you graphically why this-- this will give you the intuition.

We start at point A with P1 and Q1. We then impose imagine a first tax. Let's say we tax $0.10. OK. That causes deadweight loss ABC. Questions about that? Now, let's say I want to tax a second $0.10. That raises the deadweight loss from ABC to ADE. You add the trapezoid BCDE to your deadweight loss. You see that?

Notice BCDE is much larger than ABC even though both taxes are $0.10. Why? Why do you get this much larger trapezoid even though both taxes are the same size? What's the intuition? Come on, you guys know by now I won't bite. Let's go, Steven.

AUDIENCE: Not all trades produce the same amount of the surplus.

JON GRUBER: Yeah, in particular, which trades produce the most loss surplus?

AUDIENCE: The ones that are further left.

JON GRUBER: Yeah, the further away you get, the more loss surplus. Think about it. Think about the trade at point A. Literally, the trade right at point A, the 100 billionth gallon of gas. People are indifferent between the 99.99999 billionth. However around you have billionth of gallon of gas and the 100 billionth gallon of gas. At that price, they're roughly indifferent. So there's not a great sadness if that one last gallon doesn't get sold because that gallon was, one, delivered effectively only epsilon surplus.

But if you think about the 81st billion gallon of gas or whatever, if you think about one that's slightly to the right of Q3, that's delivering a lot of surplus. That's a gallon that people really want, and they're happy to pay way more than it cost to produce, or destroying it through this tax. So the basic intuition, the tau squared rule, is you're getting further from the equilibrium.

So you're destroying more and more valuable trades. And that's why tax, that's why you get this second effect of taxes. Not only do elasticities matter, but you have this tau squared effect, the base of taxation matters. OK. Questions about that? OK. This rule is going to become second nature for you because we're going to use it a lot. Let's start with three interesting implications of this rule.

The first implication is that, remember how we talked about how the incidence of taxes was affected by whether there was a minimum wage? The efficiency of taxes is also affected by how competitive the market is. Let's look at figure 20-6, and let's consider a tax in two markets. On the left hand side is a tax in a perfectly competitive market. In this perfectly competitive market we tax and we lower the quantity from Q1 to Q2, and we create a deadweight loss triangle ABC.

Now, imagine the market with a positive production externality. A market, or alternatively, could have been market power, either way. But imagine a market with a positive production externality. The implication of positive production externality is what? That we were already under producing. The optimum was Q1. That's where social marginal benefit equals social marginal cost.

Remember, this is a hint here, this is a example of how when we write the final, we're going to use all sorts of chapters at the same time. Here, I'm bringing back stuff back from chapter 5 and 6 into the tax chapter. So remember, if it's intuition, if you remind yourself the intuition about externalities. Remind yourself how this all relates.

Remember, in a market with a positive production externality, we are already under producing. We're producing Q2 instead of Q1. So when we add a tax to that market, it's extra inefficient. Because instead of adding a triangle, it's adding a trapezoid because they've already given up the triangle because we're already not producing enough. This is the flip side.

Remember in chapter 6. In chapter 5, we talked about corrective taxes. When there were negative externalities, a tax should get you to efficiency. Well, when there's positive externalities, the tax will get you farther from efficiency. So that's one kind of implication of the tau squared rule. What's another implication?

Well, here is the implication, which this is a hard one, hard one for we lefties. But another implication is that progressive taxes are more inefficient than flat taxes. Taxes where the marginal rate rises are more inefficient than flat taxes. I'm going to work through an example with you to show you why this is.

OK. So let's imagine that basically, there's a society with two individuals. I'm not going to write this on the board because it's too hard for me, but you just write this down. There's two individuals. One has a wage of $10 an hour. One has a wage of $20 an hour. So there's a high skill and a low skill individual. They both have an elasticity of labor supply of 1. For each of them, a 10% rise in the wage leads to 10% more labor.

And moreover, we're going to assume that they both face a elasticity of demand for labor of minus 1. They both work for employers where a 10% rise in wages leads to 10% drop in demand for labor. So both supply and demand of elasticities are one. In figure 20-7, we show the initial equilibrium in this market. On the left is the figure for the low wage worker. On the right is the figure for the high wage worker.

So the low wage worker and the high wage worker we're going to say are initially in equilibrium, working 1,000 hours a year. So the low wage worker's at point A, working 1,000 hours a year at a wage of $10. The high wage worker's at point D, on the right hand side, working 1,000 hours a year, $20 an hour. This is a complicated example. Please stop me at any point if it's not clear.

Now, we're going to roll in and levy a tax. We're going to levy a tax of 20% on all labor earnings. 20% tax, what does that do? Well, we know that lowers labor supply by 20%. We know the elasticity of labor supply is 1. So that's going to be a shift from S1 to S2. That is going to-- and the underlying math here, you can do basic elasticity. You'd be able to do this underlying math.

You can find that that lowers the hours worked from 1,000 hours to 894 hours. And that raises the equilibrium wage because there's fewer people working to $11.18. It creates a deadweight loss of ABC. So the 20% tax on low skilled workers has created a deadweight loss. Workers who are productively willing to work are no longer able to.

The same tax on high skilled workers does exactly the same thing. It creates a triangle, DEF. Why is DEF? The x-axis is the same. Change is the same. Why is DEF bigger than ABC? They both lower the labor supply by the same amount, 1,000 hours, 894. Why does that create a bigger deadweight loss for the $20 an hour worker? Why is the triangle DEF bigger than ABC? I mean, graphically, you can see it, but intuitively. Yeah.

AUDIENCE: They have a higher wage [INAUDIBLE].

JON GRUBER: Wage in a competitive labor market, your wages are marginal product. They're producing more valuable shit. That's why they're getting paid more. Your wage is a value marginal product. Therefore, you're being paid $20, you're producing something that's worth twice as much as someone being paid $10 an hour, or your labor is worth twice as much. Therefore, distorting it is twice as costly. So that's why. Yeah.

AUDIENCE: So this is saying, with same elasticity, the reduction on the labor you get, I will be the same, but the value of--

JON GRUBER: Exactly, the reduction in hours will be the same, but the hours are more valuable. Once again, really, I talked earlier today with someone about the bad economics of the dismal science. Sometimes we have to draw conclusions we don't like. One conclusion is if we're in a perfectly competitive labor market, the rich guy is producing stuff that's more-- the higher paid guy's producing stuff that's more valuable than the lower paid guy. You may not believe that, and that's because labor markets aren't perfectly competitive. But if they are, that's the implication.

Now, so if you look at the table below, if you look at the table, unfortunately, the table's-- yeah, I'm sorry. So if you look at-- so I'm now on figure 27. So look at the table on the next page, the top table. With no tax, so what we have is the tax rate, the low wage worker's hours and deadweight loss and the high wage worker's hours and deadweight loss.

You see that when you go from a zero tax rate to 20% tax rate, they each go to 894 hours. You get a deadweight loss for the rich, which is twice as big as the poor, and a total deadweight loss of $347.13. OK, questions about that?

Now, let's say I replace that with a system that in equilibrium, will be revenue neutral. But I want a progressive tax system. In other words, I want to not tax you below $10,000 a year, but tax people who earn more than $10,000 a year. I want it to be progressive. Well, in equilibrium, to raise the same amount of money, that implies a tax rate of 0, below 10,000, and 60, above 10,000.

The reason it's not just 40 is you could only tax the amount above 10,000. Even for the rich guy, you exclude the first 10,000. So it's not like the poor guy gets taxed fewer and the rich guy gets taxed 60. The poor guy gets taxed zero and the rich guy gets taxed 60 above the first 10,000. 0 on the first 10,000. It's a marginal tax rate. That's a progressive tax system.

What does that do? Well, let's go back to figure 20-7. For the poor guy, deadweight loss is 0. You've gotten rid of the tax. They're back to point A. For the rich guy, they lower their labor supply in response to this tax all the way to 837 hours of labor, creating a much bigger deadweight loss by adding the trapezoid EFIG.

So if you go back to the table, it's a 0 deadweight loss. The last row is 0 deadweight is the low wage worker, but now a $566 deadweight loss on the high wage worker, meaning the total social deadweight loss is higher. Total social deadweight loss is higher for two important reasons. First, the first is what we said before, it's mechanical. If you distort labor supply among a higher earning guy, you're costing society more because they're producing more.

But the second is because of this tau squared effect, that basically there's a trade off. The more you shrink the tax base while raising the same amount of revenues, the more you have to tax another group even more. And that's a bad deal, that the extra revenues by taxing the rich high are much more efficient. Why? Because you're already farther from that initial equilibrium. Question.

AUDIENCE: [INAUDIBLE] why do we have progressive--

JON GRUBER: Why do we have progressive taxes? Because we talked about in chap-- because progressive taxes are fair. So basically, if you think, essentially, remember, if you go back to our utilitarian social welfare function, what are we trading off? We're trading off the size of the pie and the distribution of the pie. If all you cared about was the size of the pie, you would only want proportional tax. But we don't. We care about the distribution of the pie as well. And we trade those things off, and that depends on our social welfare function. Yeah.

AUDIENCE: Would we create an efficiency gain to tax proportionally and just increase the overall level, and then redistribute from the taxes to the people who otherwise would have just had a lower tax rate and progression [INAUDIBLE]?

JON GRUBER: That is absolutely right. And we will come to that. We will come to the only time The Wall Street Journal has ever praised me.

AUDIENCE: [LAUGHS]

JON GRUBER: Keep that in mind. We'll come back to that. But basically, that's absolutely right. Basically, proportional taxes is going to be much fairer. Redistribution to progressive taxation imposes extra deadweight loss. And that imposes the trade off that we can evaluate with our social welfare function, which is we do more for the poor, but we create more inefficiency. And that's why Steven implied it may be better to not do redistribution through taxes, but through spending.

OK, so that's the second implication. There's a third implication, which is quite interesting, which is, governments should actually consider the time path of tax rates, not just the tax rate at a point in time, when deciding on how to finance spending, and do what we often call tax smoothing. So here's an example to make the point.

Imagine that a nation has to raise 20% of GDP in taxes to finance its everyday operations. And then there's a one year war, and that one year war will at one time cost 20% of GDP to fight. Should the government raise taxes to 40% of GDP for one year, then drop them back down to 20? Or should it raise taxes by, say, 1% for the next 20 years, or half a percent for 40 years, or by epsilon for infinity? And basically, do you guys ever call your parents except when you're in the car? My kid only calls me when he's in the car.

AUDIENCE: [LAUGHS]

JON GRUBER: I swear to God. Anyway, so basically, essentially, tax smoothing is saying the trade off between financing things right away and smoothing it over time, and we want to smooth it over time. Why? Because the deadweight loss goes up with the square of the tax rate. If I go up to 40 in one year, I'm causing huge deadweight loss that year. Whereas if I just have an epsilon up every year, I'm a small deadweight loss in every year and it will add up to less. So you want to smooth taxes over time.

Now, this actually is something which is quite interesting because many policies we'll talk about affect both tau squared and elasticities. There's a second reason why you wouldn't want to jack taxes up and take them down, which is that people might be more elastic to that. If you raise a tax rate half a percent forever, people might not respond. If you go to 40% before, you're back 20, people may then quit their jobs for a year or do something that's dramatic.

So oftentimes, policies can affect both the elasticity or could go the other way. It could go if you did for one year, nobody can quit. But forever, they'll take a lower paying job. It's not clear which way it goes. But the point is, often policies will affect both the tax squared and the elasticity terms. OK. So I find this an incredibly exciting and important intuition about how you think about elasticities and tax base sizes, and thinking about optimal and thinking about tax efficiency.

Now, in fact, just one side note all this presumes people understand the tax code. In fact, there's a really interesting literature showing how badly people understand the tax code. That in fact, most people make two common mistakes. First of all, first of all, people assume that the tax schedule is flatter than it is. Most people assume that rich people don't pay nearly as much more than poor people as they do.

Second of all, most people tend to assume their average rate is their marginal rate. They tend to just conflate the two. It turns out, in that world with those mistakes, that delivers lower deadweight loss from taxation. Basically, people are making mistakes since they think rates are lower than they really are. Deadweight loss of taxation is actually lower than it really is. So mistakes can maybe help government efficiency, is an interesting implication of that. Just sort of a side note.

But with that in mind, now let's go to policy. With this background in mind, let's go to policy. And let's talk about how many of you guys know who James Dean is? James Dean. Going once, going twice. Jesus, you're young. OK. James Dean-- we got one. James Dean was a very famous movie actor from the 50s, who was like the young Brad Pitt. Think Brad Pitt. Made two unbelievably important movies and then died in a car crash.

The James Dean of economists was a guy named Frank Ramsey. Frank Ramsey was an economist in England in the 1920s who wrote two of the most important articles in economic history and then died in a car crash at a very young age. And basically, one of those articles was the foundation of economic growth, the canonical growth model we use was based on Frank Ramsey's initial insights. The other is the model of how we think about optimal taxation.

And what Ramsey said was, he said, look, he asked the question very simply. He said, imagine a world where we raise taxes by taxing commodities. Let's forget income tax. This is 1920s. We didn't really have very good income taxes back then. Most taxes were raised by taxing commodities, tariffs, excise taxes, sales taxes. They were raised by taxing different goods, and often those goods were taxed at different rates. Some goods would be taxed at higher rates than other goods.

And Ramsey asked, well, what is the right way to set taxes across goods? Imagine if a commodity tax system, what is the optimal set of commodity tax rates? And the way he posed the question was quite good. He said, look, what is the way to minimize the inefficiency of taxation while raising the required number of required amount of revenues? And this is formalized in the appendix to the chapter. But Ramsey's question he posed as let's minimize the deadweight loss across all the goods we tax, subject to raising a given amount of revenues.

And he solved that problem. And what he got was Ramsey's, was what we call the Ramsey rule, which is that-- actually, I'll leave this here. He got the Ramsey rule, which is that we should set commodity taxes across goods i, to the point where the marginal deadweight loss of taxing good i of the next increment to taxation. So the deadweight loss of increasing tax by epsilon on good i over the marginal revenue raised by an epsilon tax in good i.

So if I increase tax in i, I raise some money, I cause some deadweight loss. That should be equal to a constant across all i that you should set tax across all goods such that this ratio is a constant. Not such that their tax are equal, but rather such that this relationship holds. Let's think about the intuition.

First of all, what is lambda? Lambda is a constant that represents the value of additional government revenues. Think about lambda being how much the government needs, the next dollar. If the government really needs the next dollar a lot, lambda is high. If it doesn't need the next dollar much, lambda is low. So it's just a level shifter that represents the marginal value of the next dollar of taxes to the government. That's what lambda is.

But the important thing is not what lambda represents, it's that this ratio equals a constant. What does that mean? So basically, let's recall that marginal deadweight loss is a positive function of the tax rate. Remember, marginal loss is a positive function of the tax rate. So the more you tax a good the higher the marginal deadweight loss is.

So basically, and marginal revenue, let's just assume marginal revenue is linear. Let's ignore the marginal revenue term for a second. So what this says is the more you tax a good, the higher this ratio becomes. So think about one intuition. If a good is already taxed very highly, another one's good tax is not taxed at all.

Then, what does this imply? Let's say you've got two goods. One has no tax on it, one has a very high tax on it. What does this imply about where you should put the next tax? Well, you want to set both these ratios. And let's say in both cases, you raise $1. You want to raise the next dollar of taxation. Yeah.

AUDIENCE: Is this tax taxing the gains on the commodity or sale?

JON GRUBER: The sale of the commodity. Not a commodity like, we're not doing commodity exchanges. You're in finance world. I mean, like an apple. An Apple. Buying an apple, buying a banana, buying a loaf of bread, buying a cigarette, buying a gallon of gas. Yeah.

AUDIENCE: [INAUDIBLE]

JON GRUBER: Because--

AUDIENCE: Because the marginal deadweight loss from that is lower than the commodity already taxed.

JON GRUBER: Right, right. So you want to put, on the one hand, there's that. You want to put the good on the tax on the good that's taxed less. You want to basically, that would say if that was all that existed, if that was the only issue, you'd want to basically do what with taxes? You want to set them equal. You'd want any tax that goes above, it's like the consumption smoothing argument.

Once one tax is higher-- remember with consumption smoothing, if you give $1 to a rich guy and take it away from a poor guy, that's a bad deal because the rich guy is a lower marginal utility of income. It's a similar thing here. Adding any dollar of taxes to which takes one above, it'll take one below. But going above causes more inefficiency because this tau term. So that's one argument. But there's another thing, which is why taxes shouldn't always be equal, which is what else goes into marginal deadweight loss? Yeah.

AUDIENCE: Elasticities.

JON GRUBER: Elasticities. And elasticities also turn optimal commodity taxes. So for example, what does the Ramsay tax imply? What does the Ramsay model imply about taxing luxuries versus taxing necessities? Yeah.

AUDIENCE: Tax them less.

JON GRUBER: Tax luxuries less than necessities, because necessities are more inelastic than luxuries. The Ramsay tax had some pretty nasty equity implications. It says basically things that are more elastic should get taxed less. So essentially, when you put these two things together, that's the tension that creates the optimal tax. You want to put taxes on inelastic goods, but not so much that the tau squared term comes to dominate.

Essentially, what the Ramsey rule does is trading off these two pieces, trading off this piece versus this piece. It's saying put tax on more elastic goods. But remember, the more tax you put on one good the higher the marginal deadweight loss will be. And at some point, it'd be efficient to take it off and put it on a less elastic good. If you put tax on the least elastic goods, I'm sorry, at some point, put it on more elastic good.

So Ramsey would say, if you've got a bunch of commodities lined up, let's say the world consists of apples and watches. Ramsey says, the first dollar of tax ought to go on apples because it's a lower elasticity of demand. Let's assume the elasticity applies the same. Lower elasticity of demand. You're therefore going to, essentially, create less deadweight loss.

But should the 100th dollar of taxes go on apples? Or maybe at that point, you want to keep apples on an i and do $1 on watches. Well, that, you might want to because once you have $199 of taxes, there's a big deadweight loss of that 100th tax, whereas the first incremental tax on watches doesn't cause much deadweight loss.

Another way to see this is to make a set of assumptions, and with a set of assumptions we lay out in the appendix. We generally summarize the Ramsay rule as saying that the optimal commodity tax is minus 1 over eta i times lambda. That that is the formula for the optimal commodity tax on good i. We call this the inverse elasticity rule. This is basically making some assumptions, like a perfectly elastic supply side, stuff like that.

And they're laid out in the appendix. But you really need to know those. This is just sort of a shorthand that you can generally use. Basically, it's a way of summarizing. It's saying, look, we want more taxes on inelastic goods, but don't put all the taxes on inelastic goods. You want to basically share the burden. So tax on inelastic goods should be higher than elastic goods, but shouldn't have all the tax on inelastic goods because the tau squared rule.

So it's basically the elasticity rule against-- another way to think about tau squared rule is the broad-based rule. That broad base of taxation reduce inefficiency because of the tau squared rule. And this goes back to Haig-Simons. This is why economists like Haig-Simons. Haig-Simons isn't just good for fairness, it's good for efficiency. When you tax everything, the rate can be lower.

So the idea of optimal tax theory is the broader the base, the lower the rate can be. And the lower the rate, the less damage from taxation because the tau squared rule. On the other hand, if the broad base means taxing a bunch of really elastic shit, you don't want to do that. And that's the trade off. Inherent in, optimal commodity taxation. Yeah.

AUDIENCE: Does the [INAUDIBLE] rule takes into account like externalities [INAUDIBLE]?

JON GRUBER: No, this is all perfect market stuff.

AUDIENCE: OK.

JON GRUBER: Yeah.

AUDIENCE: So you would, in this case, keep on increasing the tax on something until it reached the lambda level, and then you would go [INAUDIBLE].

JON GRUBER: You go to the next thing. I mean, you do it all simultaneously because, yeah.

AUDIENCE: Sure. In this case then, so in the economy where this is happening, your insulin, for example, super high tax. If you wanted to try and address that, then you would [INAUDIBLE]?

JON GRUBER: So then, basically, this leads to the equity question, which is as always, there's a trade off. How do economists model that? Well, they model a simple way. To every good they add an equity weight. So they say, look, yes, the Ramsay rule implies taxing inelastic things more, but poor people consume inelastic things. So let's let goods have literally a welfare weight, where the goods consumed by the poor are worth more to society.

Then, in that case, that would offset this. Not fully. Depends what welfare weights you put on it. But for example, if you said like, insulin, you could put an infinite welfare weight in insulin and then insulin would never be taxed because it'd be infinite inefficiency from taxing it. So you can always fuck around with this model by just changing the weights on the goods. Yeah.

AUDIENCE: Wouldn't that defeat the purpose?

JON GRUBER: Well, it wouldn't defeat the purpose. It depends what your purpose is. It's social welfare function. If putting an infinite weight in insulin means watches have to be taxed epsilon more, who cares? That would defeat the purpose. Yeah, it'd be a little more inefficiency, but guys wouldn't die. So maybe you're willing to make that trade off. Yeah.

AUDIENCE: Wouldn't the output [INAUDIBLE] you want to actually optimize lambda.

JON GRUBER: No, you want to reduce the deadweight loss in collecting that revenue. Or as they say, you want to pluck the goose while inducing the least squawking. So basically, the optimization function is minimized deadweight loss to raise revenue. Lambda is literally actually a Lagrange multiplier on the revenue constraint. If you do the formal math lambda is the Lagrange multiplier. That's why we use lambda.

But you don't need the formal math, you just need the intuition, which is that basically, lambda is the shifter, which tells you how you want your levels. But Ramsay isn't about that. Ramsay couldn't care about this. He's about this. Ramsay is about at a given level, how do I want to set the tax across goods?

Let me do an example of this, a very famous example, which comes to work by the Nobel Prize winning economist Angus Deaton. Angus Deaton was hired as a consultant by the country of Pakistan to help them figure out how to set their commodity taxes. In 1984, the Pakistani government had three goods it was focused on. It was focused on wheat, rice and oils and fats.

So let's go to figure 20-8. This is a complicated example, so I won't try to go through this slowly. Three markets, wheat, rice and oils and fats. The existing system was system which taxed oils and fats to subsidize wheat and rice. So look at table 20-1. You see that we have the three goods and what the existing government policy was, a 40% subsidy to wheat and rice and a 5% tax on oil and fat.

Deaton went in and estimated the elasticity of demands for these goods. And what did he found? He found that wheat was inelastic, but that rice and oil and fat were both fairly elastically demanded. So now, if you go back to your figure, figure 20-8, you can see that you have fairly inelastic demand curve for wheat, but a more elastic demand curve for rice and for oils and fats. OK. Questions so far?

Now, the initial government equilibrium, the government had put in policies such that we were at point C in the market for wheat, because of the 40% subsidy. We're at point C in the market for rice, because there was a 40% subsidy. And at point B, in the market for oils and fats, because there was a 5% tax.

Therefore, we were getting-- remember, subsidies create deadweight loss too. Anything which deviates from the market equilibrium creates deadweight loss. Let's ignore all the pre-existing externalities in these markets. So this is creating these three deadweight losses.

And they asked Deaton, how should we change our tax system? One nice thing about that question, it is way easier to figure out how to change a system than how to design a system. Changing a system means I can just look at this, and Deaton looked at this and said, well, wait a second, you've got a situation. The subsidy that basically you are taxing oil and fat to pay for these subsidies.

Well, imagine if you reduce the subsidy on wheat some. That's not going to cause that much since wheat is fairly inelastically demanded, if you reduce the subsidy on wheat. OK. If you reduce the subsidy of wheat, you don't get much social benefit. But if you reduce the subsidy on rice, you get a lot, because that's elastically demanded.

And at the same time, if you reduce the tax on oil and fats, you also get a lot of benefit because that's elastically demanded. So he said, basically what you ought to do is you ought to really reduce the tax on rice and reduce the subsidies-- reduce the tax on fat and reduce the subsidies to rice and also to wheat. But that isn't as big a deal. The welfare gain is not as big.

The real action is, why are we taxing this very elastic thing to subsidize this other very elastic thing? And he recommended that on the margin, they could reduce the tax on oil and reduce the subsidy to rice. And that overall, Pakistan will be better off. The Pakistani government said that's great, economist. But here's the problem, we also care about equity. What if people consume these goods?

Well, Deaton said, fine, I will go and measure the income distribution of people who consume these things. And what he found was that the people who consumed the wheat were the poorest, whereas the people who consume rice were higher income. Yeah, so basically, the bottom line is, wheat and oil and fat were consumed the most by lower income people, and rice more by middle income people.

So what he said is, well, when you put distributional concerns in, what that says is that further says we shouldn't reduce the subsidy to wheat, because that's poor people, but we should reduce even further the subsidy to rice. That's not the poorest people. And we should reduce the tax further on oil and fat, because that's poor people. So that's how you put equity considerations into the model.

This is not-- I don't need you to remember this example. I want you to use this as an illustration of how these abstract theory can be put in place in the real world. He can literally help the people of Pakistan through this simple Ramsey intuition by actually getting the data, estimating these demand curves and income distributions, and advising the government on what to do. Yeah.

AUDIENCE: I think I'm getting a little confused. I thought if we should tax things that are more elastic.

JON GRUBER: No. OK. I'm talking taxes and subsidies. So we should, if there's multiple goods we're taxing, we should tax things that are more inelastic.

AUDIENCE: Oh.

JON GRUBER: But this is a different example. Now, we're not taxing multiple goods, we're taxing one and subsidizing two. So it's a different example.

AUDIENCE: In curve C, the oils and gas, since it won't be elastic, that's why it's lower than your cost?

JON GRUBER: Well, no, it's just a small tax is why it's a small deadweight loss.

AUDIENCE: Oh.

JON GRUBER: It's a 40% subsidy in the first 2 and 5% tax.

AUDIENCE: The elasticity.

JON GRUBER: Yeah. All right? Yeah.

AUDIENCE: How much objectivity if you're in determining equity away from [INAUDIBLE]?

JON GRUBER: Oh, there's no objectivity. Zero. I mean, basically, it's all-- that's why the economists like focusing on efficiency. Efficiency is, in some sense, is a little more judgment free units. Not necessarily, because I have to make judgments on how perfect the market is. So there's still judgment involved in efficiency calculations, but a lot of equity is all judgment.

OK? All right. So that's an example of how you apply this method. Now, of course in the world today, most revenue in the US especially, most revenue is not raised by commodity taxes. It's raised by income taxes. So how can we take the intuition from Ramsey and apply it to income taxes? Well, let's consider the following example. Imagine the following assumptions.

First of all, everyone in society is identical. Assumption one, everyone in society is identical. Assumption two, there's diminishing marginal utility of income, which we always assume. OK. Assumption three, there's a utilitarian social welfare function. So we just care about the sum of utilities. And assumption four, is that income is not elastic, that people just sort of, what they earn is fixed. It doesn't depend on things like taxes.

Identical individuals, diminishing marginal utility of income, utilitarian social welfare function, fixed income. What's the optimal tax system? You can tell me, based on those four facts, you can tell me what the optimal tax system is in this society. Yeah.

AUDIENCE: Anyone who makes above average should get taxed to the average.

JON GRUBER: 100% tax rate above some point. Basically say, look, everybody gets X. If you're above that, you get 100%. If you're below that, we're going to bring you up to that level. Basically, equal distribution of income. Why? Why should there be equal distribution of income? Yeah.

AUDIENCE: That's the same pair of benefits.

JON GRUBER: Yeah, because basically, if you and Alec are identical, I give Alec another dollar, it makes him less happy that it makes you sad, because diminishing marginal utility of income. So the key point is a natural starting point, if you're a utilitarian, first, being a utilitarian is not a crazy left wing-- utilitarian is just you're saying, we care about everybody equally. That's not exactly left wing.

If you start with utilitarian social welfare function, you actually get that a natural starting point is a perfect distribution of income in society, which is about the leftiest outcome you can imagine. So what's wrong with this, what's wrong with it is the fourth assumption, which that incomes are not fixed. Then we tax individuals, we will create a leak in the bucket.

But let's remember, there's also the issue that, of course, individuals aren't identical. So there's two things that are wrong with this. One, individuals aren't identical. OK. Remember, economics expresses preferences. I didn't just start with James Dean. How about Scrooge McDuck? Anybody know who Scrooge McDuck is? Geez, you guys. OK, we got four at least. Four times five times many people.

Scrooge McDuck is this Disney character who loves money so much he has a swimming pool full of money he swims in. Scrooge McDuck might not actually exhibit diminishing marginal utility of consumption. He just likes more and more and more. We should respect that as economists because remember, utility social welfare function doesn't judge the structure of utility. It just adds it up.

So in that case, it might be efficient if Scrooge McDuck have a lot more money, because he actually loves money. So the bottom line is, our starting point's in equal distribution of income society. But if people are different, i.e. Different marginal utilities of consumption, and if income is endogenous, i.e. Taxes affect how much resources there are in a society, then we don't want equal distribution of income.

So now we want to say, how do we think about setting tax rates in a world where people, the amount they earn depends on the tax rate? Well, this leads to the famous Laffer curve. Arthur Laffer was a right wing economist who was very influential in the tax-- we talked about the tax revolution in America being kicked off by Proposition 13 in California, which capped property taxes.

It was a big factor in electing Ronald Reagan. And Ronald Reagan's big was, implemented some of the largest tax cuts in history. And his intellectual case was based on the work of Arthur Laffer. Laffer drew a famous Laffer curve, which was theoretically, absolutely right. What did the Laffer curve look like? The Laffer curve-- I don't have to draw it because it's already drawn for you --looked like figure 20-9.

What did Arthur Laffer point out? He pointed out, he made three observations. First, he said a tax rate of 0. So on the x-axis is tax rates, on the y-axis is revenues. He made three observations. First of all, at a tax rate of 0, you raise no revenues. Unimpeachably true. Second of all, for some positive tax rate, you do raise revenues. Unimpeachably true. And third, almost certainly true, but not unimpeachable, at 100% tax rate, you also raise no revenues.

Because if leisure is a normal good, no one would work if you couldn't keep any income. Maybe not true. There's nerds like us who'd work anyway, but it's not a crazy assumption. Given those three assumptions, you, by necessity, get a figure of the type of figure 20-8. It can be shaped very differently, but it hits the x-axis at 0 and 100%, and is positive in between. That is what we call the Laffer curve.

And Laffer argued that in a world like this, there is a revenue maximizing tax rate where basically, once you get higher than that tax rate, you actually, each additional increase in taxes lowers revenue. You might say, how can an increase in tax lower revenue? Well, it's graphically true here. There must be a point. If you accept my three premises, there must be a point at which raising the tax rate actually lowers taxes.

And that's because the tax base. The amount of tax revenues you get is what? It's the tax rate times the tax base. But the tax base is a function of the tax rate because people change how hard they work and what they earn when taxes go up.

So dR, d tau, equals, dR d tau equals B plus tau times Db d tau, which is negative. So when you raise tax rates, there's two offsetting effects. You first of all get money because you have an existing base, but you then lose money because the base shrinks.

The peak of the Laffer curve is the point at which this is 0. That is 0 at the peak of the Laffer curve. And to the right of that peak, you lose money by raising taxes. The second term dominates the first. So this is the underlying economics of the Laffer curve.

And this says that there is an optimal tax rate. The optimal tax rate is not 100%, as we implied in the previous example, in my simple example. The optimal tax rate, once again, we solved this in the appendix, but basically, the optimal tax rate is to say, how do we maximize social welfare while raising a given amount of revenue? It's more complicated. I'll throw in commodity taxes, whereas, how do I minimize deadweight loss while raising a given amount of revenue?

Optimal income taxes say, how do I maximize social welfare while raising a given amount of revenue? And the resulting optimal income tax formula is of the form-- the optimal income tax formula is of the form that you want to set income taxes such that the margin utility of individual i above the marginal revenue raised from individual i equals lambda. So it's a little bit different formula.

You raise taxes until the margin utility of income for the next dollar of revenues is equal across all people. This is sort of hard to understand. So basically, but it reflects two trade offs, a trade off two different factors. The first factor is that basically, essentially, the first factor is vertical equity that basically says you want to tax the rich more because they have a lower marginal utility of income, marginal utility of consumption.

The second factor, however, again, is behavioral responses, which is as you tax the rich more, they may respond by earning less income, and that will mean that the marginal revenue raised from them is lower. You lower marginal revenue. This is a hard intuition. So let's see it in an example. Go to figure 20-10.

We've got two people in this example OK. Dr. H, who has a high income, and Dr. L, who is a low income. I don't know why that's a doctor. Can you make a note, Valerie? I don't know why they made that a doctor. It should be Mr. L and Dr. H. Or Mrs. L and Dr. H, whatever.

OK. So basically, what you have here on the x-axis, you have the tax rate. On the y-axis, you have this ratio I talked about, which is marginal utility or marginal revenue. So the curve for Dr. L is higher than Dr. H because they have a higher marginal utility, because they're poorer.

So what that says is that we'd like to tax H more than L. The problem is that when we tax H too much, he eventually gets so poor that you want to stop taxing him. So this says that the optimal tax rate, where lambda equals that ratio, is going to be a 10% tax rate on the poor and a 20% tax rate on the rich. It's not going to be tax the rich-- to put everything on the rich.

Put everything on the rich, then, first of all, you'll lower the marginal utility so much they'll be poorer than the poor. And second of all, you won't make as much revenue because they're elastic. And so that's basically how we think about optimal income taxes. The bottom line is a trade off. Just like you want to tax the rich more, but you want to respect the fact that you don't want to shrink the size of the pie too much. And that's inherently the equity efficiency trade off. OK. Questions about that?

So the bottom line is, of all this, is essentially, we're thinking about this equity efficiency trade off. And we can implement it by thinking about what's the fairness benefit, in terms of targeting people with low marginal consumption, versus the efficiency cost, in terms of distorting people's earnings. And we'll come back to this in chapter 25 and talk about what this all implies, what the income tax schedule should look like.

Now, I want to cover one other topic before we stop, which is related to this and is near and dear to my heart, which is thinking about tax benefit linkages and the financing of social insurance programs. And this goes back to the article which most influenced my career, which is an article by the economist Larry Summers, who was my thesis advisor.

Many of you know him from other runs of life, being the Secretary of the Treasury, President of Harvard. The guy who said inflation is going to go through the roof. All the things we know about Larry Summers. But he's also a very famous academic. And he wrote an article which, by his own admission, he did not think was a big deal when he wrote it, but ended up being an incredibly influential article.

And Summers' article made basically the following point. Let's go to figure 20-11. At the time Summers wrote this article, a very hot topic in public policy was the topic of employer mandates. The idea was if we want to get-- if we can deliver social benefits by making employers deliver them. It was an idea that came out of the left. It was a liberal idea.

And this was the fundamental idea behind, for example, ideas that people like Bill Clinton were proposing for universal health care reform in the 1990s, was, let's get universal health care reform by making employers pay for it. OK. Let's make, in particular, let's mandate that employers provide health insurance to all their employees.

Now, people said that would be a disaster. It will cause mass layoffs. Why did they say that? Well, they said that because of figures like figure 20-11 a. They said, look, if you tax employers, that will raise the cost of labor now, if you tax employers. If you make employer provide health insurance, take a benefit no employers will provide and you make it easy. Or suddenly make all employers provide, I don't know, nose jobs, I don't know, whatever. Something they weren't providing.

If you suddenly make them all provide that, that's going to add a new cost. Workers will now come with a new cost that will lower the demand for labor, once again, ignoring labor market imperfections. Lower the demand for labor, reduce the amount of workers from L1 to L2, and create a huge deadweight loss. And therefore, we shouldn't mandate employers to provide benefits.

Summers said, actually, that's wrong. Because if we mandate employers to provide benefits, something else happens. We make work more valuable. Why? Because now you don't just get a wage, you get this benefit and you can only get it by working. So people are now going to want to work.

It's like, OK, so imagine the benefit we mandate is that everyone gets a vacation to Disney. You're going to mandate all employer provide a vacation to Disney. OK. Well, employers will have to pay more, but now people are like, wow, if I go work-- I'm not working but if I go work, I get a vacation to Disney. I might want to go work now. That's an extra benefit that'll entice me to work. That will shift out my labor supply curve.

That is like panel B. So now, the demand curve has shifted down but supply has shifted out. That has two effects. First of all, the reduction in employment is smaller. Instead of level of employment falling from L1 to L2, it now only falls to L3. That leads to a lower deadweight loss. The deadweight loss is lower. It also leads wages to fall more. Wages that are falling to W2 fall to W3.

Indeed, to really drive home the intuition, imagine a world where every employee valued the benefit of a trip to Disney at exactly what it costs the employer to provide that. Everybody in the world valued the trip to Disney at exactly its cost. What would that imply? That would imply figure 20-11.

Where the reduction in demand from providing the trip to Disney, from the cost of providing the trip to Disney, was the same as the increase in supply from people working because they like the trip to Disney. What happens in that case? There is no change in labor. Why? Why is there no change in labor? What's happened? What have you done with this new rule? If the government mandates every-- yeah. What have you done?

AUDIENCE: [INAUDIBLE] the compensation is the same. They just replaced it from wages to wages plus Disney.

JON GRUBER: Exactly, you haven't changed compensation at all. You've simply shifted the nature of compensation. You see wages have fallen by the entire amount of the cost of their trip to Disney, or what we call full shifting to wages. Then the world where employees and employees value benefit the most, in fact, there might be no deadweight loss for mandating benefits.

Not only might the deadweight loss be smaller, like in figure 20-11, it might be 0, like in figure or 20-12. You actually get no deadweight loss from-- yeah --from any benefits. Alec.

AUDIENCE: So if there are benefits from something like health care, why does private health not fall in this?

JON GRUBER: The University of Chicago question would be, and some is addressed figure 20-12, like, well, this is stupid. Because if this was really true, why didn't employers be doing it already? If I like Disney, and I like it as much as it costs them to provide it, why wouldn't every employer provide Disney? To which Summers had a very good response, and he didn't use Disney as an example.

He said, imagine this example was-- imagine this is the 1980s, and health insurance excludes AIDS coverage and you're going to add benefits for AIDS. And let's say that the average employee, it's very cheap because AIDS is rare. So let's say the average employee values that as exactly its cost, or maybe more. Why would employers still not do it? Yeah.

AUDIENCE: PR backlash?

JON GRUBER: No, I mean, forget the PR. I mean, today's world, that would be good, right? Why do employers not do it? Yeah.

AUDIENCE: Adverse selection.

JON GRUBER: Adverse selection. All the workers who know they're at risk for AIDS come work for you. And it doesn't cost you the average, it costs you the full cost. So no employer wants to go first and be the employer that covers AIDS, because then they get all the workers that have it. Same with worker's compensation. Why would you not want to be the one employer in 1904 that offered worker's compensation? Because all the clumsy workers are going to come work for you and you're going to go bankrupt.

Any benefit, even if it's valued fully or valued more than fully, might not be offered because no employer wants to move first. In that world, a mandate can actually be efficient. Indeed, if people value the benefit more than it's cost, a mandate could be efficiency increasing. Not only not efficiency reducing, efficiency increasing. It's a very powerful insight, a simple insight. His articles even called some simple economics and mandated benefits, but a very powerful insight.

OK. So the bottom line is that, and it comes back to Tiebout. Why? Because what was Tiebout about? He was about tax benefit linkages. Why did communities provide the efficient level of public goods in the Tiebout model? Because the Tiebout model, the taxes you paid exactly were the benefits you got.

Likewise here, something that people value, the cost of employer's efficient to provide, because they'll just pay for it in lower wages. There's a full tax benefit linkage. OK. Questions about that argument? Now, this argument raises some interesting issues. One issue is what Alec raised, which is why don't employers do it already? The other issue is, well, what's going to determine how big these tax benefit linkages are?

Well, let's go back to our Disney example. In fact, Disney is kind of fun for adults, but let's be honest, it's more fun for kids. So there's going to be a lot of difference in how employees will value that Disney benefit. Likewise, there'll be a lot of difference how people would value the AIDS benefit. If I know I'm in a heterosexual relationship and I'm never going to get a blood transfusion, I might not value AIDS coverage.

So basically, as a result, there's going to be differences across people in this tax benefit linkage, and that's going to complicate the model. His model is written for the average tax benefit linkage, but heterogeneity will make the model more complicated. OK. Here's another issue. Actually, Summers, I'm going to talk about Summers. There's a big event for him on Friday. I'll be speaking about this paper at that event. I won't mention this point. But actually, there's a mistake in the title, because this actually has nothing to do with mandated benefits.

Take unemployment insurance. There's no mandate for unemployment insurance. You're just taxed. Employers are taxed for all their employees' unemployment insurance. The analysis is the same. Employers pay a tax, employees get benefits. So because the employee unemployment insurance, I'm more willing to work because if I get laid off, I'm covered. So really, this article should be called some simple economics of benefits that are attached to work. Let me flip it around.

Actually, the reason this model doesn't really apply to Clinton health care reform is Clinton wanted to tax employers and use the money to provide insurance to everyone. Well, what happens if you provide insurance to everyone? Then you don't get the labor supply shift, even though it's a mandated benefit. If you're giving to non-workers too, there's no labor supply shift.

The key trick that makes this work is that you're restricting the benefit to the people who are paying the taxes. That's what delivers the tax benefit linkage. If the benefit is, so anything. It doesn't have to be a mandate. Anything which limits the employer cost, limits the benefit to those who actually have to go work for the employer, is going to have the Summers offsetting effect. OK. Yeah.

AUDIENCE: Insurance benefits is one of which this mandate exists, or one in which enough businesses have made the first move?

JON GRUBER: Well, no. I mean, basically, we know this mandate exists. So take workers compensation. Perfect example. This mandate exists for workers' compensation. We have mandated workers' comp coverage.

AUDIENCE: I guess I meant more for health care.

JON GRUBER: Health care, there was a small mandate in the original Affordable Care Act. There's a mandate in Massachusetts. Some states have mandates, but basically, it doesn't really exist for health insurance. But it does exist, for example, if you offer health insurance and you're one of the kinds that's regulated by the state, which is plans above x employees. X varies. You have to include certain features in your health insurance.

Now, the ACA made that national. But before the ACA, for example, in Massachusetts, you had to cover IVF. In other states, you didn't. So that was a mandate. But we didn't mandate people who had insurance. We mandated what you have to cover if you had it. Worker's comp is the best example. And indeed, my first ever published article was looking at, well, what happens if you mandate worker's comp? What is the extent of the shifting?

And I found, and worked with the late, great economist Alan Krueger, I found that, in fact, workers' comp was almost fully shifted to wages. That is when as the cost of workers' comp went up. Now, I couldn't study the introduction of workers' comp. Later articles did that and found the same thing. I was looking at a world with workers' comp. What happens when worker's comp gets more expensive or less expensive? What we found was when workers' comp gets more expensive, it lowers wages and doesn't lower employment, which is consistent with figure 20-12.

I then went on, and in my thesis asked the question, well, wait a second, what happens if it's a benefit that doesn't affect everyone, like maternity leave? What if we mandate maternity leave? Which most countries in the world mandate paid maternity leave. We mandate unpaid maternity leave. In some states, we actually mandate paid maternity leave, like Massachusetts. But in most states, we don't.

Well, that's complicated, because that's a benefit only to some workers. So if you can perfectly adjust every worker's wage, then you'd see the same model operate, but just for the workers who benefit. But if it turns out it's hard to set different wages for different workers because they'll be pissed or discrimination, then it'll be hard to do it. So I actually studied an interesting feature of the system, which is when I was born and until the mid 1970s, the way health insurance was set up is it said if you break your arm or have a burst appendix or have cancer, we'll pay your medical bills. But if you have a baby, here's 200 bucks.

The idea was that childbirth was not a medical event, it was a choice. So insurance did not pay for the cost of childbirth. This was then viewed as discriminatory and a number of states started passing laws, mandates, mandating insurance coverage childbirth. Well, this was a major increase in the cost of health insurance. So I looked at what happened to the wages in the states that mandated this maternity coverage.

Then, in 1977, the federal government mandated for the whole country. So we ran the reverse of the experiment. So ever since, some states that added it and then the rest states had to add it. And I looked at both and I looked at what happened to the wages of, for example, married women relative to say, men or very older women who weren't of childbearing age.

And what I found there was actually group specific shifting, that is, the relative wages of married women and actually, to some extent of their husbands, since their husbands provide the health insurance, actually went down relative to the wages of groups like young, single women or older women or unmarried men, which is consistent with the ability of firms to actually shift those costs even within the firm. It's consistent with this kind of model working within firms. Now, does it work perfectly? No, but the idea is that this kind of model actually works across firms and within firms as well. Yeah.

AUDIENCE: So does the labor supply for the [INAUDIBLE] I mean, for the women change?

JON GRUBER: Well, I mean, that's what's implied, that it did. The ultimate, you can only look at equilibrium outcome. Labor didn't change.

AUDIENCE: OK.

JON GRUBER: So what this says, is this good news or bad news? Well, actually, it's mixed news. Here's the good news. My evidence implies we can mandate paid maternity leave and women will not lose their jobs. That's the good news. If you want to extrapolate from my now 30-year-old paper to today. The bad news is you're not helping women. You're just redistributing the way they're compensated.

So if you think I'm going to mandate maternity leave because I want to help women, well, in fact, you're not. You're just giving them maternity leave instead of wages. But the good news, at least they won't lose their job. So in some sense, it's an either/or. You don't get the efficiency loss, but you also don't get the equity gain in some sense, of helping people out because they pay for it themselves out of their wages. OK. Yeah.

AUDIENCE: What were ways, if you can like, vertically integrate it in more ways that they would rather have indirect [INAUDIBLE]?

JON GRUBER: Yeah, great question. I couldn't investigate it in my paper. Maybe they did it through slower promotions. The other thing, this is a very high inflation environment. So they could have done it by just not giving them raises as quickly as they did other groups. So there are ways they could adjust, but I can't really get at that. It's a good question. All right. Yeah.

AUDIENCE: With-- sorry. Wouldn't fight not being laid off? Like, even with the lower wages, isn't because people value, sort of--

JON GRUBER: Right, no. But I mean, if you look at figure 20-12, they're indifferent. Maybe they're even better off because maybe they valued it more than it cost the employer. But it's not like suddenly we're delivering this massive benefit to women for free. They're paying for it in lower wages. All right.

AUDIENCE: Sure, but it's still in between what is not [INAUDIBLE] work out.

JON GRUBER: Well, I mean, basically, it's like the minimum wage argument. How do we feel about giving more wages to some low wage workers, but other people lose their jobs? So that's, the equity is harder, because it depends on that, but the efficiency is clear. It's much more efficient to not have people lose their jobs. Yeah.

AUDIENCE: [INAUDIBLE] the benefit have a floor for [INAUDIBLE].

JON GRUBER: Yeah, I mean, that's another thing. People look at the interaction with the floor and that suggests that there, you couldn't shift it. And that would cause some job losses, but then women's wages wouldn't fall as much.

So the bottom line is the efficiency is clear. The more you shift to wages, the less inefficiency. The equity is a bit hard because on the one hand, you reduce job losses, on the other hand, you lower wages.

The point is not that I have a bottom line equity. The point is it's not as good as it seems. You say women are losing their jobs, and that's wonderful for women. Not necessarily. It might not be losing their job because their wages are lower. OK, so let me stop there.